New England Common Assessment Program 2010–11 Technical Report

July 2011

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

TABLE OF CONTENTS

CHAPTER 1. OVERVIEW 1
1.1 Purpose of the New England Common Assessment Program 1
1.2 Purpose of This Report 1
1.3 Organization of This Report 2
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 5
2.2.5 Depth of Knowledge 7
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 11
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 18
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
3.3.1 Large Print and Braille 22
3.4 Administrator Training 22
3.5 Documentation of Accommodations 23
3.6 Test Security 23
3.7 Test and Administration Irregularities 24
3.8 Test Administration Window 25
3.9 NECAP Service Center 25
CHAPTER 4. SCORING 26
4.1 Scoring of Standard Test Items 26
4.1.1 Machine-Scored Items 26
4.1.2 Person-Scored Items 26
4.1.2.1 Scoring Location and Staff 27
4.1.2.2 Benchmarking Meetings with the NECAP State Specialists 28
4.1.2.3 Reader Recruitment and Qualifications 29
4.1.2.4 Methodology for Scoring Polytomous Items 29
4.1.2.5 Reader Training 30
4.1.2.6 Senior Quality Assurance Coordinator and Senior Reader Training 32
4.1.2.7 Monitoring of Scoring Quality Control and Consistency 32
CHAPTER 5. CLASSICAL ITEM ANALYSIS 37
5.1 Classical Difficulty and Discrimination Indices 37

5.2 Differential Item Functioning *40*
5.3 Dimensionality Analysis *41*
CHAPTER 6. IRT SCALING AND EQUATING 45
6.1 Item Response Theory *47*
6.2 Item Response Theory Results *49*
6.3 Equating *50*
6.4 Equating Results *51*
6.5 Achievement Standards *52*
6.6 Reported Scaled Scores *53*
CHAPTER 7. RELIABILITY 56
7.1 Reliability and Standard Errors of Measurement *57*
7.2 2010–11 Subgroup Reliability *58*
7.3 Reporting Subcategory Reliability *58*
7.4 Interrater Consistency *58*
7.5 Reliability of Achievement Level Categorization *60*
7.5.1 Accuracy and Consistency Results 61
CHAPTER 8. SCORE REPORTING 63
8.1 Teaching Year versus Testing Year Reporting *63*
8.2 Primary Reporting Deliverables *63*
8.3 Student Report *64*
8.4 Item Analysis Reports *65*
8.5 School and District Results Reports *66*
8.6 School and District Summary Reports *70*
8.7 School and District Student-Level Data Files *70*
8.8 Analysis & Reporting System *71*
8.8.1 Interactive Reports 71
8.8.1.1 Item Analysis Report 72
8.8.1.2 Achievement Level Summary 72
8.8.1.3 Released Items Summary Data 72
8.8.1.4 Longitudinal Data 72
8.8.2 User Accounts 73
8.9 Decision Rules *73*
8.10 Quality Assurance *73*
CHAPTER 9. VALIDITY 75
9.1 Questionnaire Data *76*
9.1.1 Difficulty of Assessment 76
9.1.1.1 Difficulty: Reading 76
9.1.1.2 Difficulty: Mathematics 78
9.1.1.3 Difficulty: Writing 79
9.1.2 Content 79
9.1.1.4 Content: Reading 79
9.1.1.5 Content: Mathematics 81
9.1.1.6 Content: Writing 83
9.1.3 Homework 84
9.1.1.7 Homework: Reading 85
9.1.1.8 Homework: Mathematics 86
9.1.4 Performance in Courses 88

REFERENCES 91

APPENDICES 93

Appendix A *Committee Membership*
Appendix B *Participation Rates*
Appendix C *Accommodation Frequencies by Content Area*
Appendix D *Table of Standard Accommodations*
Appendix E *Nimble Accommodations*
Appendix F *Item-Level Classical Statistics*
Appendix G *Item-Level Score Point Distributions*
Appendix H *Differential Item Functioning Results*
Appendix I *Item Response Theory Calibration Results*
Appendix J *TCC and TIF Plots*
Appendix K *Delta Analyses and Rescore Analyses*
Appendix L *a-Plots and b-Plots*

Appendix M *2010-11 NECAP Standard Setting Report*
Appendix N *Performance Level Score Distributions*
Appendix O *Raw to Scaled Score Look-up Tables*
Appendix P *Scaled Score Distributions*
Appendix Q *Classical Reliabilities*
Appendix R *Interrater Agreement*
Appendix S *Decision Accuracy and Consistency Results*
Appendix T *Sample Reports*
Appendix U *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 as follows: (1) provide data on student achievement in reading/language arts and mathematics to meet the requirements of NCLB; (2) provide information to support program evaluation and improvement; and (3) provide information regarding student and school performance to both parents and the public. The tests are constructed to meet rigorous technical criteria, to include universal design elements and accommodations to allow all students access to test content, and to gather reliable student demographic information for accurate reporting. School improvement is supported by

- providing a transparent test design through the elementary and middle school grade level expectations (GLEs), the high school grade span expectations (GSEs), distributions of emphasis, and practice tests;
- reporting results by GLE/GSE subtopics, released items, and subgroups; 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 2010–11 NECAP. In October 2010, students in grades 3 through 8 and 11 participated in the administration of the NECAP in reading and mathematics. Students in grades 5, 8, and 11 also participated in writing. This report provides information about the technical quality of those tests, including a description of the processes used to develop, administer, and score the tests and to analyze the test results. This report is intended to serve as a guide for replicating and/or improving the procedures in subsequent years.

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

1.3 Organization of This Report

The organization of this report is based on the conceptual flow of a test's life span. The report begins with the initial test specification and addresses all the intermediate steps that lead to final score reporting. Chapters 2 through 4 provide a description of the NECAP test 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 test are developed specifically for those states participating in the NECAP and are directly linked to the NECAP Grade Level Expectations/Grade Span Expectations. 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 items are 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, to provide breadth of coverage of the GLEs/GSEs. Because they require approximately one minute for most students to answer, these items make efficient use of limited testing time and allow coverage of a wide range of knowledge and skills, including, for example, word identification and vocabulary skills.

Short-answer 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. Short-answer items require approximately two to five minutes for most students to answer. The advantage of this item type is that it requires students to demonstrate knowledge and skills by generating, rather than merely selecting, an answer.

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

A single common **writing prompt** and one additional matrix writing prompt per form were administered in grade 11. Students were given 45 minutes (plus additional time if necessary) to compose an extended response for the common prompt that was scored by two independent readers both on quality of the stylistic and rhetorical aspects of the writing and on the use of standard English conventions.

Approximately 25% of the common NECAP items were released to the public in 2009–10. The released NECAP items are posted on a Web site hosted by Measured Progress and on the Department of Education Web sites. Schools are encouraged to incorporate the use of released items in their instructional activities so that students will be familiar with the types of questions found on the NECAP test.

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. The number of test forms varies by content 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 test takers 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 all items on the NECAP test were designed to measure a specific GLE. The test framework for reading in grade 11 was based on the NECAP GSEs, and all items on the NECAP test were designed to measure a specific GSE.

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. Constructed-response items are more complex, requiring 5 to 10 minutes of response time. 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. NECAP 2010–11: Reading Item Types

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

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

2.2.3 Test Design

Table 2-2 summarizes the number and types of items that were used in the 2010–11 NECAP reading test for grades 3 through 8. Note that in reading, all students received the common items and one of either the equating or field-test forms. Each multiple-choice item was worth one point, and each constructed-response item was worth four points.

Table 2-2. 2010–11 NECAP: Item Type and Number of Items—Reading Grades 3–8

	Long passages	*Short passages*	*Stand-alone MC*	*Total MC*	*Total CR*
Common	2	2	4	28	6
Matrix—Equating					
Forms 1–3	1	1	2	14	3
Matrix—FT					
Forms 4–7	1	1	2	14	3
Forms 8–9	0	3	2	14	3
Total per Student					
Forms 1–7	3	3	6	42	9
Forms 8–9	2	5	6	42	9

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

Table 2-3 summarizes the numbers and types of items that were used in the 2010–11 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. 2010–11 NECAP: Item Type and Number of Items—Reading Grade 11

	Long passages	*Short passages*	*Stand-alone MC*	*Total MC*	*Total CR*
Common	2	2	4	28	6
Matrix—Equating					
Forms 1–2	1	1	2	14	3
Matrix—FT					
Forms 3–8	1	1	2	14	3
Total per Student	3	3	6	42	9

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

2.2.4 Blueprints

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

Table 2-4. 2010–11 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 category reporting structure for reading and the maximum possible number of raw score points that students could earn. (With the exception of word identification/vocabulary items, reading items were reported in two ways: type of text and level of comprehension.) Note: because only common items are counted toward students' scaled scores, only common items are reflected in this table.

Table 2-5. 2010–11 NECAP: Reporting Subcategories and Possible Raw Score Points by Grade—Reading Grades 3–8 and 11

Subcategory	*Grade tested*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Word ID/Vocabulary	20	18	9	10	10	10	10
Type of Text							
Literary	16	18	22	21	21	21	21
Informational	16	16	21	21	21	21	21
Level of Comprehension							
Initial Understanding	18	20	22	20	19	18	16
Analysis and Interpretation	14	14	21	22	23	24	26
Total	52	52	52	52	52	52	52

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

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

Table 2-6. 2010–11 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	56%	65%	32%	23%	23%	26%	41%
Level 2	44%	35%	59%	74%	74%	71%	53%
Level 3	0%	0%	9%	3%	3%	3%	6%
Total	100%	100%	100%	100%	100%	100%	100%

2.2.5 Depth of Knowledge

Each item on the NECAP test in reading 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 Table 2-7:

Table 2-7. 2010–11 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.

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

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

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

Additionally, problem solving, reasoning, connections, and communication are embedded throughout the GLEs/GSEs.

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 simple problem, requiring two to five minutes of response time. Constructed-response items are more complex, requiring 8 to 10 minutes of response time. Each type of item is worth a specific number of points in the student's total mathematics score, as shown in Table 2-8.

Table 2-8. NECAP 2010–11: 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; OR = open-response

2.3.3 Test Design

Table 2-9 summarizes the numbers and types of items that were used in the 2010–11 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 either one or two points, and each constructed-response item four points. Score points within a grade level were evenly divided, so that multiple-choice items represented approximately 50% of possible score points, and short-answer and constructed-response items together represented approximately 50% of score points.

Table 2-9. 2010–11 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
Mathematics 3–4	35	10	10		6	2	2		3	1	1		44	13	13	
Mathematics 5–8	32	6	6	4	6	2	2	1	3	1	1	1	41	9	9	6
Mathematics 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
For grades 3–4 and 5–8, total of nine forms; six contained unique matrix-equating items while Forms 7–9 contained the same matrix-equating items as Forms 1–3.
For grade 11, total of eight forms; six contained unique matrix-equating items while Forms 7–8 contained the same matrix-equating items as Forms 1–2.

2.3.4 Blueprints

The distribution of emphasis for NECAP content strands for mathematics is shown in Table 2-10.

Table 2-10. 2010–11 NECAP: Distribution of Emphasis in Terms of Target Percentage of Test by Grade—Mathematics Grades 3–8 and 11

Subcategory	*Grade*						
	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 2-11 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 2-11. 2010–11 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	12	9
Geometry and Measurement	10	13	13	17	15	17	19
Functions and Algebra	10	10	13	13	20	26	26
Data, Statistics, and Probability	10	10	10	10	10	10	10
Total	65	65	66	66	64	65	64

Table 2-12 lists the percentage of total score points assigned to each level of DOK in mathematics.

Table 2-12. 2010–11 NECAP: Depth of Knowledge in Terms of Targeted Percentage of Test by Grade—Mathematics Grades 3–8 and 11

DOK	*Grade*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Level 1	23%	22%	35%	26%	27%	29%	27%
Level 2	68%	71%	65%	64%	67%	62%	70%
Level 3	9%	8%	0%	11%	6%	9%	3%
Total	100%	100%	100%	100%	100%	100%	100%

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 Table 2-13.

Table 2-13. 2010–11 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.

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 test 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 high school. 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 3–6. 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 receive rulers for use on the mathematics test. Students may keep the rulers 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 all items on the NECAP test were designed to measure a specific GLE. The content standards for grades 5 and 8 writing identify four major genres that are assessed in the writing portion of the NECAP test each year:

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

Grade 11

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

- 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 (MC) items, constructed-response (CR) items, and extended-response (ER) 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. Constructed-response items are more complex, requiring 5 to 10 minutes of response time. At grades 5, 8, and 11, students are required to answer an extended-response item, receiving 45 minutes (plus additional time if necessary) to compose a response. Each type of item is worth a specific number of points in the student's total writing score, as shown in Table 2-14.

Table 2-14. NECAP 2010–11: 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

2.4.3 Test Design

Table 2-15 summarizes the numbers and types of items that were used in the 2010–11 NECAP writing test for grades 5 and 8. Note that all items on the grades 5 and 8 writing tests were common. Each MC item was worth one point, each CR item four points, and the ER writing prompt 12 points.

Table 2-15. 2010–11 NECAP: Number of Items by Item Type (All Items Common) and Number of Items—Writing Grades 5 and 8

MC	*CR*	*ER*
10	3	1

Table 2-16 summarizes the test design used in the 2010–11 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 ER writing prompts, one common and either one matrix-equating or one field-test prompt. The common prompt was worth 12 points.

Table 2-16. 2010–11 NECAP: Number of Items by Item Type and Number of Items—Writing Grade 11 (8 Forms)

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

2.4.4 Blueprints

Grades 5 and 8

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

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

The category reporting structure for grades 5 and 8 writing is shown in Table 2-17. The table provides the maximum possible number of raw score points that students could earn. The content category "Short Responses" lists the total raw score points from the three CR items; the reporting category "Extended Response" lists the total raw score points from the writing prompt.

Table 2-17. 2010–11 NECAP: Reporting Subcategory and Possible Raw Score Points Possible by Grade—Writing Grades 5 and 8

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

Short response = CR 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 Table 2-18:

Table 2-18. 2010–11 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 2-19 lists the percentage of actual score points assigned to each level of DOK in writing for grades 5 and 8.

Table 2-19. 2010–11 NECAP: Depth of Knowledge by Grade (in Percentage of Test)—Writing Grades 5 and 8

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

Table 2-20 lists the percentage of actual score points assigned to each level of DOK in writing for grade 11.

Table 2-20. 2010–11 NECAP: Depth of Knowledge in Terms of Percentage of Test, by Grade—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 and are consistent with NECAP GLE and GSE content standards. Measured Progress test developers work with Rhode Island, Vermont, Maine, and New Hampshire 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. All items used on the common portions of the NECAP tests were reviewed by a committee of content experts and by a committee of bias experts. Tables 2-21 through 2-24 show the number of items developed within each content area for the 2010–2011 NECAP tests.

Table 2-21. 2010–11 NECAP: Annual English Language Arts Item Development—Grades 3–8

Passages	*MC*	*CR*
4 long passages (divided by literary and informational)	64	12
7 short passages (divided by literary and informational)	56	14
Standalones	20	0
11 total passages	140	26

MC = multiple-choice; CR = constructed response

Table 2-22. 2010–11 NECAP: Annual English Language Arts Item Development—Grade 11

Passages	*MC*	*CR*
5 long passages (divided by literary and informational)	80	15
5 short passages (divided by literary and informational)	40	10
Standalones	20	0
10 total passages	140	25

MC = multiple-choice; CR = constructed response

Table 2-23. 2010–11 NECAP: Annual Writing Item Development—Grade 11

Grades	*ER*
11	6

ER = extended response writing prompt

Table 2-24. 2009–10 NECAP: Annual Mathematics Item Development—Grades 3–8 and 11

Grades	*MC*	*SA1*	*SA2*	*CR*
3	27	9	9	0
4	27	9	9	0
5	27	9	9	9
6	27	9	9	9
7	27	9	9	9
8	27	9	9	9
11	46	24	20	14

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 item, open-response scoring guide, 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)
- Ensuring that, for each item, there was only one correct answer
- Consideration of scorability and evaluation as to whether the scoring guide adequately addressed performance on the item

Fundamental questions the lead developer considered, but was not limited to, included the following:

- What is the item asking?
- Is the key the only possible key? (Is there only one correct answer?)
- Is the 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 committee's role is to review test items for the NECAP, provide feedback, and make recommendations about which items should be selected for program use. The 2010–11 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 2010–11 operational test and made recommendations about selecting, revising, or eliminating specific items from the item pool. Members reviewed each item against the following criteria:

- Grade-Level/Grade-Span Expectation Alignment
  - Is the test item aligned to the appropriate GLE/GSE?
  - If not, which GLE/GSE or grade level is more appropriate?
- Correctness
  - Are the items and distractors correct with respect to content accuracy and developmental appropriateness?
  - Are the scoring guides consistent with GLE/GSE wording and developmental appropriateness?
- Depth of Knowledge[1]
  - Are the items coded to the appropriate DOK?
  - If consensus cannot be reached, is there clarity around why the item might be on the borderline of two levels?
- Language
  - Is the item language clear?
  - Is the item language accurate (syntax, grammar, conventions)?
- Universal Design
  - Is there an appropriate use of simplified language? (Does 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 Brailled?

[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.

2.5.4 Bias and Sensitivity Review

Bias review is an essential part of the development process. During the bias review 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 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 subject matter, regardless of reading ability;
- exhibited high technical quality in terms of psychometric characteristics;
- had appropriate answer options or score-point descriptors; and
- were free of potentially sensitive content.

2.5.7 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 2010–11 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 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 for grades 3 through 8 and 11 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 adapted into a large-print version.

Chapter 3. TEST ADMINISTRATION

3.1 Responsibility for Administration

The 2010 *NECAP Principal/Test Coordinator Manual* indicated that principals and/or their designated NECAP test coordinators were responsible for the proper administration of the NECAP. Uniformity of administration procedures from school to school was ensured by using manuals that contained explicit directions and scripts to be read aloud to students by test administrators.

3.2 Administration Procedures

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

Schools were required to return a Student Answer Booklet for every enrolled student in the grade level, with the exception of students who took an alternate assessment in the previous school year. Students who were alternately assessed in the 2009–10 school year were not required to participate in the NECAP in 2010–11. On those occasions when it was deemed impossible to test a particular student, school personnel were required to inform their Department of Education. A grid was included on the Student Answer Booklets that listed the approved reasons why a booklet could be returned blank for one or more sessions of the test:

- Student is new to the United States after October 1, 2009, 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 2010; however, these students were required to take the mathematics test in 2010.

- Student withdrew from school after October 1, 2010
  B. If a student withdrew after October 1, 2010, 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, 2010
  C. If a student enrolled after October 1, 2010, 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, 2010, 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 three states combined in reading, mathematics, and writing.

3.3.1 Large Print and Braille

All Form 1s of the test in grades 3 through 8 and 11 were enlarged to 20-point font for visually impaired students. In addition, common items in each grade-level test were translated into Braille by National Braille Press, a subcontractor that specializes in test materials for blind students.

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 regional locations in each state to inform school personnel about the NECAP and to provide training on the policies and procedures regarding administration of the tests. A test administration workshop was also conducted via an online webinar for each state. These live webinars were recorded so that test coordinators and test administrators could view them at a time that was convenient for them. A link was provided to each state for their recorded workshop presentation in order

for it to be added to the Department of Education Web site for school personnel to access. Lastly, an audio PowerPoint workshop presentation was pre-recorded 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 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 new 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 2010 NECAP test administration. The accommodation codes (T1-4, S1-2, P1-11, R1-7, O1, M1, and M3) are defined in the Table of Standard Test Accommodations, which can be found in Appendix D. Appendix C also shows the accommodation codes N01 to N07 which were available to only grade 11 students who participated in the reading and mathematics testing online using NimbleTools. These codes are defined in Appendix E.

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 and/or principal were responsible for immediately reporting the concern to the District Superintendent and the State Assessment Director at the Department of Education. Test security was also strongly emphasized at test administration workshops that were conducted in all 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 several irregularities that occurred during the 2010 NECAP test administration. Some of the irregularities involved items in the assessments while others were attributed to printing issues. These irregularities as well as how they were addressed are described in the following list of bullets.

- **Irregularity:** Several items appeared in the 2010–11 test forms in grades 7 and 8 that had been released to the public with the results from the 2007–08 test administration.
- **Solution:** These items were excluded using the process approved by the NECAP Technical Advisory Committee to calculate scores when an item cannot included for some reason.
- **Irregularity:** A field-test item was incorrectly included in Session 1 of the grade 5 mathematics test. The item only appeared in Form 1. The item required the use of the Mathematics Tool Kit. However, in Session 1 students do not have access to their tool kits.
- **Solution:** A notice was immediately sent to all elementary school principals and test coordinators. The notice included instructions to inform their grade 5 test administrators of the issue and have them instruct students that were using a Form 1 test booklet to skip the item.
- **Irregularity:** In grade 5 mathematics, constructed-response common item 62 was questioned by people who believed that students did not have sufficient answer space for all four parts of the item.
- **Solution:** The item had been field-tested previously and statistics were strong. It was decided during testing to examine the responses once they were scored and make a determination on whether to count the item. Additional concern was raised that students would continue their answer outside of the answer space. Scoring selected a random sample of 200 answer booklets

and none of the booklets contained work outside of the answer space. Scoring statistics on the item were appropriate, and the item was counted toward students' scores.

- **Irregularity:** A school reported that the shapes in the Mathematics Tool Kit for the grade 4 large-print test were not scaled correctly. One item required the student to cover a preprinted shape in the large-print test with the shapes from the tool kit with no gaps or overlaps. Because the tool kit shapes were not enlarged to scale, there were gaps when the shape was covered.
- **Solution:** The item is multiple-choice and no other answer is possible due to this error, so the item was scored and included in student results. In future years, Measured Progress will evaluate exact size of the large-print tool kit pieces.
- **Irregularity:** There were a total of thirteen test booklets (12 grade 3 and 1 grade 4) that contained multiple- choice bubbles that were printed very lightly. This led to a concern that if some students could not see the faint multiple-choice bubbles they may have circled the multiple-choice option for their responses instead.
- **Solution:** To address this concern, Measured Progress implemented a data check for multiple-choice items that were not answered by filling in a bubble and then performed a visual check to see if students circled their response. If it was determined that a student had circled their response, they were credited with that response. As result of this step, no students lost credit due to this printing error. Measured Progress discussed this issue with the print vendor and is continuing to work with the company to ensure the print quality of the test booklets in the future.

3.8 Test Administration Window

The test administration window was October 1–22, 2010.

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. It provides a centralized location to which individuals in the field can call using a toll-free number to ask specific questions or report any problems they may be experiencing. Representatives are responsible for receiving, responding to, and tracking calls, then routing issues to the appropriate person(s) for resolution. All calls are logged into a database 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 AM to 4:00 PM 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 backwards, and identification of critical data that are missing (e.g., a student ID number), test forms that are out of range or missing, and page or document sequence errors. When a problem is detected, the scanner stops and displays an error message directing the operator to investigate and 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 onsite. (Chief Reader and other scoring roles are described in the section that follows.)

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

4.1.2.1 Scoring Location and Staff

Scoring Location

The iScore database, its operation, and its administrative controls are all based in Dover, New Hampshire. Table 4-1 presents the locations where 2010–11 NECAP test item responses by grade and content area were scored.

Table 4-1. 2010–11 NECAP: Operational Scoring Locations by Content Area and Grade

Content area	*Grade*	*Louisville, KY*	*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 2010–11 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 a later section) for Quality Assurance Coordinators. Chief Readers are employees of Measured Progress.
- Numerous Quality Assurance Coordinators (QACs), selected from a pool of experienced Senior Readers for their ability to score accurately and their ability to instruct and train Readers, participated in benchmarking activities for each specific grade and content area. QACs provided read-behind activities (defined in a later section) for Senior Readers at their sites. The ratio of QACs and Senior Readers to Readers was approximately 1:11.
- Numerous Senior Readers, 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 2010–11 student responses. Recruitment of Readers is described in Section 5.1.2.3.

4.1.2.2 Benchmarking Meetings with the NECAP State Specialists

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

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

This repeated use of NECAP-approved sets of 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 2010–11 NECAP, 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 about Readers (e.g., gender, race, educational background) was electronically captured for reporting.

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

Table 4-2 summarizes the qualifications of the 2010–11 NECAP scoring leadership and Readers.

Table 4-2. 2010–11 NECAP: Qualifications of Scoring Leadership and Readers—Fall Administration

Scoring Responsibility	*Educational Credentials*				*Total*
	Doctorate	*Master's*	*Bachelor's*	*Other*	
Scoring Leadership	5.9%	26.5%	60.0%	7.6%	100.0%
Readers	4.2%	30.1%	54.8%	10.9%	100.0%

Scoring Leadership = Chief Readers, QACs, and Senior Readers
*4 QACs/Senior Readers had an associate's degree and 10 had at least 48+ college credits.
**81 Readers had an associate's degree and 70 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. 2010–11 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
Non-scorable items	0

Non-Scorable Items

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

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

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

Scoring Procedures

Scoring procedures for polytomous items included both single scoring and double scoring. Single-scored items were scored by one Reader. Double-scored items were scored independently by two Readers, whose scores were tracked for "interrater agreement." (For further discussion of double scoring and interrater agreement, see Section 5.1.2.7 and Appendix Q.)

4.1.2.5 Reader Training

Reader training began with an introduction of the onsite scoring staff and providing 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 reviewing or scoring the particular response set organized for that training: Anchor Sets, Training Sets, and Qualifying Sets. (These are defined below.)

During training, Readers could highlight or mark hard copies of the Anchor 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 via computer.

Anchor Set

Readers first reviewed an Anchor Set of exemplary responses for an item. This is a set approved by the reading, writing, and mathematics content specialists representing the 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 contains one client-approved sample response per score point considered to be a midrange exemplar. The set includes a second sample response if there is more than one plausible way to illustrate the merits and intent of a score point.

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

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

Training Set

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

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

Qualifying Set

After the Training Set had been completed, Readers were required to score responses accurately and reliably in Qualifying Sets assembled for constructed-response items, writing prompts, and all two-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 8 of the 10 responses) and at least 90% exact or adjacent agreement (i.e., to exactly match or be within one score point of the predetermined score on 9 or 10 of the 10 responses), except 70% and 90%, respectively, for six-point writing-prompt responses. In other words, Readers were allowed one discrepant score (i.e., one score of 10 that was more than one score point from the predetermined score) provided they had at least eight exact scores (seven for writing-prompt items).

To be eligible to score one-point short-answer mathematics items (which were benchmarked "right" or "wrong") and two-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 discussing 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 latter.

4.1.2.7 Monitoring of Scoring Quality Control and Consistency

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

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

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

Furthermore, the accuracy rate required of a Reader to *qualify* to score responses live was 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 Committee-Reviewed Responses (CRRs)

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

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

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

Embedded CRRs were employed for all constructed-response items. They were not used for writing six-point extended-response items, because these are 100% double-blind scored (defined below). Embedded CRRs were also not used for math two-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-ended item types.

Responses placed into the read-behind queue were randomly selected by scoring leadership; Readers were not aware which of their responses would be reviewed by their 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 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. 2010–11 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-ended item types. Table 4-5 shows by which method(s) both common and equating open-ended item responses for each operational test were scored.

Table 4-5. 2010–11 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	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 Q provides the NECAP 2010–11 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, Readers were required to take an online Recalibration Set at the start and midpoint of the shift of their resumption of scoring.

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

- Readers who were discrepant on two of five responses of the first Recalibration Set, or exact on two or fewer, were not permitted to score on that item that day and were either assigned to a different item or dismissed for the day.
- Readers who were discrepant on only one of five responses of the first Recalibration Set, and/or exact on three, were retrained by their Senior Reader by discussing the Recalibration Set responses in terms of the score point descriptions and the original Anchor Set. After this retraining, such Readers began scoring operational responses under the proviso that the Reader's scores for that day and that item would be kept only if the Reader was exact on all five of five responses of the second Recalibration Set administered at the shift midpoint. The QAC determined whether or when these Readers had received enough retraining to resume scoring operational responses. Scoring leadership also carefully monitored the accuracy of such Readers by significantly increasing the number of their read-behinds.

- Readers who were not discrepant on any response of the first Recalibration Set, and exact on at least four, were allowed to begin scoring operational responses immediately, under the proviso that this Recalibration Set performance would be combined with that of the second Recalibration Set administered at the shift midpoint.

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

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

Recalibration Sets were employed for all constructed-response items. They were not used for writing six-point extended-response items, which were 100% double-blind scored. They were also not used for two-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 2010. 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 two-point short-answer items, for which students can receive scores of 0, 1, or 2, and constructed-response items, which are worth four 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 they 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 (0.25 for four-

option multiple-choice items or essentially zero for constructed-response items) to 0.90, with the majority of items generally falling between around 0.4 and 0.7. 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 0.6.

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 subject 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, because only a single writing prompt is administered in grades 5 and 8, 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 5-1.2010–11 NECAP: Summary of Item Difficulty and Discrimination Statistics by Subject and Grade

Subject	*Grade*	*Item type*	*Number of items*	p-*Value*		*Discrimination*	
				Mean	*Standard deviation*	*Mean*	*Standard deviation*
Mathematics	3	ALL	55	0.67	0.17	0.44	0.08
		MC	35	0.71	0.16	0.43	0.07
		SA	20	0.60	0.18	0.46	0.10
	4	ALL	55	0.65	0.19	0.40	0.10
		MC	35	0.67	0.19	0.37	0.09
		SA	20	0.61	0.17	0.46	0.09
	5	ALL	48	0.59	0.17	0.44	0.11
		CR	4	0.53	0.19	0.61	0.08
		MC	32	0.63	0.17	0.40	0.09
		SA	12	0.53	0.12	0.50	0.07
	6	ALL	48	0.58	0.16	0.46	0.10
		CR	4	0.38	0.11	0.62	0.07
		MC	32	0.63	0.13	0.42	0.08
		SA	12	0.51	0.17	0.52	0.05

Subject	Grade	Item type	Number of items	p-Value		Discrimination	
				Mean	Standard deviation	Mean	Standard deviation
Mathematics	7	ALL	46	0.53	0.15	0.44	0.13
		CR	4	0.36	0.04	0.68	0.02
		MC	30	0.59	0.13	0.37	0.08
		SA	12	0.43	0.14	0.53	0.08
	8	ALL	47	0.53	0.16	0.45	0.11
		CR	4	0.35	0.06	0.68	0.04
		MC	32	0.58	0.12	0.41	0.08
		SA	11	0.45	0.19	0.51	0.07
	11	ALL	46	0.44	0.19	0.45	0.13
		CR	4	0.32	0.13	0.68	0.06
		MC	24	0.53	0.16	0.38	0.10
		SA	18	0.34	0.17	0.51	0.09
Reading	3	ALL	34	0.71	0.15	0.44	0.09
		CR	6	0.48	0.13	0.51	0.10
		MC	28	0.77	0.10	0.43	0.08
	4	ALL	34	0.72	0.14	0.42	0.06
		CR	6	0.59	0.22	0.47	0.05
		MC	28	0.75	0.10	0.41	0.06
	5	ALL	34	0.69	0.15	0.41	0.09
		CR	6	0.44	0.03	0.57	0.06
		MC	28	0.74	0.10	0.38	0.06
	6	ALL	34	0.70	0.15	0.46	0.08
		CR	6	0.45	0.03	0.61	0.04
		MC	28	0.75	0.10	0.43	0.05
	7	ALL	34	0.66	0.14	0.40	0.13
		CR	6	0.45	0.05	0.65	0.03
		MC	28	0.70	0.11	0.35	0.07
	8	ALL	34	0.70	0.14	0.40	0.11
		CR	6	0.50	0.06	0.59	0.04
		MC	28	0.74	0.11	0.35	0.07
	11	ALL	34	0.66	0.15	0.40	0.14
		CR	6	0.50	0.04	0.64	0.01
		MC	28	0.70	0.14	0.34	0.08
Writing	5	ALL	14	0.72	0.17	0.38	0.13
		CR	3	0.48	0.05	0.55	0.03
		MC	10	0.82	0.05	0.31	0.07
		WP	1	0.46		0.58	
	8	ALL	14	0.69	0.12	0.42	0.14
		CR	3	0.63	0.10	0.61	0.08
		MC	10	0.72	0.12	0.34	0.04
		WP	1	0.55		0.66	
	11	ALL	1	0.52			
		WP	1	0.52			

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 theirs 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 F; 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 G; 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. *Standards for Educational and Psychological Testing* (AERA et al., 1999) includes similar guidelines. As part of the effort to identify such problems, NECAP items were evaluated 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, there must be a minimum of 200 students 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 2010–11 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 H present the number of items classified as either "low" or "high" DIF, overall and by group favored.

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 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 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

[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.

and (b) the nature of the multidimensionality. Findings from dimensionality analyses performed on the 2010–11 NECAP common items for mathematics, reading, and writing are reported below. (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 2010–11 NECAP. 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 16,000 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. NECAP 2010–2011:
Multidimensionality Effect Sizes by Grade and Subject

Subject	*Grade*	*Multidimensionality Effect Size*	
		*Prior Administration**	*2010–11*
Reading	3	0.18	0.13
	4	0.18	0.19
	5	0.18	0.24
	6	0.19	0.23
	7	0.20	0.21
	8	0.32	0.34
	11	0.28	0.29
	Average	0.22	0.23
Mathematics	3	0.17	0.16
	4	0.13	0.13
	5	0.15	0.16
	6	0.16	0.18
	7	0.16	0.14
	8	0.16	0.11
	11	0.12	0.13
	Average	0.15	0.15
Writing	5	0.20	0.24
	8	0.18	0.28
	Average	0.19	0.26

* 2009–10 for reading and mathematics; 2008–09 for writing

All of the DETECT values indicated very weak to weak multidimensionality, except for grade 8 reading whose value of 0.34 is slightly more than halfway between weak and moderate. The two writing test forms (average DETECT value of 0.26, weak multidimensionality) displayed slightly greater multidimensionality than the reading test forms (average of 0.22, weak multidimensionality), which in turn had slightly greater multidimensionality than mathematics (average of 0.15, very weak multidimensionality). Also shown in Table 5-2 are the values reported in last year's dimensionality analyses (except for writing, for

which the 2008–09 values are given since writing was not assessed in 2009–10). The individual values for the different grade levels as well as the averages for both mathematics and reading are seen to be very similar to those from last year, whereas the writing tests displayed slightly higher DETECT values in comparison to 2008–09, the most recent school year in which they were administered.

The way in which DETECT divided the tests into clusters was also investigated to determine whether there were any discernable patterns with respect to the multiple-choice (MC) and constructed-response (CR) item types. Inspection of the DETECT clusters indicated that MC-CR separation occurred much more strongly with reading and writing than with mathematics, a pattern that has been consistent across all four years of dimensionality analyses for the NECAP fall tests. Specifically, for mathematics, only grade 5 mathematics showed some evidence of MC-CR separation in that one cluster was totally composed of 21 MC items and a second cluster was composed of 10 CR items accounting for 25 points. Thus, two clusters that displayed strong MC-CR separation accounted for 46 of the 66 points on the grade 5 mathematics test. Each of the remaining grade 5 mathematics clusters displayed a mix of MC and CR items. No other grade levels in mathematics displayed separation of any substantial numbers of MC and CR items into separate clusters. In reading, however, grades 5, 6, 7, 8, and 11 all displayed very strong MC-CR separation, with DETECT indicating a two-cluster solution in every case where one cluster was all MC and the other was all CR. In grade 4 reading, 32 out of 52 points appeared in two such clusters while the remaining points occurred in clusters that were a mix of MC and CR items. In reading, only grade 3 displayed no evidence of any MC-CR separation. For writing, both grades displayed a strong MC-CR two-cluster solution in the same manner as occurred with reading. Despite this multidimensionality between the multiple-choice items and remaining items for reading, the effect sizes were not strong enough to warrant further investigation.

Thus, a tendency is suggested for MC and CR items to sometimes measure statistically separable dimensions, especially in regard to the reading and writing tests. This has been consistent across all four 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 small 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 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 2010–11 analyses, the violations of local independence, as evidenced by the DETECT effect sizes, were weak or very weak in all cases. Thus, these effects do not seem to warrant any changes in test design or scoring. In addition, the magnitude of the violations of local independence have been consistently low over the years, and the patterns with respect to the MC and CR items have also been consistent, with reading and writing tending to display more separation than mathematics.

Chapter 6. IRT SCALING AND EQUATING

This chapter describes the procedures used to calibrate, equate, and scale NECAP. 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 *a*-plots and *b*-plots for reasonableness); and evaluation of the scaling results (e.g., parallel processing by the Psychometrics and Research and Data Analysis departments; 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 the Item Response Theory Results and Equating Results sections below.

Table 6-1. 2010–11 NECAP:
Items That Required Intervention During IRT Calibration and Equating

Item number	*Subject*	*Grade*	*Reasons*	*Action*
124433	MAT	03	c parameter	c = 0
119821	MAT	03	c parameter	c = 0
201312	MAT	03	Delta Analysis	Removed from equating
119896	MAT	03	Delta Analysis	Removed from equating
145070	MAT	04	c parameter	c = 0
139477	MAT	04	c parameter	c = 0
144648	MAT	04	c parameter	c = 0
124522	MAT	04	c parameter	c = 0
124592	MAT	04	Delta Analysis	Removed from equating
232445	MAT	04	c parameter	c = 0
255664	MAT	04	Delta Analysis	Removed from equating
124866	MAT	05	c parameter	c = 0
120799	MAT	05	c parameter	c = 0
139399	MAT	05	c parameter	c = 0
255761	MAT	05	c parameter	c = 0
119311	MAT	06	c parameter	c = 0
145608	MAT	06	c parameter	c = 0
119288	MAT	06	c parameter	c = 0
139217	MAT	06	c parameter	c = 0
122249	MAT	06	c parameter	c = 0
228071	MAT	06	c parameter	c = 0
225428	MAT	06	Delta Analysis	Removed from equating
123513	MAT	06	Delta Analysis	Removed from equating
120329	MAT	07	IRT Plot Outlier	Removed from equating

Item number	*Subject*	*Grade*	*Reasons*	*Action*
154775	MAT	07	Delta Analysis/ IRT Plot Outlier	Removed from equating
140025	MAT	07	Delta Analysis	Removed from equating
234459	MAT	07	IRT Plot Outlier	Removed from equating
139845	MAT	08	c parameter	c = 0
206256	MAT	08	Delta Analysis	Removed from equating
120932	MAT	08	Delta Analysis	Removed from equating
199768	MAT	08	Delta Analysis	Removed from equating
139869	MAT	08	Delta Analysis	Removed from equating
140155	MAT	11	c parameter	c = 0
119494	MAT	11	Delta Analysis	Removed from equating
117732	REA	03	c parameter	c = 0
148024	REA	03	c parameter	c = 0
147674	REA	03	c parameter	c = 0
148198	REA	03	c parameter	c = 0
117661	REA	03	c parameter	c = 0
147877	REA	04	c parameter	c = 0
147902	REA	04	c parameter	c = 0
147915	REA	04	a parameter	a set to initial value
118207	REA	05	c parameter	c = 0
148344	REA	05	c parameter	c = 0
118083	REA	05	c parameter	c = 0
118073	REA	05	c parameter	c = 0
149112	REA	05	c parameter	c = 0
148719	REA	05	c parameter	c = 0
148763	REA	05	c parameter	c = 0
118165	REA	05	Delta Analysis	Removed from equating
118268	REA	06	c parameter	c = 0
148978	REA	06	c parameter	c = 0
118347	REA	06	c parameter	c = 0
256674	REA	06	c parameter	c = 0
147663	REA	07	c parameter	c = 0
128125	REA	07	c parameter	c = 0
118536	REA	07	c parameter	c = 0
129224	REA	07	c parameter	c = 0
118546	REA	07	c parameter	c = 0
118547	REA	07	c parameter	c = 0
147203	REA	07	c parameter	c = 0
147210	REA	07	c parameter	c = 0
147215	REA	07	c parameter	c = 0
118573	REA	07	c parameter	c = 0
201482	REA	07	c parameter	c = 0
147546	REA	07	IRT Plot Outlier	Removed from equating
147546	REA	07	c parameter	c = 0
147549	REA	07	Delta Analysis	Removed from equating
129217	REA	07	c parameter	c = 0
147223	REA	08	c parameter	c = 0
147239	REA	08	c parameter	c = 0
118666	REA	08	c parameter	c = 0
118669	REA	08	c parameter	c = 0
118601	REA	08	c parameter	c = 0
118603	REA	08	c parameter	c = 0
118604	REA	08	c parameter	c = 0
147516	REA	08	c parameter	c = 0
147611	REA	08	Delta Analysis	Removed from equating
147934	REA	11	c parameter	c = 0
118758	REA	11	c parameter	c = 0

Item number	Subject	Grade	Reasons	Action
118764	REA	11	c parameter	c = 0
147551	REA	11	c parameter	c = 0
147850	REA	11	c parameter	c = 0
147860	REA	11	c parameter	c = 0
147695	REA	11	c parameter	c = 0
147716	REA	11	c parameter	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, it is assumed that all items are independent measures of the same construct (i.e., of the same θ). Another way to think of θ is as a mathematical representation of the latent trait of interest. Several common IRT models are used to specify the relationship between θ and p (Hambleton & 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 2010–11 NECAP, 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:

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

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

In the GRM for polytomous items, an item is scored in k + 1 graded categories that can be viewed as a set of k dichotomies. At each point of dichotomization (i.e., at each threshold), a two-parameter model can

be used. This implies that a polytomous item with k + 1 categories can be characterized by k item category threshold curves (ICTCs) of the two-parameter logistic form:

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

where
i indexes the items,
j indexes students,
k indexes threshold,
a represents item discrimination,
b represents item difficulty,
d represents threshold, and
D is a normalizing constant equal to 1.701.

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

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

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.

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

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

6.2 Item Response Theory Results

The tables in Appendix I give the IRT item parameters of all common items on the 2010–11 NECAP tests by grade and content area. In addition, Appendix J 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 7.1, the expected raw score at a given value of θ_j is

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

where
i indexes the items (and n is the number of items contributing to the raw score),
j indexes students (here, θ_j runs from –4 to 4), and
$E(X \mid \theta_j)$ is the expected raw score for a student of ability θ_j.

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

The TIF displays the amount of statistical information 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 the 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. 2010–11 NECAP: Number of Newton Cycles Required for Convergence

Subject	*Grade*	*Cycles*
Mathematics	3	42
	4	71
	5	74
	6	58
	7	77
	8	53
	11	105
Reading	3	61
	4	150
	5	55
	6	49
	7	48
	8	49
	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 2010–11 administration of NECAP 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 2010–11 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 2010–11 were placed on the 2009–10 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 2009–10 and 2010–11 NECAP tests should have the same item parameters. After the item parameters for each 2010–11 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 2010–11 TCC for the equating items was as close as possible to that of 2009–10.

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 K 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 K are the results from the rescore analysis. With this analysis, 200 random papers from the previous year were interspersed with this year's papers to evaluate scorer consistency from one year to the next. All effect sizes were well below the criterion value for excluding an item as an equating item, 0.80 in absolute value.

Finally, *a*-plots and *b*-plots, which show IRT parameters for 2010–11 plotted against the values for 2009–10, are presented in Appendix L. 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. 2010–11 NECAP: Stocking and Lord Transformation Constants

Content area	*Grade*	a-*slope*	b-*intercept*
Mathematics	3	0.989	0.123
	4	1.059	0.012
	5	1.014	0.095
	6	1.078	0.196
	7	1.034	0.192
	8	0.933	0.241
	11	1.015	0.158
Reading	3	0.953	0.027
	4	1.029	0.256
	5	0.981	0.133
	6	1.024	0.060
	7	1.088	-0.009
	8	1.034	0.221
	11	1.074	0.278

The next administration of NECAP (2011–12) will be scaled to the 2010–11 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 2010; for complete details of the standard setting, please see the *2010–11 New England Common Assessment Program Standard Setting Report* (Measured Progress, 2011). The report is included as Appendix M

The cuts on the theta scale that were established via standard setting and used for reporting in fall 2010 are presented in Table 6-4 below. Also shown in the table are the cutpoints on the reporting score scale (described below). These cuts will remain fixed throughout the assessment program unless standards are reset for any reason.

Table 6-4. 2010–11 NECAP: Cut Scores on the Theta Metric and Reporting Scale by Subject and Grade

Subject	Grade	Theta			Scaled Score				
		Cut 1	*Cut 2*	*Cut 3*	*Minimum*	*Cut 1*	*Cut 2*	*Cut 3*	*Maximum*
Mathematics	3	–1.0381	-0.2685	0.9704	300	332	340	353	380
	4	–1.1504	-0.3779	0.9493	400	431	440	455	480
	5	–0.9279	-0.2846	1.0313	500	533	540	554	580
	6	–0.8743	-0.2237	1.0343	600	633	640	653	680
	7	–0.7080	-0.0787	1.0995	700	734	740	752	780
	8	–0.6444	-0.0286	1.1178	800	834	840	852	880
	11	–0.1169	0.6190	2.0586	1100	1134	1140	1152	1180
Reading	3	–1.3229	-0.4970	1.0307	300	331	340	357	380
	4	–1.1730	-0.3142	1.1473	400	431	440	456	480
	5	–1.3355	-0.4276	1.0404	500	530	540	556	580
	6	–1.4780	-0.5180	1.1255	600	629	640	659	680
	7	–1.4833	-0.5223	1.2058	700	729	740	760	780
	8	–1.5251	-0.5224	1.1344	800	828	840	859	880
	11	–1.2071	-0.3099	1.0038	1100	1130	1140	1154	1180
Writing	5	–1.2835	-0.0087	1.5244	500	527	540	555	580
	8	–1.3486	-0.1059	1.2682	800	827	840	854	880

Table N-1 in Appendix N shows achievement level distributions by subject 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, results are shown only for the 2010–11 administration.

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 2010–11 NECAP tests were translated to scaled scores using a data analysis process called *scaling*. Scaling simply converts from one scale to another. In the same way that a given temperature can be expressed on either Fahrenheit or Celsius scales, or the same distance can be expressed in either miles or kilometers, student scores on the 2010–11 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 x40. 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/content 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 x40 and the x00 values (e.g., 440 and 400 for grade 4). The x00 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 x00. The maximum possible raw score is assigned a scaled score of x80 (e.g., 480 in the case of grade 4). Because only two points within the θ scaled score space are fixed, the scaled score cuts between Substantially Below Proficient and Partially Proficient and between Proficient and Proficient with Distinction are free to vary across the grade/content combinations.

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

Table 6-5. 2010–11 NECAP: Scaled Score Slope and Intercept by Subject and Grade

Subject	*Grade*	*Slope*	*Intercept*
Mathematics	3	10.7195	342.8782
	4	11.0432	444.1727
	5	10.7659	543.0634
	6	10.5922	642.3690
	7	10.2007	740.8028
	8	10.0720	840.2881
	11	8.6600	1134.6399
Reading	3	11.4188	345.6751
	4	10.8525	443.4098
	5	11.1970	544.7878
	6	11.4875	645.9499
	7	11.5019	746.0074
	8	11.5022	846.0087
	11	10.8399	1143.3595
Writing	5	10.0217	540.0869
	8	10.2719	841.0878

Appendix O contains raw score to scaled score look-up tables for the 2010–11 NECAP tests. These are the actual tables used to determine student scaled scores, error bands, and achievement levels.

Appendix P 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 impact a student's score are referred to as "measurement error." Any assessment includes some amount of measurement error; that is, no measurement is perfect. This is true of all academic assessments—some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. When tests have a high amount of measurement error, student scores are very unstable. Students with high ability may get low scores, or vice versa. Consequently, one cannot reliably measure a student's true level of ability with such a test. Assessments that have less measurement error (i.e., errors made are small on average and student scores on such a test will consistently represent their ability) are described as reliable.

There are a number of ways to estimate an assessment's reliability. One possible approach is to give the same test to the same students at two different points in time. If students receive the same scores on each test, 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), that 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 2010–11 NECAP:

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

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

7.1 Reliability and Standard Errors of Measurement

Table 7-1 presents descriptive statistics, Cronbach's α coefficient, and raw score standard errors of measurement (SEMs) for each 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. 2010–11 NECAP: Raw Score Descriptive Statistics, Cronbach's Alpha, and Standard Errors of Measurement (SEM) by Subject and Grade

Subject	*Grade*	*Number of students*	*Raw score*			*Alpha*	*SEM*
			Maximum	*Mean*	*Standard deviation*		
Mathematics	3	43893	65	42.14	13.02	0.93	3.43
	4	44350	65	41.67	12.08	0.92	3.39
	5	44207	66	37.54	14.05	0.92	3.97
	6	44477	66	35.16	14.52	0.93	3.94
	7	46536	64	30.56	13.79	0.92	3.90
	8	46567	65	31.29	13.88	0.93	3.72
	11	32526	64	25.72	13.91	0.93	3.78
Reading	3	43736	52	32.93	8.82	0.89	2.96
	4	44206	52	35.18	8.64	0.88	3.04
	5	44031	52	31.23	8.56	0.89	2.90
	6	44329	52	31.72	9.24	0.91	2.85
	7	46409	52	30.47	9.16	0.89	3.10
	8	46456	52	32.75	8.39	0.88	2.92
	11	32527	52	31.48	8.80	0.88	3.02
Writing	5	43956	34	19.37	4.72	0.73	2.48
	8	46274	34	21.28	5.32	0.78	2.52
	11	32409	12	6.30	1.79		

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 2010–11 Subgroup Reliability

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2010–11 NECAP test. Appendix Q 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 R 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 Q. Once again as expected, because they 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 subtests.

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. 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 NECAP. 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 percent of responses that required a third score. This same information is provided at the item level in Appendix R.

Table 7-2. 2010–11 NECAP: Summary of Interrater Consistency Statistics Collapsed Across Items by Subject and Grade

Subject	*Grade*	*Number of score categories*	*Number of included scores*	*Percent exact*	*Percent adjacent*	*Correlation*	*Percent of third scores*
Mathematics	3	2	8677	98.12	1.88	0.96	0.00
		3	8680	95.52	4.27	0.96	0.21
	4	2	8807	98.67	1.33	0.97	0.00
		3	8997	92.42	7.29	0.93	0.29
	5	2	5344	96.76	3.24	0.93	0.00
		3	5330	87.58	11.67	0.90	0.75
		5	3690	84.15	13.85	0.95	2.06
	6	2	5342	97.77	2.23	0.95	0.00
		3	5315	91.01	8.32	0.93	0.68
		5	3852	85.49	13.32	0.95	1.22
	7	2	5592	98.48	1.52	0.97	0.00
		3	5538	92.34	7.51	0.93	0.14
		5	3922	81.44	15.60	0.93	2.93
	8	2	4595	97.37	2.63	0.95	0.00
		3	6517	87.45	12.14	0.89	0.52
		5	3946	80.31	18.32	0.92	1.34
	11	2	7317	96.68	3.32	0.93	0.00
		3	3516	94.17	5.20	0.95	0.63
		5	2416	87.38	11.05	0.95	1.57
Reading	3	5	5297	76.01	22.09	0.87	1.87
	4	5	5497	74.99	23.41	0.90	1.58
	5	5	5483	64.33	33.65	0.75	1.92
	6	5	5437	64.13	33.68	0.77	2.12
	7	5	5631	61.75	36.14	0.79	1.92
	8	5	5631	63.74	34.24	0.76	1.90
	11	5	3922	65.45	32.43	0.76	2.01
Writing	5	5	2830	66.57	31.80	0.75	1.63
		7	43423	56.70	38.97	0.70	4.10
	8	5	2870	66.24	32.13	0.79	1.57
		7	45173	64.84	33.67	0.77	1.36
	11	7	31548	64.12	34.32	0.76	1.50

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 2010–11 NECAP because it is easily adaptable to all types of testing formats, including mixed format tests.

The accuracy and consistency estimates reported in Appendix S 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 2010–11 NECAP, 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:

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

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 S-1 of Appendix S. 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 0.83 for Substantially Below Proficient for grade 3 mathematics. This figure indicates that among the students whose true scores placed them in this classification, 83% would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.76 indicates that 76% 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 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 2010–11 NECAP, Table S-2 in Appendix S 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 cut and whose true scores were below the cut. A false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.)

The above indices are derived from Livingston and Lewis's (1995) method of estimating the accuracy and consistency of classifications. It should be noted that Livingston and Lewis discuss two versions of the

accuracy and consistency tables. A standard version performs calculations for forms parallel to the form taken. An "adjusted" version adjusts the results of one form to match the observed score distribution obtained in the data. The tables 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 S should be interpreted with caution. In addition, it is important to remember that it is inappropriate to compare DAC statistics between grades and content areas.

Chapter 8. SCORE REPORTING

8.1 Teaching Year versus Testing Year Reporting

The data used for the NECAP reports are the results of the fall 2010 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 2010–11 NECAP:

- 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 T.

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 2010 NECAP Reports*. These resources help foster proper use and interpretation of NECAP results.

The *Guide* 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 ($\sigma_{\bar{x}_1-\bar{x}_2}$) 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:

$$\sigma^2_{\bar{x}_1 - x_2} = \sigma^2_w (\frac{1}{n_1} + \frac{1}{n_2})$$

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

$$\sigma^2_{\bar{x}_1 - \bar{x}_2} = \sigma^2 (\frac{1}{n} + \frac{1}{n}) = \frac{2\sigma^2}{n}, \text{ and}$$
$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{2\sigma^2 / n}$$

Therefore, if the difference in scaled scores of two equally sized groups is greater than or equal to $\sigma_{\bar{x}_1 - \bar{x}_2}$, 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* 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:

$$s_p^2 = \frac{s^2}{n} = \frac{p(1-p)}{n}, \text{ and}$$
$$s_p = \sqrt{p(1-p) / n}$$

Together, these two tables in the *Guide to Using the 2010 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 Report

The *NECAP Student Report* is a single-page, double-sided report printed on 8.5”-by-14” paper. The front of the report includes informational text about the design and uses of the assessment. The front of the report also contains text describing the three corresponding sections on the reverse side of the report and the achievement level descriptions. The reverse side of the *Student Report* provides a complete picture of an individual student’s performance on the NECAP, divided into three sections. The first section provides the student’s overall performance for each content area. The student’s achievement levels are provided, and scaled scores are presented numerically as well as in a graphic that depicts the scaled score with the standard

error of measurement bar constructed about it, set within the full range of possible scaled scores demarcated into the four achievement levels.

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

The third section shows the student's raw score performance in content area reporting categories relative to possible points; gives the average points earned for the school, district, and state; and gives the average points earned by students at the Proficient level on the overall content area test. For reading, with the exception of Word ID/Vocabulary items, items are reported by Type of Text (Literary, Informational) and Level of Comprehension (Initial Understanding, Analysis and Interpretation). For mathematics, the reporting subcategories are Numbers and Operations; Geometry and Measurement; Functions and Algebra; and Data, Statistics, and Probability. 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.

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

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

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 percentage 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, 2010. 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 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 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 testing program.

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

- **Alternate Assessment**: Students in this category completed an alternate test for the 2009–10 school year.
- **First-Year LEP**: Students in this category are defined as being new to the United States after October 1, 2009, and were not required to take the NECAP tests in reading and writing. Students in this category were expected to take the mathematics portion of the NECAP.
- **Withdrew after October 1**: Students withdrawing from a school after October 1, 2010, 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, 2010, 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 2008-09, 2009-10, and 2010-2011 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 2008-09 and the 2010-11 test administrations as well as cumulative data for the two years. Data do not exist for the 2009-10 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 includes:

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

Information about each content area reporting category for reading and mathematics in all grades and item type for writing in grades 5 and 8 includes the following:

- The total possible points for that reporting or item type category. In order to provide as much information as possible for each category, the total number of points includes both the common items used to calculate scores and additional items in each category used for equating the test from year to year.
- A graphic display of the percent of total possible points for the school, district, and state. In this graphic display, there are symbols representing school, district, and state performance. In addition, there is a line representing the standard error of measurement. This statistic indicates how much a student's score could vary if the student were examined repeatedly with the same test (assuming that no learning were to occur between test administrations).

In an effort to provide more information on all the types of writing that are assessed by the NECAP grade 11 writing test, the content area results page was modified and two new additional pages were created. The first content area results page provides data for the 2008-09, 2009-10, and 2010-2011 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 the first content area results page includes a table that lists the type of writing for the common prompt (i.e., the prompt on which the results in the top half of the page are based) for each of the last three test administrations. The type of writing (genre) and a description of that type is included for each year.

The second page of the grade 11 writing content area results reports lists the types of writing that are assessed in the grade 11 writing test. The types of writing are made up of 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 a separate row for the current year (2010-11) and the previous year that each type of writing was assessed. The symbol (C) indicates the type of writing that was common in the fall 2010 test. The number tested and the mean raw score are provided for the school, district, and state. 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. The range of 0 to 12 on the graphic display represents the possible score range for the writing prompt. The 0 to 12 range is a result of adding the two scores assigned to the student's response from the 6-point rubric. The 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 score 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 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 percent of students (%) for each score on the 0 to 12 scale for the school and district. The last column provides the percent (%) 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
- Primary 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 2010–11 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 include 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/content area. A more detailed description of the information included on this report can be found in section 9.4 of this document. Please note that when advanced filtering criteria are applied by the user, the School and District Percent Correct/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/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/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/percent 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 this 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 2010 NECAP scores and achievement levels are provided for each tested student in reading, mathematics, and writing. In addition, fall NECAP 2008 and 2009 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 and 2010 NECAP scores and achievement levels in reading and mathematics, since this is only the second 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 in their school such as teachers, curriculum coordinators, or special education coordinators. 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 2010–11 NECAP are accurate relative to collected data and other pertinent information, a document that delineates analysis and reporting rules was created. These decision rules were observed in the analyses of NECAP test data and in reporting the test results. Moreover, these rules are the main reference for quality assurance checks.

The decision rules document used for reporting results of the October 2010 administration of the NECAP is found in Appendix V.

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.

8.10 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 and Psychometrics 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 receives scaled scores and/or is included in different levels of aggregation are also parallel processed. Using the decision rules document, two data analysts independently write a computer program that assigns students' exclusions. For each content area and grade combination, the exclusions assigned by each data analyst are compared across all students. Only when 100% agreement is achieved can the rest of data analysis be completed.

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

The first set includes those that satisfy the following criteria:

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

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

- Private school
- Small school that receives no school report
- Small district that receives no district report
- District that receives a report but 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 *2010–11 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.

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 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* and *Test Administrator Manuals*.

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, reliability, and scaling and equating 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/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 2010 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/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 impact 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. 2010–11 NECAP: Reading Grade 8 Questionnaire Responses—Difficulty

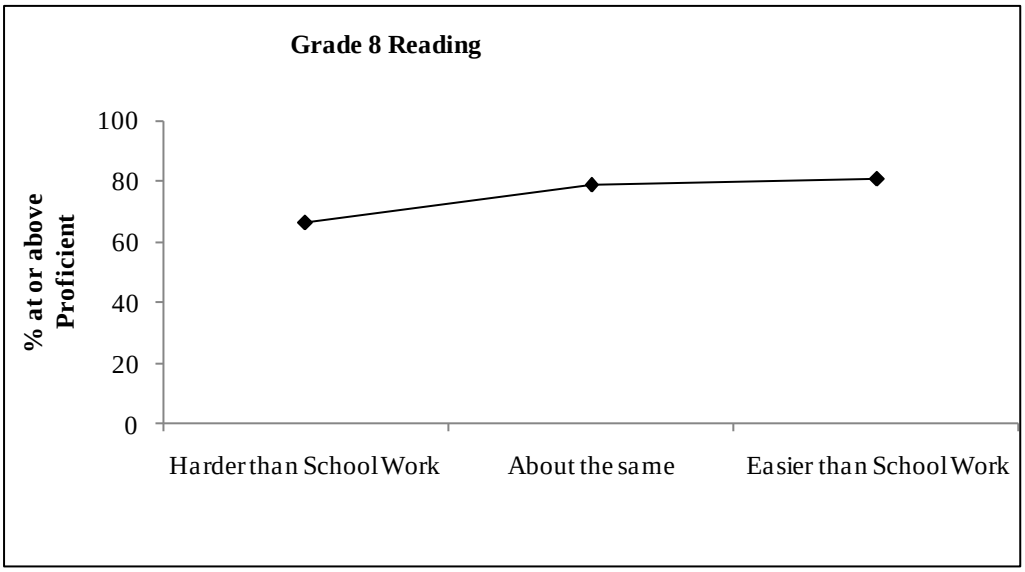
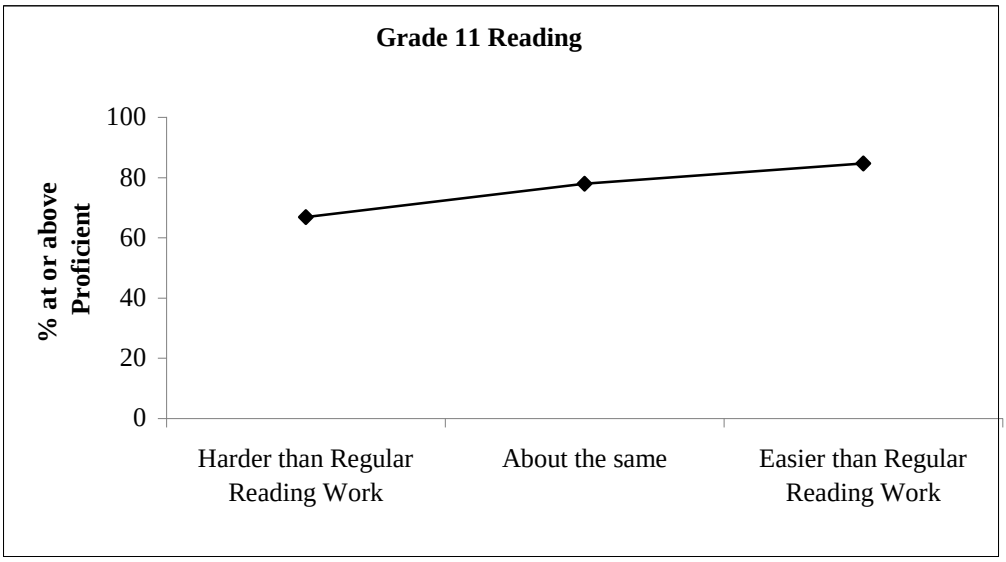


Figure 9-2. 2010–11 NECAP: Reading Grade 11 Questionnaire Responses—Difficulty



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. 2010–11 NECAP: Mathematics Grade 8 Questionnaire Responses—Difficulty

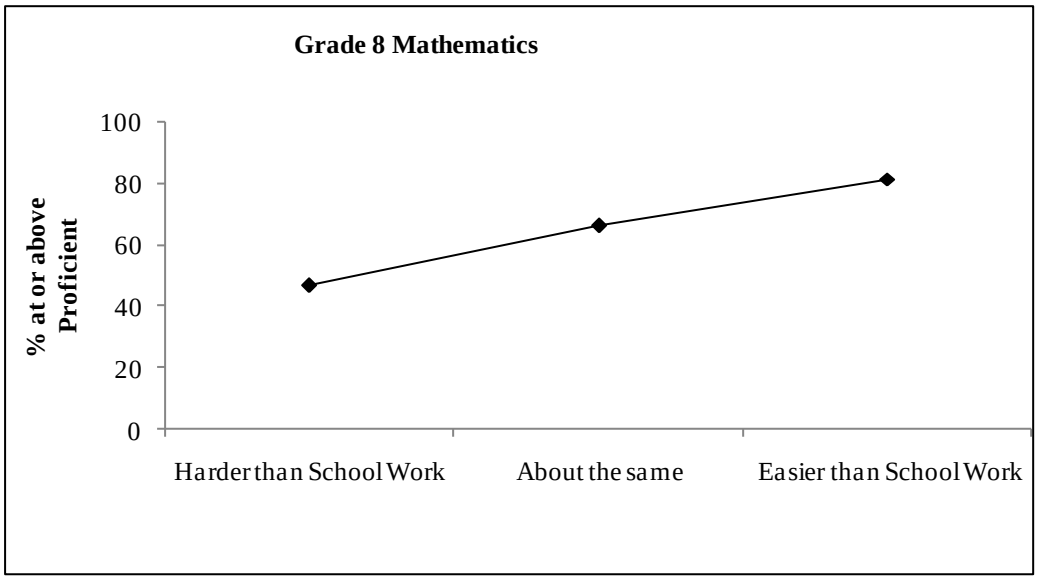
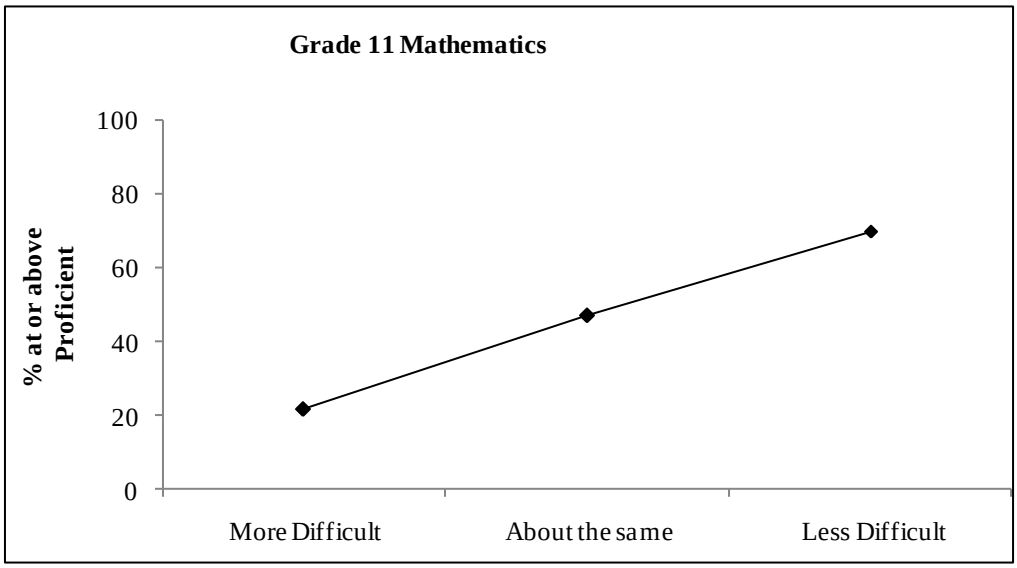


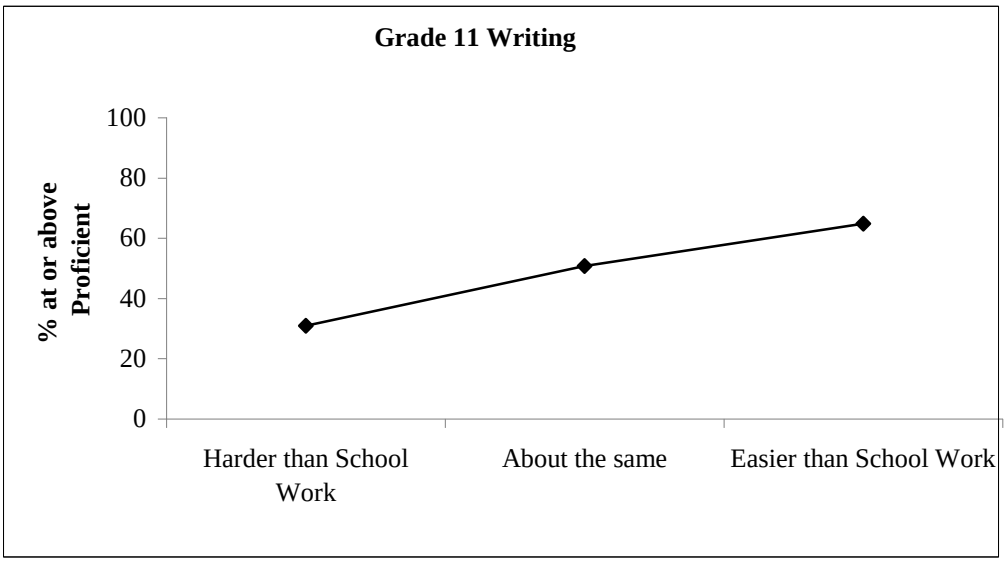
Figure 9-4. 2010–11 NECAP: Mathematics Grade 11 Questionnaire Responses—Difficulty



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. 2010–11 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/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. 2010–11 NECAP: Grade 3 Reading Questionnaire Responses—Content

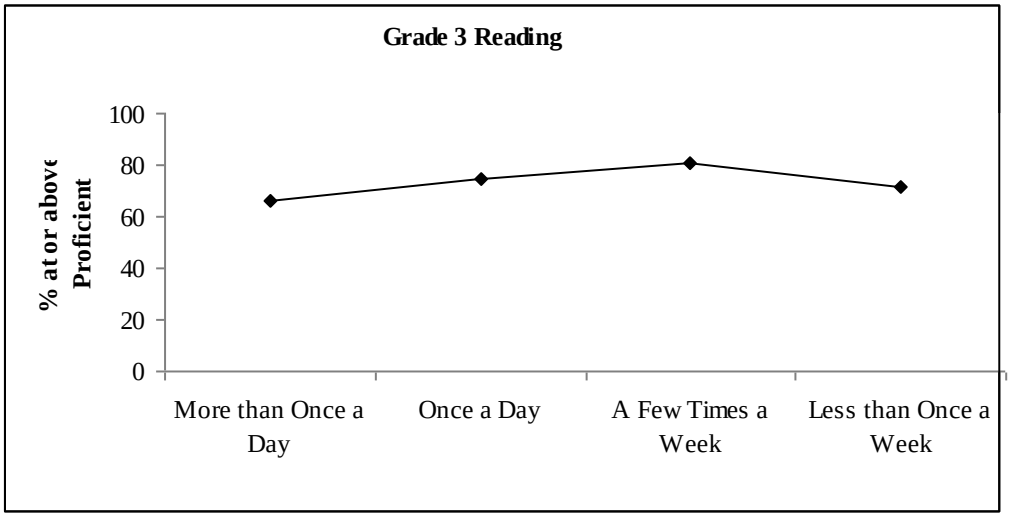


Figure 9-7. 2010–11 NECAP: Grade 4 Reading Questionnaire Responses—Content

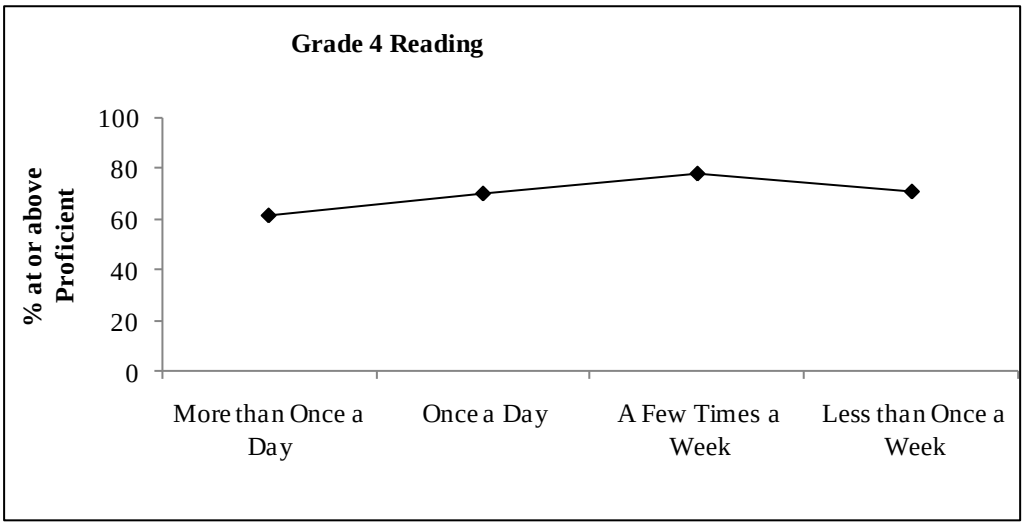


Figure 9-8. 2010–11 NECAP: Grade 7 Reading Questionnaire Responses—Content

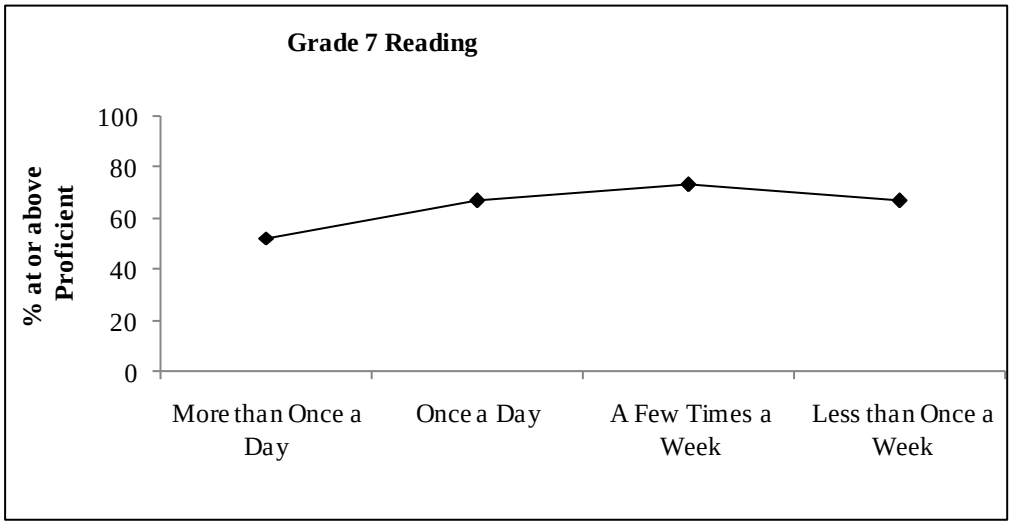
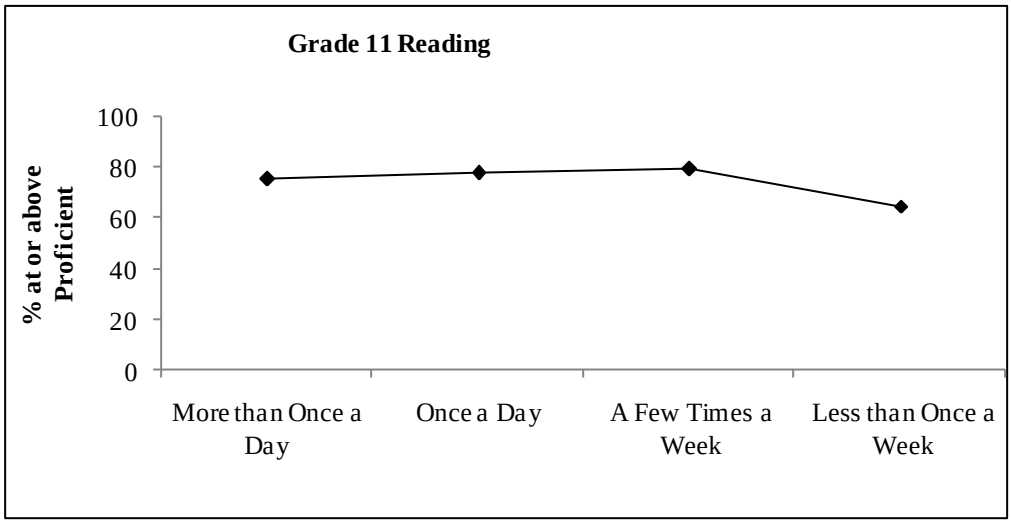


Figure 9-9. 2010–11 NECAP: Grade 11 Reading Questionnaire Responses—Content



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. Grade 7 Mathematics Questionnaire Responses—Content

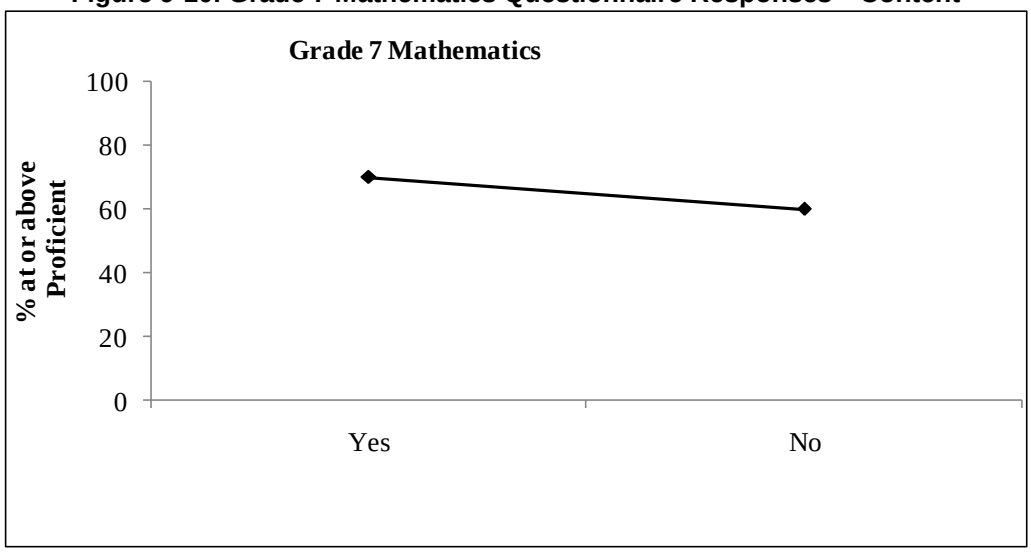


Figure 9-11. Grade 8 Mathematics Questionnaire Responses—Content

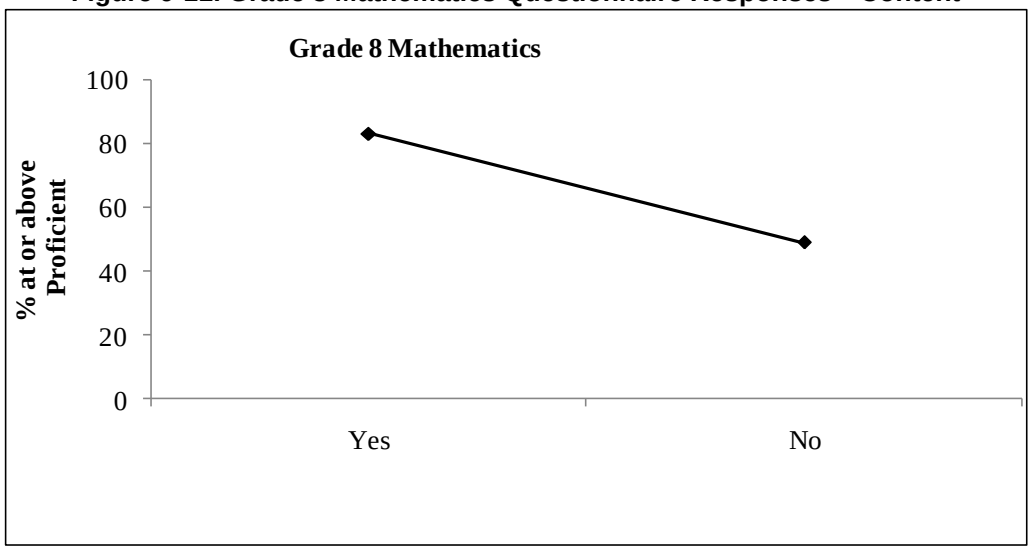
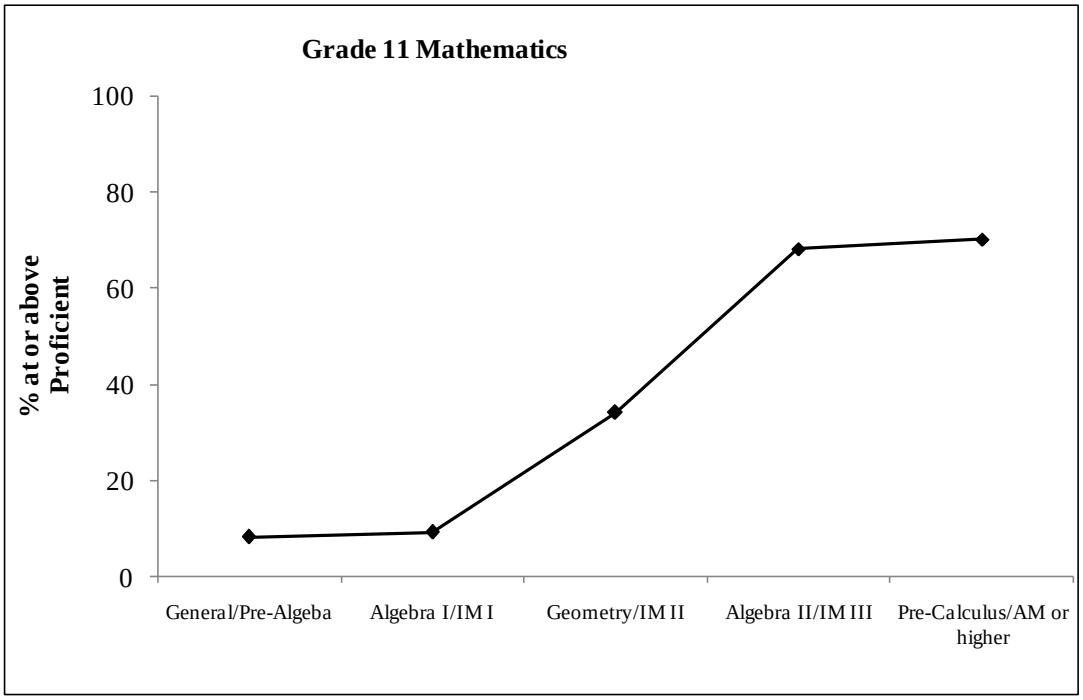


Figure 9-12. Grade 11 Mathematics Questionnaire Responses—Content



9.1.2.3 Content: Writing

Examinees in writing were asked how often they write more than one draft. Figures 9-13 through 10-15 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. 2010–11 NECAP: Grade 5 Writing Questionnaire Responses—Content

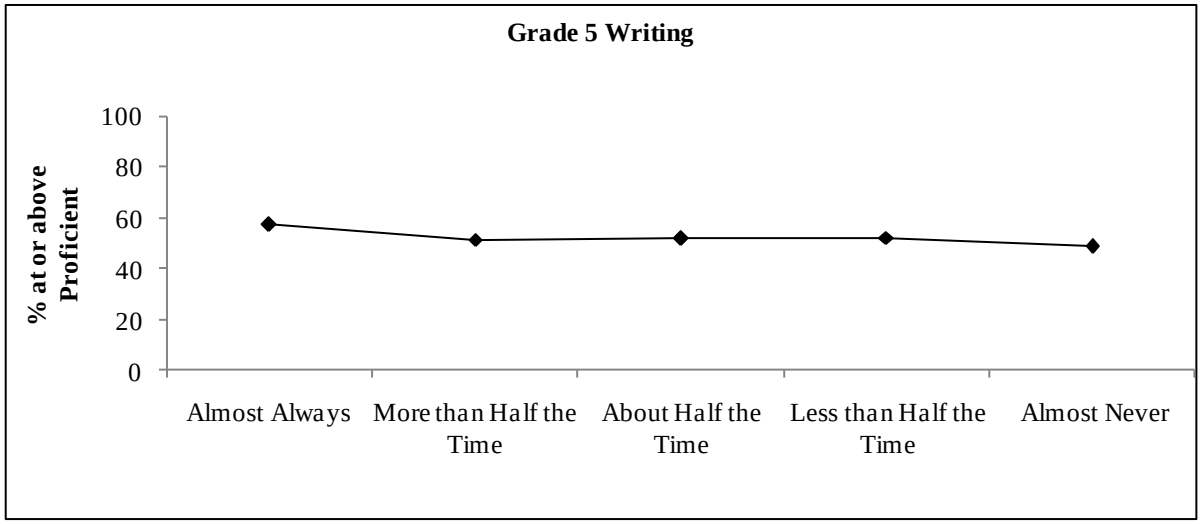


Figure 9-14. 2010–11 NECAP: Grade 8 Writing Questionnaire Responses—Content

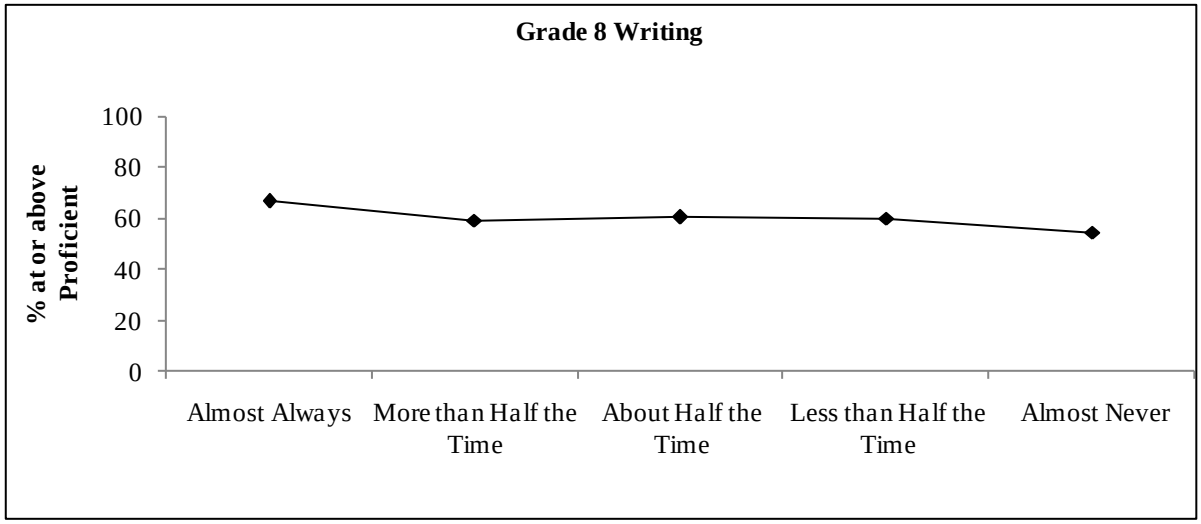
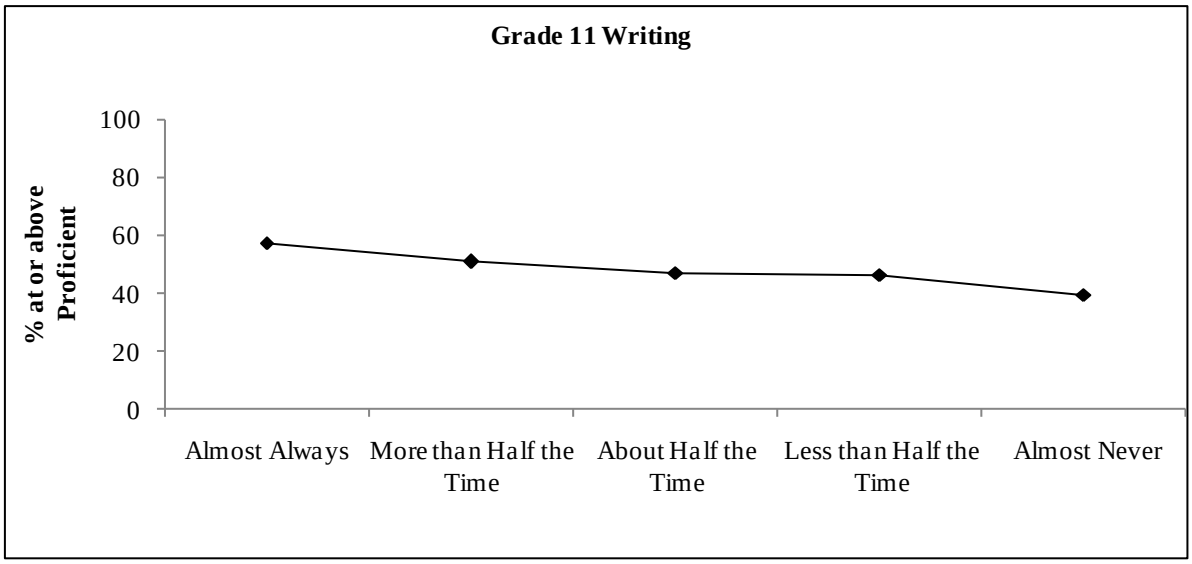


Figure 9-15. 2010–11 NECAP: Grade 11 Writing Questionnaire Responses—Content



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. 2010–11 NECAP: Grade 4 Reading Questionnaire Responses—Homework

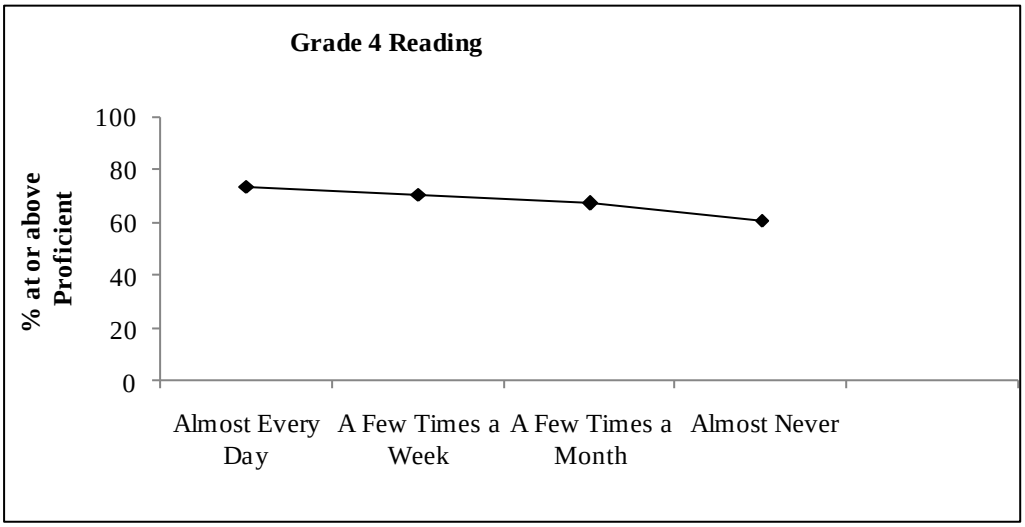


Figure 9-17. 2010–11 NECAP: Grade 7 Reading Questionnaire Responses—Homework

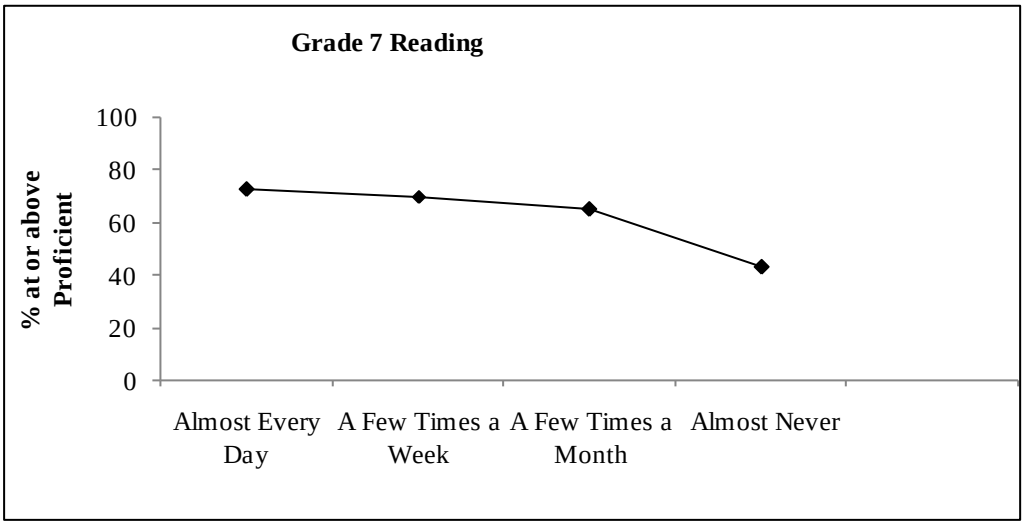
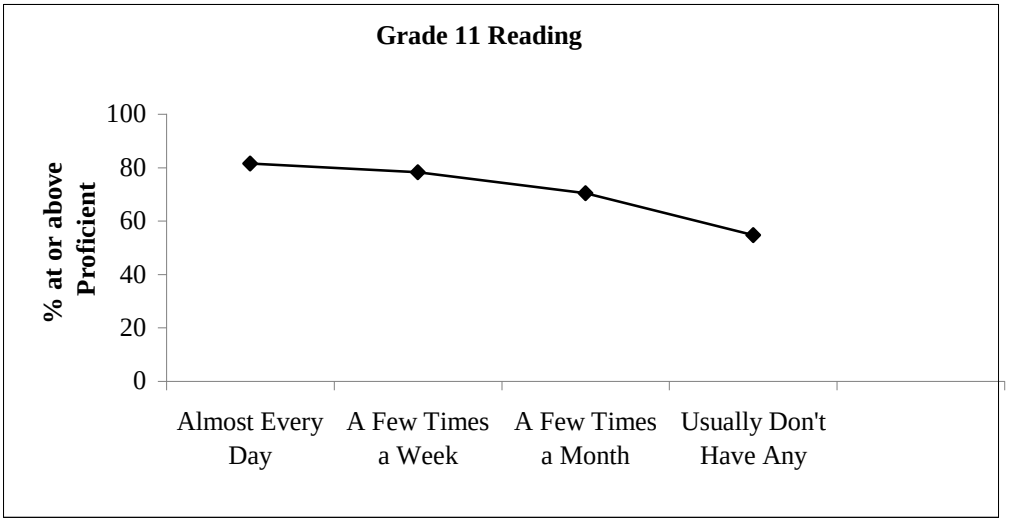


Figure 9-18. 2010–11 NECAP: Grade 11 Reading Questionnaire Responses—Homework



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 that 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. 2010–11 NECAP: Grade 4 Mathematics Questionnaire Responses—Homework

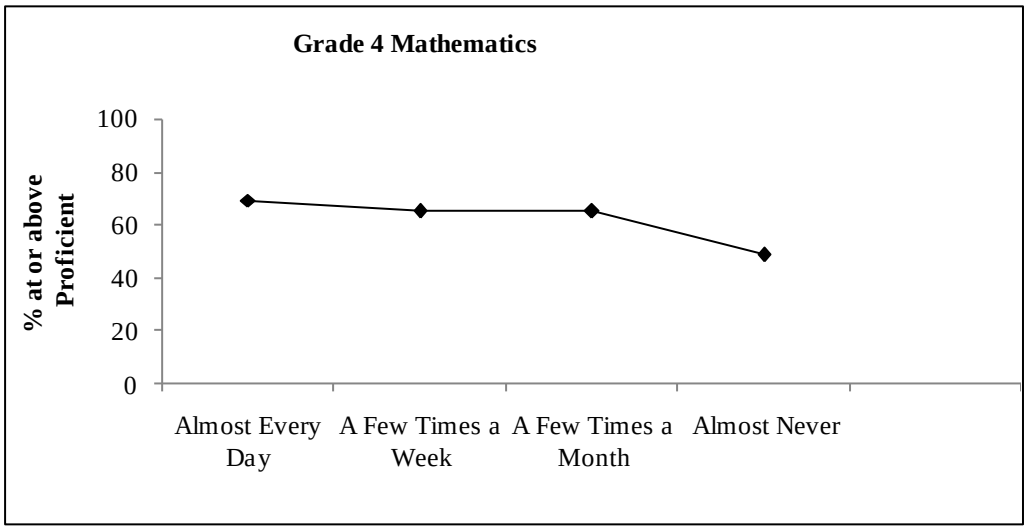


Figure 9-20. 2010–11 NECAP: Grade 5 Mathematics Questionnaire Responses—Homework

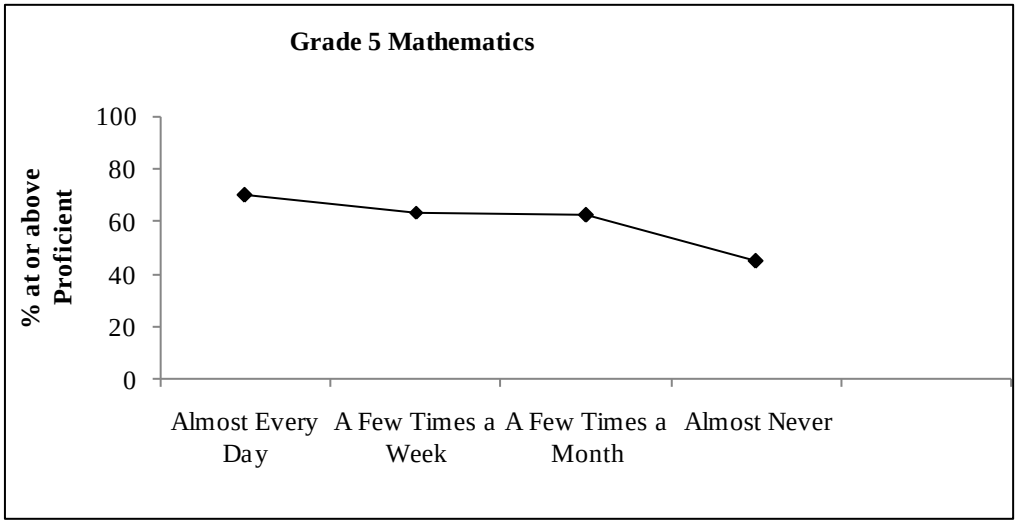


Figure 9-21. 2010–11 NECAP: Grade 8 Mathematics Questionnaire Responses—Homework

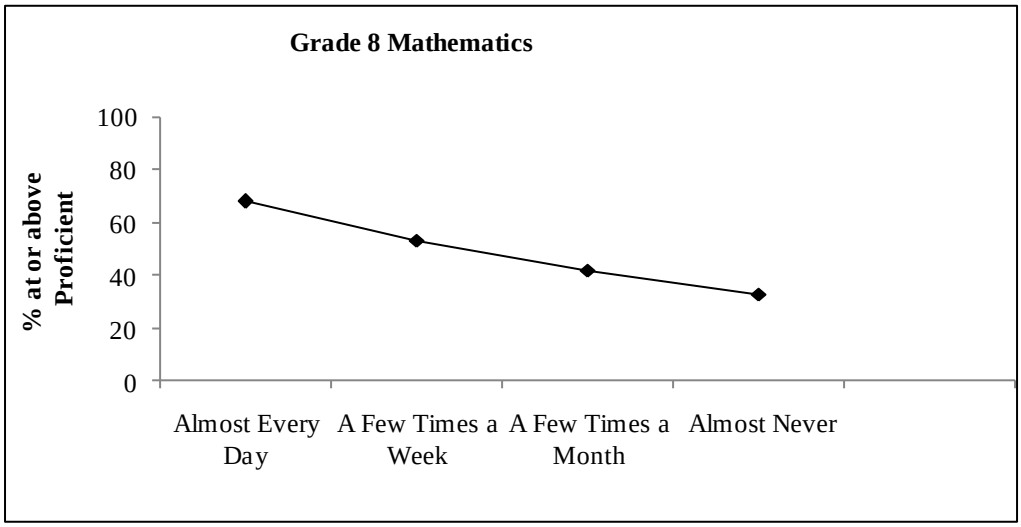
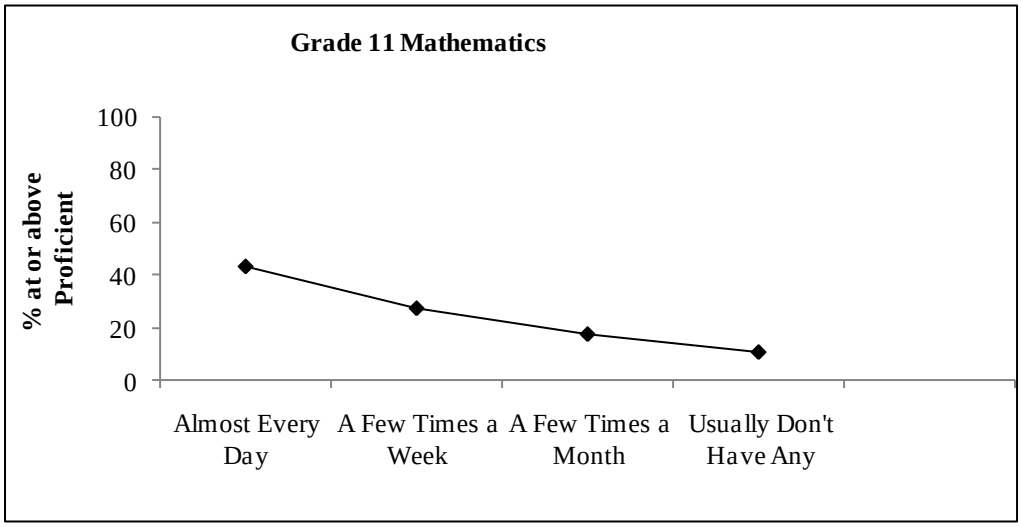


Figure 9-22. 2010–11 NECAP: Grade 11 Mathematics Questionnaire Responses—Homework



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 subject area.

Figure 9-23. 2010–11 NECAP: Grade 11 Questionnaire Responses—Grade in Reading

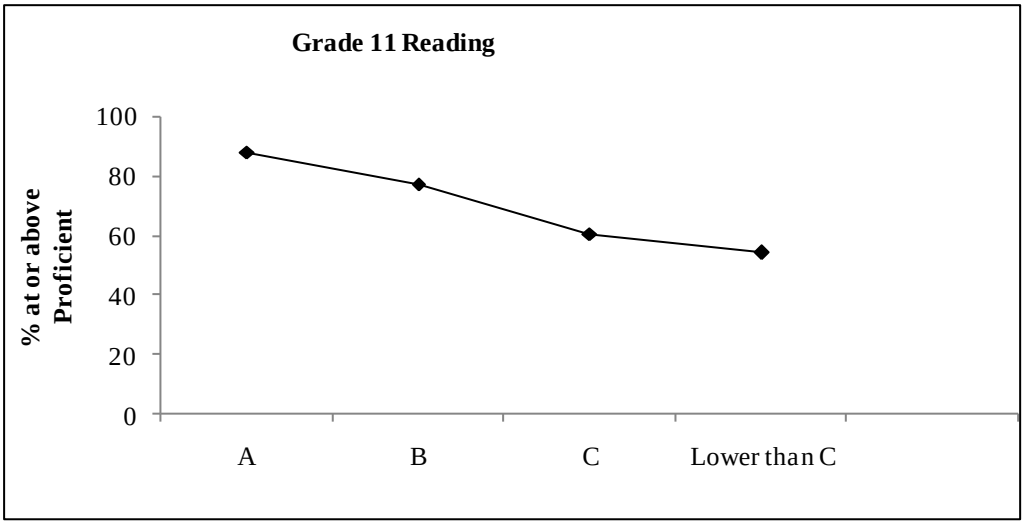
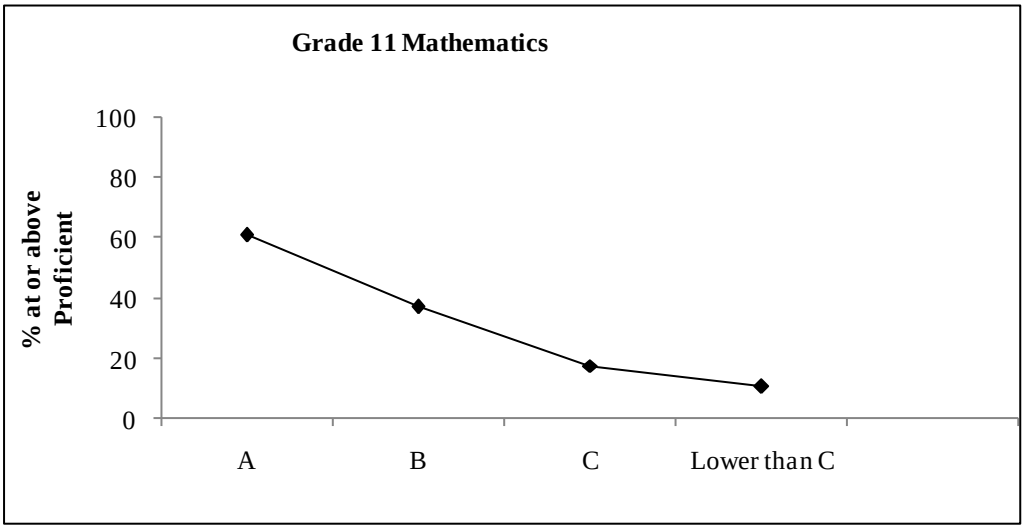


Figure 9-24. 2010–11 NECAP: Questionnaire Responses—Grade in Mathematics



The evidence presented in this report supports inferences made about student achievement on the content represented in the NECAP standards. As such, the evidence provided also supports the use of NECAP results for the purposes of program and instructional improvement and as a component of school accountability.

REFERENCES

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

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

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.

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

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, NJ: 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.

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

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

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

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

Joint Committee on Testing Practices (2004). *Code of fair testing practices in education*. Washington, DC: Joint Committee on Testing Practices. Available from www.apa.org/science/programs/testing/fair-code.aspx.

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, MA: Addison-Wesley.

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

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

APPENDICES

Appendix A	*Committee Membership*
Appendix B	*Participation Rates*
Appendix C	*Accommodation Frequencies by Content Area*
Appendix D	*Table of Standard Accommodations*
Appendix E	*Nimble Accommodations*
Appendix F	*Item-Level Classical Statistics*
Appendix G	*Item-Level Score Point Distributions*
Appendix H	*Differential Item Functioning Results*
Appendix I	*Item Response Theory Calibration Results*
Appendix J	*TCC and TIF Plots*
Appendix K	*Delta Analyses and Rescore Analyses*
Appendix L	*a-Plots and b-Plots*
Appendix M	*2010-11 NECAP Standard Setting Report*
Appendix N	*Performance Level Score Distributions*
Appendix O	*Raw to Scaled Score Look-up Tables*
Appendix P	*Scaled Score Distributions*
Appendix Q	*Classical Reliabilities*
Appendix R	*Interrater Agreement*
Appendix S	*Decision Accuracy and Consistency Results*
Appendix T	*Sample Reports*
Appendix U	*Decision Rules*

Appendix A—Committee Membership

TECHNICAL ADVISORY COMMITTEE MEMBERS
MARCH 24–25, 2010

Name	*Association/Affiliation*
Dale Carlson	NAEP Coach NAEO-Westat
Bill Erpenbach	WJE Consulting Ltd.
Jeff Nellhaus	Massachusetts Department of Elementary and Secondary Education
Jim Pellegrino	University of Illinois
Steve Slater	Oregon Department of Education
Martha Thurlow	Director, National Center on Educational Outcomes
Lauress Wise	Principal Scientist, HumRRO

MAINE ITEM REVIEW COMMITTEE
MARCH 3–5, 2010

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Beth	Alholm	Medomak Middle School	Grades 7 & 8 ELA and Social Studies Teacher
Melissa	Beihn	Hermon Middle School	Literacy Coach
Mary	Belisle	MSAD 51	Grades 6 & 7 Math Teacher
Stephen	Belleau	Pettingill Elementary School	Grade 5 Teacher
Susan	Caler	MSAD 60	Grade 3 Teacher
Wendy	Dunbar	Mt. Jefferson Junior High School	ELA Teacher and Literacy Specialist
Patricia	Friesland	Camden Rockport Middle	Teacher
Michele	Gee	St. George School	Grades 6–8 Math Teacher
Melissa	Goeke	Hope Elementary School	Grade 5 Teacher
Kate	Greeley	MSAD 75	District Math Leadership
Robert	Haskell	Gray-New Gloucester Middle	Grade 4 Teacher
Peg	Hillman	Nathan Clifford School	Grade 4 Elementary Teacher Leader
Ellen	Honan	Yarmouth Elementary School	Grades K–4 Language Arts Teacher
Karin	Howe	Caribou School Department	Literacy Looping Teacher
Karen	Mayo	James Bean School	Grade 3 Teacher
Eleanor	Merrick	Willard School	Principal
Jo Ellen	Merry	Houlton Southside	Grade 5 ELA Teacher
Louise	Moses	Portland District Office	District Math Coordinator & Consultant
Tammy	Schoff	Marshwood Middle School	Literacy Coach
Jane	Smith	Woodland Elementary School	Teaching Principal
Beth	Soules	Benjamin Wentworth	Math Support Teacher
Terry	Taiani	Jordan-Small Middle School	Math Teacher
Amy	Tucker	Winthrop Middle School	Technology Integration Specialist
Gloria	Westhrin	Glenburn School	Grade 4 Teacher

NEW HAMPSHIRE ITEM REVIEW COMMITTEE
MARCH 3–5, 2010

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Keli	Barry	Woodbury Middle School	Grade 6 Math Teacher
Patricia	Barry	Conant High School	English Teacher
Tara	Beauchemin	Chichester Central School	Grade 4 Teacher
Lee	Brazell	Maple Ave. Elementary School	Reading Teacher
Coleen	Bridle	Iber Holmes Gove Middle	Grade 7 Math Teacher
Susan	Dean-Olson	Kingswood Regional High	English Teacher
Linda	Donaldson Guidi	Pennichuck Middle School	Grades K–8 Math Peer Coach
Deborah	Doscher	Pittsfield Middle HS	Math Teacher
Cynthia	Douglass	Oyster River Middle School	Reading Specialist
Denise	Duchesne	Hooksett Memorial School	Retired Grades 4 &5 Teacher
Jack	Finley	Franklin High School	Teacher
Kathy	Fowler	Sandown Central School	TRSD Elementary Math Coach
Katrina	Hall	Hollis Brookline Middle	Grade 7 Math Teacher
Susan	Hewey	Winnisquam Reg Middle	Grade 6 ELA Teacher
Maggie	Holm	Chester Academy	Director of Academic Programs
Gretchen	Kimball	Timberlane Regional Middle	Enrichment
Jeanine	King	Hanover High School	Department Head, Math
Sabrina	Kirwan	Spaulding High School	Math Para
Christian	McDonald	Newport Middle High	Grade 8 English Teacher
Susan	Megas-Russell	Cutler School	Reading Specialist
Kristen	Morey	Barron Elementary School	Grade K–5 Special Education Teacher
Dawn	Mulvey	Josiah Bartlett Elementary	Grades 3 & 4 Teacher
Mary Lou	Noonan	Hollis Upper Elementary	Grade 6 Teacher
Sarah	Potter	Deerfield Community School	Grade 5 Teacher
Mitzi	Sinclair	Richards Elementary School	Elementary Teacher
Greg	Superchi	Lisbon High School	Math & Engineering Teacher
Jeanne	Sturges	Souhegan High School	Writing Teacher
Valerie	Tarbell	Jewett Street School	Grade 4 Teacher
Michelle	Trafton	Milton Elementary School	Grade 5 Teacher
Kristi	Upschulte	Newport Middle High School	High School Math Teacher
Maryann	Vollaro	Effingham Elementary School	Grade 6 Teacher
Brenda	Zarnowski	Moharimet School	Reading Specialist

RHODE ISLAND ITEM REVIEW COMMITTEE
MARCH 3–5, 2010

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Nicholas	Alfred	Jamestown School-Lawn	Grades 7 & 8 Teacher
Carrie	Appel	Beacon Charter High School	Literacy Coach/English Teacher
John	Butler	Beacon Charter High, Woonsocket	Math Teacher
Stephanie	Ciccanesi	Broad Street Elementary School	Grade 5 Reading Coach
Linda	Davis	Hugh B. Bain Middle School	Grades 7 & 8 Reading Consultant
Michelle	Diaz	Westerly Middle School	Grade 6 English Teacher
Cindy	Englehardt	Windmill School, Providence	Math Intervention Specialist
Donna	Gattinella	Dr. Earl F. Calcutt Middle School	Math Coach
Jenny	Gaynor	Hampden Meadows	Grades 4 & 5 Loop Teacher
Arzinia	Gill	Providence	ELA Intervention Specialist
Mariann	Hayward	Peace Dale Elementary School	Math Coach
Lauren	Hopkins	Coventry High School	English/Literacy Teacher
Shelley	Kenny	Chariho Regional High School	High School English Teacher
Melissa	Kerins	J. H. Gaudet, Middletown	Grade 7 Title I Math Teacher
Gina	Kilday	Metcalf Elementary School	Math Teacher
Jessica	Lavallee	Providence Public Schools	ELA Specialist
Kim	Mather	South Kingstown District	Math Coach
Carolyn	Melillo	Dr. Earl F. Calcutt Middle School	Literacy Coach
Barbara	Montecalvo	Fishing Cove Elementary School	Grade 3 Teacher
Barbara	Mooadian	Dr. Earl F. Calcutt Middle School	Literacy Coach
Rebecca	Moore	Mt. Hope High School	English Teacher
Tracey	Nangle	North Smithfield Middle School	Grade 7 Teacher
Patricia	Pora	Leo A. Savoie	Grade 4 Teacher
Margaret	Prior	Portsmouth High School	Math Teacher
Karyn	Rosenfield	Lima School, Providence	Non-Teacher/Assistance Team
Paul	Ruhle	Lincoln High School, Lincoln	Math Teacher
Leslie	Sawicki	Exeter-W. Greenwich Junior High	Math Teacher
Kevin	Seekell	A.S. Feinstein Middle School	Math Curriculum Coordinator
Kristina	Sparfven	Chariho Middle School	Math Teacher
Julee	Thoman	Winman Junior High School	Grades 7 & 8 ELA Teacher
Patricia	Vecchione	Winsor Hill	Grade 5 Literacy Coach
Jennifer	Vorro	Central Office, Providence	Math Intervention Specialist

VERMONT ITEM REVIEW COMMITTEE
MARCH 3–5, 2010

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Carol	Amos	Twinfield Union	Vice Principal K–6, Math Network Leader
Lorraine	Bargmann	Rutland Intermediate School	Reading Coach and Enrichment Teacher
Lisa	Besso	Ensoburg Falls High School	Math Teacher
Carolyn	Bruckner	Oxbow High School	Math Teacher
Suzanne	Buck	Vergennes Union High School	Special Education Teacher
Sharon	Buffey	Thetford Elementary School	Grades K–6 Reading Specialist
Janis	Chamberlain	Caledonia North SU	Literacy Specialist
Shane	Clifford	Grafton	VMI Student, Grades 5 & 6 Teacher
Elaine	Collins	Brownington Central School	Grades 5–8 Teacher, Literacy Leader
Maggie	Eaton	U-32 Middle High School	Language Arts/Teacher Leader
Sandy	Frizzell	North Country High School	English Teacher, Literacy Coach
Janice	Garrow	Rutland Intermediate School	Grades 5 & 6 Literacy Leader
Kelly	Green	Rockingham Central Elementary	Grade 3 Teacher
Dana	Johnson	Rutland City Public Schools	Math Coach
Lindy	Johnson	Washington Central SU	Grades K–12 SU Reading Curriculum Leader
Zachary	Kramer	Colchester High School	Math Teacher
Geoff	Lawrence	Proctor Junior/Senior High School	MS Literacy Teacher
Karen	Lawson	Concord School	Math Recovery Teacher & Support Specialist
Loralyn	Leblanc	Rumney School	Special Education & Literary Consultant
Abigail	Locke	Lyndon Town School	Math Coordinator for District Grades 3–5
Andrea	McLaughlin	Barre Town School	Reading Recovery Literacy Network Leader
Cathy	Newton	Hartford Elementary School	District Literacy Leader & Grade 2 Teacher
Linda	Parker	Windsor Elementary School	Teaches at Anticoch, Castleton
Travis	Redman	Rutland Town School	Grades 6 and 8 Math Teacher
Katy	Schonbeck	Mt. Anthony High School	District Data Committee
Carol	Tina-Sprecht	Hinesburg	Grades 6–8 Math Coordinator
Heather	Stalling	Barre City Elementary School	Grade 2 Teacher
Penny	Sterns	Burlington SU	Grades K–6 Math Coordinator
Catherine	Stoddert	Middlebury Union High School	Grades 9 & 10 English Lead Teacher
Sean	Therot	Franklin NESU	Grades K–12 Math Consultant
Loretta	Whitehead	Caledonia No SU	Grades K–8 Math Coach
Marilyn	Woodard	Mt. Anthony High School	Literature and Composition Teacher

BIAS AND SENSITIVITY COMMITTEE
MARCH 3–5, 2010

MAINE

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Anne	Boucher	Old Town Regional Program	Special Education
Indriani	Demers	Riverton Elementary School	ESL Teacher
Sheree	Granger	The School at Sweetser-Saco	Special Education
Linda	Parkin	Maine Department of Education	Retired
Bruno	Yomoah	Portland Public Schools	ESL Teacher

NEW HAMPSHIRE

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Diane	Bush	Jaffrey-Rindge Middle School	School Counselor
Enchi	Chen	Farmington High School	ESL Teacher
Kathaleen	Cobb	Cutler Elementary School	Grade 4 Teacher
Christine	Leach	Nashua High School-South	Guidance Counselor
Lisa	Witte	Pembroke Academy	Assistant Principal

RHODE ISLAND

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Colleen	Driscoll	Vartan Gregorian Elementary	Reading Coach
Luciana	Lancellotta	Westerly Middle School	Grade 6 Reading, Writing, Special Education
Denise	Perron	Calcutt Middle School	Special Education
Ricardo	Pimentel	Shea High School	ELL Reading/Writing
Emily	Santelises	Woonsocket Middle School	ESL Teacher

VERMONT

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Jackie	Fleming	Rutland Town School	Teacher Language Arts LA/SS
Darlene	Petke	Central Elementary School	Special Educator
Nichole	Pfister	Flood Brook Union School	Science and Language Arts Teacher, Literacy Network Leader
Brenda	Seitz	Vermont Center for the Deaf and Hard of Hearing	Director of Special Education & Admissions
Rebekah	Thomas	J.J. Flynn Elementary School	Grades 2–5 Math & Literacy Teacher

BIAS AND SENSITIVITY COMMITTEE
AUGUST 2–4, 2010

MAINE

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Bruno	Yomoah	Portland Public Schools	ESL Teacher
Sheree	Granger	The School at Sweetser-Saco	Special Educator
Indriana	Demers	East End Community School	Title I and Read 180 Teacher
Melvin	Curtis	Retired	Special Educator
Deborah	Howard	Governor Baxter School for the Deaf	Specialist for Deaf

NEW HAMPSHIRE

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Enchi	Chen	Farmington High School	ESL Teacher
Ashley	Meehan	James Mastricola School	Grade 5 Teacher
Kathaleen	Cobb	Cutler School	Grade 4 Teacher
Alexander	Markowsky	Franklin/Hill Inter-Lakes	School Psychologist
Lisa	Witte	Pembroke Academy	Assistant Principal

RHODE ISLAND

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Sharon	D'Antuono	Greenwood Elementary School	Speech Language Pathologist
Socorro	Gomez-Potter	Reservoir Ave. School	Principal
Kimberly	Lebrun-Steinberg	Hugh B. Bain Middle School	ESL Teacher
Kristen	Twardowski	East Greenwich High School	Teacher
Marybeth	Vierra	Rogers High School	Special Educator

VERMONT

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Colleen	Fiore	Long Trail School	Special Services Director
Sharon	Hunt	Gilman Middle School	Special Educator
Darlene	Petke	Central Elementary School	Special Educator
Susan	Locke		
Susan	Steiner	Bellows Falls Union High School	Biology Teacher

STANDARD-SETTING COMMITTEE MEMBERS
DECEMBER 8–9, 2010

MAINE

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Beth	Ahlholm	Medomak Middle School	Grades 7 & 8 LA/SS Teacher
Wendy	Dunbar	Mt. Jefferson Junior High School	Literacy Specialist/ ELA Teacher
Patricia	Friesland	Camden Rockport Middle School	Teacher
Kathy	Kauffman	Oak Hill Middle School	Grade 8 ELA Teacher
Jo Ellen	Merry	Houlton Southside	Grade 5 Teacher
Jan	Obery	Gilbert Elementary School	Literacy Specialist
Gloria	Westhrin	Glenburn School	Grade 4 Teacher

NEW HAMPSHIRE

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Alison	Bryant	Belmont High School	English Teacher
Cynthia	Douglass	Oyster River Middle School	Reading Specialist
Jill	Duffield	New Durham School	Literacy Teacher
Bonnie	Robinson	Newport Middle High School	Reading Specialist
Michelle	Trafton	Milton Elementary School	Grade 5 Teacher
Kathleen	Woodbury	Gilsum School	Reading Teacher, Special Education Care Manager, Assessment Coordinator

RHODE ISLAND

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Brenda	Aspelund	Aldrich Junior High School	ELA Teacher
Sally	Caruso	Kickemuit Middle School	Grade 8 ELA Teacher
Kathleen	Desrosiers	Warwick School District	Coordinator/ELA/ K–12
Elizabeth	Ferguson	Curtis Corner Middle School	Teacher/Curriculum Coordinator
Melissa	Lourenco	Segue Institute for Learning	Literacy Coach/Title 1
Kevin	Montoya	Segue Institute for Learning	Grade 6 ELA Teacher
Janice	Place	Pilgrim High and Aldrich Junior High Schools	English Dept. Chairperson/Teacher
Alison	Santerre	Kingston Hill Academy	Grade 5 Teacher
Patricia	Vecchione	Winsor Hill School	Literacy Coach/ Reading Specialist

STANDARD-SETTING COMMITTEE MEMBERS
DECEMBER 8–9, 2010

VERMONT

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Elaine	Collins	Browington Central School	Grades 6–8 Humanities Teacher
Maggie	Eaton	U-32 Middle/High School	Teacher/Curriculum Leader
Janice	Garrow	Rutland Intermediate School	Grade 6 Teacher
Kathryn	Grace	Founders School	Literacy Teacher Leader
Geoff	Lawrence	Proctor Junior/Senior High School	MS English Teacher
Kimberly	Means	Stern Center	Special Educator
Carol	Owen	Orange East SU	Literacy Teacher Leader

Appendix B—Participation Rates

Table B-1. 2010–11 NECAP: Summary of Participation by Demographic Category – Reading

Description	*Number:*			*Percent Tested*
	Enrolled	*Approved Exemptions*	*Tested*	
All Students	309214	5033	301694	97.57
Male	159830	3110	155255	97.14
Female	149348	1923	146405	98.03
Gender Not Reported	36	0	34	94.44
Hispanic or Latino	21153	601	20307	96.00
American Indian or Alaskan Native	1988	41	1919	96.53
Asian	6785	225	6477	95.46
Black or African American	11114	412	10575	95.15
Native Hawaiian or Pacific Islander	262	10	248	94.66
White (non-Hispanic)	262891	3618	257374	97.90
Two or More Races (non-Hispanic)	3365	51	3287	97.68
No Primary Race/Ethnicity Reported	1656	75	1507	91.00
Currently Receiving LEP services	9028	930	7924	87.77
Former LEP Student - Monitoring Year 1	1276	6	1264	99.06
Former LEP Student - Monitoring Year 2	587	2	581	98.98
LEP: All Other Students	298323	4095	291925	97.86
Students with an IEP	49199	3631	44546	90.54
IEP: All Other Students	260015	1402	257148	98.90
Economically Disadvantaged Students	112855	2756	109032	96.61
SES: All Other Students	196359	2277	192662	98.12
Migrant Students	101	3	98	97.03
Migrant: All Other Students	309113	5030	301596	97.57
Students Receiving Title 1 Services	48926	934	47691	97.48
Title 1: All Other Students	260288	4099	254003	97.59
Plan 504	2884	15	2857	99.06
Plan 504: All Other Students	306330	5018	298837	97.55

Table B-2. 2010–11 NECAP: Summary of Participation by Demographic Category – Mathematics

Description	*Number:*			*Percent Tested*
	Enrolled	*Approved Exemptions*	*Tested*	
All Students	309214	4092	302556	97.85
Male	159830	2583	155704	97.42
Female	149348	1509	146818	98.31
Gender Not Reported	36	0	34	94.44
Hispanic or Latino	21153	246	20714	97.92
American Indian or Alaskan Native	1988	30	1929	97.03
Asian	6785	78	6661	98.17
Black or African American	11114	165	10823	97.38
Native Hawaiian or Pacific Islander	262	5	252	96.18
White (non-Hispanic)	262891	3471	257357	97.89
Two or More Races (non-Hispanic)	3365	43	3291	97.80
No Primary Race/Ethnicity Reported	1656	54	1529	92.33
Currently Receiving LEP services	9028	87	8855	98.08
Former LEP Student - Monitoring Year 1	1276	2	1268	99.37
Former LEP Student - Monitoring Year 2	587	2	581	98.98
LEP: All Other Students	298323	4001	291852	97.83
Students with an IEP	49199	3534	44532	90.51

Description	Number:			Percent Tested
	Enrolled	Approved Exemptions	Tested	
IEP: All Other Students	260015	558	258024	99.23
Economically Disadvantaged Students	112855	2109	109574	97.09
SES: All Other Students	196359	1983	192982	98.28
Migrant Students	101	3	97	96.04
Migrant: All Other Students	309113	4089	302459	97.85
Students Receiving Title 1 Services	46967	475	46204	98.38
Title 1: All Other Students	262247	3617	256352	97.75
Plan 504	2884	15	2854	98.96
Plan 504: All Other Students	306330	4077	299702	97.84

Table B-3. 2010–11 NECAP: Summary of Participation by Demographic Category – Writing

Description	Number:			Percent Tested
	Enrolled	Approved Exemptions	Tested	
All Students	126625	2020	122639	96.85
Male	65542	1218	63134	96.33
Female	61065	802	59487	97.42
Gender Not Reported	18	0	18	100.00
Hispanic or Latino	8556	231	8121	94.92
American Indian or Alaskan Native	738	15	710	96.21
Asian	2740	91	2592	94.60
Black or African American	4581	158	4325	94.41
Native Hawaiian or Pacific Islander	123	6	113	91.87
White (non-Hispanic)	107920	1462	104938	97.24
Two or More Races (non-Hispanic)	1300	22	1256	96.62
No Primary Race/Ethnicity Reported	667	35	584	87.56
Currently Receiving LEP services	3131	343	2677	85.50
Former LEP Student - Monitoring Year 1	421	0	418	99.29
Former LEP Student - Monitoring Year 2	299	2	293	97.99
LEP: All Other Students	122774	1675	119251	97.13
Students with an IEP	20677	1379	18481	89.38
IEP: All Other Students	105948	641	104158	98.31
Economically Disadvantaged Students	42961	1011	41104	95.68
SES: All Other Students	83664	1009	81535	97.46
Migrant Students	35	1	34	97.14
Migrant: All Other Students	126590	2019	122605	96.85
Students Receiving Title 1 Services	16248	328	15703	96.65
Title 1: All Other Students	110377	1692	106936	96.88
Plan 504	1237	8	1218	98.46
Plan 504: All Other Students	125388	2012	121421	96.84

Appendix C—Accommodation Frequencies by Content Area

Table C-1. 2010–11 NECAP: Numbers of Students Tested with and without Accommodations by Subject and Grade

Subject	*Grade*	*Number of Students Tested:*	
		with Accommodations	*without Accommodations*
Mathematics	3	35283	8610
	4	35266	9084
	5	34699	9508
	6	36523	7954
	7	38695	7841
	8	39140	7427
	11	28584	5003
Reading	3	35377	8359
	4	35437	8769
	5	35015	9016
	6	36591	7738
	7	38805	7604
	8	39094	7362
	11	28684	4216
Writing	5	35329	8627
	8	39298	6976
	11	29295	3491

Table C-2. 2010–11 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	3581	3882	4926	3837	4175	3659	2738
mataccomT2	166	179	186	231	158	150	144
mataccomT3	3144	3313	3364	2603	2415	2128	667
mataccomT4	189	255	215	202	143	155	100
mataccomS1	5865	6384	6272	4546	3874	3886	2037
mataccomS2	43	36	41	41	41	41	29
mataccomP1	1206	1262	1186	897	727	639	210
mataccomP2	5962	6329	6360	5334	5058	5059	1783
mataccomP3	5013	5019	4771	2975	2466	2234	379
mataccomP4	530	451	443	451	524	468	221
mataccomP5	2669	2855	2682	1960	1663	1621	264
mataccomP6	17	24	21	12	15	19	29
mataccomP7	3071	3331	3258	2038	1636	1500	554
mataccomP8	52	64	48	44	40	44	12
mataccomP9	2	2	1	3	1	4	1
mataccomP10	23	28	27	45	48	66	7
mataccomP11	304	246	198	91	39	34	18
mataccomR1	1725	1754	1739	1046	696	516	32
mataccomR2	25	25	26	25	16	16	8
mataccomR3	370	429	266	151	119	88	28
mataccomR4	22	29	95	88	128	179	25
mataccomR5	108	107	131	84	67	53	19
mataccomR6	23	14	16	11	12	11	2
mataccomR7	26	12	22	41	40	38	11
mataccomO1	0	1	1	1	2	1	5

Accommodation Code	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
mataccomM1	13	7	34	19	34	28	54
mataccomM3	3	2	1	2	3	14	2
mataccomN01							285
mataccomN02							355
mataccomN03							29
mataccomN04							197
mataccomN05							52
mataccomN06							31
mataccomN07							40
mataccomN08							222

Table C-3. 2010–11 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	3627	3906	4619	3735	3977	3667	1951
reaaccomT2	173	186	180	234	161	153	62
reaaccomT3	3177	3337	3361	2611	2428	2136	657
reaaccomT4	185	256	205	209	144	156	94
reaaccomS1	5746	6213	6108	4562	3866	3918	2038
reaaccomS2	49	34	38	41	42	43	25
reaaccomP1	1214	1248	1184	915	728	625	201
reaaccomP2	5791	6104	6180	5311	5065	5081	1788
reaaccomP4	781	608	600	531	515	494	225
reaaccomP5	2608	2743	2577	1936	1601	1563	271
reaaccomP6	18	21	20	19	15	19	31
reaaccomP7	2972	3224	3171	2006	1604	1468	554
reaaccomP8	50	65	47	45	42	43	13
reaaccomP9	2	2	1	3	2	4	1
reaaccomP11	365	323	229	107	52	43	17
reaaccomR1	1967	2102	1989	1395	942	637	53
reaaccomR2	27	29	26	38	23	23	11
reaaccomR3	421	451	281	152	123	90	28
reaaccomR4	30	54	134	169	284	383	78
reaaccomR5	118	110	133	93	66	50	19
reaaccomR6	24	24	15	14	14	14	2
reaaccomO1	9	8	1	1	2	2	7
reaaccomM2	12	6	8	16	3	5	0
reaaccomM3	1	1	1	1	3	1	2
reaaccomN01							307
reaaccomN02							29
reaaccomN03							212
reaaccomN04							56
reaaccomN05							32
reaaccomN06							40
reaaccomN07							237

Table C-4. 2010–11 NECAP: Numbers of Students Tested with Accommodations by Accommodation Type and Subject – Writing

Accommodation Code	*Grade 5*	*Grade 8*	*Grade 11*
wriaccomT1	4105	3257	1343
wriaccomT2	151	131	49
wriaccomT3	3062	2004	691
wriaccomT4	194	140	94
wriaccomS1	5794	3708	2066
wriaccomS2	40	40	26
wriaccomP1	1039	575	203
wriaccomP2	5957	4857	1816
wriaccomP3	4290	2067	404
wriaccomP4	392	462	237
wriaccomP5	2480	1560	267
wriaccomP6	21	20	30
wriaccomP7	3064	1405	566
wriaccomP8	46	42	11
wriaccomP9	1	4	1
wriaccomP10	15	35	4
wriaccomP11	174	33	15
wriaccomR3	245	81	31
wriaccomR4	242	508	146
wriaccomR5	117	45	14
wriaccomR6	12	9	2
wriaccomR7	11	22	7
wriaccomO1	6	1	5
wriaccomO2	50	50	13
wriaccomM3	30	24	4

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—NIMBLE ACCOMMODATIONS

NimbleTools® Accessibility Tools and Accommodations

Accessibility Tool Descriptions

Read Aloud (accommodation code - mataccomN01)
NimbleTools® links pre-recorded human voice recordings to test items presented to students. Students benefit by listening to a fully approved, standardized human voice, assuring correct pronunciation of words, symbols, and equations. NimbleTools® empowers students to decide when they want to hear the text read to them, and allows them to play sound clips repeatedly. All buttons and directions have human-read sound clips associated with them. The Text & Graphics version describes graphics and diagrams (Grade 11 only).

Auditory Calming (Background Music) (accommodation code - mataccomN02 and reaaccomN01)
For students who focus better when receiving auditory input, music or sounds can be provided during testing. These sounds are embedded into the system, so no extra hardware is needed, and there are no concerns about monitoring the content. The player is simple to use, without distracting visuals of many commercial computer MP3 players.

Magnifier (accommodation code - mataccomN03 and reaaccomN02)
The Magnifier Tool allows students to enlarge the entire test interface. Students have control over when and where they use this tool. The tool options are shown in enlarged, high contrast text. This should only be assigned to students who need the entire test enlarged throughout the test, as it can be disorienting for students who are not used to working this way.

Color Overlay (accommodation code - mataccomN04 and reaaccomN03)
Students can choose from a variety of color tints which are placed over the questions and directions of the test. Many students find their reading accuracy and speed increases with the use of color overlays.

Reverse Contrast (accommodation code - mataccomN05 and reaaccomN04)
Students can choose to reverse the colors for the entire test interface. You also have the option of adding a color tint to the question text using the Color Overlay Tool (see above) which is automatically included when you choose Reverse Contrast.

Color Chooser (accommodation code - mataccomN06 and reaaccomN05)
Students can change the font and background colors for the test questions and direction. Students pick the font and background color combinations from a palette of colors proven to help students. This differs from Color Overlay in that only the text and background colors change. Lines and graphics are not affected by the color changes.

Custom & Answer Masking (accommodation code - mataccomN07 and reaaccomN06)
A common technique for focusing a student's attention on a specific part of a test item is provided by the Masking Tools. Two masking tools are currently available: Answer Masking and Custom Masking. Answer Masking hides the answers until students have an opportunity to solve the problem and then allows students to reveal answer choices individually or all at once. Custom Masking allows students to create and place 'sheets' on top of any part of the test question, masking those parts of the question they don't want to focus on. Both Masking options increase students' focus on the test question by

temporarily hiding all other test elements. Masking can be turned on and off at the student's discretion.

Test Accommodations

Allow Secure and Supervised Breaks

Students are allowed to log out of a test session, take a break, then continue with the same session by logging back in. Answers will be saved and restored when they return. This should only be assigned to students who specifically require breaks, as their test session will require more careful supervision (to ensure test security).

NimbleTools®

Appendix F—Item-Level Classical Statistics

Table F-1. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Mathematics Grade 3

Item Number	*Item Type*	*Difficulty*	*Discrimination*	*Percent Omitted*
119683	MC	0.69	0.46	1
119685	MC	0.80	0.50	1
119688	SA	0.72	0.43	1
119692	MC	0.74	0.32	1
119696	MC	0.89	0.44	1
119706	MC	0.84	0.49	2
119731	SA	0.82	0.49	1
119768	MC	0.73	0.43	1
119810	MC	0.78	0.49	1
119811	MC	0.86	0.46	1
119821	MC	0.78	0.36	1
119825	MC	0.77	0.46	1
119838	SA	0.40	0.54	1
119839	MC	0.64	0.53	1
119840	MC	0.70	0.43	3
119848	MC	0.92	0.38	0
119869	MC	0.58	0.44	3
119882	MC	0.74	0.45	0
119885	MC	0.77	0.46	0
119889	SA	0.51	0.52	1
119891	MC	0.46	0.23	1
119918	SA	0.64	0.59	1
122862	MC	0.63	0.51	0
122872	SA	0.95	0.29	1
124370	MC	0.86	0.43	1
124431	MC	0.57	0.39	1
124433	MC	0.87	0.40	0
124436	MC	0.79	0.50	1
124459	SA	0.75	0.27	1
139548	MC	0.73	0.34	1
139560	MC	0.89	0.36	1
139563	MC	0.55	0.49	2
139569	MC	0.87	0.46	1
139583	SA	0.61	0.49	1
139590	SA	0.33	0.44	1
139603	SA	0.69	0.59	2
139641	MC	0.50	0.51	1
139645	SA	0.65	0.45	1
139653	SA	0.48	0.46	1
139655	SA	0.48	0.57	1
139663	MC	0.36	0.40	1
139665	MC	0.56	0.37	1
139669	SA	0.62	0.40	1
139673	SA	0.66	0.39	1
139687	MC	0.56	0.51	0
139689	SA	0.54	0.47	1
144612	SA	0.26	0.39	2
144617	SA	0.83	0.29	1
144623	SA	0.42	0.56	1
145249	SA	0.70	0.52	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
145509	MC	0.39	0.33	0
145526	MC	0.73	0.57	1
145528	MC	0.42	0.35	1
145678	MC	0.91	0.37	0
145683	MC	0.90	0.41	0

Table F-2. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Mathematics Grade 4

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
120075	SA	0.62	0.56	1
120087	MC	0.82	0.33	0
120119	SA	0.88	0.29	1
120127	SA	0.81	0.35	0
120144	MC	0.84	0.45	1
120209	MC	0.78	0.36	2
120211	MC	0.53	0.47	1
120215	MC	0.28	0.36	2
120219	MC	0.51	0.43	2
120220	MC	0.82	0.50	2
120221	SA	0.77	0.38	1
120222	SA	0.50	0.52	2
120240	SA	0.61	0.50	2
120251	MC	0.87	0.42	2
120255	MC	0.71	0.47	2
120261	MC	0.82	0.29	2
120262	SA	0.44	0.46	1
120276	MC	0.39	0.22	1
120286	SA	0.78	0.47	1
120288	MC	0.95	0.26	0
123185	MC	0.76	0.47	1
124522	MC	0.91	0.26	2
124525	MC	0.87	0.43	1
124527	MC	0.51	0.29	1
124560	MC	0.74	0.42	1
124616	MC	0.82	0.43	1
124661	SA	0.45	0.53	1
124723	SA	0.67	0.55	1
124727	MC	0.73	0.39	1
139424	MC	0.72	0.44	1
139428	MC	0.43	0.45	1
139431	MC	0.70	0.44	3
139437	MC	0.77	0.43	2
139447	SA	0.43	0.55	2
139452	SA	0.56	0.56	1
139455	MC	0.57	0.41	1
139466	MC	0.67	0.39	1
139477	MC	0.81	0.26	1
139479	SA	0.28	0.34	1
139493	MC	0.39	0.47	1
139495	MC	0.27	0.21	2
139517	SA	0.47	0.48	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
139523	SA	0.67	0.33	1
139527	SA	0.67	0.51	1
139529	SA	0.58	0.39	1
144648	MC	0.68	0.34	1
145070	MC	0.89	0.32	0
145085	SA	0.77	0.43	1
145088	MC	0.39	0.31	2
145545	SA	0.80	0.47	1
145554	MC	0.34	0.28	1
145559	SA	0.38	0.57	1
145804	MC	0.87	0.41	1
145847	MC	0.66	0.44	3
145860	MC	0.69	0.17	2

Table F-3. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Mathematics Grade 5

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
120624	MC	0.24	0.23	0
120626	MC	0.53	0.43	0
120628	MC	0.86	0.32	0
120634	MC	0.67	0.45	1
120642	MC	0.85	0.37	0
120678	MC	0.58	0.41	0
120706	MC	0.45	0.39	0
120725	SA	0.50	0.58	1
120726	SA	0.41	0.56	3
120728	MC	0.67	0.36	0
120736	MC	0.51	0.51	1
120748	MC	0.55	0.19	0
120795	MC	0.51	0.35	0
120799	MC	0.90	0.24	0
120804	MC	0.87	0.38	1
124804	SA	0.65	0.45	0
124858	CR	0.32	0.63	1
124862	MC	0.36	0.40	0
124866	MC	0.76	0.37	0
124881	MC	0.53	0.43	1
139086	MC	0.40	0.49	0
139129	SA	0.58	0.53	1
139136	SA	0.53	0.57	1
139141	SA	0.44	0.51	1
139161	MC	0.78	0.36	0
139181	MC	0.68	0.53	0
139203	SA	0.25	0.46	3
139283	CR	0.77	0.49	1
139344	MC	0.64	0.45	0
139384	CR	0.59	0.68	1
139399	MC	0.88	0.39	0
139403	MC	0.52	0.33	1
139414	MC	0.55	0.47	0
144657	MC	0.69	0.45	0

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
144659	MC	0.59	0.48	1
145096	MC	0.38	0.47	1
145571	SA	0.50	0.54	1
145573	MC	0.86	0.42	0
145576	CR	0.44	0.64	1
145584	MC	0.57	0.43	0
145585	MC	0.43	0.48	0
145879	SA	0.69	0.41	0
145887	SA	0.62	0.41	1
145888	SA	0.57	0.53	2
145893	MC	0.79	0.50	0
145910	MC	0.68	0.42	0
145922	SA	0.58	0.41	1
150808	MC	0.77	0.15	0

Table F-4. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Mathematics Grade 6

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
119195	SA	0.28	0.55	2
119221	MC	0.53	0.42	0
119222	MC	0.40	0.53	1
119230	SA	0.67	0.55	3
119231	MC	0.64	0.32	0
119245	MC	0.79	0.45	0
119279	MC	0.45	0.39	0
119288	MC	0.55	0.24	0
119295	MC	0.82	0.42	0
119306	SA	0.75	0.50	0
119311	MC	0.53	0.38	0
119317	MC	0.52	0.32	1
119334	MC	0.90	0.35	0
119361	MC	0.68	0.47	0
119375	MC	0.57	0.36	0
123489	CR	0.47	0.62	1
125021	MC	0.56	0.53	0
125022	MC	0.53	0.47	0
139212	MC	0.58	0.53	0
139217	MC	0.50	0.43	0
139219	MC	0.74	0.45	0
139274	MC	0.71	0.45	0
139305	MC	0.78	0.53	0
139306	MC	0.55	0.43	0
139310	SA	0.44	0.54	1
139312	SA	0.43	0.47	0
139320	CR	0.23	0.52	2
139335	MC	0.58	0.30	0
139356	CR	0.38	0.67	1
139358	MC	0.81	0.47	0
139377	SA	0.49	0.54	1
139387	CR	0.45	0.68	1
144707	SA	0.23	0.56	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
145128	MC	0.63	0.53	0
145133	SA	0.51	0.46	2
145326	MC	0.48	0.39	0
145328	MC	0.75	0.31	0
145330	MC	0.53	0.52	0
145346	MC	0.69	0.43	2
145347	SA	0.67	0.42	0
145608	MC	0.82	0.39	0
145969	MC	0.57	0.38	0
145978	MC	0.47	0.32	1
145988	MC	0.55	0.39	0
145995	SA	0.50	0.58	1
146011	SA	0.39	0.46	1
146037	MC	0.82	0.41	0
197550	SA	0.76	0.55	2

Table F-5. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Mathematics Grade 7

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
120345	MC	0.55	0.26	1
120353	MC	0.60	0.31	1
120357	MC	0.68	0.25	0
120361	MC	0.74	0.45	0
120362	MC	0.57	0.29	0
120363	MC	0.52	0.42	1
120381	MC	0.77	0.42	1
120387	SA	0.59	0.43	0
120406	SA	0.49	0.54	2
120419	MC	0.60	0.43	1
120427	SA	0.33	0.53	3
120431	MC	0.54	0.45	0
120446	MC	0.47	0.50	0
120450	MC	0.75	0.31	0
120468	SA	0.41	0.47	2
120470	SA	0.31	0.66	2
120485	MC	0.44	0.41	1
120486	MC	0.43	0.37	0
120494	MC	0.40	0.34	0
120496	MC	0.46	0.23	1
120509	MC	0.75	0.41	0
120531	MC	0.66	0.24	1
122332	MC	0.53	0.23	0
125382	SA	0.40	0.63	1
125397	SA	0.68	0.45	0
139925	MC	0.54	0.45	0
139980	SA	0.27	0.59	3
139987	CR	0.39	0.64	2
140223	CR	0.31	0.68	1
140233	MC	0.70	0.47	0
140241	SA	0.30	0.56	2
140251	CR	0.39	0.69	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
140259	MC	0.68	0.39	0
140279	MC	0.35	0.36	0
140280	SA	0.64	0.40	1
140292	CR	0.36	0.69	1
144726	MC	0.66	0.32	0
144756	SA	0.40	0.50	2
145376	MC	0.43	0.45	0
146230	SA	0.33	0.58	2
181194	MC	0.43	0.43	1
181196	MC	0.52	0.38	0
181199	MC	0.81	0.27	0
181200	MC	0.59	0.45	0
181202	MC	0.87	0.39	0
181205	MC	0.55	0.43	0

Table F-6. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Mathematics Grade 8

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
120882	MC	0.82	0.48	1
120887	MC	0.62	0.50	1
120889	SA	0.70	0.56	1
120896	MC	0.59	0.41	0
120910	MC	0.43	0.34	1
120923	MC	0.69	0.47	0
120929	MC	0.47	0.41	0
120933	SA	0.71	0.48	1
120942	CR	0.35	0.68	1
120946	MC	0.62	0.38	0
120958	SA	0.23	0.49	4
120960	MC	0.65	0.50	0
120961	MC	0.60	0.39	0
120987	MC	0.78	0.39	0
120989	MC	0.66	0.35	0
121008	MC	0.55	0.38	1
121038	MC	0.32	0.28	1
121039	SA	0.29	0.53	3
121044	MC	0.35	0.22	0
121078	MC	0.68	0.34	0
122688	MC	0.45	0.44	1
123801	MC	0.52	0.36	0
125503	MC	0.63	0.35	0
125563	SA	0.52	0.62	2
125580	MC	0.59	0.33	0
139766	MC	0.60	0.41	0
139795	CR	0.43	0.72	1
139813	SA	0.28	0.44	2
139827	CR	0.30	0.63	2
139845	MC	0.66	0.23	0
139847	MC	0.58	0.44	0
139853	MC	0.61	0.51	0
139863	SA	0.71	0.52	2

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
139875	MC	0.41	0.41	0
144760	MC	0.67	0.43	1
144778	CR	0.30	0.67	2
145176	MC	0.53	0.53	1
145660	SA	0.42	0.53	3
145661	SA	0.30	0.43	3
146286	MC	0.34	0.45	0
146291	MC	0.55	0.53	0
146325	MC	0.75	0.40	1
146362	MC	0.67	0.38	0
146422	SA	0.27	0.43	3
146483	SA	0.51	0.61	2
152804	MC	0.49	0.49	0
194560	MC	0.71	0.47	0

Table F-7. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Mathematics Grade 11

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
119419	MC	0.82	0.32	0
119420	MC	0.64	0.37	1
119451	MC	0.30	0.38	2
119457	SA	0.41	0.53	6
119481	SA	0.14	0.50	8
119492	SA	0.80	0.30	2
119534	MC	0.87	0.27	0
119543	SA	0.35	0.55	11
119569	MC	0.42	0.19	2
119576	MC	0.53	0.33	1
119606	SA	0.38	0.58	7
119616	SA	0.48	0.33	5
125772	SA	0.29	0.47	10
130121	SA	0.20	0.55	9
140029	MC	0.58	0.46	1
140031	SA	0.40	0.45	8
140035	SA	0.28	0.59	5
140042	MC	0.49	0.54	1
140044	MC	0.37	0.40	1
140051	MC	0.52	0.32	1
140079	MC	0.61	0.41	1
140135	SA	0.15	0.50	11
140155	MC	0.77	0.32	0
140156	MC	0.44	0.28	1
140181	MC	0.58	0.51	1
140203	SA	0.37	0.56	6
141248	MC	0.55	0.45	1
141291	CR	0.30	0.71	11
141307	CR	0.16	0.59	12
141322	CR	0.46	0.71	7
141365	CR	0.37	0.71	8
144800	MC	0.44	0.52	1
144804	MC	0.75	0.48	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
145227	SA	0.21	0.53	10
145233	MC	0.53	0.52	0
145238	MC	0.50	0.31	1
145444	MC	0.40	0.50	1
145456	MC	0.36	0.32	1
145465	SA	0.32	0.43	5
145484	SA	0.11	0.46	14
145492	SA	0.54	0.52	2
145498	SA	0.54	0.66	10
146944	SA	0.18	0.59	18
181297	MC	0.33	0.27	1
181299	MC	0.59	0.31	1
181301	MC	0.29	0.28	1

Table F-8. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Reading Grade 3

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
117641	MC	0.65	0.50	1
117642	MC	0.71	0.49	1
117643	MC	0.91	0.42	1
117644	MC	0.68	0.37	1
117650	CR	0.62	0.55	1
117661	MC	0.85	0.36	1
117665	MC	0.63	0.37	1
117666	MC	0.58	0.33	1
117667	MC	0.58	0.26	1
117668	MC	0.81	0.39	1
117669	CR	0.36	0.57	2
117732	MC	0.66	0.42	0
117735	MC	0.90	0.46	0
117757	MC	0.80	0.42	0
117760	MC	0.93	0.48	1
117762	MC	0.86	0.51	1
117763	MC	0.59	0.39	1
117766	CR	0.39	0.47	1
147671	MC	0.82	0.54	1
147673	MC	0.82	0.47	1
147674	MC	0.77	0.39	1
147679	MC	0.72	0.55	2
147682	MC	0.83	0.39	1
147719	MC	0.74	0.50	1
147733	MC	0.80	0.52	2
147739	MC	0.89	0.48	2
147742	CR	0.43	0.64	1
147744	CR	0.67	0.40	1
148024	MC	0.78	0.30	1
148039	MC	0.84	0.54	1
148041	MC	0.66	0.54	1
148054	MC	0.82	0.32	1
148057	CR	0.39	0.41	1
148198	MC	0.80	0.36	0

Table F-9. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Reading Grade 4

Item Number	*Item Type*	*Difficulty*	*Discrimination*	*Percent Omitted*
117942	MC	0.62	0.40	1
117943	MC	0.72	0.44	1
117944	MC	0.77	0.41	1
117945	MC	0.66	0.42	2
117946	CR	0.36	0.49	1
117956	MC	0.90	0.46	0
117958	MC	0.83	0.46	1
117959	MC	0.62	0.41	0
117960	MC	0.62	0.31	1
117961	MC	0.71	0.46	3
117964	MC	0.77	0.37	1
117965	MC	0.73	0.40	2
117968	MC	0.74	0.51	1
117970	CR	0.47	0.48	1
117976	MC	0.80	0.44	1
118000	MC	0.90	0.39	1
118002	MC	0.90	0.44	0
118005	CR	0.83	0.49	0
147869	MC	0.56	0.37	1
147877	MC	0.69	0.35	1
147878	MC	0.84	0.45	2
147883	MC	0.68	0.42	1
147888	MC	0.82	0.52	1
147891	MC	0.74	0.48	1
147899	MC	0.84	0.44	1
147902	MC	0.90	0.40	2
147914	CR	0.36	0.40	1
147915	CR	0.78	0.44	0
147965	MC	0.84	0.42	0
147973	MC	0.71	0.33	1
147976	MC	0.65	0.27	1
147980	MC	0.64	0.42	1
147988	CR	0.73	0.53	1
148444	MC	0.85	0.29	1

Table F-10. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Reading Grade 5

Item Number	*Item Type*	*Difficulty*	*Discrimination*	*Percent Omitted*
118042	MC	0.73	0.35	0
118043	MC	0.76	0.41	1
118044	MC	0.63	0.45	0
118045	MC	0.79	0.48	1
118046	MC	0.62	0.43	1
118047	MC	0.60	0.42	1
118048	CR	0.40	0.64	1
118049	CR	0.47	0.57	1
118064	MC	0.62	0.32	0
118067	MC	0.77	0.45	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
118073	MC	0.61	0.29	0
118074	MC	0.92	0.40	0
118075	MC	0.72	0.41	0
118076	CR	0.42	0.59	1
118077	MC	0.60	0.36	1
118078	MC	0.80	0.47	1
118079	MC	0.65	0.33	1
118080	CR	0.43	0.59	1
118083	MC	0.89	0.32	0
118207	MC	0.79	0.29	0
118210	MC	0.71	0.39	0
148344	MC	0.73	0.27	0
148348	MC	0.70	0.39	0
148351	MC	0.73	0.29	0
148362	MC	0.81	0.39	1
148369	CR	0.45	0.56	1
148719	MC	0.93	0.34	0
148726	MC	0.70	0.43	0
148744	MC	0.77	0.35	0
148763	MC	0.93	0.39	1
148767	CR	0.46	0.46	0
149099	MC	0.75	0.44	0
149104	MC	0.66	0.41	0
149112	MC	0.82	0.35	0

Table F-11. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Reading Grade 6

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
118266	MC	0.82	0.52	0
118267	MC	0.75	0.45	0
118268	MC	0.69	0.33	0
118269	MC	0.85	0.47	1
118270	MC	0.75	0.50	1
118271	MC	0.87	0.49	1
118272	MC	0.51	0.39	1
118273	MC	0.76	0.46	1
118274	CR	0.44	0.62	1
118275	CR	0.47	0.65	1
118345	MC	0.77	0.46	0
118347	MC	0.94	0.38	0
118349	MC	0.64	0.45	0
118352	MC	0.64	0.43	0
118353	CR	0.43	0.59	1
126896	MC	0.67	0.43	0
126897	MC	0.74	0.36	0
126898	MC	0.56	0.44	0
126899	MC	0.79	0.48	0
126900	MC	0.71	0.37	0
126901	MC	0.64	0.43	1
126902	MC	0.77	0.45	1
126903	MC	0.69	0.47	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
126904	CR	0.45	0.60	1
126905	CR	0.41	0.63	1
148825	MC	0.84	0.35	0
148865	MC	0.79	0.49	0
148874	MC	0.88	0.42	0
148876	MC	0.75	0.45	0
148891	CR	0.48	0.54	1
148950	MC	0.76	0.43	0
148956	MC	0.79	0.35	0
148966	MC	0.75	0.42	0
148978	MC	0.90	0.34	0

Table F-12. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Reading Grade 7

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
118477	MC	0.59	0.33	0
118535	MC	0.87	0.30	0
118536	MC	0.71	0.32	0
118538	MC	0.81	0.36	0
118540	MC	0.80	0.46	0
118542	CR	0.51	0.66	1
118544	MC	0.74	0.40	0
118546	MC	0.72	0.37	0
118547	MC	0.87	0.35	0
118548	MC	0.60	0.31	1
118549	MC	0.56	0.35	1
118550	MC	0.79	0.50	1
118551	MC	0.62	0.46	1
118554	MC	0.56	0.38	1
118557	CR	0.49	0.64	1
118558	CR	0.46	0.70	1
118563	MC	0.74	0.38	0
128125	MC	0.57	0.26	0
128216	MC	0.63	0.41	0
129221	MC	0.64	0.22	1
129222	MC	0.60	0.26	0
129223	MC	0.61	0.24	0
129224	MC	0.87	0.35	0
129225	MC	0.57	0.36	1
129226	MC	0.67	0.40	1
129227	MC	0.70	0.44	0
129228	CR	0.46	0.63	1
129229	CR	0.41	0.67	1
147203	MC	0.91	0.32	0
147210	MC	0.72	0.27	0
147213	MC	0.78	0.36	1
147215	MC	0.74	0.29	0
147224	CR	0.39	0.61	1
147663	MC	0.60	0.28	0

Table F-13. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Reading Grade 8

Item Number	*Item Type*	*Difficulty*	*Discrimination*	*Percent Omitted*
118599	MC	0.65	0.17	0
118600	MC	0.79	0.32	0
118601	MC	0.82	0.33	0
118602	MC	0.74	0.40	1
118603	MC	0.88	0.26	1
118604	MC	0.81	0.26	1
118605	MC	0.43	0.33	1
118606	MC	0.89	0.41	1
118607	CR	0.55	0.57	1
118608	CR	0.48	0.53	1
118662	MC	0.81	0.40	0
118663	MC	0.78	0.45	0
118664	MC	0.78	0.30	0
118665	MC	0.74	0.48	0
118666	MC	0.72	0.35	0
118667	MC	0.86	0.45	0
118668	MC	0.75	0.42	1
118669	MC	0.89	0.36	1
118670	CR	0.46	0.64	1
118671	CR	0.44	0.64	1
118745	MC	0.57	0.30	0
127358	MC	0.79	0.39	0
147223	MC	0.73	0.26	0
147227	MC	0.64	0.35	0
147232	MC	0.70	0.38	0
147239	MC	0.91	0.31	0
147242	CR	0.59	0.56	0
147447	MC	0.53	0.33	0
147459	MC	0.65	0.42	0
147469	MC	0.65	0.32	0
147476	MC	0.71	0.39	0
147481	CR	0.50	0.58	1
147605	MC	0.65	0.41	0
147619	MC	0.79	0.38	0

Table F-14. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Reading Grade 11

Item Number	*Item Type*	*Difficulty*	*Discrimination*	*Percent Omitted*
118758	MC	0.70	0.25	0
118759	MC	0.48	0.39	1
118760	MC	0.95	0.36	0
118761	MC	0.64	0.30	1
118762	MC	0.76	0.33	0
118763	MC	0.52	0.30	1
118764	MC	0.89	0.31	0
118765	MC	0.86	0.47	0
118766	CR	0.57	0.63	1
118767	CR	0.50	0.65	2
147550	MC	0.71	0.34	0
147551	MC	0.91	0.41	0

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
147552	MC	0.57	0.33	1
147556	MC	0.80	0.48	1
147564	CR	0.48	0.62	2
147695	MC	0.80	0.23	1
147698	MC	0.71	0.41	1
147703	MC	0.57	0.34	1
147707	MC	0.47	0.24	1
147713	MC	0.67	0.48	1
147716	MC	0.69	0.24	1
147721	MC	0.59	0.48	1
147723	MC	0.48	0.28	1
147726	CR	0.52	0.65	3
147732	CR	0.44	0.66	4
147850	MC	0.94	0.33	1
147855	MC	0.59	0.37	1
147860	MC	0.74	0.35	1
147874	MC	0.75	0.43	1
147885	CR	0.49	0.64	2
147897	MC	0.64	0.29	0
147912	MC	0.61	0.22	1
147927	MC	0.70	0.35	1
147934	MC	0.72	0.32	0

Table F-15. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Writing Grade 5

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
148023	MC	0.76	0.32	0
148062	MC	0.74	0.31	1
148081	MC	0.80	0.32	0
148126	MC	0.77	0.13	0
148167	MC	0.90	0.23	0
148246	MC	0.85	0.37	0
148263	MC	0.86	0.33	0
148278	MC	0.85	0.31	0
148316	MC	0.84	0.37	0
148327	MC	0.78	0.36	0
150134	CR	0.42	0.57	1
150165	WP	0.46	0.58	0
150198	CR	0.51	0.57	1
150237	CR	0.51	0.52	1

Table F-16. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Writing Grade 8

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
148434	MC	0.63	0.36	0
148557	MC	0.64	0.38	0
148588	MC	0.84	0.28	0
148610	MC	0.82	0.39	0
148662	MC	0.74	0.35	0
148700	MC	0.78	0.35	0

Item Number	*Item Type*	*Difficulty*	*Discrimination*	*Percent Omitted*
148715	MC	0.89	0.37	0
148762	MC	0.60	0.33	1
148773	MC	0.72	0.30	0
148812	MC	0.50	0.28	0
150433	CR	0.57	0.66	1
150789	CR	0.74	0.52	1
150803	CR	0.58	0.64	2
150854	WP	0.55	0.66	0

Table F-17. 2010–11 NECAP: Item-Level Classical Test Theory Statistics – Writing Grade 11

Item Number	*Item Type*	*Difficulty*	*Discrimination*	*Percent Omitted*
121154	WP	0.52		

Appendix G—Item-Level Score Point Distributions

Table G-1. 2010–11 NECAP: Item-Level Score Distributions for Constructed-Response Items by Subject and Grade – Mathematics and Reading

Subject	*Grade*	*Item Number*	*Total Possible Points*	*Percent of Students at Score Point*				
				0	*1*	*2*	*3*	*4*
Mathematics	3	119688	2	15.41	21.56	61.60		
		119731	1	16.66	82.42			
		119838	2	44.07	29.33	25.35		
		119889	2	33.50	28.53	37.23		
		119918	2	28.75	11.62	58.22		
		122872	1	4.28	94.61			
		124459	1	24.23	75.23			
		139583	1	37.42	61.24			
		139590	1	65.36	33.19			
		139603	2	11.46	35.09	51.87		
		139645	1	33.68	65.00			
		139653	2	38.04	27.19	33.94		
		139655	2	33.70	34.41	31.22		
		139669	1	36.57	62.23			
		139673	2	8.04	49.57	41.70		
		139689	1	44.10	54.50			
		144612	2	68.12	7.89	22.12		
		144617	1	16.17	82.88			
		144623	2	39.75	34.77	24.18		
		145249	1	28.79	69.90			
	4	120075	2	14.45	44.83	39.82		
		120119	1	11.14	88.02			
		120127	2	4.10	29.20	66.25		
		120221	1	22.58	76.59			
		120222	2	13.10	71.23	13.98		
		120240	1	37.39	60.82			
		120262	1	54.55	44.39			
		120286	2	12.53	16.74	69.64		
		124661	2	39.65	27.29	31.84		
		124723	2	17.61	28.75	52.92		
		139447	1	55.37	42.84			
		139452	2	30.03	26.95	42.19		
		139479	1	71.71	27.50			
		139517	1	51.47	47.49			
		139523	1	31.42	67.24			
		139527	2	12.05	40.12	47.21		
		139529	2	19.84	41.48	37.49		
		145085	1	22.33	76.70			
		145545	1	19.12	79.98			
		145559	2	51.47	19.23	28.42		
	5	120725	2	27.71	42.12	28.85		
		120726	2	50.98	11.96	34.56		
		124804	1	34.09	65.49			
		124858	4	37.54	24.45	16.80	9.42	10.56
		139129	1	41.96	57.51			
		139136	1	45.82	53.34			
		139141	2	35.71	39.22	24.48		

Subject	Grade	Item Number	Total Possible Points	Percent of Students at Score Point				
				0	1	2	3	4
		139203	2	56.95	29.47	10.37		
		139283	4	10.99	4.82	5.70	17.99	59.74
		139384	4	14.56	14.94	15.54	25.71	28.44
		145571	2	40.05	17.61	40.91		
		145576	4	23.70	18.07	30.96	11.72	14.75
		145879	1	30.30	69.20			
		145887	1	37.17	62.23			
		145888	2	30.27	20.61	47.07		
		145922	1	41.22	58.25			
	6	119195	2	55.51	29.13	13.53		
		119230	2	22.65	15.36	59.44		
		119306	1	24.29	75.25			
		123489	4	22.18	13.37	31.69	15.52	16.61
		139310	1	54.96	44.32			
		139312	1	56.42	43.27			
		139320	4	43.61	29.46	15.59	7.22	2.54
		139356	4	38.23	17.48	11.75	14.38	17.36
		139377	2	39.83	19.29	39.48		
		139387	4	19.63	24.19	27.88	9.33	18.15
		144707	2	72.63	6.35	19.67		
		145133	1	47.32	50.65			
		145347	1	32.13	67.43			
		145995	1	49.45	49.90			
		146011	2	49.49	20.88	28.43		
		197550	2	16.49	11.00	70.43		
	7	120387	1	40.15	59.49			
		120406	1	49.16	49.30			
		120427	2	56.22	15.75	25.31		
		120468	1	57.05	41.19			
		120470	2	48.99	36.63	12.33		
		125382	2	47.53	23.42	28.17		
		125397	1	31.29	68.22			
		139980	2	55.89	29.44	12.03		
		139987	4	38.26	16.72	10.23	11.82	20.55
		140223	4	43.01	19.49	14.65	8.94	12.43
		140241	2	45.20	45.72	7.42		
		140251	4	30.10	18.72	20.47	20.36	9.30
		140280	1	35.59	63.71			
		140292	4	19.43	38.86	26.76	4.59	9.53
		144756	2	25.79	64.12	8.28		
		146230	1	65.23	32.90			
	8	120889	1	29.28	69.72			
		120933	1	28.23	70.87			
		120942	4	27.38	24.56	29.88	10.58	6.47
		120958	2	69.21	8.25	18.99		
		121039	1	68.33	29.00			
		125563	2	28.38	34.96	34.32		
		139795	4	25.05	16.58	28.57	16.05	12.49
		139813	2	58.98	22.52	16.33		
		139827	4	31.78	32.54	22.10	5.14	6.73

Subject	Grade	Item Number	Total Possible Points	Percent of Students at Score Point				
				0	1	2	3	4
		139863	1	27.64	70.81			
		144778	4	32.48	27.85	24.03	11.22	2.81
		145660	2	38.86	33.32	25.07		
		145661	2	50.51	33.10	13.74		
		146422	1	70.29	26.55			
		146483	2	24.07	45.38	28.64		
	11	119457	1	53.04	41.25			
		119481	2	70.79	14.35	7.02		
		119492	1	18.27	80.20			
		119543	2	51.61	3.19	33.72		
		119606	1	54.41	38.15			
		119616	1	47.16	47.96			
		125772	1	60.82	29.23			
		130121	1	71.08	20.09			
		140031	1	52.16	39.57			
		140035	2	65.36	3.87	25.68		
		140135	1	74.43	14.87			
		140203	1	57.05	36.51			
		141291	4	39.97	5.49	28.14	2.47	12.89
		141307	4	50.60	21.83	8.27	4.19	3.16
		141322	4	19.51	2.15	42.54	17.48	11.34
		141365	4	26.37	24.52	12.90	14.24	14.21
		145227	2	53.46	31.54	4.88		
		145465	1	63.74	31.69			
		145484	1	74.89	11.02			
		145492	1	43.46	54.46			
		145498	2	24.71	22.27	42.98		
		146944	2	61.17	5.73	14.79		
Reading	3	117650	4	5.52	18.91	21.47	25.45	27.91
		117669	4	15.70	35.14	34.20	10.40	2.80
		117766	4	6.50	43.50	37.59	9.63	1.81
		147742	4	11.13	23.16	46.86	14.22	3.30
		147744	4	0.78	5.24	47.35	13.82	32.02
		148057	4	11.43	32.90	42.32	10.26	2.02
	4	117946	4	20.40	33.28	30.04	13.57	2.09
		117970	4	3.89	34.08	35.13	22.35	3.90
		118005	4	0.67	5.99	13.95	16.31	62.67
		147914	4	13.94	41.90	31.30	9.75	2.40
		147915	4	0.24	2.06	30.40	17.36	49.58
		147988	4	4.23	9.10	17.84	22.54	45.29
	5	118048	4	15.52	28.64	36.85	14.51	3.66
		118049	4	4.42	25.97	49.37	14.97	4.18
		118076	4	9.76	31.32	40.43	14.36	3.21
		118080	4	8.81	31.59	41.63	13.78	3.50
		148369	4	6.70	28.50	43.65	16.67	3.84
		148767	4	2.81	31.74	47.82	13.60	3.59
	6	118274	4	6.02	31.59	43.49	15.00	2.89
		118275	4	7.40	25.24	42.28	20.22	4.16
		118353	4	11.19	28.39	39.96	16.49	3.18
		126904	4	6.33	31.59	42.00	15.20	4.26

Subject	Grade	Item Number	Total Possible Points	Percent of Students at Score Point				
				0	1	2	3	4
		126905	4	13.67	31.01	34.83	14.33	5.06
		148891	4	3.18	25.54	47.97	18.97	3.83
	7	118542	4	6.76	21.28	38.67	24.08	8.55
		118557	4	3.31	23.99	47.80	18.99	4.82
		118558	4	8.39	29.24	35.10	19.16	6.88
		129228	4	8.58	29.12	36.81	17.78	6.83
		129229	4	12.50	32.21	34.99	15.18	4.01
		147224	4	14.60	32.83	33.89	14.75	3.28
	8	118607	4	2.90	19.53	41.40	24.44	10.78
		118608	4	4.15	26.03	45.33	18.45	4.86
		118670	4	5.06	30.16	42.33	17.59	4.09
		118671	4	4.43	34.12	42.87	15.38	2.62
		147242	4	1.28	12.88	42.64	33.35	9.51
		147481	4	8.03	20.17	40.98	22.46	7.73
	11	118766	4	2.56	13.96	43.26	28.60	10.40
		118767	4	3.69	16.98	53.44	17.37	6.46
		147564	4	6.75	22.37	42.20	21.55	5.28
		147726	4	4.20	21.01	39.26	24.03	8.97
		147732	4	8.09	27.40	39.18	16.37	5.00
		147885	4	4.41	20.19	48.27	19.77	5.07

Table G-2. 2010–11 NECAP: 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	150134	4	5.37	36.53	44.40	11.53	1.57								
	150165	12	1.25		4.78	5.59	19.44	15.58	26.19	13.00	8.72	3.59	1.50	0.10	0.02
	150198	4	2.12	25.15	42.98	22.28	6.91								
	150237	4	5.18	14.40	54.80	21.46	3.45								
8	150433	4	2.23	14.45	43.24	28.65	10.52								
	150789	4	1.80	5.07	21.09	36.49	34.67								
	150803	4	1.29	12.71	43.19	31.85	9.40								
	150854	12	0.75		2.06	1.80	8.14	6.93	28.20	16.05	25.73	7.18	2.50	0.18	0.07
11	121154	12	1.50		2.82	2.08	9.11	8.29	28.17	18.92	24.83	3.21	0.91	0.11	0.04

Appendix H—Differential Item Functioning Results

Table H-1. 2010–11 NECAP: Number of Items Classified as "Low" or "High" DIF Overall and by Grade and Group Favored – Mathematics

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
3	Male	Female	MC	35	0	0	0	0	0	0
	Male	Female	SA	20	2	1	1	0	0	0
	No Disability	Disability	MC	35	3	3	0	0	0	0
	No Disability	Disability	SA	20	0	0	0	0	0	0
	Not Low-income	Low-income	MC	35	0	0	0	0	0	0
	Not Low-income	Low-income	SA	20	0	0	0	0	0	0
	Non-LEP	LEP	MC	35	7	6	1	2	2	0
	Non-LEP	LEP	SA	20	2	2	0	0	0	0
	White	Asian	MC	35	9	6	3	2	2	0
	White	Asian	SA	20	2	1	1	0	0	0
	White	Black	MC	35	1	1	0	0	0	0
	White	Black	SA	20	1	1	0	0	0	0
	White	Hispanic	MC	35	7	6	1	0	0	0
	White	Hispanic	SA	20	2	2	0	0	0	0
4	Male	Female	MC	35	2	1	1	1	1	0
	Male	Female	SA	20	1	1	0	0	0	0
	No Disability	Disability	MC	35	0	0	0	0	0	0
	No Disability	Disability	SA	20	1	1	0	0	0	0
	Not Low-income	Low-income	MC	35	0	0	0	0	0	0
	Not Low-income	Low-income	SA	20	0	0	0	0	0	0
	Non-LEP	LEP	MC	35	4	4	0	0	0	0
	Non-LEP	LEP	SA	20	2	2	0	0	0	0
	White	Asian	MC	35	4	2	2	0	0	0
	White	Asian	SA	20	4	3	1	0	0	0
	White	Black	MC	35	4	2	2	0	0	0
	White	Black	SA	20	2	2	0	0	0	0
	White	Hispanic	MC	35	6	4	2	0	0	0
	White	Hispanic	SA	20	6	4	2	0	0	0
5	Male	Female	CR	4	1	0	1	0	0	0
	Male	Female	MC	32	4	4	0	0	0	0
	Male	Female	SA	12	1	0	1	0	0	0
	No Disability	Disability	CR	4	0	0	0	0	0	0
	No Disability	Disability	MC	32	3	3	0	0	0	0
	No Disability	Disability	SA	12	0	0	0	0	0	0

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
	Not Low-income	Low-income	CR	4	0	0	0	0	0	0
	Not Low-income	Low-income	MC	32	0	0	0	0	0	0
	Not Low-income	Low-income	SA	12	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	0	0	0	0	0	0
	Non-LEP	LEP	MC	32	4	4	0	0	0	0
	Non-LEP	LEP	SA	12	1	1	0	0	0	0
	White	Asian	CR	4	0	0	0	0	0	0
	White	Asian	MC	32	3	1	2	0	0	0
	White	Asian	SA	12	0	0	0	0	0	0
	White	Black	CR	4	0	0	0	0	0	0
	White	Black	MC	32	1	1	0	0	0	0
	White	Black	SA	12	1	1	0	0	0	0
	White	Hispanic	CR	4	0	0	0	0	0	0
	White	Hispanic	MC	32	4	2	2	1	1	0
	White	Hispanic	SA	12	1	1	0	0	0	0
6	Male	Female	CR	4	0	0	0	0	0	0
	Male	Female	MC	32	5	4	1	1	1	0
	Male	Female	SA	12	1	1	0	0	0	0
	No Disability	Disability	CR	4	0	0	0	0	0	0
	No Disability	Disability	MC	32	0	0	0	0	0	0
	No Disability	Disability	SA	12	2	2	0	0	0	0
	Not Low-income	Low-income	CR	4	0	0	0	0	0	0
	Not Low-income	Low-income	MC	32	0	0	0	0	0	0
	Not Low-income	Low-income	SA	12	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	0	0	0	0	0	0
	Non-LEP	LEP	MC	32	7	6	1	0	0	0
	Non-LEP	LEP	SA	12	1	1	0	0	0	0
	White	Asian	CR	4	0	0	0	0	0	0
	White	Asian	MC	32	4	2	2	0	0	0
	White	Asian	SA	12	1	1	0	0	0	0
	White	Black	CR	4	0	0	0	0	0	0
	White	Black	MC	32	5	4	1	0	0	0
	White	Black	SA	12	2	1	1	0	0	0
	White	Hispanic	CR	4	0	0	0	0	0	0
	White	Hispanic	MC	32	4	3	1	0	0	0
	White	Hispanic	SA	12	1	1	0	0	0	0

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
	Male	Female	CR	4	0	0	0	0	0	0
	Male	Female	MC	30	3	3	0	1	1	0
	Male	Female	SA	12	3	2	1	1	0	1
	No Disability	Disability	CR	4	0	0	0	0	0	0
	No Disability	Disability	MC	30	2	2	0	0	0	0
	No Disability	Disability	SA	12	0	0	0	0	0	0
	Not Low-income	Low-income	CR	4	0	0	0	0	0	0
	Not Low-income	Low-income	MC	30	0	0	0	0	0	0
	Not Low-income	Low-income	SA	12	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	0	0	0	0	0	0
7	Non-LEP	LEP	MC	30	7	5	2	0	0	0
	Non-LEP	LEP	SA	12	1	1	0	0	0	0
	White	Asian	CR	4	0	0	0	0	0	0
	White	Asian	MC	30	5	2	3	0	0	0
	White	Asian	SA	12	2	1	1	0	0	0
	White	Black	CR	4	0	0	0	0	0	0
	White	Black	MC	30	0	0	0	2	2	0
	White	Black	SA	12	2	2	0	0	0	0
	White	Hispanic	CR	4	0	0	0	0	0	0
	White	Hispanic	MC	30	4	4	0	0	0	0
	White	Hispanic	SA	12	3	2	1	0	0	0
	Male	Female	CR	4	0	0	0	0	0	0
	Male	Female	MC	32	6	4	2	0	0	0
	Male	Female	SA	11	2	0	2	0	0	0
	No Disability	Disability	CR	4	0	0	0	0	0	0
	No Disability	Disability	MC	32	1	1	0	0	0	0
	No Disability	Disability	SA	11	3	3	0	0	0	0
	Not Low-income	Low-income	CR	4	0	0	0	0	0	0
8	Not Low-income	Low-income	MC	32	0	0	0	0	0	0
	Not Low-income	Low-income	SA	11	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	0	0	0	0	0	0
	Non-LEP	LEP	MC	32	3	2	1	0	0	0
	Non-LEP	LEP	SA	11	4	4	0	0	0	0
	White	Asian	CR	4	0	0	0	0	0	0
	White	Asian	MC	32	4	2	2	0	0	0
	White	Asian	SA	11	1	0	1	0	0	0

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
	White	Black	CR	4	0	0	0	0	0	0
	White	Black	MC	32	1	1	0	0	0	0
	White	Black	SA	11	1	1	0	0	0	0
	White	Hispanic	CR	4	0	0	0	0	0	0
	White	Hispanic	MC	32	1	0	1	0	0	0
	White	Hispanic	SA	11	0	0	0	0	0	0
	Male	Female	CR	4	1	0	1	0	0	0
	Male	Female	MC	24	6	5	1	0	0	0
	Male	Female	SA	18	1	1	0	0	0	0
	No Disability	Disability	CR	4	2	2	0	0	0	0
	No Disability	Disability	MC	24	0	0	0	0	0	0
	No Disability	Disability	SA	18	2	2	0	0	0	0
	Not Low-income	Low-income	CR	4	0	0	0	0	0	0
	Not Low-income	Low-income	MC	24	0	0	0	0	0	0
	Not Low-income	Low-income	SA	18	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	1	1	0	1	1	0
11	Non-LEP	LEP	MC	24	4	3	1	3	2	1
	Non-LEP	LEP	SA	18	3	2	1	3	1	2
	White	Asian	CR	4	0	0	0	0	0	0
	White	Asian	MC	24	2	1	1	1	0	1
	White	Asian	SA	18	3	0	3	0	0	0
	White	Black	CR	4	0	0	0	0	0	0
	White	Black	MC	24	3	2	1	0	0	0
	White	Black	SA	18	0	0	0	0	0	0
	White	Hispanic	CR	4	0	0	0	0	0	0
	White	Hispanic	MC	24	3	2	1	0	0	0
	White	Hispanic	SA	18	0	0	0	0	0	0

Table H-2. 2010–11 NECAP: Number of Items Classified as "Low" or "High" DIF Overall and by Grade and Group Favored – Reading

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
3	Male	Female	CR	6	1	1	0	0	0	0
	Male	Female	MC	28	0	0	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
	No Disability	Disability	MC	28	3	3	0	0	0	0
	Not Low-income	Low-income	CR	6	0	0	0	0	0	0
	Not Low-income	Low-income	MC	28	0	0	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	5	5	0	1	1	0
	White	Asian	CR	6	0	0	0	0	0	0
	White	Asian	MC	28	1	1	0	0	0	0
	White	Black	CR	6	1	1	0	0	0	0
	White	Black	MC	28	3	3	0	0	0	0
	White	Hispanic	CR	6	0	0	0	0	0	0
	White	Hispanic	MC	28	4	4	0	0	0	0
4	Male	Female	CR	6	1	0	1	0	0	0
	Male	Female	MC	28	0	0	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
	No Disability	Disability	MC	28	3	3	0	0	0	0
	Not Low-income	Low-income	CR	6	0	0	0	0	0	0
	Not Low-income	Low-income	MC	28	0	0	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	4	4	0	1	1	0
	White	Asian	CR	6	0	0	0	0	0	0
	White	Asian	MC	28	1	1	0	0	0	0
	White	Black	CR	6	0	0	0	0	0	0
	White	Black	MC	28	3	3	0	0	0	0
	White	Hispanic	CR	6	0	0	0	0	0	0
	White	Hispanic	MC	28	5	5	0	0	0	0
5	Male	Female	CR	6	1	0	1	0	0	0
	Male	Female	MC	28	2	2	0	1	1	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
	No Disability	Disability	MC	28	4	4	0	0	0	0
	Not Low-income	Low-income	CR	6	0	0	0	0	0	0
	Not Low-income	Low-income	MC	28	1	1	0	0	0	0

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	7	7	0	3	3	0
	White	Asian	CR	6	0	0	0	0	0	0
	White	Asian	MC	28	3	3	0	0	0	0
	White	Black	CR	6	0	0	0	0	0	0
	White	Black	MC	28	8	7	1	0	0	0
	White	Hispanic	CR	6	1	0	1	0	0	0
	White	Hispanic	MC	28	7	7	0	2	2	0
6	Male	Female	CR	6	1	0	1	0	0	0
	Male	Female	MC	28	5	5	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
	No Disability	Disability	MC	28	2	2	0	0	0	0
	Not Low-income	Low-income	CR	6	0	0	0	0	0	0
	Not Low-income	Low-income	MC	28	0	0	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	4	4	0	5	5	0
	White	Asian	CR	6	0	0	0	0	0	0
	White	Asian	MC	28	5	5	0	0	0	0
	White	Black	CR	6	0	0	0	0	0	0
	White	Black	MC	28	5	5	0	0	0	0
	White	Hispanic	CR	6	0	0	0	0	0	0
	White	Hispanic	MC	28	5	5	0	3	3	0
7	Male	Female	CR	6	1	0	1	0	0	0
	Male	Female	MC	28	3	3	0	1	1	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
	No Disability	Disability	MC	28	2	2	0	0	0	0
	Not Low-income	Low-income	CR	6	0	0	0	0	0	0
	Not Low-income	Low-income	MC	28	2	2	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	6	6	0	5	5	0
	White	Asian	CR	6	0	0	0	0	0	0
	White	Asian	MC	28	7	7	0	1	1	0
	White	Black	CR	6	0	0	0	0	0	0
	White	Black	MC	28	3	3	0	1	1	0
	White	Hispanic	CR	6	0	0	0	0	0	0
	White	Hispanic	MC	28	6	6	0	2	2	0

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
8	Male	Female	CR	6	3	0	3	0	0	0
	Male	Female	MC	28	8	8	0	1	1	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
	No Disability	Disability	MC	28	3	3	0	0	0	0
	Not Low-income	Low-income	CR	6	0	0	0	0	0	0
	Not Low-income	Low-income	MC	28	1	1	0	0	0	0
	Non-LEP	LEP	CR	6	2	0	2	0	0	0
	Non-LEP	LEP	MC	28	10	10	0	6	6	0
	White	Asian	CR	6	3	0	3	0	0	0
	White	Asian	MC	28	6	6	0	0	0	0
	White	Black	CR	6	0	0	0	0	0	0
	White	Black	MC	28	8	8	0	0	0	0
	White	Hispanic	CR	6	1	0	1	0	0	0
	White	Hispanic	MC	28	13	13	0	2	2	0
11	Male	Female	CR	6	1	0	1	0	0	0
	Male	Female	MC	28	6	6	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
	No Disability	Disability	MC	28	1	1	0	0	0	0
	Not Low-income	Low-income	CR	6	0	0	0	0	0	0
	Not Low-income	Low-income	MC	28	6	6	0	0	0	0
	Non-LEP	LEP	CR	6	2	0	2	0	0	0
	Non-LEP	LEP	MC	28	10	10	0	7	7	0
	White	Asian	CR	6	1	0	1	0	0	0
	White	Asian	MC	28	8	8	0	3	3	0
	White	Black	CR	6	1	0	1	0	0	0
	White	Black	MC	28	9	8	1	2	2	0
	White	Hispanic	CR	6	1	0	1	0	0	0
	White	Hispanic	MC	28	10	9	1	2	2	0

Table H-3. 2010–11 NECAP: Number of Items Classified as "Low" or "High" DIF Overall and by Grade and Group Favored – Writing

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
5	Male	Female	CR	3	1	0	1	0	0	0
	Male	Female	MC	10	0	0	0	0	0	0
	Male	Female	WP	1	0	0	0	0	0	0
	No Disability	Disability	CR	3	0	0	0	0	0	0
	No Disability	Disability	MC	10	3	3	0	1	1	0
	No Disability	Disability	WP	1	1	1	0	0	0	0
	Not Low-income	Low-income	CR	3	0	0	0	0	0	0
	Not Low-income	Low-income	MC	10	0	0	0	0	0	0
	Not Low-income	Low-income	WP	1	0	0	0	0	0	0
	Non-LEP	LEP	CR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	10	3	3	0	1	1	0
	Non-LEP	LEP	WP	1	0	0	0	0	0	0
	White	Asian	CR	3	0	0	0	0	0	0
	White	Asian	MC	10	1	0	1	0	0	0
	White	Asian	WP	1	0	0	0	0	0	0
	White	Black	CR	3	0	0	0	0	0	0
	White	Black	MC	10	2	2	0	1	1	0
	White	Black	WP	1	0	0	0	0	0	0
	White	Hispanic	CR	3	0	0	0	0	0	0
	White	Hispanic	MC	10	4	4	0	0	0	0
	White	Hispanic	WP	1	0	0	0	0	0	0
8	Male	Female	CR	3	1	0	1	0	0	0
	Male	Female	MC	10	1	1	0	0	0	0
	Male	Female	WP	1	0	0	0	0	0	0
	No Disability	Disability	CR	3	0	0	0	0	0	0
	No Disability	Disability	MC	10	7	7	0	3	3	0
	No Disability	Disability	WP	1	0	0	0	0	0	0
	Not Low-income	Low-income	CR	3	0	0	0	0	0	0
	Not Low-income	Low-income	MC	10	2	2	0	0	0	0
	Not Low-income	Low-income	WP	1	0	0	0	0	0	0
	Non-LEP	LEP	CR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	10	3	3	0	5	5	0
	Non-LEP	LEP	WP	1	0	0	0	0	0	0
	White	Asian	CR	3	0	0	0	0	0	0

Grade	Reference Group	Focal Group	Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring Reference	Favoring Focal	Total	Favoring Reference	Favoring Focal
	White	Asian	MC	10	4	2	2	0	0	0
	White	Asian	WP	1	0	0	0	0	0	0
	White	Black	CR	3	0	0	0	0	0	0
	White	Black	MC	10	6	6	0	1	1	0
	White	Black	WP	1	0	0	0	0	0	0
	White	Hispanic	CR	3	0	0	0	0	0	0
	White	Hispanic	MC	10	4	4	0	3	3	0
	White	Hispanic	WP	1	0	0	0	0	0	0
	Male	Female	WP	1	0	0	0	0	0	0
	No Disability	Disability	WP	1	0	0	0	1	1	0
	Not Low-income	Low-income	WP	1	1	1	0	0	0	0
11	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	1	1	0	0	0	0
	White	Hispanic	WP	1	0	0	0	0	0	0

Appendix I—Item Response Theory Calibration Results

Table I-1. 2010–11 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)*
119681	0.72807	0.0427	-1.95066	0.16421	0.19498	0.07433
119683	0.79961	0.01361	-0.55795	0.02636	0.06513	0.01198
119685	0.97818	0.01365	-1.00243	0.0195	0.03261	0.00987
119692	0.59552	0.01719	-0.62235	0.07134	0.24671	0.02327
119696	1.03256	0.01796	-1.52407	0.0329	0.0927	0.02022
119699	0.68686	0.02366	-2.29952	0.06647	0	0
119701	0.4841	0.0161	-1.5876	0.05589	0	0
119702	1.43238	0.05457	0.53067	0.02165	0.21288	0.00996
119704	0.76463	0.0486	0.32981	0.06607	0.15017	0.02531
119706	1.19615	0.02091	-0.97845	0.024	0.1901	0.01381
119715	1.09776	0.05345	-0.98525	0.06862	0.15965	0.03698
119731	0.96125	0.01115	-1.21209	0.01241	0	0
119747	0.74466	0.02515	-0.45297	0.03075	0	0
119752	0.79763	0.03944	-1.4702	0.10299	0.12526	0.04787
119768	0.92655	0.01862	-0.38042	0.02789	0.2402	0.01204
119770	1.07024	0.04281	-0.74961	0.04492	0.0636	0.02131
119775	0.65567	0.01659	-0.57482	0.02524	0	0
119808	0.82782	0.02892	-1.21682	0.06852	0.10153	0.03322
119810	1.03337	0.01754	-0.73751	0.0238	0.14608	0.01244
119811	1.16062	0.02213	-1.01393	0.02849	0.25676	0.01538
119821	0.56752	0.00802	-1.42077	0.02112	0	0
119825	0.87413	0.01581	-0.80336	0.03094	0.12706	0.01503
119839	1.12222	0.01736	-0.1601	0.01445	0.1113	0.00724
119840	0.84822	0.01717	-0.31501	0.02903	0.19582	0.01227
119842	0.94694	0.06962	0.1579	0.07909	0.38121	0.02717
119848	0.93267	0.01488	-1.94202	0.03052	0.03858	0.01634
119869	0.97853	0.01853	0.21081	0.01732	0.17484	0.00754
119873	0.85777	0.02723	0.18289	0.02333	0	0
119877	0.5355	0.03197	-2.30842	0.1894	0.1457	0.064
119880	0.72421	0.04108	-0.83482	0.11269	0.15463	0.04667
119882	1.01782	0.01999	-0.37425	0.02497	0.26584	0.0111
119885	0.91085	0.01684	-0.76293	0.03041	0.16209	0.01471
119886	0.72193	0.02454	0.04948	0.02697	0	0
119891	0.57509	0.02376	1.27678	0.03448	0.25028	0.01026
119894	0.7359	0.02933	0.42672	0.03776	0.0642	0.0147

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
119896	1.04221	0.03768	0.84269	0.02148	0.06118	0.00777
119901	1.07282	0.0613	0.4618	0.04282	0.19723	0.01818
119903	0.77159	0.01903	0.79436	0.02068	0	0
119911	0.73014	0.03384	-0.90471	0.07968	0.08401	0.0326
119924	1.38096	0.04297	-0.9736	0.03222	0.09129	0.01953
119926	1.07201	0.04786	0.4297	0.03472	0.2597	0.01424
119930	1.27164	0.05483	0.74211	0.02465	0.2305	0.01013
122862	1.07217	0.01698	-0.1416	0.01535	0.11242	0.00752
122872	0.75633	0.0136	-2.64985	0.03686	0	0
124370	0.87264	0.01323	-1.49568	0.02855	0.03919	0.01478
124431	0.72409	0.01599	0.13187	0.02785	0.13866	0.0108
124433	0.7928	0.01059	-1.6919	0.01915	0	0
124436	1.12231	0.01914	-0.71926	0.02214	0.17721	0.01184
124459	0.3839	0.007	-1.72034	0.03441	0	0
139548	0.57626	0.01527	-0.78476	0.07065	0.16141	0.02512
139560	0.82836	0.01927	-1.47866	0.06253	0.26091	0.02942
139563	0.89268	0.01423	0.05798	0.01546	0.0544	0.0068
139569	1.05672	0.01854	-1.30892	0.03068	0.12486	0.0182
139575	1.13023	0.09489	1.16775	0.04745	0.55043	0.01038
139583	0.81325	0.0088	-0.29373	0.00892	0	0
139590	0.82406	0.00944	0.81896	0.00941	0	0
139606	0.86013	0.04971	0.98387	0.03985	0.23941	0.01369
139641	1.10945	0.01706	0.30685	0.01098	0.07774	0.005
139645	0.68089	0.00804	-0.52581	0.01129	0	0
139661	0.83653	0.04238	-1.76905	0.10707	0.12682	0.05184
139663	1.39394	0.02541	0.9118	0.0089	0.12832	0.00348
139665	0.84635	0.01977	0.43767	0.02212	0.22813	0.00838
139667	0.80596	0.03581	-0.46794	0.07271	0.21806	0.02952
139669	0.59784	0.00751	-0.449	0.01216	0	0
139687	1.16192	0.01857	0.1652	0.0123	0.12312	0.00591
139689	0.74511	0.00833	-0.03426	0.00889	0	0
144609	0.76991	0.03396	-0.1003	0.06051	0.17357	0.02379
144617	0.47432	0.00807	-2.06541	0.03475	0	0
145244	1.34789	0.05661	1.06785	0.02045	0.14117	0.00727
145249	0.90622	0.00968	-0.61925	0.00941	0	0
145509	0.86026	0.02083	1.02452	0.0159	0.15249	0.00573
145526	1.17619	0.01473	-0.62737	0.012	0.02348	0.00578

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
145528	1.28843	0.02669	0.88395	0.01096	0.20285	0.00433
145678	0.9798	0.02296	-1.4927	0.05415	0.33239	0.02645
145683	0.96217	0.01872	-1.58768	0.04359	0.14692	0.02564
145720	1.02406	0.05181	-0.5988	0.06729	0.17577	0.03338
198283	0.78593	0.04224	-1.34566	0.11829	0.16184	0.05375
198315	0.64039	0.03318	-1.67342	0.12834	0.11793	0.0501
201312	0.85595	0.02707	-1.53815	0.05785	0.0662	0.02743
201416	0.88007	0.02579	-0.68509	0.03777	0.0469	0.01668
201450	1.41542	0.06872	0.54549	0.02645	0.13137	0.01221
201465	0.48154	0.02266	-1.55798	0.07885	0	0
201510	0.95122	0.03149	-1.00568	0.03282	0	0
201806	0.89307	0.03448	-0.01805	0.04236	0.14	0.01849
201811	0.72518	0.02906	-0.84103	0.08079	0.12465	0.03469
223883	0.73694	0.03223	-0.75714	0.08806	0.17229	0.03669
255679	0.53019	0.03378	-1.72623	0.21337	0.18597	0.0728
255686	1.00023	0.03888	-0.66289	0.05624	0.20776	0.0273
255900	1.32153	0.06	0.60427	0.02471	0.07821	0.01043
255929	0.67932	0.02222	-2.10879	0.05937	0	0
255932	1.05914	0.0226	0.3363	0.01411	0	0

Table I-2. 2010–11 NECAP: IRT Parameters for Polytomous Items – Mathematics Grade 3

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
202089	0.85515	0.02279	0.38359	0.02218	0	0	0.17321	0.0232	-0.17321	0.02355	0	0				
144612	0.68166	0.00702	1.24142	0.01192	0	0	0.21452	0.00992	-0.21452	0.0108	0	0				
119918	0.99154	0.00814	-0.35256	0.00692	0	0	0.22483	0.0075	-0.22483	0.00699	0	0				
119920	0.81302	0.01788	-0.2826	0.02234	0	0	0.48823	0.02685	-0.48823	0.02405	0	0				
198521	0.78839	0.01212	0.3918	0.01559	0	0	0.50808	0.01759	-0.50808	0.0184	0	0				
145259	1.03656	0.02127	0.78601	0.01824	0	0	0.58079	0.01996	-0.58079	0.02438	0	0				
119838	0.87156	0.00598	0.54352	0.00675	0	0	0.58234	0.00764	-0.58234	0.00841	0	0				
119889	0.7499	0.00524	0.08864	0.00755	0	0	0.60115	0.00901	-0.60115	0.00875	0	0				
119853	1.01624	0.0166	-1.14094	0.01654	0	0	0.60332	0.02322	-0.60332	0.01553	0	0				
139653	0.63548	0.0046	0.27476	0.00878	0	0	0.63539	0.01007	-0.63539	0.01026	0	0				
119688	0.61057	0.00488	-1.03746	0.01115	0	0	0.64799	0.01325	-0.64799	0.01044	0	0				
139655	0.91142	0.00596	0.21662	0.00631	0	0	0.65273	0.00773	-0.65273	0.00773	0	0				
144623	0.93791	0.00613	0.48788	0.00625	0	0	0.66345	0.00732	-0.66345	0.00806	0	0				

119761	0.8416	0.01681	0.61334	0.02084	0	0	0.68539	0.0236	-0.68539	0.02692	0	0
124490	0.71204	0.01004	0.16597	0.01663	0	0	0.75728	0.02013	-0.75728	0.01996	0	0
198507	0.5993	0.00995	-1.26034	0.02454	0	0	0.79604	0.03148	-0.79604	0.02264	0	0
139603	0.94654	0.00643	-0.73604	0.00675	0	0	0.81377	0.0104	-0.81377	0.00715	0	0
119806	0.8751	0.01338	-1.10173	0.01683	0	0	0.84798	0.02667	-0.84798	0.01651	0	0
119909	0.61572	0.01239	0.67771	0.02751	0	0	0.87592	0.03075	-0.87592	0.03522	0	0
119708	0.75484	0.01538	-0.92321	0.02474	0	0	1.00396	0.0394	-1.00396	0.02559	0	0
139651	0.65615	0.00888	0.81753	0.01855	0	0	1.02981	0.02063	-1.02981	0.02596	0	0
139673	0.50727	0.00319	-1.1448	0.01133	0	0	1.76156	0.02033	-1.76156	0.01213	0	0

Table I-3. 2010–11 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)*
120071	0.8038	0.02697	-0.18916	0.02391	0	0
120072	0.52096	0.02599	-2.14524	0.08627	0	0
120074	0.68674	0.01783	-0.58309	0.0215	0	0
120087	0.55773	0.01413	-1.65792	0.07859	0.12628	0.03237
120098	0.8508	0.03048	0.97434	0.02799	0	0
120102	0.64149	0.0242	-1.78403	0.08012	0.08075	0.03392
120108	0.45737	0.03462	-0.66636	0.19585	0.26738	0.05276
120119	0.51975	0.00935	-2.57706	0.03552	0	0
120126	1.1106	0.05386	-1.37961	0.05799	0.10867	0.03221
120141	1.50341	0.04914	0.76228	0.014	0.04796	0.00474
120144	0.85034	0.01332	-1.52966	0.02418	0.03361	0.01211
120146	0.35926	0.02293	0.20803	0.11808	0.07847	0.03083
120159	0.87157	0.0312	-1.09083	0.03286	0	0
120173	0.77647	0.04882	0.11698	0.06417	0.16499	0.026
120174	0.56534	0.0161	0.1729	0.02175	0	0
120183	0.75878	0.04379	-2.38426	0.12721	0.14231	0.0594
120187	0.81162	0.03856	-0.54025	0.06602	0.26122	0.02701
120197	0.62696	0.0237	-0.39797	0.03059	0	0
120209	0.60064	0.0148	-1.25408	0.06143	0.14352	0.02505
120211	0.99839	0.01795	0.18273	0.01332	0.12689	0.00615
120214	0.75391	0.03566	-1.42996	0.07476	0.07739	0.0322
120215	1.44442	0.02934	1.12927	0.0085	0.10697	0.00271
120219	0.85813	0.01719	0.30029	0.01611	0.12973	0.0069
120220	1.04933	0.01719	-1.15963	0.02014	0.07962	0.01152
120221	0.57316	0.00819	-1.42507	0.01784	0	0
120232	1.00445	0.04651	-1.33785	0.05782	0.08707	0.02952

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
120236	0.95509	0.07147	-0.1686	0.07987	0.41238	0.02857
120240	0.79203	0.00907	-0.39226	0.00847	0	0
120251	0.91994	0.01932	-1.41153	0.03706	0.21039	0.01951
120253	0.72069	0.03335	-1.40886	0.10092	0.19347	0.04514
120255	1.0856	0.02098	-0.33626	0.01828	0.24857	0.00884
120257	0.83948	0.04059	-0.27042	0.04852	0.07874	0.02162
120261	0.45836	0.01126	-2.00458	0.09095	0.08559	0.03368
120262	0.73796	0.00873	0.29058	0.00835	0	0
120266	0.73675	0.01846	-0.41577	0.01913	0	0
120269	0.75063	0.04652	0.11227	0.06289	0.14795	0.025
120276	0.44368	0.02184	1.5192	0.04016	0.16985	0.01212
120280	0.94765	0.02202	-0.55327	0.01658	0	0
120288	0.65238	0.01489	-3.0498	0.07584	0.09237	0.04052
120291	0.73227	0.03791	-1.34638	0.09021	0.10925	0.04077
123185	0.89879	0.01652	-0.84557	0.02486	0.13803	0.01253
124522	0.50115	0.01001	-3.00924	0.04566	0	0
124525	0.86411	0.01313	-1.74414	0.02266	0.02523	0.01051
124526	0.88049	0.04728	-0.57956	0.06474	0.15629	0.0299
124527	0.80657	0.02447	0.86452	0.02096	0.2928	0.00698
124560	0.65939	0.0126	-1.04216	0.0355	0.05923	0.01565
124592	0.86688	0.03171	-0.79797	0.04751	0.10781	0.02311
124616	0.92477	0.01926	-1.0428	0.03174	0.24281	0.01546
124727	0.63744	0.01454	-0.89299	0.04505	0.1271	0.01858
139424	0.71225	0.0135	-0.83206	0.03052	0.07299	0.01377
139428	1.1714	0.02154	0.59962	0.00994	0.12806	0.00431
139431	0.69488	0.01244	-0.82707	0.02801	0.04864	0.01238
139437	0.71228	0.01316	-1.16809	0.03365	0.06223	0.01578
139447	1.04836	0.01111	0.31353	0.00638	0	0
139455	0.76588	0.01676	0.11584	0.0221	0.15629	0.00903
139465	0.74161	0.04489	-0.45918	0.08505	0.16721	0.03547
139466	0.70597	0.01654	-0.33931	0.0334	0.19555	0.01312
139477	0.39631	0.00768	-2.36256	0.03867	0	0
139479	0.56158	0.00836	1.25388	0.01604	0	0
139493	1.21592	0.02109	0.64662	0.00883	0.09512	0.0037
139495	1.24545	0.03833	1.64742	0.01456	0.17807	0.00303
139517	0.77507	0.00893	0.15689	0.00792	0	0
139521	0.57209	0.01688	-1.23147	0.03413	0	0

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
139523	0.44794	0.00705	-1.03506	0.01798	0	0
144648	0.47126	0.00718	-1.00803	0.01697	0	0
145070	0.60572	0.01001	-2.36626	0.02861	0	0
145079	0.76554	0.03803	-0.18088	0.0614	0.2452	0.02376
145085	0.67739	0.00889	-1.27394	0.0144	0	0
145088	1.30857	0.03177	1.0602	0.01086	0.22507	0.00382
145268	0.82065	0.04581	-1.12665	0.09193	0.16354	0.04398
145281	0.51391	0.02494	-1.15043	0.11889	0.12	0.04217
145545	0.83377	0.01038	-1.30318	0.01248	0	0
145554	0.46648	0.01638	1.26695	0.02923	0.06457	0.00968
145563	0.38617	0.02275	-3.75429	0.23368	0.13626	0.06179
145804	0.809	0.01432	-1.71986	0.03402	0.05678	0.01823
145847	0.78515	0.01542	-0.40192	0.02468	0.128	0.01098
145854	1.09769	0.03883	0.20741	0.02314	0.10757	0.01084
145860	0.24562	0.01519	-1.21611	0.33582	0.17744	0.0601
198351	0.81309	0.04943	-2.41511	0.13137	0.15898	0.06556
198385	0.51796	0.02856	-0.01214	0.08388	0.11012	0.02822
198396	0.76777	0.04754	-0.09104	0.0711	0.16968	0.02913
202346	0.70299	0.02702	-1.86754	0.08179	0.09484	0.03773
202390	0.63613	0.03231	-1.90277	0.1476	0.21379	0.06244
227100	0.61795	0.02016	-2.08732	0.05053	0	0
227853	0.61283	0.06111	0.62003	0.10343	0.2873	0.03136
232445	0.60516	0.02005	-2.08726	0.05213	0	0
232636	0.43721	0.02137	0.91324	0.0506	0	0
255664	0.82535	0.04778	-1.55963	0.10915	0.17919	0.05458
255685	0.93067	0.0295	-1.19939	0.03913	0.05443	0.01906
255692	0.66193	0.0286	-0.94505	0.07832	0.12354	0.03256
255694	0.60527	0.02295	-0.99092	0.06583	0.06732	0.02559

Table I-4. 2010–11 NECAP: IRT Parameters for Polytomous Items – Mathematics Grade 4

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
144645	1.05615	0.02029	0.59652	0.01312	0	0	0.17537	0.0133	-0.17537	0.01395	0	0				
270069	0.7451	0.01431	-0.21386	0.01625	0	0	0.26714	0.01765	-0.26714	0.01707	0	0				
145559	1.05485	0.0083	0.49358	0.00564	0	0	0.35859	0.00617	-0.35859	0.00665	0	0				

120301	0.91112	0.01485	-0.06887	0.013	0	0	0.43518	0.01531	-0.43518	0.0148	0	0
120286	0.68391	0.00614	-1.40338	0.01076	0	0	0.55164	0.01225	-0.55164	0.00937	0	0
139452	0.81691	0.00608	-0.18783	0.00665	0	0	0.56075	0.0081	-0.56075	0.00755	0	0
124661	0.77276	0.00578	0.24226	0.00692	0	0	0.57814	0.00799	-0.57814	0.00826	0	0
120290	0.755	0.01719	-0.83137	0.02344	0	0	0.69465	0.03083	-0.69465	0.02415	0	0
124723	0.77101	0.00576	-0.75647	0.0075	0	0	0.70735	0.00996	-0.70735	0.00787	0	0
120299	0.68526	0.01639	-1.16641	0.02771	0	0	0.75989	0.0368	-0.75989	0.02663	0	0
198431	0.64353	0.00996	0.43263	0.01739	0	0	0.78668	0.0197	-0.78668	0.02166	0	0
120083	0.72483	0.0165	-1.06599	0.02503	0	0	0.82303	0.03516	-0.82303	0.02513	0	0
198442	0.72969	0.0188	-1.76938	0.0311	0	0	0.87421	0.04656	-0.87421	0.02704	0	0
120203	0.69345	0.01023	-0.54872	0.01633	0	0	0.91938	0.02237	-0.91938	0.01821	0	0
139482	0.65126	0.01403	-0.78243	0.0254	0	0	0.95272	0.03561	-0.95272	0.02707	0	0
120075	0.81195	0.00545	-0.58761	0.00676	0	0	1.0548	0.01029	-1.0548	0.00767	0	0
139527	0.71333	0.005	-0.93286	0.00785	0	0	1.08884	0.01233	-1.08884	0.00839	0	0
139529	0.47051	0.0032	-0.52081	0.01084	0	0	1.31302	0.01445	-1.31302	0.01229	0	0
227082	0.90293	0.0139	-1.15826	0.01398	0	0	1.42086	0.03133	-1.42086	0.0149	0	0
120222	0.86745	0.00538	0.04483	0.00718	0	0	1.66803	0.00999	-1.66803	0.00976	0	0
120127	0.45693	0.00395	-2.60148	0.01494	0	0	1.67007	0.028	-1.67007	0.01296	0	0
124700	0.83107	0.016	-0.51733	0.02168	0	0	1.67061	0.03771	-1.67061	0.02537	0	0

Table I-5. 2010–11 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)*
120624	0.81565	0.02782	1.87451	0.02331	0.12808	0.00413
120626	1.17032	0.02339	0.50766	0.01313	0.23108	0.00556
120628	0.61709	0.01301	-1.82427	0.0706	0.09256	0.03333
120634	0.91577	0.01818	-0.18699	0.02351	0.2075	0.01022
120640	0.64458	0.02846	-2.00186	0.14197	0.1703	0.06477
120642	0.74761	0.01502	-1.48362	0.05155	0.10607	0.02662
120644	1.12017	0.04057	-0.59093	0.04172	0.1826	0.02205
120653	0.73609	0.03012	-0.76699	0.07565	0.13164	0.03307
120657	1.27101	0.06175	0.03393	0.03647	0.17376	0.01811
120667	0.68963	0.01746	0.61237	0.02109	0	0
120678	0.66981	0.0146	-0.03763	0.0305	0.09657	0.01197
120682	0.7449	0.04842	-0.08408	0.08929	0.19531	0.03427
120686	0.7876	0.01823	-0.13854	0.01791	0	0
120698	0.89448	0.02035	0.52908	0.01654	0	0
120700	1.28259	0.07913	0.23216	0.04554	0.33942	0.01903
120706	0.69867	0.01582	0.58146	0.02067	0.09091	0.00797
120713	1.17102	0.05315	1.12424	0.02295	0.13412	0.00745

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
120720	1.03188	0.05737	-0.06707	0.05519	0.22611	0.02451
120728	0.68463	0.01794	-0.15146	0.04184	0.24667	0.0146
120733	0.89334	0.04685	-1.96122	0.10274	0.1296	0.05336
120736	1.32793	0.0225	0.4181	0.01009	0.15893	0.00471
120748	0.97145	0.03771	1.33537	0.02155	0.43261	0.00526
120762	1.04797	0.05822	0.62402	0.03871	0.37338	0.01301
120764	1.01284	0.05707	-0.29444	0.06485	0.23365	0.02955
120769	0.57917	0.03004	0.36224	0.06557	0.09659	0.02299
120795	1.0502	0.02531	0.77829	0.01536	0.27662	0.00566
120799	0.5196	0.01007	-2.78978	0.04591	0	0
120803	1.24678	0.04278	-0.49783	0.03344	0.16816	0.01847
120804	0.80467	0.01444	-1.64152	0.04197	0.07175	0.02349
120805	0.80614	0.03965	-0.17599	0.06799	0.27334	0.02562
120808	0.5976	0.02256	0.49971	0.03199	0	0
120810	0.67013	0.02443	0.65625	0.03017	0	0
120822	0.70828	0.02525	-0.76786	0.03578	0	0
120830	0.5199	0.03352	0.5425	0.08567	0.13278	0.02744
120838	0.75367	0.02504	0.15345	0.02537	0	0
120850	0.87068	0.02746	-0.2641	0.02419	0	0
120855	0.88795	0.12723	1.56212	0.08553	0.51608	0.01557
124036	0.8601	0.03716	-0.63174	0.06889	0.22489	0.03038
124760	0.97608	0.03454	-0.60958	0.04544	0.12765	0.02294
124804	0.68997	0.00811	-0.57709	0.01086	0	0
124862	1.03413	0.02113	0.95821	0.01156	0.12133	0.00428
124866	0.59267	0.00808	-1.29979	0.01783	0	0
124881	1.02044	0.02056	0.43494	0.01524	0.20136	0.00638
139086	1.41829	0.02407	0.72415	0.00835	0.12153	0.00353
139129	0.89588	0.00933	-0.15651	0.00768	0	0
139136	1.03593	0.01035	0.01538	0.0067	0	0
139161	0.71675	0.01853	-0.74796	0.05473	0.29413	0.02011
139181	1.05466	0.01619	-0.41115	0.01602	0.08408	0.00824
139344	0.88223	0.01743	-0.08705	0.02264	0.17868	0.00975
139399	0.87156	0.01169	-1.71053	0.0169	0	0
139403	0.82709	0.02256	0.74839	0.02144	0.27036	0.00743
139414	0.93416	0.01685	0.19174	0.01603	0.12006	0.00707
144657	0.76802	0.01376	-0.57619	0.02812	0.07037	0.01283
144659	1.13045	0.02054	0.15782	0.01449	0.19892	0.00665

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
144684	0.78567	0.04271	-0.99332	0.10063	0.15639	0.04534
145096	1.34503	0.02365	0.80793	0.00874	0.11914	0.00353
145573	1.07349	0.02118	-1.18067	0.03229	0.22965	0.01812
145584	0.81809	0.01666	0.13415	0.02173	0.145	0.009
145585	1.15339	0.02021	0.62829	0.01037	0.11717	0.0044
145879	0.62915	0.00787	-0.81574	0.01311	0	0
145887	0.59749	0.0075	-0.47837	0.01175	0	0
145893	1.15917	0.01965	-0.7877	0.02043	0.16255	0.01159
145910	0.79112	0.01671	-0.33993	0.03117	0.18259	0.01302
145922	0.59594	0.00743	-0.27977	0.01097	0	0
150808	0.2231	0.00948	-2.6915	0.32973	0.13277	0.05876
198645	0.69609	0.03537	0.15828	0.06226	0.17539	0.02283
203258	0.66848	0.03287	-1.72079	0.1568	0.22949	0.06815
203301	0.8184	0.0725	1.27628	0.0555	0.21969	0.01718
203556	0.35545	0.01891	-0.77781	0.06412	0	0
203588	0.77365	0.06704	0.6245	0.0824	0.33371	0.02597
225021	0.82966	0.02128	1.12196	0.0228	0	0
225316	0.94811	0.06204	0.41544	0.05596	0.25215	0.02191
226748	1.13929	0.06627	0.44406	0.04105	0.2293	0.01736
226810	0.89409	0.0589	1.27763	0.03947	0.08401	0.01151
255145	0.60285	0.016	0.2793	0.02161	0	0
255761	0.49486	0.01582	-1.4839	0.04711	0	0
255763	1.13922	0.05062	0.65393	0.02701	0.22406	0.0108
269369	0.87248	0.01972	0.36726	0.0161	0	0
269393	1.2151	0.04015	0.03678	0.02496	0.12586	0.01242

Table I-6. 2010–11 NECAP: IRT Parameters for Polytomous Items – Mathematics Grade 5

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
120631	0.84771	0.02237	0.15088	0.02155	0	0	0.19265	0.0231	-0.19265	0.02292	0	0				
120639	0.73606	0.01422	-0.351	0.01801	0	0	0.21239	0.0189	-0.21239	0.01818	0	0				
120726	0.95978	0.00805	0.4901	0.00655	0	0	0.22262	0.00692	-0.22262	0.00715	0	0				
269405	0.97723	0.0231	0.05511	0.01871	0	0	0.2785	0.021	-0.2785	0.02053	0	0				
255766	0.79931	0.01982	0.14145	0.02199	0	0	0.29336	0.02419	-0.29336	0.02406	0	0				
145571	0.82498	0.00652	0.1341	0.00703	0	0	0.34982	0.00789	-0.34982	0.00783	0	0				
228544	0.88448	0.02206	1.20412	0.02451	0	0	0.40767	0.02308	-0.40767	0.02809	0	0				

225389	0.93516	0.01538	0.8555	0.01453	0	0	0.41194	0.01501	-0.41194	0.01723	0	0				
145888	0.78699	0.00605	-0.19505	0.00739	0	0	0.43398	0.0086	-0.43398	0.00805	0	0				
198603	0.69885	0.01716	-0.60406	0.02695	0	0	0.44144	0.03048	-0.44144	0.02696	0	0				
203621	0.88609	0.0143	0.80239	0.01487	0	0	0.45909	0.01561	-0.45909	0.01787	0	0				
120704	0.79148	0.01259	-0.36443	0.01581	0	0	0.50228	0.01899	-0.50228	0.01695	0	0				
255255	0.8217	0.01199	0.60547	0.01482	0	0	0.68122	0.01659	-0.68122	0.0191	0	0				
120694	0.41454	0.00744	0.53392	0.02764	0	0	0.72078	0.02981	-0.72078	0.0314	0	0				
139283	0.67065	0.00477	-1.28665	0.00941	0	0	0.76672	0.01359	0.32747	0.01182	-0.08133	0.01065	-1.01286	0.0093	0	0
269314	0.96922	0.01746	0.44808	0.01686	0	0	0.83526	0.02145	0.48166	0.02053	-0.51018	0.02192	-0.80674	0.02365	0	0
120725	0.87335	0.00558	0.10688	0.00638	0	0	0.85321	0.0082	-0.85321	0.00798	0	0				
230748	1.01142	0.01905	1.1158	0.01769	0	0	0.86273	0.01984	0.37104	0.02034	-0.41584	0.02484	-0.81793	0.02959	0	0
139203	0.71658	0.0053	1.42292	0.00908	0	0	0.86547	0.00883	-0.86547	0.01329	0	0				
120793	1.18806	0.01446	0.84634	0.0101	0	0	0.92439	0.01261	0.27588	0.01248	-0.26766	0.01392	-0.93261	0.01845	0	0
139141	0.67962	0.00443	0.37338	0.00799	0	0	0.92972	0.00951	-0.92972	0.01019	0	0				
272113	0.74557	0.01388	0.1752	0.02219	0	0	1.06759	0.02857	-1.06759	0.02814	0	0				
139384	1.07544	0.00578	-0.23638	0.0049	0	0	1.09332	0.00852	0.33046	0.00674	-0.26147	0.00625	-1.16231	0.0068	0	0
124858	1.08824	0.00607	0.84003	0.00504	0	0	1.11809	0.00648	0.27131	0.00629	-0.41644	0.00729	-0.97296	0.0092	0	0
120731	0.85612	0.00987	0.15665	0.01277	0	0	1.16617	0.0186	0.56483	0.01641	-0.52988	0.01622	-1.20112	0.01868	0	0
145576	0.98783	0.00519	0.39667	0.00528	0	0	1.28922	0.00779	0.54899	0.00676	-0.62195	0.0073	-1.21628	0.00883	0	0
120834	0.95259	0.01049	0.40713	0.01149	0	0	1.32106	0.01675	0.34934	0.01447	-0.35323	0.01517	-1.31717	0.02028	0	0
120767	1.16567	0.01845	0.41373	0.01394	0	0	1.73216	0.0248	0.01813	0.01792	-0.65337	0.01982	-1.09692	0.02319	0	0

Table I-7. 2010–11 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)*
119175	0.72127	0.05048	0.14654	0.08008	0.20515	0.03045
119184	0.97589	0.05618	-0.6925	0.07212	0.21604	0.03632
119186	0.53121	0.02865	-1.56723	0.17008	0.1933	0.06449
119200	0.81311	0.01973	-0.53153	0.01902	0	0
119219	0.98828	0.05444	-0.03568	0.04944	0.17155	0.02335
119221	0.71979	0.0165	0.37749	0.02239	0.13636	0.00893
119222	1.3102	0.02154	0.72995	0.00793	0.08784	0.00337
119231	0.57828	0.01871	0.12725	0.04892	0.26379	0.01534
119232	0.40742	0.02293	-1.2857	0.18343	0.14018	0.05553
119238	0.82482	0.03003	0.23849	0.03039	0.06324	0.01288
119245	0.81802	0.01526	-1.01678	0.03074	0.08274	0.0163
119258	1.2236	0.04173	0.2605	0.02092	0.10906	0.01044
119262	0.4909	0.03779	0.75466	0.09359	0.18949	0.02833
119277	1.19629	0.03865	-0.52988	0.0203	0	0
119279	0.74295	0.01779	0.77287	0.01822	0.13654	0.00707

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
119288	0.29041	0.00612	-0.24098	0.01975	0	0
119295	0.77508	0.01199	-1.35067	0.02454	0.02988	0.01219
119306	0.91282	0.01058	-0.86167	0.00951	0	0
119311	0.50081	0.00707	0.03528	0.01136	0	0
119313	1.05944	0.06556	0.69311	0.03746	0.19324	0.01577
119316	0.8618	0.03167	-0.106	0.03655	0.08531	0.01703
119317	0.89667	0.02593	0.9831	0.01836	0.31501	0.00618
119319	0.71678	0.02612	-0.71214	0.03196	0	0
119328	0.73219	0.01947	1.08352	0.02178	0	0
119332	0.33811	0.0269	-0.9052	0.22523	0.12844	0.05544
119334	0.94079	0.02435	-1.34294	0.04941	0.36867	0.02337
119349	0.99477	0.05147	0.5583	0.03304	0.09206	0.01405
119351	0.85994	0.05662	-0.10064	0.07375	0.26009	0.03039
119356	1.06149	0.03617	1.0227	0.02209	0	0
119361	0.90453	0.01768	-0.26211	0.0216	0.17255	0.01015
119364	1.36035	0.05232	1.19575	0.01659	0.06857	0.0051
119368	0.72154	0.02538	-0.16378	0.02618	0	0
119375	0.69917	0.01894	0.4485	0.02773	0.23051	0.00993
119377	0.79065	0.03846	0.48033	0.04188	0.18934	0.01631
119379	0.89352	0.05756	1.50683	0.03407	0.21383	0.01013
119393	0.47818	0.03382	0.7057	0.08801	0.13644	0.02763
119396	0.81054	0.01937	0.08959	0.0162	0	0
122249	0.29154	0.01274	0.04079	0.0391	0	0
123513	0.61174	0.01709	0.91452	0.0235	0	0
125021	1.14441	0.0198	0.27251	0.01149	0.14139	0.00555
125022	1.05078	0.02051	0.48015	0.01305	0.1791	0.00578
139212	1.14426	0.01983	0.18329	0.01203	0.14889	0.00591
139217	0.58243	0.00755	0.20403	0.00987	0	0
139219	0.7445	0.01153	-0.90439	0.02228	0.02781	0.01034
139274	0.73533	0.01302	-0.70355	0.0266	0.04817	0.01256
139305	1.21813	0.02026	-0.71942	0.01573	0.11868	0.00946
139306	1.24651	0.02704	0.64183	0.01226	0.28171	0.00516
139310	0.87694	0.00967	0.42699	0.0072	0	0
139312	0.69544	0.00838	0.49911	0.00873	0	0
139335	1.14125	0.03373	0.98793	0.01602	0.41413	0.00509
139358	0.91604	0.01292	-1.13031	0.01691	0.02123	0.00833
145128	0.93682	0.01525	-0.21705	0.01538	0.0654	0.00754

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
145133	0.66285	0.00805	0.16843	0.00888	0	0
145326	1.23052	0.02739	0.87268	0.01141	0.24736	0.00458
145328	0.60391	0.02023	-0.3572	0.06402	0.36351	0.01888
145330	0.99029	0.01668	0.27465	0.01215	0.08534	0.0056
145346	0.91841	0.02043	-0.04647	0.02299	0.27841	0.00958
145347	0.61341	0.00794	-0.68012	0.012	0	0
145608	0.67381	0.00937	-1.48376	0.01674	0	0
145969	0.61746	0.01582	0.17908	0.03289	0.1351	0.0123
145978	0.64553	0.01981	0.95202	0.02498	0.19763	0.00866
145988	1.01529	0.02378	0.64409	0.01566	0.27762	0.00613
145995	1.02274	0.01076	0.21196	0.00637	0	0
146037	0.7941	0.01685	-1.18186	0.04132	0.14236	0.02152
198608	0.68744	0.02734	-0.81476	0.06717	0.09473	0.02975
198622	0.81403	0.02722	0.04239	0.02295	0	0
198709	0.7147	0.04616	-0.58008	0.10502	0.18901	0.044
198715	1.10349	0.03565	0.79141	0.01973	0	0
203460	0.82826	0.07088	1.45771	0.0478	0.17351	0.01459
203483	0.76018	0.02027	-1.13071	0.02644	0	0
225273	0.34677	0.05121	1.59314	0.17942	0.16643	0.04446
225313	1.15178	0.04086	0.11597	0.02475	0.13863	0.01214
225428	0.84072	0.03364	0.70109	0.02691	0.0705	0.0109
228068	0.76936	0.03458	-0.51608	0.06697	0.1722	0.02991
228071	0.83185	0.02083	-0.96255	0.02246	0	0
255369	1.52725	0.08248	0.74257	0.02399	0.15996	0.01081
255371	0.69448	0.01764	-0.02281	0.01859	0	0
270370	1.22523	0.05336	-0.14562	0.03002	0.07618	0.01548
270584	1.9307	0.11736	1.74189	0.01987	0.16182	0.00487

Table I-8. 2010–11 NECAP: IRT Parameters for Polytomous Items – Mathematics Grade 6

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
144707	1.2928	0.01284	1.18451	0.00619	0	0	0.13646	0.00584	-0.13646	0.0063	0	0				
119370	0.9911	0.01949	0.46226	0.0131	0	0	0.13839	0.01369	-0.13839	0.01387	0	0				
225287	0.8596	0.01584	0.31906	0.01391	0	0	0.27134	0.01522	-0.27134	0.01529	0	0				
197550	0.96588	0.00882	-0.87075	0.00764	0	0	0.28191	0.0082	-0.28191	0.00713	0	0				

225370	1.0634	0.03245	1.7794	0.0264	0	0	0.33487	0.02207	-0.33487	0.02887	0	0				
119230	0.82541	0.0071	-0.533	0.00748	0	0	0.35677	0.0084	-0.35677	0.00751	0	0				
119229	0.94765	0.02278	-0.22775	0.01846	0	0	0.37655	0.02167	-0.37655	0.01993	0	0				
139377	0.75217	0.0062	0.23506	0.00714	0	0	0.41612	0.00806	-0.41612	0.00804	0	0				
203255	1.0511	0.02412	0.72424	0.01651	0	0	0.45242	0.01834	-0.45242	0.02038	0	0				
125111	1.16524	0.02412	0.36005	0.0141	0	0	0.48684	0.01737	0.24043	0.01693	-0.18503	0.01704	-0.54225	0.01802	0	0
198727	0.93684	0.02244	1.13023	0.01994	0	0	0.52168	0.02037	-0.52168	0.02524	0	0				
146011	0.60351	0.00512	0.74471	0.00894	0	0	0.53827	0.00949	-0.53827	0.01029	0	0				
139351	0.78027	0.01793	0.58735	0.02059	0	0	0.56225	0.02309	-0.56225	0.02485	0	0				
255575	0.98663	0.02227	1.08416	0.01826	0	0	0.65233	0.01942	-0.65233	0.02532	0	0				
119195	0.86319	0.00656	1.19827	0.00693	0	0	0.7179	0.00717	-0.7179	0.00965	0	0				
119281	0.81185	0.0122	-0.34585	0.01409	0	0	0.84086	0.0195	-0.84086	0.0159	0	0				
139356	1.08603	0.00668	0.64729	0.00477	0	0	0.8632	0.00616	0.23206	0.0059	-0.21907	0.00617	-0.8762	0.00738	0	0
255384	0.99756	0.01426	0.09613	0.01158	0	0	0.91713	0.0157	-0.91713	0.01469	0	0				
139322	1.03662	0.02129	1.53628	0.01722	0	0	1.20569	0.01859	1.00124	0.01869	-0.88301	0.03419	-1.32392	0.04498	0	0
119203	0.66516	0.00922	0.06386	0.01646	0	0	1.22202	0.02176	-1.22202	0.02049	0	0				
119341	0.57157	0.00813	0.79278	0.01903	0	0	1.24437	0.02176	-1.24437	0.02636	0	0				
123489	0.87451	0.00488	0.31275	0.00555	0	0	1.30571	0.00826	0.68368	0.00725	-0.5955	0.00728	-1.39389	0.00887	0	0
139387	1.02082	0.00569	0.38514	0.00481	0	0	1.4166	0.00765	0.39107	0.00624	-0.67683	0.00671	-1.13085	0.00763	0	0
123448	1.00871	0.01191	0.86792	0.01062	0	0	1.60223	0.01521	0.14606	0.01367	-0.37113	0.01501	-1.37716	0.02228	0	0
139320	0.70489	0.00425	1.90402	0.00741	0	0	1.92935	0.00839	0.55555	0.00922	-0.58905	0.01296	-1.89585	0.02376	0	0
203632	0.70721	0.00782	0.6156	0.01423	0	0	2.00745	0.02219	0.59593	0.01749	-0.64928	0.0191	-1.95411	0.02828	0	0
256902	0.90753	0.01505	1.17561	0.01706	0	0	2.04646	0.02356	1.23817	0.02067	-0.96716	0.03099	-2.31747	0.06532	0	0

Table I-9. 2010–11 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)*
120328	1.04193	0.04776	-0.06968	0.04098	0.10864	0.02003
120329	0.44927	0.02391	-0.32212	0.12308	0.10505	0.03773
120334	0.71541	0.03876	-0.303	0.07873	0.11691	0.03253
120345	0.54935	0.02093	0.87577	0.04402	0.2798	0.01252
120353	0.69079	0.02056	0.55743	0.03311	0.30824	0.01041
120357	0.34959	0.00644	-1.18217	0.0263	0	0
120361	0.9307	0.01658	-0.59377	0.02425	0.13858	0.01249
120362	0.38776	0.00633	-0.25295	0.01548	0	0
120363	1.21717	0.02526	0.71021	0.01256	0.26485	0.00508
120366	0.70319	0.02397	0.33666	0.0257	0	0
120380	0.79587	0.04253	0.9802	0.03889	0.19482	0.01386
120381	0.85128	0.0162	-0.7777	0.03143	0.14241	0.01581
120385	0.85388	0.03912	0.92227	0.0314	0.13604	0.01187

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
120387	0.66631	0.0078	-0.24221	0.00962	0	0
120406	0.89235	0.00923	0.22155	0.00709	0	0
120409	1.12512	0.03313	0.37819	0.01803	0	0
120411	1.12149	0.06135	-1.54181	0.08518	0.17502	0.05346
120419	0.78781	0.01593	0.08102	0.02356	0.14729	0.00993
120431	0.83239	0.01572	0.26462	0.01836	0.11665	0.00791
120434	0.97599	0.02944	0.14961	0.02005	0	0
120435	0.91445	0.02875	-0.28739	0.02289	0	0
120444	0.54848	0.04325	-0.11414	0.15029	0.19562	0.0487
120446	1.16848	0.01997	0.57881	0.01062	0.1372	0.00473
120450	0.4902	0.00742	-1.32381	0.02106	0	0
120455	0.87344	0.0526	0.38825	0.05521	0.18161	0.02275
120468	0.7331	0.00826	0.56954	0.00867	0	0
120479	0.83723	0.03725	-0.36232	0.06435	0.24102	0.02748
120482	0.76853	0.08265	1.61613	0.06751	0.29631	0.01737
120485	1.32593	0.02764	0.96892	0.01059	0.22712	0.00406
120486	0.74705	0.01824	0.93176	0.01887	0.15206	0.00695
120487	0.70949	0.01739	-0.439	0.02083	0	0
120494	0.75478	0.01982	1.15587	0.01819	0.16369	0.00633
120496	0.35699	0.01801	1.02858	0.08492	0.12559	0.02196
120501	0.97917	0.05228	-0.39167	0.06225	0.1897	0.03009
120509	0.79225	0.01549	-0.69327	0.03336	0.13356	0.01591
120511	0.80517	0.04192	-0.24723	0.07756	0.34813	0.02791
120517	0.65815	0.01678	0.77239	0.02155	0	0
120531	0.3262	0.0063	-1.1213	0.02716	0	0
122332	0.31976	0.01156	0.16819	0.09787	0.06172	0.0236
122346	0.933	0.03502	-0.07653	0.03929	0.15222	0.01821
125286	0.37692	0.01978	-1.1233	0.07226	0	0
125349	0.96608	0.05525	1.25059	0.03313	0.25609	0.01067
125397	0.76943	0.00872	-0.59109	0.00981	0	0
139925	1.02677	0.02008	0.46603	0.01492	0.20717	0.00636
139994	0.81414	0.03915	0.56146	0.0436	0.20692	0.01671
140025	0.39719	0.01473	1.69224	0.0543	0	0
140026	0.65683	0.01714	0.94679	0.02337	0	0
140233	0.97579	0.01692	-0.35216	0.01996	0.14366	0.01009
140259	0.60439	0.00989	-0.64757	0.0282	0.0285	0.01124
140279	0.67205	0.01704	1.21175	0.01812	0.08976	0.00633

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
140280	0.61114	0.00757	-0.4771	0.01128	0	0
144726	0.45522	0.0068	-0.75845	0.01665	0	0
145151	0.85521	0.05855	1.05442	0.04603	0.16666	0.01616
145376	0.83849	0.01616	0.70483	0.01449	0.09298	0.00585
146220	0.83228	0.08925	1.97625	0.06911	0.21184	0.01286
146230	1.16157	0.01175	0.79044	0.00651	0	0
146263	0.52843	0.05452	2.19267	0.07849	0.18333	0.01617
146266	0.84485	0.05568	0.48658	0.06066	0.2183	0.02351
154775	1.06426	0.0764	0.8438	0.0476	0.32424	0.01668
181194	1.20923	0.02443	0.93007	0.01101	0.19546	0.00428
181196	0.90179	0.02112	0.71159	0.01812	0.24896	0.00675
181199	0.45821	0.00781	-1.90368	0.03054	0	0
181200	0.94916	0.01855	0.24059	0.0179	0.20072	0.00775
181202	0.96965	0.01683	-1.37456	0.03027	0.07813	0.01881
181205	0.78824	0.0159	0.27988	0.02084	0.13242	0.00867
199804	0.63148	0.03193	-0.32917	0.09244	0.18088	0.03438
199900	1.45518	0.09883	1.75947	0.0286	0.2586	0.00634
199905	0.53438	0.04108	1.28005	0.06755	0.0682	0.02064
206092	1.39044	0.05237	0.53532	0.02113	0.20343	0.0098
206107	0.88898	0.05473	1.25858	0.03786	0.08847	0.01165
206112	0.77392	0.01829	0.75965	0.01874	0	0
206144	1.47917	0.06446	0.40253	0.02589	0.35213	0.01132
224764	0.84102	0.0307	-1.61657	0.07326	0.10681	0.04049
224796	1.25662	0.05159	0.80205	0.02273	0.20366	0.00922
234445	1.20995	0.05377	0.83066	0.02545	0.24667	0.00988
234458	0.67054	0.02441	-0.7971	0.03683	0	0
234459	0.37055	0.02085	1.91745	0.09131	0	0
256141	0.9582	0.05335	0.35792	0.0469	0.16631	0.02043
256152	0.92018	0.04586	1.50056	0.02864	0.10135	0.00788
269013	0.79337	0.05238	-0.04662	0.08452	0.25352	0.03248
269118	1.40722	0.10313	1.38165	0.03253	0.23786	0.01006

Table I-10. 2010–11 NECAP: IRT Parameters for Polytomous Items – Mathematics Grade 7

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)

224856	0.81765	0.01769	1.30674	0.02075	0	0	0.1468	0.0175	-0.1468	0.01872	0	0				
154762	1.02429	0.02645	0.48813	0.01833	0	0	0.15226	0.01911	-0.15226	0.01953	0	0				
150893	0.85046	0.01779	1.2496	0.01952	0	0	0.16743	0.01688	-0.16743	0.01824	0	0				
120523	1.10699	0.02838	1.01919	0.01915	0	0	0.20944	0.01879	-0.20944	0.02052	0	0				
225135	0.84612	0.01661	1.29947	0.01901	0	0	0.25939	0.01677	-0.25939	0.019	0	0				
120427	0.84313	0.00689	0.89841	0.00735	0	0	0.34295	0.00741	-0.34295	0.00821	0	0				
120492	0.95965	0.01575	0.77805	0.01348	0	0	0.35347	0.01413	-0.35347	0.01561	0	0				
125382	1.0557	0.00756	0.56599	0.00549	0	0	0.43235	0.00618	-0.43235	0.00669	0	0				
120417	1.15535	0.02416	0.58458	0.01521	0	0	0.47473	0.01745	-0.47473	0.01926	0	0				
255899	1.36804	0.02089	1.02602	0.01018	0	0	0.52313	0.01106	-0.52313	0.01442	0	0				
199932	0.88539	0.0196	1.14632	0.02113	0	0	0.60256	0.02146	-0.60256	0.0278	0	0				
139980	0.99627	0.00711	1.1855	0.0063	0	0	0.69237	0.00653	-0.69237	0.00916	0	0				
146128	1.12877	0.0173	1.28767	0.01244	0	0	0.72175	0.01279	-0.72175	0.01972	0	0				
120470	1.23296	0.00842	0.97128	0.00499	0	0	0.74435	0.00557	-0.74435	0.00778	0	0				
139987	0.98233	0.00592	0.6179	0.00535	0	0	0.79451	0.0066	0.18614	0.00643	-0.21882	0.00672	-0.76184	0.00767	0	0
140223	1.17011	0.00681	0.90979	0.00465	0	0	0.91808	0.00575	0.26411	0.00577	-0.34053	0.00657	-0.84166	0.008	0	0
140251	1.15553	0.0061	0.65218	0.00443	0	0	1.12503	0.00617	0.46089	0.00568	-0.26139	0.0061	-1.32454	0.00899	0	0
140241	0.88465	0.00585	1.23878	0.00662	0	0	1.17496	0.00712	-1.17496	0.01223	0	0				
269083	1.36738	0.02403	1.35125	0.0127	0	0	1.1871	0.01553	0.32174	0.01683	-0.54446	0.02351	-0.96437	0.03054	0	0
206198	1.14299	0.02014	1.53643	0.01514	0	0	1.3324	0.01753	0.32616	0.01987	-0.40045	0.02629	-1.25811	0.04458	0	0
120393	1.07533	0.01666	0.81486	0.01416	0	0	1.41789	0.02	0.3895	0.0181	-0.44739	0.021	-1.35999	0.03072	0	0
120504	0.95485	0.01213	1.75282	0.01273	0	0	1.51339	0.01406	0.94552	0.01462	-0.55408	0.02405	-1.90483	0.05621	0	0
144756	0.79394	0.00478	0.82853	0.0074	0	0	1.62105	0.00859	-1.62105	0.0128	0	0				
256095	0.62196	0.0119	1.48197	0.02687	0	0	1.66688	0.02875	-1.66688	0.05286	0	0				
140292	1.18212	0.00615	0.77631	0.00439	0	0	1.71038	0.00695	0.29803	0.0057	-0.84296	0.00751	-1.16545	0.00874	0	0
120355	1.18297	0.01279	0.76249	0.00934	0	0	1.80789	0.0154	0.26128	0.01203	-0.26384	0.01308	-1.80532	0.02691	0	0
234455	0.59088	0.01163	1.70025	0.02843	0	0	1.91365	0.02993	-1.91365	0.06461	0	0				
256118	0.88562	0.01031	1.68064	0.01262	0	0	2.13561	0.01568	1.27389	0.01491	-0.92508	0.0287	-2.48442	0.07593	0	0

Table I-11. 2010–11 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)*
120877	0.94129	0.04788	0.1257	0.06123	0.16313	0.02601
120882	1.34936	0.01827	-0.8423	0.01657	0.05683	0.01063
120886	1.36083	0.05361	-0.67012	0.0434	0.09787	0.0259
120887	1.36753	0.02147	0.1789	0.01311	0.21013	0.00644
120889	1.28743	0.01182	-0.39665	0.00688	0	0
120895	1.25514	0.07154	0.66009	0.04478	0.53169	0.01227
120896	0.87931	0.01753	0.26913	0.02444	0.19951	0.00967
120910	1.09997	0.02508	1.10698	0.01454	0.23713	0.00493
120912	1.46787	0.07577	0.23339	0.04124	0.3039	0.01923

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
120915	0.97932	0.04646	-1.11122	0.09646	0.1562	0.05221
120919	1.05569	0.05065	-0.07856	0.05797	0.17267	0.02697
120921	0.97726	0.05078	-0.75414	0.09709	0.20934	0.0479
120923	1.01029	0.01574	-0.301	0.02167	0.11271	0.01084
120929	1.05817	0.02056	0.78568	0.01463	0.1906	0.0057
120931	0.8835	0.01794	0.32026	0.01646	0	0
120932	0.53142	0.02043	-0.61365	0.04666	0	0
120933	0.94252	0.00937	-0.54372	0.00936	0	0
120946	0.68858	0.01516	-0.00804	0.04006	0.13731	0.01504
120960	1.64819	0.02686	0.2049	0.01156	0.27709	0.00581
120961	0.96242	0.02031	0.43312	0.02216	0.27274	0.0083
120974	1.21787	0.05658	0.37008	0.03812	0.16155	0.01719
120976	1.10968	0.05892	0.3689	0.04948	0.22838	0.02061
120987	0.78611	0.01392	-0.87535	0.043	0.08699	0.02093
120989	0.69663	0.01683	-0.02339	0.04565	0.21487	0.01599
120998	1.19989	0.06055	0.817	0.03365	0.14502	0.01326
121004	0.41089	0.02478	0.31994	0.14218	0.10022	0.03684
121008	0.90489	0.01972	0.58948	0.02169	0.23866	0.00801
121025	1.29187	0.06246	-0.04291	0.04856	0.24472	0.02353
121027	0.90535	0.01845	-0.0958	0.01705	0	0
121038	0.86743	0.02408	1.54661	0.01849	0.16228	0.00497
121039	1.10593	0.01069	0.95708	0.00789	0	0
121044	1.07774	0.03354	1.6874	0.01882	0.24326	0.00404
121048	0.97192	0.02111	1.09206	0.0199	0	0
121073	1.02508	0.06553	1.66208	0.03987	0.22809	0.00912
121078	0.60925	0.01501	-0.37026	0.06322	0.14511	0.02255
121081	1.45542	0.05502	0.61125	0.02435	0.26428	0.01015
121083	1.23348	0.03276	0.33983	0.0182	0	0
121087	1.05351	0.03876	0.85325	0.02545	0.11161	0.00965
121089	1.22604	0.04242	1.02237	0.02047	0.0865	0.00717
121091	1.04824	0.03817	0.02393	0.04344	0.19581	0.01993
122522	0.71793	0.02845	-0.03698	0.06709	0.10434	0.0266
122667	1.13798	0.04515	0.86551	0.02706	0.17379	0.01071
122688	0.82541	0.01475	0.57569	0.01743	0.07186	0.00702
123801	0.94185	0.02135	0.77023	0.01957	0.25393	0.00701
125503	0.6112	0.01479	-0.12724	0.05342	0.1271	0.0189
125580	0.52975	0.01157	-0.12141	0.05063	0.05246	0.01708

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
139760	0.77829	0.04087	1.50312	0.03781	0.09158	0.01052
139766	0.73843	0.01431	0.00012	0.03158	0.09971	0.01276
139771	0.92827	0.02035	-0.65099	0.02188	0	0
139782	0.7147	0.07848	1.75295	0.08365	0.25432	0.01921
139837	0.58035	0.02468	-0.30194	0.09883	0.1005	0.03446
139845	0.35833	0.00627	-0.92273	0.0269	0	0
139847	1.04134	0.01911	0.34779	0.01826	0.21479	0.00764
139853	1.21601	0.01833	0.09716	0.01422	0.15037	0.00702
139863	1.08593	0.01036	-0.488	0.00815	0	0
139875	1.17156	0.02234	0.96515	0.0121	0.17232	0.00454
139880	1.67009	0.12378	1.60554	0.0342	0.16916	0.00778
144760	0.83858	0.01524	-0.2214	0.03015	0.13188	0.01323
145176	1.38224	0.02071	0.41843	0.01064	0.15323	0.00504
145640	1.17231	0.05873	1.54068	0.02963	0.13976	0.00683
146286	1.68033	0.02811	1.04453	0.00805	0.13382	0.00294
146287	1.73277	0.06375	1.0852	0.01717	0.14942	0.00623
146291	1.29479	0.01897	0.31242	0.01148	0.13335	0.00552
146325	0.83775	0.01605	-0.58356	0.03993	0.16694	0.01812
146358	1.00073	0.06115	1.04784	0.04214	0.15258	0.01514
146362	1.1832	0.02523	0.34936	0.02018	0.39068	0.00732
146422	0.81814	0.00901	1.19016	0.01156	0	0
146455	1.05709	0.09762	1.4428	0.05555	0.32575	0.01446
152804	1.34575	0.02193	0.60048	0.011	0.17854	0.00485
194560	1.04467	0.01693	-0.29814	0.02238	0.15421	0.01112
199731	0.96451	0.0492	0.65656	0.04292	0.11076	0.01737
199756	0.68508	0.03488	0.0734	0.09199	0.21772	0.03132
199768	0.87173	0.02525	0.4661	0.02395	0	0
206256	1.28122	0.06571	0.73428	0.03434	0.18716	0.01401
206323	1.02106	0.02836	0.25479	0.02083	0	0
224873	1.74121	0.0679	1.10775	0.01803	0.17277	0.00629
224878	1.0096	0.04895	1.07783	0.03337	0.21888	0.01169
224936	1.04266	0.02945	0.77228	0.02275	0	0
225437	0.99099	0.02972	1.05827	0.02715	0	0
233602	0.95668	0.04074	0.41029	0.04388	0.24535	0.01667
233737	1.38252	0.11969	1.80412	0.04829	0.18594	0.00845
242392	1.24261	0.02363	0.55315	0.01308	0	0
242467	0.69018	0.01667	1.0047	0.02575	0	0

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
256121	1.32085	0.07724	1.09976	0.03353	0.19655	0.0116
256511	0.52649	0.02709	-0.78752	0.17743	0.17105	0.05751

Table I-12. 2010–11 NECAP: IRT Parameters for Polytomous Items – Mathematics Grade 8

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
256320	1.47253	0.04381	1.53755	0.02495	0	0	0.09321	0.02067	-0.09321	0.02261	0	0				
139777	1.26069	0.02203	0.87775	0.01338	0	0	0.10233	0.01299	-0.10233	0.01355	0	0				
139869	1.07282	0.01821	0.49975	0.01362	0	0	0.11645	0.01409	-0.11645	0.01431	0	0				
120958	0.988	0.0085	1.26141	0.00887	0	0	0.18803	0.00772	-0.18803	0.00858	0	0				
120936	1.17299	0.02488	-0.01988	0.01723	0	0	0.22289	0.01928	-0.22289	0.01857	0	0				
120985	1.14218	0.02414	0.5997	0.01766	0	0	0.24983	0.01894	-0.24983	0.01996	0	0				
269098	1.17173	0.02355	0.77974	0.01736	0	0	0.35653	0.01866	-0.35653	0.02096	0	0				
224947	1.03326	0.01439	0.56424	0.01282	0	0	0.40538	0.01437	-0.40538	0.01536	0	0				
139885	0.80764	0.01091	0.02501	0.01537	0	0	0.5619	0.01853	-0.5619	0.01763	0	0				
120890	0.74925	0.0106	-0.36387	0.01752	0	0	0.58164	0.02155	-0.58164	0.01853	0	0				
269172	0.95337	0.03636	2.913	0.06762	0	0	0.5838	0.03886	-0.5838	0.08281	0	0				
125563	1.08268	0.00643	0.15934	0.00557	0	0	0.60047	0.00706	-0.60047	0.0068	0	0				
139813	0.63952	0.0046	1.3112	0.01061	0	0	0.63957	0.01014	-0.63957	0.0128	0	0				
145660	0.82496	0.00505	0.57515	0.00714	0	0	0.67453	0.0082	-0.67453	0.00906	0	0				
206324	0.6332	0.00874	-0.07904	0.01913	0	0	0.68427	0.02305	-0.68427	0.02154	0	0				
120876	1.15241	0.02009	-0.09722	0.01621	0	0	0.73661	0.02294	-0.73661	0.01931	0	0				
146483	1.06811	0.00606	0.18985	0.00564	0	0	0.80303	0.00744	-0.80303	0.00719	0	0				
121085	1.05281	0.01232	0.96205	0.01217	0	0	0.81453	0.01415	0.29008	0.01429	-0.3823	0.01661	-0.72232	0.01899	0	0
121002	1.338	0.02059	0.71747	0.01314	0	0	0.89855	0.01742	0.19043	0.01675	-0.38921	0.01883	-0.69977	0.02125	0	0
145661	0.65643	0.0042	1.2885	0.00963	0	0	0.90461	0.00978	-0.90461	0.01363	0	0				
139795	1.34508	0.00634	0.51193	0.00419	0	0	1.04913	0.00625	0.48291	0.00557	-0.42759	0.00597	-1.10446	0.00775	0	0
256329	1.16704	0.01181	0.67025	0.01018	0	0	1.18808	0.0143	0.35881	0.01293	-0.55507	0.01516	-0.9918	0.01817	0	0
120942	1.23987	0.00578	0.82965	0.00457	0	0	1.30893	0.00649	0.52258	0.00585	-0.5554	0.00725	-1.27611	0.01057	0	0
144778	1.18566	0.00562	1.1366	0.00485	0	0	1.43926	0.00644	0.54356	0.00614	-0.45026	0.00809	-1.53256	0.01583	0	0
139827	1.04638	0.00503	1.1072	0.00544	0	0	1.4636	0.00709	0.36452	0.00686	-0.68526	0.00942	-1.14284	0.01201	0	0
121071	1.15358	0.01825	1.64912	0.0164	0	0	1.52104	0.01888	0.22681	0.02305	-0.66937	0.03671	-1.07848	0.04928	0	0
206331	1.04915	0.00975	0.12328	0.01118	0	0	1.55783	0.02171	0.76415	0.01594	-0.37714	0.01414	-1.94485	0.02298	0	0
121099	1.17188	0.01818	1.63754	0.01711	0	0	2.3641	0.02195	-0.08384	0.02618	-1.0455	0.04688	-1.23476	0.05424	0	0

Table I-13. 2010–11 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)*
119419	0.82429	0.02478	-0.67517	0.05631	0.39261	0.02006
119420	0.65376	0.01789	-0.18301	0.04566	0.14373	0.01739
119430	0.96328	0.07695	1.44725	0.045	0.11803	0.01205
119437	1.08288	0.06212	0.75493	0.03795	0.33914	0.01337
119442	0.91937	0.07461	1.34582	0.04787	0.14707	0.01415
119450	1.85144	0.09913	0.759	0.0219	0.13072	0.00972
119451	1.5191	0.03968	1.33799	0.01105	0.16258	0.00336
119452	0.81881	0.05215	0.02866	0.07435	0.16489	0.03044
119457	0.92902	0.01119	0.52202	0.00851	0	0
119464	1.14544	0.02644	0.6033	0.01479	0	0
119476	1.59595	0.0949	0.96445	0.02551	0.14407	0.00999
119484	1.63785	0.08693	0.32565	0.02777	0.18064	0.01422
119488	0.74103	0.05584	-0.05532	0.10681	0.24186	0.03893
119492	0.62067	0.00985	-1.42338	0.02158	0	0
119494	1.06603	0.03002	1.42441	0.02388	0	0
119516	1.57687	0.13601	1.5655	0.03414	0.41323	0.00771
119520	1.01389	0.03307	-0.01975	0.02223	0	0
119534	0.60853	0.01386	-1.88532	0.07345	0.08443	0.03462
119542	1.09306	0.03603	0.61798	0.02177	0	0
119553	1.00362	0.03435	0.701	0.02411	0	0
119562	0.99236	0.03586	1.04774	0.02808	0	0
119569	0.49615	0.03207	1.81891	0.04664	0.25348	0.01258
119573	0.76136	0.03716	-0.63059	0.08857	0.19996	0.03729
119576	0.79281	0.02588	0.78762	0.02684	0.2782	0.00906
119603	0.93783	0.03592	1.23548	0.03334	0	0
119605	1.22564	0.02715	0.02319	0.01368	0	0
119606	1.11034	0.01295	0.60893	0.0076	0	0
119616	0.48364	0.0079	0.30195	0.01412	0	0
119630	0.97701	0.05023	0.33474	0.04574	0.29272	0.01754
125740	1.17618	0.04876	1.58299	0.03517	0	0
125772	0.81512	0.01103	1.05143	0.01171	0	0
130081	1.32999	0.07007	0.49176	0.03412	0.39176	0.01293
130121	1.28754	0.01697	1.25155	0.00917	0	0
139984	0.91738	0.07222	1.36854	0.04568	0.11907	0.013
140029	1.04567	0.02243	0.25368	0.01775	0.1839	0.00789
140031	0.72948	0.0097	0.63794	0.01068	0	0

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
140042	1.47603	0.02837	0.52426	0.01022	0.15992	0.00479
140044	1.46021	0.03662	1.13597	0.01116	0.19859	0.00395
140051	1.33905	0.04209	1.08279	0.01543	0.36754	0.00497
140079	0.99721	0.02469	0.29602	0.02189	0.26951	0.00877
140135	1.28511	0.01849	1.50436	0.01102	0	0
140155	0.57368	0.00923	-1.31105	0.02159	0	0
140156	1.21884	0.04121	1.3403	0.01629	0.30654	0.00473
140181	1.42472	0.02895	0.33878	0.01242	0.22178	0.00595
140196	0.79193	0.07268	0.7604	0.07857	0.30939	0.02536
140203	1.08103	0.01277	0.67042	0.00789	0	0
141248	1.10162	0.02477	0.45618	0.01619	0.21143	0.00692
141256	1.29972	0.04811	0.38675	0.02289	0.13351	0.01085
144800	1.25453	0.02445	0.65862	0.01092	0.11909	0.00469
144804	1.39875	0.0281	-0.32582	0.01708	0.2479	0.00928
145233	0.99542	0.01692	0.15901	0.01358	0.0442	0.00604
145238	0.75363	0.02669	0.98476	0.02694	0.26725	0.00884
145444	1.3497	0.02752	0.81991	0.01029	0.13236	0.00421
145456	1.05781	0.03335	1.39325	0.01642	0.20787	0.0048
145465	0.74145	0.01025	0.98092	0.01222	0	0
145484	1.35073	0.0212	1.70143	0.01253	0	0
145492	0.9429	0.01104	0.02701	0.00824	0	0
181297	0.95211	0.03494	1.6325	0.02045	0.20181	0.00491
181299	0.88801	0.03016	0.78438	0.02617	0.36904	0.0082
181301	0.65113	0.02659	1.78295	0.02767	0.11771	0.00676
259790	0.8978	0.06572	1.50223	0.04541	0.0637	0.01026
259814	1.12907	0.0385	0.87476	0.02352	0	0
259828	1.28632	0.05207	0.52789	0.02433	0.1738	0.01096
259873	1.14467	0.10714	1.39201	0.04663	0.26441	0.01325
259914	0.93857	0.06089	1.31229	0.03752	0.23263	0.0114
259927	1.12867	0.08849	1.17467	0.04084	0.2107	0.01385
259947	1.01034	0.07903	1.01557	0.04595	0.22317	0.01634
259989	0.90128	0.02244	0.74837	0.01859	0	0
270801	1.22267	0.04508	1.20744	0.02623	0	0
270853	0.88133	0.06736	1.05326	0.04822	0.15523	0.01711
270856	0.67264	0.07123	1.6379	0.0728	0.1497	0.01908
270868	0.87917	0.05843	0.76696	0.04694	0.12835	0.01789

Table I-14. 2010–11 NECAP: IRT Parameters for Polytomous Items – Mathematics Grade 11

IREF	*Parameters and Measures of Standard Error*															
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
119543	1.01062	0.01138	0.72601	0.00816	0	0	0.05981	0.00789	-0.05981	0.00802	0	0				
140035	1.25112	0.01397	0.96635	0.0075	0	0	0.07278	0.00705	-0.07278	0.00729	0	0				
146944	1.46588	0.01684	1.31769	0.00765	0	0	0.1304	0.00689	-0.1304	0.00763	0	0				
260001	1.19516	0.03346	1.35261	0.02275	0	0	0.34168	0.02111	-0.34168	0.02685	0	0				
259921	0.81309	0.02489	1.58609	0.03445	0	0	0.36275	0.02844	-0.36275	0.03533	0	0				
145498	1.32232	0.01078	0.06326	0.00542	0	0	0.36375	0.00659	-0.36375	0.00632	0	0				
260675	0.59161	0.02214	2.36292	0.06447	0	0	0.43006	0.0416	-0.43006	0.05411	0	0				
119481	1.03146	0.01074	1.73436	0.01035	0	0	0.47045	0.00873	-0.47045	0.01294	0	0				
259958	1.64714	0.03848	1.24767	0.01466	0	0	0.52984	0.01598	0.37747	0.01634	-0.30207	0.02087	-0.60523	0.02515	0	0
119620	0.89155	0.02284	1.3098	0.02579	0	0	0.53686	0.02459	-0.53686	0.0326	0	0				
119546	1.13366	0.02301	0.46935	0.01621	0	0	0.59979	0.02013	0.35252	0.01962	-0.29811	0.02027	-0.65419	0.02201	0	0
141291	1.50591	0.01134	0.90865	0.00489	0	0	0.65876	0.00577	0.48304	0.00577	-0.51368	0.00738	-0.62814	0.00781	0	0
119472	0.83555	0.01321	0.74429	0.01589	0	0	0.69674	0.01767	-0.69674	0.02081	0	0				
119502	1.04599	0.01644	0.93372	0.01362	0	0	0.71351	0.01491	-0.71351	0.01982	0	0				
141322	1.42291	0.00922	0.3743	0.00463	0	0	0.9592	0.00672	0.8723	0.00657	-0.53982	0.00642	-1.2917	0.00856	0	0
145227	0.90679	0.00784	1.68394	0.00891	0	0	0.98553	0.00856	-0.98553	0.01718	0	0				
141365	1.19047	0.00756	0.65328	0.00524	0	0	1.02476	0.00712	0.19181	0.00668	-0.28074	0.00715	-0.93582	0.00896	0	0
119621	1.29242	0.0167	1.0724	0.01023	0	0	1.09917	0.01288	0.4187	0.0129	-0.32716	0.01579	-1.19071	0.02553	0	0
141307	1.00903	0.00766	1.81369	0.00762	0	0	1.19869	0.00778	0.23609	0.00967	-0.41048	0.01306	-1.02429	0.01912	0	0
119589	1.21928	0.02613	1.80219	0.0181	0	0	1.24544	0.01907	0.10165	0.02608	-0.43866	0.03566	-0.90842	0.0506	0	0
119460	1.43046	0.02741	1.47821	0.01398	0	0	1.33702	0.01699	0.40467	0.0188	-0.52003	0.02921	-1.22165	0.05129	0	0

Table I-15. 2010–11 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)*
117641	1.34048	0.0213	-0.12543	0.01323	0.18868	0.00676
117642	1.40062	0.02331	-0.27476	0.01478	0.24492	0.00761
117643	1.13076	0.02296	-1.46643	0.04066	0.26791	0.02226
117644	0.67354	0.01564	-0.49366	0.0467	0.16309	0.01727
117661	0.65721	0.00917	-1.84894	0.02444	0	0
117665	0.78722	0.0173	-0.1036	0.03069	0.20164	0.01169
117666	0.72	0.0178	0.16661	0.03212	0.20381	0.01138
117667	0.48302	0.01805	0.16364	0.07522	0.19864	0.0205
117668	0.69716	0.01272	-1.38218	0.04937	0.06978	0.02228
117732	0.66217	0.00784	-0.6589	0.01234	0	0
117735	1.15114	0.01962	-1.57432	0.0315	0.10568	0.01982

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
117757	0.83937	0.01601	-1.03582	0.04028	0.14735	0.01876
117760	1.56509	0.02846	-1.60986	0.02383	0.13559	0.01733
117762	1.27095	0.02042	-1.20028	0.02291	0.12836	0.0139
117763	0.72396	0.01476	-0.12482	0.02937	0.10389	0.01156
117787	0.57985	0.03857	-0.56058	0.14456	0.16058	0.04713
117789	1.24158	0.06176	-0.43832	0.05243	0.23787	0.02575
117790	0.44098	0.03375	-1.16201	0.26689	0.19491	0.07067
117793	0.81438	0.05396	-0.37058	0.1044	0.29523	0.03716
117794	0.50075	0.03412	-0.58639	0.16351	0.13717	0.04858
117796	1.34565	0.07377	0.58518	0.03131	0.18376	0.01329
117797	1.31305	0.07453	0.40926	0.03684	0.24448	0.01592
117798	0.97968	0.05194	-0.76384	0.0815	0.22298	0.03668
117808	0.89313	0.04523	-0.7794	0.08173	0.16157	0.0362
147671	1.39009	0.02162	-0.91674	0.01794	0.15391	0.01069
147673	1.02345	0.01757	-1.07212	0.02982	0.14589	0.01572
147674	0.64183	0.0082	-1.28827	0.01788	0	0
147679	1.47569	0.02136	-0.45666	0.01254	0.14293	0.00715
147682	0.82816	0.01798	-1.11213	0.05079	0.24278	0.02178
147719	1.21524	0.01926	-0.55792	0.01794	0.16809	0.00941
147733	1.41951	0.02286	-0.74312	0.01732	0.20617	0.00991
147739	1.1762	0.01903	-1.49041	0.02745	0.08665	0.01709
147819	1.05382	0.04447	-0.89486	0.05761	0.09481	0.02794
147835	0.7038	0.03258	-0.60173	0.07463	0.07392	0.02784
147847	0.86222	0.04816	-0.36045	0.07986	0.18918	0.03247
147849	0.74464	0.05828	0.96589	0.05987	0.15909	0.02014
147851	1.79688	0.0834	-1.09571	0.0413	0.14158	0.0267
147856	0.45651	0.03717	-0.38326	0.20925	0.1683	0.05596
147865	1.03497	0.052	-0.20434	0.05476	0.17323	0.02449
147870	1.28395	0.05842	-0.58481	0.04868	0.17068	0.02536
148024	0.48205	0.00754	-1.72131	0.02895	0	0
148039	1.43892	0.02283	-1.00923	0.01842	0.1604	0.01131
148041	1.38003	0.01976	-0.25227	0.01181	0.12404	0.00634
148054	0.54975	0.0114	-1.7266	0.07764	0.07799	0.03043
148178	0.92444	0.0459	-0.93918	0.08546	0.15811	0.03932
148198	0.60547	0.00827	-1.56651	0.02204	0	0
148219	0.91296	0.04018	-1.23349	0.07489	0.09777	0.03437
148224	0.86508	0.04069	-1.48964	0.0982	0.12414	0.0459

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
202178	1.37782	0.0652	-0.64395	0.04975	0.21144	0.02659
202179	0.72622	0.03827	-1.55129	0.13769	0.15768	0.05783
202180	0.97484	0.05429	-0.54822	0.08074	0.24818	0.0343
202183	0.72862	0.04149	-1.24238	0.14161	0.18853	0.057
270034	0.53181	0.04521	0.18931	0.14092	0.17924	0.0415
270042	1.06879	0.05431	0.25128	0.0401	0.14019	0.01753
270043	0.70531	0.03389	-1.37751	0.10618	0.10591	0.04196
270062	0.99272	0.05227	-0.53757	0.07149	0.216	0.03179

Table I-16. 2010–11 NECAP: IRT Parameters for Polytomous Items – Reading Grade 3

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
117650	0.80438	0.00391	-0.64092	0.00679	0	0	1.83893	0.01632	0.3593	0.00945	-0.58727	0.00823	-1.61095	0.00906	0	0
117801	0.70351	0.00985	-1.22148	0.02453	0	0	1.86587	0.06359	1.11711	0.04617	-0.23389	0.03083	-2.74908	0.0346	0	0
147742	1.13086	0.00524	0.44884	0.00512	0	0	1.93478	0.00945	0.82705	0.00662	-0.79898	0.0079	-1.96284	0.01537	0	0
117669	0.96348	0.00447	0.81249	0.00585	0	0	2.08742	0.00936	0.64724	0.00714	-0.77925	0.00987	-1.95543	0.01919	0	0
230973	0.81641	0.01099	0.32681	0.02009	0	0	2.09721	0.03623	0.72506	0.02527	-0.74378	0.02727	-2.07849	0.04566	0	0
147875	0.67092	0.01123	-2.08808	0.02674	0	0	2.27771	0.11983	0.87671	0.0611	-0.87789	0.03376	-2.27653	0.02899	0	0
147889	0.95558	0.01372	1.2524	0.01809	0	0	2.42987	0.02747	0.99205	0.02169	-0.84537	0.03792	-2.57655	0.1235	0	0
117802	0.49323	0.00812	-3.26275	0.03532	0	0	2.61986	0.17879	1.96326	0.13784	-0.79105	0.05419	-3.79206	0.03849	0	0
147744	0.57408	0.00303	-1.62627	0.00994	0	0	3.04074	0.04163	1.37233	0.02086	-1.85831	0.01106	-2.55476	0.0117	0	0
117766	0.73662	0.00337	0.85366	0.00777	0	0	3.27052	0.01614	0.74084	0.00892	-1.16892	0.01322	-2.84244	0.0304	0	0
148057	0.56845	0.00254	1.07743	0.00975	0	0	3.36481	0.01647	1.25468	0.01118	-1.27121	0.01631	-3.34829	0.03717	0	0
270066	0.66969	0.00923	1.3773	0.02532	0	0	3.44732	0.0429	1.15876	0.02894	-1.19157	0.05142	-3.41452	0.15773	0	0

Table I-17. 2010–11 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)
117924	0.45168	0.03134	-0.66286	0.17142	0.13362	0.05049
117925	1.31777	0.066	-1.13271	0.05557	0.15288	0.03321
117926	1.00038	0.05158	-0.47593	0.06111	0.17872	0.02927
117927	0.82939	0.04735	-0.05318	0.06667	0.17206	0.02765
117942	0.87524	0.01848	0.24736	0.02177	0.21846	0.00895
117943	0.80085	0.0155	-0.45537	0.02958	0.13767	0.01344
117944	0.74943	0.01545	-0.72068	0.03862	0.14722	0.01733

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
117945	0.71005	0.01387	-0.26346	0.02956	0.08948	0.01256
117956	1.29983	0.02615	-1.147	0.02601	0.25352	0.01592
117958	1.10639	0.02134	-0.7544	0.02544	0.24751	0.01343
117959	0.68025	0.01304	-0.14057	0.02757	0.06404	0.01138
117960	0.56002	0.01595	0.1129	0.04914	0.17838	0.01619
117961	0.92405	0.01679	-0.31392	0.02251	0.15341	0.01079
117964	0.9719	0.0235	-0.12778	0.02927	0.40899	0.0107
117965	0.77084	0.01678	-0.40497	0.03535	0.20866	0.0147
117968	1.31257	0.02234	-0.2992	0.01495	0.20437	0.00834
117976	0.82974	0.01561	-0.93897	0.03406	0.1173	0.01712
118000	0.95581	0.02185	-1.28902	0.04571	0.2983	0.02281
118002	1.04949	0.01927	-1.5173	0.03169	0.09806	0.02001
147869	0.98117	0.02156	0.61084	0.01718	0.25332	0.00684
147877	0.51469	0.00731	-0.78334	0.01634	0	0
147878	0.91439	0.01711	-1.06113	0.03266	0.13643	0.01755
147883	0.84799	0.01712	-0.11636	0.02549	0.1924	0.01102
147888	1.46342	0.02612	-0.60064	0.01607	0.23738	0.00973
147891	1.02329	0.01782	-0.42364	0.02044	0.15673	0.0105
147899	1.06589	0.02167	-0.77443	0.02865	0.28331	0.01433
147902	0.85431	0.01237	-1.82277	0.01988	0	0
147965	0.81276	0.01612	-1.21117	0.04184	0.12961	0.02154
147973	0.57032	0.01564	-0.41633	0.06072	0.18964	0.02088
147976	0.40311	0.01389	-0.43547	0.1045	0.11707	0.02948
147980	0.88512	0.01785	0.09034	0.02206	0.20097	0.00946
148096	1.06142	0.0581	-1.45226	0.08838	0.16839	0.04999
148097	0.47163	0.03164	-1.00838	0.18369	0.14701	0.05739
148099	0.98745	0.05049	-0.90554	0.07289	0.15146	0.03764
148109	0.77268	0.04082	0.00557	0.06139	0.10277	0.02558
148398	0.88532	0.04611	-0.55401	0.07057	0.15473	0.0322
148403	0.80817	0.04608	-0.35158	0.07899	0.17897	0.03306
148411	0.64929	0.03999	-1.12413	0.14491	0.18044	0.05777
148444	0.53329	0.0157	-1.59252	0.11599	0.20096	0.04329
203668	0.61366	0.03525	0.64822	0.05449	0.05762	0.01843
203670	0.45877	0.03604	0.67099	0.10576	0.09506	0.03128
203673	0.79314	0.04833	-0.15703	0.08012	0.20661	0.03186
203675	1.05823	0.05333	-0.23987	0.05069	0.17256	0.0244
203678	1.18949	0.0796	1.0136	0.03621	0.24317	0.01361

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
232576	1.03727	0.05087	-1.15504	0.0693	0.13453	0.03679
232579	1.02412	0.0481	-0.36279	0.04913	0.12167	0.02381
232585	0.92644	0.04521	-0.49085	0.05967	0.12562	0.02789
270467	0.90109	0.04607	-0.84871	0.077	0.13916	0.03761
270468	0.84272	0.04298	-0.78839	0.07875	0.12727	0.03679
270473	0.41882	0.03088	-0.09419	0.15554	0.11063	0.0432
270477	0.43404	0.0469	0.594	0.18876	0.21066	0.04896
270479	0.74793	0.03559	-0.52359	0.06691	0.07905	0.02786
270485	0.67788	0.03878	-0.41174	0.09353	0.12723	0.03703
270502	0.91508	0.04937	-0.66059	0.07751	0.17759	0.03703
270506	0.87416	0.05049	-0.00188	0.06555	0.19196	0.02815
270923	0.65644	0.0528	0.06053	0.12456	0.29566	0.0399

Table I-18. 2010–11 NECAP: IRT Parameters for Polytomous Items – Reading Grade 4

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
205951	0.7713	0.0141	-1.44513	0.02227	0	0	1.55369	0.06015	0.63749	0.03997	-0.53811	0.02759	-1.65306	0.02411	0	0
147988	0.70426	0.00413	-1.07537	0.0077	0	0	1.64018	0.01823	0.51689	0.01184	-0.57871	0.00913	-1.57835	0.00868	0	0
118005	0.6991	0.00481	-2.02043	0.00938	0	0	2.17705	0.03774	0.40433	0.0161	-0.8529	0.01055	-1.72848	0.00903	0	0
148116	0.49176	0.00777	-0.97134	0.03038	0	0	2.29545	0.06805	1.0402	0.04745	-0.86279	0.0355	-2.47285	0.03833	0	0
203684	0.80205	0.01159	0.47105	0.01887	0	0	2.31769	0.03647	0.68855	0.0239	-0.8527	0.02598	-2.15354	0.04141	0	0
117946	0.6932	0.00346	1.38353	0.00732	0	0	2.52708	0.01062	0.88927	0.00875	-0.70092	0.01162	-2.71545	0.02797	0	0
270516	0.84653	0.0123	0.53144	0.0182	0	0	2.54083	0.03963	0.64446	0.02284	-0.92576	0.02569	-2.25953	0.04381	0	0
270511	0.59019	0.0098	-1.31136	0.02619	0	0	2.63304	0.09111	0.51447	0.04156	-0.93409	0.03104	-2.21343	0.03083	0	0
117931	0.64929	0.00927	0.36596	0.02345	0	0	3.28554	0.05801	0.73875	0.02873	-1.19293	0.0317	-2.83136	0.0547	0	0
117970	0.64911	0.00314	0.4734	0.00787	0	0	3.48812	0.02079	0.68766	0.00948	-0.97902	0.01032	-3.19675	0.02214	0	0
147914	0.52374	0.00253	1.58282	0.00971	0	0	3.58137	0.01536	0.92259	0.01122	-1.23615	0.01642	-3.26781	0.03405	0	0
147915	0.3402	0	-3.95127	0.01588	0	0	4.94574	0.10345	2.29679	0.04959	-2.9254	0.0175	-4.31712	0.01655	0	0

Table I-19. 2010–11 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)*
118042	0.6202	0.01489	-0.7251	0.05839	0.14151	0.02236
118043	0.7747	0.01497	-0.83421	0.03845	0.11473	0.01778
118044	0.91277	0.01621	-0.11736	0.0206	0.12605	0.00949
118045	1.02527	0.01612	-0.88594	0.0227	0.07928	0.01251
118046	0.8996	0.0169	-0.04652	0.02177	0.15451	0.00969
118047	0.8334	0.01583	-0.00931	0.02281	0.12517	0.00989
118064	0.47918	0.01058	-0.44706	0.05162	0.04542	0.01683
118067	1.05754	0.0193	-0.57356	0.02445	0.21356	0.01228
118073	0.40596	0.00664	-0.59792	0.01851	0	0
118074	1.02127	0.01818	-1.75409	0.0367	0.07759	0.02312
118075	0.76952	0.01523	-0.61429	0.03599	0.12963	0.01593
118077	0.65857	0.01545	-0.02935	0.03713	0.1394	0.01388
118078	0.99256	0.01665	-0.91149	0.02633	0.10875	0.01431
118079	0.5285	0.01335	-0.45608	0.06061	0.08896	0.02093
118083	0.67451	0.01054	-2.09792	0.02896	0	0
118142	0.63823	0.03574	-1.10693	0.13199	0.13559	0.05135
118144	1.06513	0.05414	-1.35706	0.08621	0.15483	0.0482
118145	1.11078	0.05288	-0.71969	0.05973	0.14247	0.03184
118148	1.32844	0.06308	-1.12871	0.05783	0.13372	0.03521
118151	1.26897	0.07496	-1.39297	0.09315	0.26791	0.05488
118152	1.48078	0.07522	-1.47581	0.06038	0.1269	0.03969
118154	0.89546	0.04226	-1.3381	0.08484	0.10786	0.04154
118155	1.31818	0.06275	-0.72482	0.05132	0.16666	0.02977
118159	0.71638	0.03784	-2.26258	0.13263	0.12692	0.05587
118160	0.82599	0.04225	-1.92567	0.11646	0.1403	0.05663
118163	0.94712	0.04427	-0.6397	0.06053	0.1134	0.02841
118165	0.52985	0.03019	-0.60335	0.11497	0.09367	0.03727
118170	0.6906	0.03961	-1.83999	0.16086	0.17352	0.06917
118172	0.43989	0.03478	-0.02685	0.17176	0.1299	0.04744
118175	0.81713	0.04115	-1.78951	0.10864	0.11927	0.05056
118176	0.72398	0.03904	-0.46226	0.08749	0.11462	0.03558
118199	0.56013	0.04236	0.22941	0.11077	0.14337	0.03547
118205	0.63383	0.04491	-0.49653	0.14057	0.21773	0.04755
118207	0.46213	0.00774	-1.75574	0.03113	0	0
118210	0.65521	0.01256	-0.71698	0.03978	0.06201	0.01659
147820	0.83259	0.03862	-1.08357	0.07913	0.09873	0.03662

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
147825	1.02562	0.05058	-0.69452	0.06738	0.15979	0.03472
147986	0.75588	0.0396	-1.24836	0.11191	0.14083	0.05001
147991	0.6838	0.04051	-0.66941	0.11835	0.16138	0.04652
147992	0.90654	0.04444	-1.13152	0.08521	0.13304	0.0427
147997	1.12793	0.05563	-0.3144	0.05224	0.17464	0.02669
147999	0.99247	0.05147	-0.6114	0.07234	0.18466	0.03568
148008	1.06331	0.05436	-0.17094	0.05359	0.17652	0.02591
148344	0.3871	0.00698	-1.49896	0.03137	0	0
148348	0.71745	0.01478	-0.53032	0.03838	0.12056	0.01617
148351	0.50619	0.01571	-0.76711	0.09653	0.18148	0.03044
148362	0.76521	0.01565	-1.09654	0.04768	0.13809	0.0224
148595	1.06661	0.05387	-0.68859	0.06405	0.1884	0.03151
148614	1.11062	0.06551	-0.24133	0.06144	0.31194	0.02604
148616	1.32182	0.06736	-0.96542	0.05984	0.22768	0.03366
148617	0.92948	0.05113	-0.2605	0.06486	0.19103	0.02794
148618	1.79592	0.0867	-1.01102	0.03952	0.176	0.02652
148639	0.74523	0.04487	-0.38204	0.09319	0.17447	0.03615
148667	0.98495	0.05713	-0.27474	0.06752	0.25516	0.02849
148675	1.2091	0.07146	-1.26306	0.09039	0.31888	0.04687
148719	0.90322	0.01409	-2.16331	0.02528	0	0
148726	0.82427	0.01559	-0.44785	0.02934	0.13128	0.01326
148744	0.60159	0.01354	-1.09124	0.06422	0.10093	0.02597
148763	1.08481	0.01577	-1.91654	0.01901	0	0
149086	0.51607	0.03033	-1.89263	0.17465	0.13082	0.05729
149099	0.89981	0.01657	-0.62526	0.02896	0.15196	0.01394
149104	0.77502	0.01528	-0.29528	0.03001	0.12658	0.01287
149112	0.62033	0.00883	-1.56089	0.02213	0	0
149116	0.92461	0.04585	-1.57589	0.09593	0.12691	0.04925
149127	0.85688	0.04537	-0.68875	0.08628	0.16109	0.0397
201746	0.48909	0.03203	-2.5698	0.2224	0.1501	0.06655
201752	0.57067	0.04761	0.6159	0.10716	0.16544	0.03415
201757	0.79312	0.05459	-0.00591	0.09398	0.26633	0.0345
201760	0.69125	0.05489	0.83754	0.07473	0.18595	0.02539
269423	1.09669	0.06406	-0.066	0.05937	0.28127	0.02633

Table I-20. 2010–11 NECAP: IRT Parameters for Polytomous Items – Reading Grade 5

IREF	Parameters and Measures of Standard Error															
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
118156	0.93657	0.01322	0.49293	0.0171	0	0	1.77862	0.02856	0.66914	0.02162	-0.59563	0.02381	-1.85213	0.03992	0	0
118048	1.13785	0.00545	0.683	0.00484	0	0	1.82549	0.00817	0.68379	0.00616	-0.64937	0.00764	-1.85991	0.01468	0	0
201769	1.02822	0.01442	0.07081	0.01568	0	0	1.88203	0.03418	0.59929	0.0213	-0.68778	0.02088	-1.79355	0.03049	0	0
148698	1.20564	0.01755	0.55176	0.01416	0	0	1.98593	0.02548	0.84805	0.01824	-0.83772	0.02378	-1.99625	0.04739	0	0
148680	1.12412	0.01636	0.76373	0.01502	0	0	2.19682	0.02635	0.68827	0.0187	-0.88547	0.02762	-1.99963	0.05494	0	0
148158	0.83895	0.01217	1.18653	0.01943	0	0	2.25124	0.028	0.80796	0.02342	-0.91907	0.0364	-2.14014	0.06916	0	0
118076	0.97398	0.00456	0.63605	0.00563	0	0	2.27491	0.01077	0.75367	0.007	-0.81911	0.00872	-2.20948	0.01771	0	0
118080	0.97	0.00454	0.59255	0.00567	0	0	2.3397	0.01133	0.73937	0.00705	-0.88342	0.0088	-2.19566	0.01708	0	0
118158	0.95691	0.01363	0.78703	0.01738	0	0	2.3935	0.03232	0.48281	0.02102	-0.87983	0.02857	-1.99649	0.0513	0	0
148137	0.82667	0.01215	1.46535	0.01988	0	0	2.47614	0.02787	0.87769	0.02411	-0.9094	0.04133	-2.44442	0.09861	0	0
118178	0.78973	0.011	0.84015	0.02046	0	0	2.52683	0.03578	0.70061	0.02443	-0.85379	0.03177	-2.37365	0.06378	0	0
148369	0.88058	0.00407	0.44366	0.00621	0	0	2.53207	0.0136	0.82327	0.00779	-0.92237	0.00898	-2.43298	0.01784	0	0
118049	0.94887	0.00443	0.32905	0.00591	0	0	2.55511	0.0144	0.85821	0.00759	-1.04917	0.00863	-2.36416	0.0161	0	0
118167	0.85937	0.01218	0.45224	0.01965	0	0	3.11327	0.05011	0.77871	0.0235	-1.25372	0.03178	-2.63826	0.06402	0	0
148767	0.70314	0.00326	0.36607	0.00788	0	0	3.56521	0.02328	0.85378	0.00943	-1.3968	0.01154	-3.02218	0.02258	0	0

Table I-21. 2010–11 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)*
118266	1.25035	0.02125	-0.95601	0.0193	0.14662	0.01208
118267	0.82646	0.01594	-0.84356	0.03248	0.12094	0.01607
118268	0.46269	0.00716	-1.08588	0.01913	0	0
118269	1.0007	0.01586	-1.36934	0.02425	0.04891	0.01456
118270	0.99387	0.01691	-0.73356	0.0218	0.10755	0.0118
118271	1.34232	0.02451	-1.24056	0.02202	0.17134	0.01501
118272	0.77042	0.01713	0.38406	0.02111	0.14358	0.00842
118273	0.87555	0.0159	-0.86823	0.02856	0.10302	0.01482
118282	0.83696	0.05151	-0.91014	0.1046	0.24026	0.04595
118284	0.58131	0.04249	-0.54368	0.14797	0.19942	0.05012
118285	0.6962	0.03962	-1.41223	0.12694	0.15807	0.05534
118287	0.6129	0.02935	-1.39072	0.08888	0.0733	0.03248

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
118327	0.92847	0.05381	-0.52179	0.07268	0.22438	0.03242
118328	0.90016	0.04457	-1.18147	0.07733	0.12383	0.03908
118330	1.0987	0.05332	-1.73638	0.06968	0.10597	0.04076
118333	0.55575	0.03343	-0.65681	0.11649	0.11119	0.04077
118335	1.34037	0.06302	-1.47085	0.05097	0.0975	0.03193
118338	0.52919	0.02999	-1.30984	0.13572	0.11331	0.0475
118339	0.80766	0.04942	-1.54834	0.13668	0.22885	0.06463
118341	1.06259	0.0523	-1.05748	0.06361	0.14041	0.03468
118345	0.8411	0.01487	-0.96544	0.02947	0.07585	0.01532
118347	1.10085	0.01754	-2.12033	0.01909	0	0
118349	0.85899	0.01645	-0.22848	0.02265	0.13538	0.01033
118352	0.7192	0.01388	-0.40846	0.02868	0.07113	0.01257
118384	0.77568	0.04304	-0.86955	0.09531	0.14678	0.04355
118387	0.64467	0.0418	-0.71402	0.124	0.17476	0.04675
118393	0.46217	0.04753	0.39795	0.17236	0.21548	0.0462
126896	0.82258	0.01687	-0.34984	0.02774	0.16693	0.01227
126897	0.56813	0.0111	-1.15696	0.0479	0.05164	0.01958
126898	0.94073	0.01829	0.1945	0.01695	0.16115	0.00747
126899	0.97281	0.01772	-0.94592	0.02726	0.13704	0.01498
126900	0.68815	0.0174	-0.47196	0.04557	0.22536	0.01723
126901	0.90115	0.01856	-0.07628	0.02232	0.20573	0.00968
126902	0.84933	0.0162	-0.91309	0.0323	0.12197	0.01643
126903	0.85973	0.01587	-0.53327	0.02521	0.1138	0.01218
129078	0.762	0.04383	-1.20839	0.1142	0.17423	0.05137
129081	0.61309	0.03402	-1.03809	0.11353	0.11614	0.04384
129082	0.7261	0.03642	-1.35348	0.09614	0.10987	0.04278
129083	0.72846	0.03633	-1.99108	0.10148	0.10001	0.04426
148388	0.82071	0.03895	-1.90055	0.08085	0.07821	0.03484
148393	0.71525	0.03462	-0.58463	0.06729	0.0766	0.02751
148399	0.71201	0.03812	-1.83849	0.12058	0.13318	0.05452
148401	0.87787	0.05113	-0.68224	0.08362	0.21094	0.03768
148410	0.98882	0.05477	-1.21701	0.08749	0.2023	0.04614
148416	0.26774	0.02775	-0.265	0.34158	0.15176	0.06293
148425	1.02644	0.05355	-1.51277	0.08292	0.15351	0.04728
148428	0.61243	0.03326	-1.36566	0.12077	0.11857	0.04763
148549	0.93811	0.05554	-0.92138	0.09331	0.23723	0.04573
148563	0.78368	0.04629	-1.97323	0.14214	0.18443	0.07165

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
148566	0.56098	0.03683	-0.27354	0.11528	0.12118	0.04026
148573	0.94794	0.05396	-0.79327	0.08253	0.21165	0.04081
148673	0.92611	0.04879	-1.64339	0.09545	0.1402	0.05278
148684	1.29994	0.07164	-1.27663	0.0684	0.20477	0.04408
148692	0.68342	0.03696	-1.5605	0.1205	0.13066	0.05329
148695	0.79412	0.04338	-1.2044	0.10363	0.15035	0.04993
148699	0.65403	0.03244	-1.03769	0.08858	0.08626	0.03584
148703	1.5641	0.07838	-1.44424	0.04662	0.09913	0.03197
148709	0.52685	0.03193	-1.91524	0.17281	0.14392	0.06222
148711	1.0083	0.05381	-1.90171	0.09074	0.12792	0.05219
148825	0.67459	0.0161	-1.54991	0.07112	0.15228	0.0337
148865	0.99747	0.01675	-0.99339	0.02383	0.08763	0.01364
148874	0.92191	0.01734	-1.61237	0.03874	0.09527	0.02395
148876	0.81743	0.01524	-0.83366	0.03088	0.09636	0.01535
148950	0.8128	0.01659	-0.85466	0.03604	0.15298	0.01728
148951	0.67078	0.0492	-0.81823	0.15123	0.27652	0.05363
148956	0.65494	0.01746	-1.04207	0.06751	0.23302	0.02645
148966	0.73307	0.01524	-0.8757	0.04043	0.1174	0.01867
148978	0.71539	0.01147	-2.20307	0.02631	0	0
148984	0.82361	0.04508	-1.70339	0.11418	0.15321	0.05881
256674	0.33675	0.01901	-1.31307	0.0824	0	0

Table I-22. 2010–11 NECAP: IRT Parameters for Polytomous Items – Reading Grade 6

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
126905	1.03248	0.0051	0.55143	0.00504	0	0	1.97868	0.00881	0.64092	0.00642	-0.73453	0.00776	-1.88506	0.0133	0	0
118343	0.99596	0.01452	0.22242	0.0156	0	0	2.04306	0.03038	0.76348	0.02065	-0.79845	0.02236	-2.00808	0.03663	0	0
118289	1.02958	0.01534	0.69314	0.01528	0	0	2.14	0.02627	0.63754	0.01926	-0.73516	0.02487	-2.04237	0.04808	0	0
118275	1.07762	0.0053	0.298	0.00491	0	0	2.1958	0.01066	0.76855	0.00653	-0.75172	0.00707	-2.21263	0.01389	0	0
118353	0.92303	0.0045	0.57725	0.00563	0	0	2.27327	0.01028	0.85923	0.00709	-0.76969	0.00847	-2.36282	0.01797	0	0
148714	0.94157	0.0137	0.30508	0.0169	0	0	2.42816	0.03723	0.96385	0.02215	-0.9321	0.02429	-2.45991	0.04817	0	0
129084	0.94496	0.01372	0.29837	0.01659	0	0	2.44273	0.03571	0.82627	0.02144	-0.85292	0.02401	-2.41609	0.04833	0	0
148455	0.99387	0.0146	0.31702	0.01593	0	0	2.47595	0.0363	0.68369	0.02036	-0.89048	0.02362	-2.26916	0.04396	0	0
118344	0.98328	0.01446	0.42085	0.0162	0	0	2.51283	0.03449	0.87662	0.02068	-0.90501	0.02479	-2.48442	0.05453	0	0
126904	0.95227	0.00465	0.37245	0.00554	0	0	2.52619	0.01252	0.72678	0.00702	-0.95678	0.00827	-2.29618	0.01532	0	0
118274	1.01659	0.00503	0.46346	0.00526	0	0	2.52844	0.0118	0.76963	0.00666	-0.91484	0.00811	-2.38324	0.01716	0	0

148440	1.04011	0.01562	0.56309	0.01551	0	0	2.56536	0.03331	0.59877	0.01925	-0.90068	0.02511	-2.26346	0.05101	0	0
148718	0.99489	0.01462	0.06893	0.01613	0	0	2.56842	0.04521	0.86891	0.02207	-0.9476	0.02211	-2.48973	0.04274	0	0
148579	0.7759	0.01107	0.10873	0.02002	0	0	2.90068	0.05286	0.89879	0.026	-0.98521	0.02655	-2.81427	0.05214	0	0
148891	0.81699	0.00394	0.14372	0.00648	0	0	3.04686	0.01858	0.91413	0.00837	-1.11298	0.00892	-2.84801	0.01823	0	0

Table I-23. 2010–11 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)*
118477	0.46356	0.01436	-0.24827	0.06145	0.1046	0.02019
118526	0.54446	0.02975	-1.49684	0.11198	0.09964	0.04263
118528	0.61181	0.04112	-0.92539	0.12886	0.17882	0.05019
118535	0.5599	0.01448	-2.11944	0.09357	0.13597	0.04367
118536	0.43201	0.00711	-1.39989	0.02047	0	0
118538	0.6257	0.01578	-1.44178	0.06542	0.16073	0.02952
118540	0.85918	0.01598	-1.22375	0.02925	0.09446	0.01659
118544	0.60747	0.01015	-1.21184	0.02924	0.03065	0.01262
118546	0.52908	0.00768	-1.26009	0.01596	0	0
118547	0.66297	0.01049	-2.14472	0.02223	0	0
118548	0.39375	0.00874	-0.60934	0.04673	0.03286	0.01418
118549	0.47328	0.01217	-0.17456	0.04484	0.05378	0.01546
118550	1.08701	0.01897	-1.01727	0.01956	0.12333	0.01199
118551	0.82003	0.01638	-0.21185	0.02126	0.14197	0.00964
118554	0.56997	0.01467	-0.01397	0.03595	0.10577	0.01338
118563	0.57233	0.01072	-1.22122	0.03842	0.0429	0.01649
118572	0.77735	0.03998	-1.60104	0.08729	0.11615	0.04431
118573	0.55775	0.02422	-1.60972	0.05273	0	0
128125	0.31207	0.00612	-0.55401	0.01863	0	0
128216	0.7104	0.01628	-0.2272	0.02923	0.16673	0.01202
129210	0.42393	0.03099	-0.60357	0.15702	0.12078	0.04686
129211	0.70354	0.03876	-2.07349	0.11002	0.12326	0.0519
129212	0.7484	0.04163	-1.47865	0.10045	0.1436	0.04909
129213	0.7558	0.04792	-0.56759	0.08686	0.20136	0.03689
129214	0.68862	0.03697	-1.3814	0.09592	0.11707	0.04372
129215	0.97114	0.05607	-0.3889	0.05661	0.2149	0.02687
129216	0.35953	0.03482	0.49876	0.16319	0.11459	0.04184
129217	0.34027	0.02008	-1.92663	0.09604	0	0
129221	0.46793	0.02384	0.57238	0.07336	0.39821	0.01624
129222	0.41506	0.01821	0.10997	0.08757	0.21793	0.02305

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
129223	0.3037	0.00905	-0.73741	0.09032	0.05117	0.0223
129224	0.64393	0.01018	-2.12252	0.02237	0	0
129225	0.50812	0.01402	-0.14304	0.04631	0.09042	0.01634
129226	0.56598	0.01139	-0.80514	0.03726	0.04647	0.01529
129227	0.72314	0.01462	-0.75457	0.03126	0.09903	0.01472
147154	1.16015	0.06748	-1.94851	0.08005	0.15314	0.05619
147157	0.9521	0.04999	-1.62887	0.0795	0.12832	0.04747
147161	0.99124	0.05284	-1.45577	0.07639	0.14169	0.04601
147162	0.52595	0.02772	-1.16463	0.09406	0.07539	0.03353
147164	0.82964	0.04655	-1.62543	0.10209	0.1555	0.05591
147171	0.75029	0.03587	-1.67598	0.07405	0.07616	0.03385
147172	1.04126	0.05244	-1.14819	0.06022	0.12226	0.03551
147174	0.64456	0.04007	-0.28654	0.08392	0.11661	0.0335
147203	0.67849	0.01166	-2.45511	0.02667	0	0
147210	0.35571	0.00679	-1.69688	0.02825	0	0
147213	0.57876	0.01393	-1.37186	0.0625	0.10772	0.02742
147215	0.39847	0.00711	-1.70899	0.02577	0	0
147526	0.82879	0.04231	-1.98411	0.08136	0.09556	0.0407
147539	0.8305	0.04807	-2.07395	0.11028	0.15887	0.06023
147546	0.24136	0.01761	-1.51779	0.10731	0	0
147549	0.8325	0.04319	-1.82479	0.08509	0.11267	0.04473
147641	0.60734	0.03168	-1.48152	0.09794	0.09314	0.04051
147643	0.45904	0.03239	-0.87913	0.1629	0.1344	0.0528
147646	0.65138	0.0361	-0.59539	0.0804	0.09482	0.03335
147647	0.57315	0.03166	-1.04916	0.10144	0.0965	0.03975
147663	0.33508	0.00626	-0.75312	0.01887	0	0
147666	0.82263	0.0435	-1.69532	0.09053	0.13201	0.04851
147668	0.59439	0.03769	-0.96029	0.12517	0.14371	0.04987
147677	0.61843	0.03706	-0.85499	0.10649	0.12424	0.04357
147681	0.69217	0.04855	0.59197	0.05783	0.13667	0.02136
147686	0.69353	0.04609	0.13226	0.06753	0.14631	0.0261
147690	0.47734	0.02996	-1.37595	0.14959	0.12428	0.05124
201466	0.68733	0.03828	-1.1226	0.09544	0.12664	0.04258
201468	0.9996	0.05125	-1.32746	0.06636	0.13605	0.03855
201470	0.97303	0.05118	-2.12682	0.0758	0.10306	0.04331
201472	0.67886	0.03416	-1.67683	0.08966	0.09538	0.04035
201476	1.34673	0.06863	-1.38222	0.04838	0.13578	0.03317

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
201479	0.86395	0.04246	-1.07131	0.06391	0.10359	0.03273
201482	0.43485	0.02001	-0.88468	0.04558	0	0
201487	0.87998	0.04615	-1.97918	0.08541	0.11913	0.04787

Table I-24. 2010–11 NECAP: IRT Parameters for Polytomous Items – Reading Grade 7

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
129218	1.09583	0.01759	0.51973	0.01321	0	0	1.52153	0.01976	0.64435	0.01719	-0.53114	0.01994	-1.63473	0.03138	0	0
147179	1.20476	0.01887	0.12555	0.01202	0	0	1.65003	0.02295	0.62202	0.01684	-0.55213	0.01701	-1.71992	0.02554	0	0
147181	1.14637	0.01803	0.50802	0.01269	0	0	1.78695	0.02147	0.59172	0.01672	-0.62399	0.01956	-1.75469	0.03246	0	0
129219	1.11166	0.01719	-0.04553	0.0129	0	0	1.90779	0.02798	0.48771	0.01769	-0.58839	0.01774	-1.80711	0.02584	0	0
118558	1.25881	0.00659	0.19159	0.0039	0	0	1.96936	0.00828	0.56959	0.00536	-0.68819	0.00583	-1.85077	0.00937	0	0
118542	1.08368	0.00555	-0.06005	0.00445	0	0	2.038	0.01017	0.73365	0.00631	-0.71585	0.00613	-2.0558	0.0096	0	0
129229	1.11427	0.0058	0.51	0.00437	0	0	2.07627	0.00794	0.66138	0.00571	-0.72185	0.00692	-2.0158	0.01278	0	0
201492	1.04397	0.01593	0.04302	0.01382	0	0	2.18215	0.0303	0.74713	0.01888	-0.67842	0.01921	-2.25085	0.03488	0	0
129228	1.00536	0.00512	0.22111	0.00477	0	0	2.18722	0.0098	0.63645	0.00627	-0.8063	0.00693	-2.01737	0.01106	0	0
201490	0.9791	0.01491	-0.21839	0.01465	0	0	2.2014	0.03535	0.74904	0.02073	-0.88255	0.01981	-2.06789	0.02847	0	0
147649	1.07144	0.01653	-0.05499	0.01352	0	0	2.21484	0.03419	0.56906	0.0187	-0.7423	0.01845	-2.04159	0.02858	0	0
147224	0.92561	0.00475	0.66369	0.00516	0	0	2.25178	0.00873	0.7516	0.00651	-0.74206	0.00811	-2.26132	0.01632	0	0
118532	0.8763	0.01332	0.55568	0.0162	0	0	2.32511	0.02793	0.70937	0.0203	-0.74634	0.02461	-2.28815	0.04807	0	0
118557	1.078	0.00562	0.0364	0.00465	0	0	2.5351	0.01268	0.84981	0.00641	-0.98716	0.00668	-2.39774	0.01216	0	0
147555	0.82122	0.01233	-0.3365	0.01746	0	0	2.70041	0.05061	0.87906	0.02486	-0.94135	0.02228	-2.63811	0.03643	0	0

Table I-25. 2010–11 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)
118599	0.28254	0.02211	0.16958	0.28774	0.29969	0.04726
118600	0.54772	0.01527	-1.16005	0.0903	0.17348	0.03435
118601	0.5363	0.00843	-1.70746	0.02514	0	0
118602	0.68525	0.01458	-0.71225	0.04328	0.119	0.01928
118603	0.48414	0.00923	-2.47564	0.04164	0	0
118604	0.41574	0.00775	-2.05627	0.03682	0	0
118605	0.54352	0.01479	0.83881	0.02985	0.07683	0.01051

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
118606	0.95541	0.01832	-1.48221	0.0381	0.12004	0.02416
118662	0.84005	0.01824	-0.80911	0.03981	0.23654	0.01866
118663	0.8765	0.01532	-0.83521	0.02792	0.09076	0.01508
118664	0.47673	0.01051	-1.39724	0.07237	0.06721	0.02704
118665	0.97503	0.01718	-0.49907	0.02232	0.14361	0.01178
118666	0.51913	0.0075	-1.04573	0.01772	0	0
118667	1.00708	0.01693	-1.27675	0.02696	0.07299	0.01699
118668	0.71613	0.01268	-0.83709	0.03302	0.05649	0.0158
118669	0.77041	0.01151	-1.86438	0.0215	0	0
118704	0.64945	0.04094	-0.78228	0.13325	0.18501	0.0528
118706	0.64982	0.03708	-1.00573	0.12309	0.14794	0.0512
118707	0.50241	0.03618	0.14066	0.12211	0.12035	0.03896
118708	0.55812	0.0414	0.32365	0.10681	0.1419	0.03577
118742	0.41966	0.04251	-0.27262	0.27651	0.25338	0.06843
118745	0.42051	0.01153	-0.07813	0.06306	0.057	0.01924
127358	0.74627	0.01643	-0.8845	0.04632	0.18194	0.02135
147223	0.35503	0.00678	-1.5537	0.03228	0	0
147227	0.62261	0.01643	-0.02388	0.04424	0.1982	0.01574
147232	0.58571	0.01117	-0.71078	0.0398	0.04793	0.01631
147239	0.66988	0.01149	-2.29179	0.03101	0	0
147318	1.04376	0.05486	-0.3697	0.05744	0.1976	0.02864
147319	0.73502	0.04074	0.13103	0.06358	0.10713	0.0256
147322	0.46485	0.02724	-0.54654	0.1234	0.09122	0.03858
147324	0.80738	0.04374	-0.86402	0.09293	0.15752	0.04425
147332	0.89098	0.04542	-0.61371	0.07001	0.14002	0.03441
147333	0.71563	0.03516	-1.24894	0.09425	0.10706	0.0431
147334	0.7772	0.03829	-0.99577	0.08456	0.11244	0.04028
147340	0.8559	0.05267	0.67807	0.04891	0.16358	0.01877
147447	0.50632	0.01408	0.30887	0.04557	0.08674	0.01538
147459	0.78371	0.01603	-0.12045	0.02743	0.15604	0.01191
147469	0.48759	0.01413	-0.38596	0.07203	0.11701	0.02402
147476	0.60961	0.01142	-0.73659	0.03825	0.04906	0.01618
147492	1.094	0.05217	-0.72126	0.0546	0.13447	0.03066
147496	0.73257	0.03614	-1.81798	0.09509	0.095	0.04236
147498	0.53128	0.03597	0.2876	0.09944	0.10307	0.03283
147505	0.72971	0.04223	-0.97657	0.11661	0.17518	0.05206
147516	0.40626	0.01984	-0.99397	0.06267	0	0

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
147521	1.00453	0.04793	-1.57451	0.07088	0.09901	0.03945
147523	0.55338	0.04772	0.71972	0.10467	0.17589	0.03335
147531	0.58907	0.03639	-0.78475	0.13628	0.15254	0.05131
147537	0.78805	0.04169	-1.06392	0.09803	0.1468	0.04723
147540	0.45124	0.02893	-1.67747	0.20024	0.14949	0.06357
147545	0.43954	0.02791	-1.32943	0.18315	0.13191	0.05624
147547	1.13832	0.06204	-1.33353	0.08098	0.19745	0.05047
147600	0.56525	0.03323	-0.93456	0.13457	0.133	0.05014
147605	0.66892	0.014	-0.26611	0.03435	0.09588	0.01442
147611	0.63993	0.03898	-0.50713	0.11386	0.14989	0.04479
147616	0.64155	0.04529	-0.49569	0.14151	0.23173	0.05185
147619	0.65124	0.01156	-1.19891	0.03988	0.05055	0.0186
147621	0.58177	0.03549	-2.14	0.1769	0.168	0.07042

Table I-26. 2010–11 NECAP: IRT Parameters for Polytomous Items – Reading Grade 8

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
147557	1.11	0.01633	-0.06817	0.01346	0	0	1.62927	0.02925	0.60793	0.01977	-0.54878	0.0177	-1.68842	0.02338	0	0
147346	1.21103	0.0178	0.16377	0.01249	0	0	1.73155	0.02581	0.62346	0.0177	-0.54586	0.01743	-1.80914	0.02657	0	0
147562	1.10016	0.01601	0.23047	0.01361	0	0	1.7569	0.02711	0.66046	0.01886	-0.61204	0.01867	-1.80534	0.02817	0	0
147351	1.02933	0.0151	0.64195	0.01451	0	0	1.83078	0.0239	0.64971	0.01879	-0.60483	0.02167	-1.87568	0.03704	0	0
147509	1.20086	0.01786	0.62836	0.01283	0	0	1.9388	0.0232	0.69453	0.01702	-0.72722	0.02055	-1.90611	0.03649	0	0
147481	0.91657	0.00437	0.2407	0.00544	0	0	2.09254	0.01133	0.84347	0.0074	-0.76492	0.00729	-2.17108	0.01186	0	0
118711	1.0147	0.0147	-0.16873	0.01499	0	0	2.27257	0.04475	0.77783	0.02266	-0.82424	0.01931	-2.22617	0.02934	0	0
118670	1.14969	0.00569	0.42991	0.00455	0	0	2.38382	0.01135	0.68141	0.006	-0.86316	0.00678	-2.20207	0.01294	0	0
118607	0.85956	0.0041	-0.12626	0.00581	0	0	2.50718	0.01692	0.75918	0.00827	-0.95459	0.00741	-2.31176	0.01092	0	0
118671	1.17252	0.00593	0.54609	0.00455	0	0	2.58596	0.0119	0.67793	0.00585	-0.91396	0.0071	-2.34993	0.01546	0	0
118608	0.81455	0.00385	0.35006	0.0062	0	0	2.77714	0.01524	0.89035	0.00798	-1.03972	0.0086	-2.62777	0.016	0	0
147242	0.86126	0.00417	-0.47	0.00593	0	0	2.84832	0.02512	0.92389	0.00966	-1.00273	0.00722	-2.76948	0.01148	0	0

Table I-27. 2010–11 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)
118758	0.31621	0.0076	-1.36688	0.0371	0	0
118759	0.59098	0.01572	0.55863	0.02797	0.05734	0.01054

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
118760	0.94891	0.02511	-2.08302	0.05907	0.14713	0.03941
118761	0.38199	0.01125	-0.52428	0.07408	0.05254	0.02159
118762	0.46827	0.01348	-1.1741	0.08599	0.08699	0.03092
118763	0.47983	0.01969	0.68426	0.05598	0.15543	0.0174
118764	0.58541	0.01176	-2.24994	0.03569	0	0
118765	0.99075	0.02116	-1.16954	0.03189	0.12098	0.01889
118874	0.70761	0.04429	-1.52496	0.13209	0.1607	0.05989
118876	0.54068	0.03156	-0.87597	0.1043	0.0839	0.036
118878	0.58082	0.03095	-0.59319	0.07601	0.06008	0.02603
118880	0.96719	0.05632	-0.26238	0.0607	0.1585	0.02945
147381	0.68843	0.04789	0.04009	0.08794	0.14892	0.03464
147382	0.69029	0.03829	-0.37391	0.07401	0.08447	0.03011
147383	0.69364	0.04471	-0.42116	0.09841	0.14456	0.0406
147385	0.92505	0.05001	-0.35972	0.05943	0.11556	0.02894
147423	0.9716	0.05622	-1.16656	0.08367	0.15392	0.04539
147433	0.83073	0.04463	-0.96261	0.07762	0.10484	0.03694
147435	0.61942	0.03603	-0.49162	0.08784	0.08753	0.03344
147439	0.96123	0.05238	-1.22871	0.07621	0.11833	0.04063
147450	0.59951	0.06535	1.1892	0.09005	0.22145	0.02887
147456	0.64017	0.03846	-0.4365	0.09	0.09933	0.03522
147463	0.61243	0.03536	-1.1175	0.10755	0.10133	0.04182
147473	0.63719	0.03678	-0.56251	0.08824	0.09107	0.03443
147550	0.51605	0.01634	-0.6458	0.07596	0.12824	0.02718
147551	0.85069	0.01481	-1.84904	0.02206	0	0
147552	0.43054	0.01313	0.01679	0.05813	0.05321	0.01826
147556	0.91912	0.0193	-0.84855	0.02988	0.11806	0.01614
147695	0.33448	0.00849	-2.33702	0.05524	0	0
147698	0.68485	0.01829	-0.38213	0.04378	0.16863	0.01795
147703	0.57176	0.01963	0.41836	0.04545	0.18239	0.01539
147707	0.33485	0.01785	0.96522	0.0929	0.08885	0.02403
147713	0.92998	0.021	-0.0251	0.02286	0.18286	0.01062
147716	0.30594	0.00751	-1.33567	0.03766	0	0
147721	0.86946	0.01966	0.23504	0.02084	0.14011	0.0093
147723	0.41093	0.01932	0.87367	0.06795	0.12203	0.01982
147735	0.81081	0.05041	-1.89853	0.11876	0.14565	0.05814
147736	0.98789	0.05885	-1.75355	0.08936	0.12487	0.04828
147747	0.80507	0.05249	-0.42524	0.09126	0.19274	0.04026

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
147749	0.67642	0.03949	0.0538	0.06626	0.07834	0.02595
147753	0.56193	0.03358	-1.01975	0.11535	0.10158	0.04216
147764	1.14373	0.06581	0.03913	0.04587	0.17975	0.02315
147766	0.94722	0.05487	-0.66064	0.07426	0.16557	0.03782
147767	0.82467	0.04561	-0.72025	0.07725	0.1183	0.03662
147850	0.82025	0.01659	-2.33343	0.03067	0	0
147855	0.57117	0.01749	0.1824	0.04443	0.12647	0.01601
147860	0.47708	0.00865	-1.15459	0.02329	0	0
147874	0.67839	0.01421	-0.81055	0.0362	0.05393	0.01634
147897	0.50205	0.02219	0.29481	0.07517	0.28287	0.02076
147900	0.47523	0.04351	0.50952	0.12993	0.13172	0.0401
147908	0.48556	0.0397	0.01745	0.13918	0.1303	0.04394
147912	0.2709	0.01107	-0.47195	0.14039	0.06978	0.03022
147926	0.48688	0.04846	-0.55189	0.24258	0.27431	0.06916
147927	0.58794	0.01959	-0.21334	0.0608	0.22909	0.02075
147929	0.62453	0.05029	-0.35205	0.1426	0.23116	0.05088
147934	0.42118	0.00824	-1.18973	0.02639	0	0

Table I-28. 2010–11 NECAP: IRT Parameters for Polytomous Items – Reading Grade 11

IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
147488	1.48734	0.02652	0.15344	0.01154	0	0	1.64808	0.02724	0.69501	0.01843	-0.54062	0.01598	-1.80246	0.02467	0	0
147774	1.2335	0.02127	0.106	0.01338	0	0	1.66877	0.03017	0.68402	0.02055	-0.5371	0.01787	-1.81569	0.02595	0	0
147484	1.43571	0.02547	0.18499	0.01195	0	0	1.78126	0.02909	0.69857	0.0187	-0.683	0.01672	-1.79683	0.02533	0	0
147388	1.30578	0.0228	0.62049	0.013	0	0	1.86864	0.02562	0.69074	0.01788	-0.76316	0.01997	-1.79621	0.03391	0	0
147775	1.26529	0.02203	0.42316	0.01347	0	0	1.95011	0.02948	0.76731	0.01925	-0.83461	0.0196	-1.88282	0.03181	0	0
147732	1.16049	0.00705	0.64279	0.00507	0	0	1.97862	0.00983	0.67882	0.00678	-0.74963	0.00767	-1.9078	0.01335	0	0
147726	1.16109	0.00703	0.19352	0.00504	0	0	2.01663	0.01231	0.67005	0.00734	-0.73768	0.00689	-1.949	0.01055	0	0
118883	1.04401	0.0177	0.62805	0.01572	0	0	2.02886	0.02996	0.73026	0.02077	-0.70048	0.02271	-2.05865	0.04195	0	0
147564	1.02948	0.00615	0.42764	0.00567	0	0	2.15782	0.01218	0.8255	0.00778	-0.76942	0.00788	-2.2139	0.01436	0	0
118767	1.16255	0.00722	0.24513	0.00531	0	0	2.19582	0.01313	0.94618	0.00782	-1.04045	0.00747	-2.10154	0.01179	0	0
147885	1.131	0.0069	0.3603	0.00533	0	0	2.2155	0.01259	0.88585	0.00757	-0.89475	0.00754	-2.2066	0.01355	0	0
118766	1.05566	0.00641	-0.13135	0.00559	0	0	2.24022	0.01675	0.85701	0.00897	-0.87234	0.0072	-2.22487	0.01068	0	0

Table I-29. 2010–11 NECAP: IRT Parameters for Dichotomous Items – Writing Grade 5

IREF	*Parameters and Measures of Standard Error*					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
148023	0.58375	0.01499	-1.08744	0.0762	0.13484	0.03033
148062	0.50882	0.00774	-1.32602	0.02093	0	0
148081	0.58542	0.01313	-1.42463	0.07147	0.09221	0.03091
148126	0.13461	0.00567	0	0.04207	0	0
148167	0.52471	0.01293	-2.57237	0.11775	0.12285	0.05202
148246	0.8187	0.01517	-1.51673	0.04302	0.07926	0.02454
148263	0.69499	0.01208	-1.77477	0.04266	0.04972	0.02124
148278	0.6239	0.01136	-1.83946	0.04994	0.05307	0.0231
148316	0.80458	0.01573	-1.42384	0.04657	0.09911	0.02575
148327	0.69961	0.01261	-1.20886	0.04067	0.05681	0.01966

Table I-30. 2010–11 NECAP: IRT Parameters for Polytomous Items – Writing Grade 5

IREF	*Parameters and Measures of Standard Error*															
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
150134	1.05811	0.00527	0.62836	0.00535	0	0	2.70706	0.01216	0.82888	0.00657	-0.98779	0.0091	-2.54815	0.02268	0	0
150165	0.9698	0.00409	1.13011	0.00518	0	0	3.89183	0.01815	3.18996	0.01209	2.6924	0.00959	1.7261	0.00719	1.14982	0.0068
150198	1.04808	0.00509	-0.09222	0.00522	0	0	2.52766	0.01682	0.64208	0.00714	-0.92029	0.00707	-2.24945	0.01185	0	0
150237	0.92306	0.00447	0.12095	0.0061	0	0	2.37516	0.01363	1.28607	0.00865	-0.92847	0.00815	-2.73276	0.01782	0	0

IREF (continued)	*Parameters and Measures of Standard Error*													
	D5	*SE (D5)*	*D6*	*SE (D6)*	*D7*	*SE (D7)*	*D8*	*SE (D8)*	*D9*	*SE (D9)*	*D10*	*SE (D10)*	*D11*	*SE (D11)*
150165	0.17383	0.00757	-0.52412	0.00948	-1.3364	0.0141	-2.2138	0.02459	-3.82449	0.08334	-4.92515	0.20405	0	0

Table I-31. 2010–11 NECAP: IRT Parameters for Dichotomous Items – Writing Grade 8

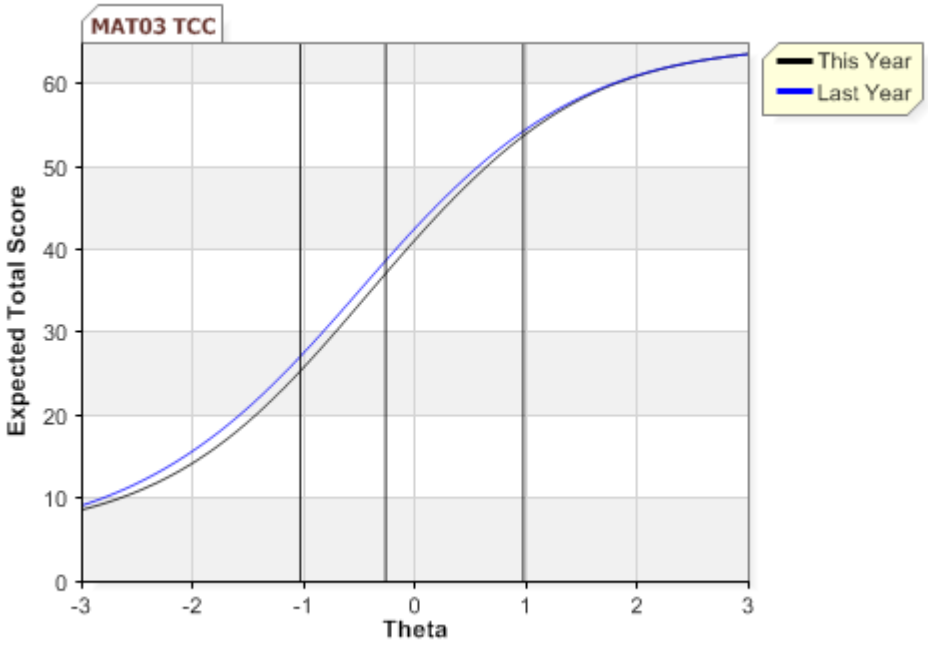
IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
148434	0.63196	0.0161	-0.17838	0.04277	0.18512	0.01509
148557	0.73171	0.01722	-0.15512	0.03404	0.22332	0.01267
148588	0.50512	0.01625	-1.72695	0.1418	0.23409	0.0477
148610	0.7958	0.01854	-1.03379	0.04971	0.2926	0.02046
148662	0.68662	0.01806	-0.5809	0.05242	0.29558	0.01803
148700	0.69327	0.01852	-0.78184	0.05879	0.32242	0.0202
148715	0.89802	0.02142	-1.49589	0.05627	0.33981	0.02584
148762	0.57021	0.01664	0.00789	0.04962	0.19434	0.01596
148773	0.50918	0.01644	-0.73628	0.09166	0.2239	0.02774
148812	0.51286	0.01798	0.61343	0.04656	0.17988	0.01411

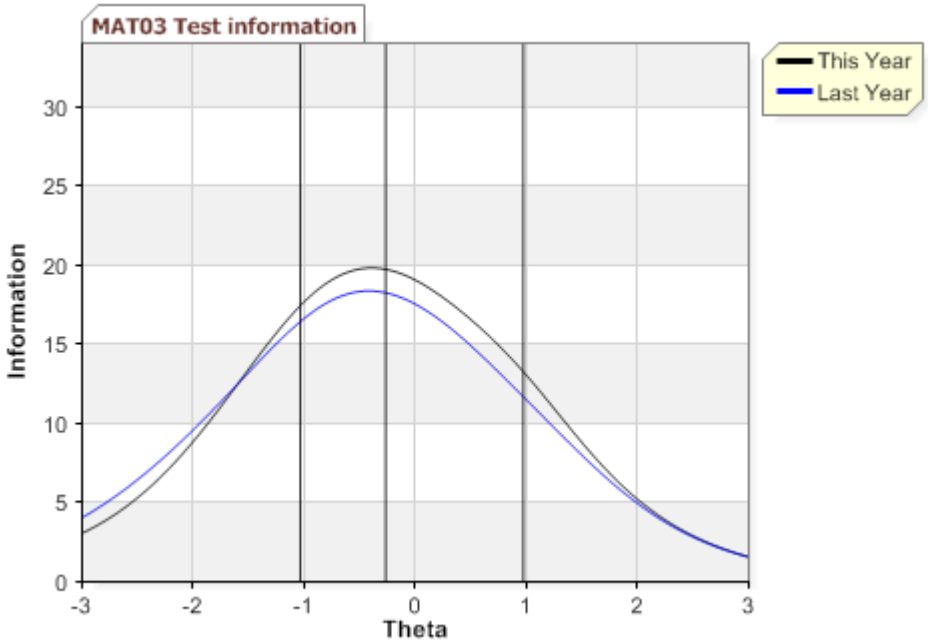
Table I-32. 2010–11 NECAP: IRT Parameters for Polytomous Items – Writing Grade 8

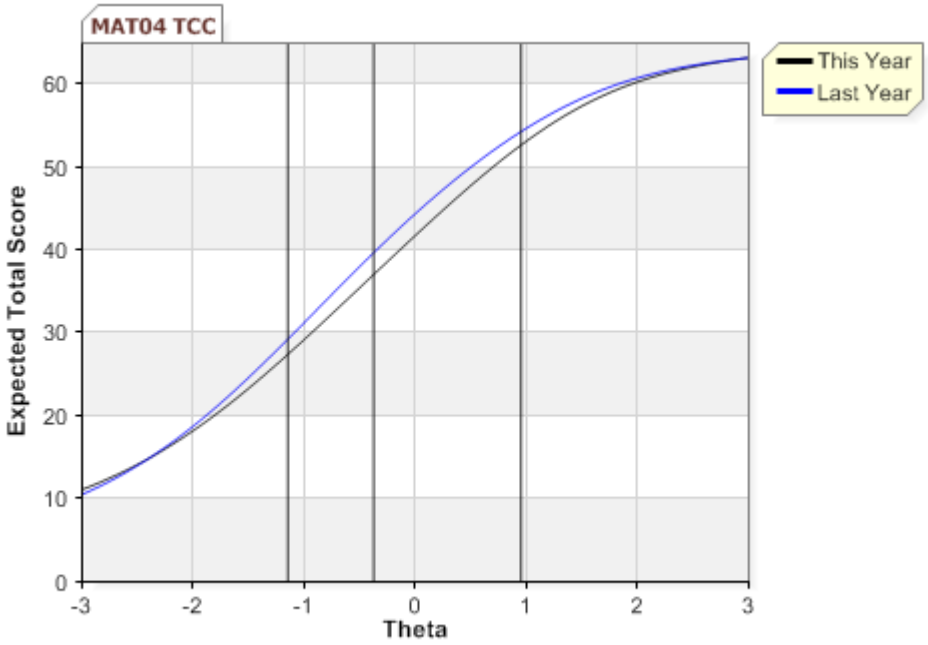
IREF	Parameters and Measures of Standard Error															
	a	SE (a)	b	SE (b)	c	SE (c)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
150789	0.87664	0.00454	-1.29715	0.0061	0	0	1.60765	0.01822	0.80956	0.01211	-0.5097	0.00782	-1.90752	0.00758	0	0
150433	1.50304	0.00751	-0.35582	0.00371	0	0	1.84962	0.01162	0.72512	0.00622	-0.71137	0.00502	-1.86336	0.00739	0	0
150803	1.38036	0.00683	-0.39925	0.00402	0	0	1.92209	0.01271	0.81647	0.00685	-0.70391	0.00529	-2.03465	0.00821	0	0
150854	1.31598	0.00559	0.33706	0.00382	0	0	2.88164	0.01473	2.53815	0.01161	2.29911	0.01	1.68862	0.00728	1.33053	0.00634

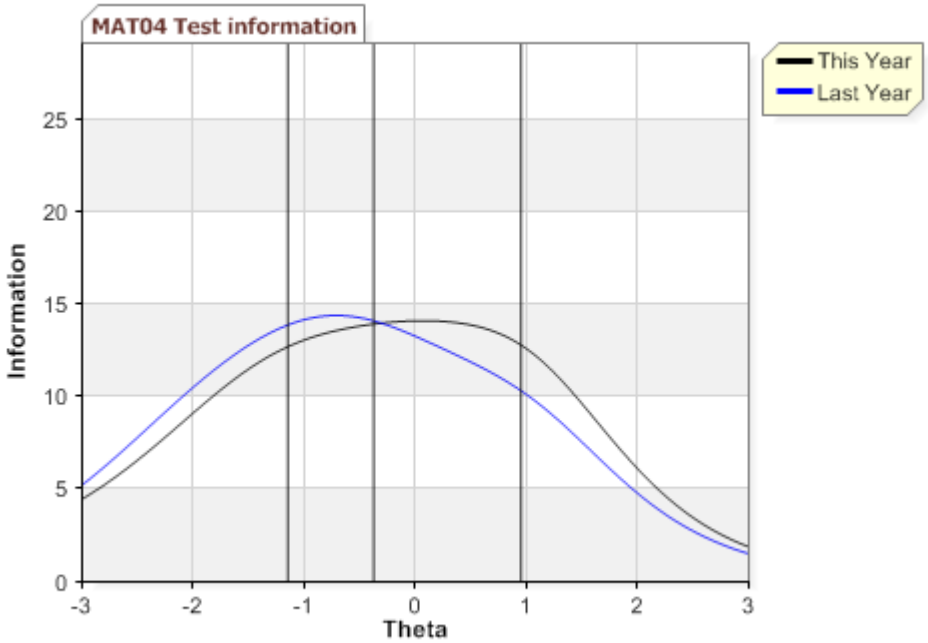
IREF (continued)	Parameters and Measures of Standard Error													
	D5	SE (D5)	D6	SE (D6)	D7	SE (D7)	D8	SE (D8)	D9	SE (D9)	D10	SE (D10)	D11	SE (D11)
150854	0.35194	0.0053	-0.16077	0.0055	-1.28751	0.00819	-2.12515	0.01425	-3.37971	0.04431	-4.13685	0.0979	0	0

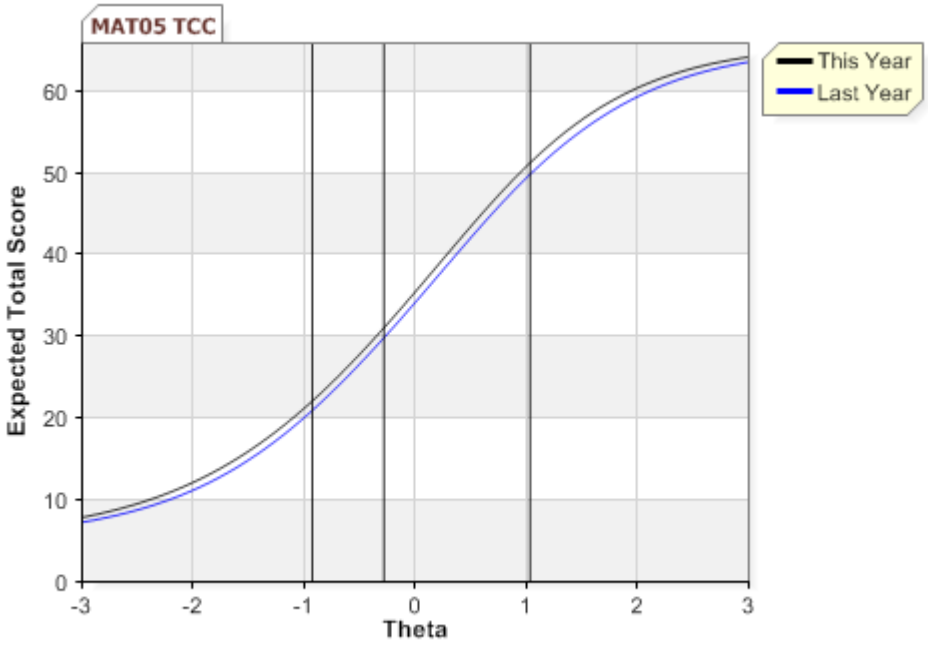
Appendix J—TCC AND TIF PLOTS

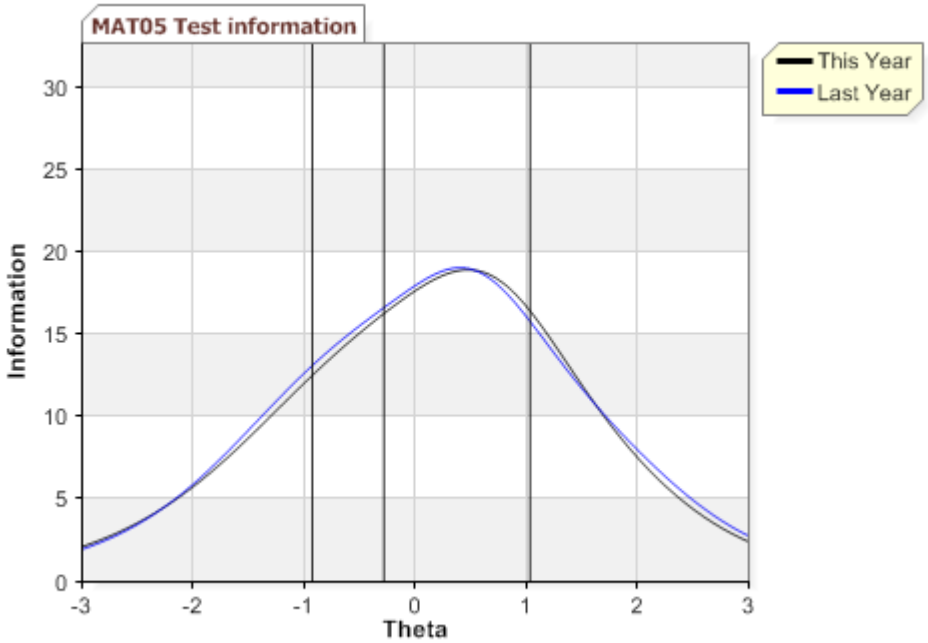

MAT03 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
60
-3
-2
-1
0
1
2
3
Theta

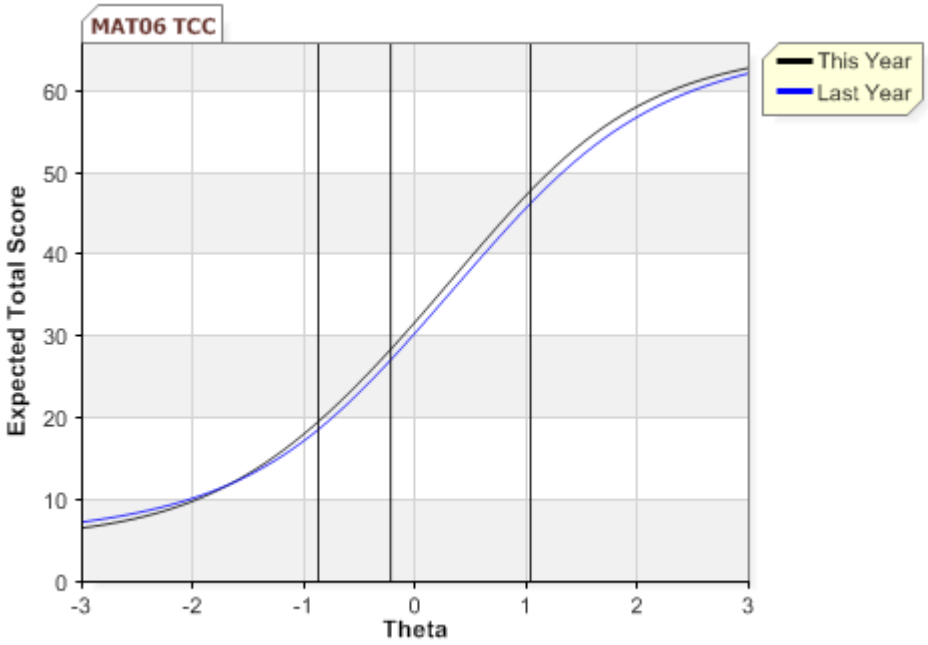

MAT03 Test information
This Year
Last Year
Information
0
5
10
15
20
25
30
-3
-2
-1
0
1
2
3
Theta

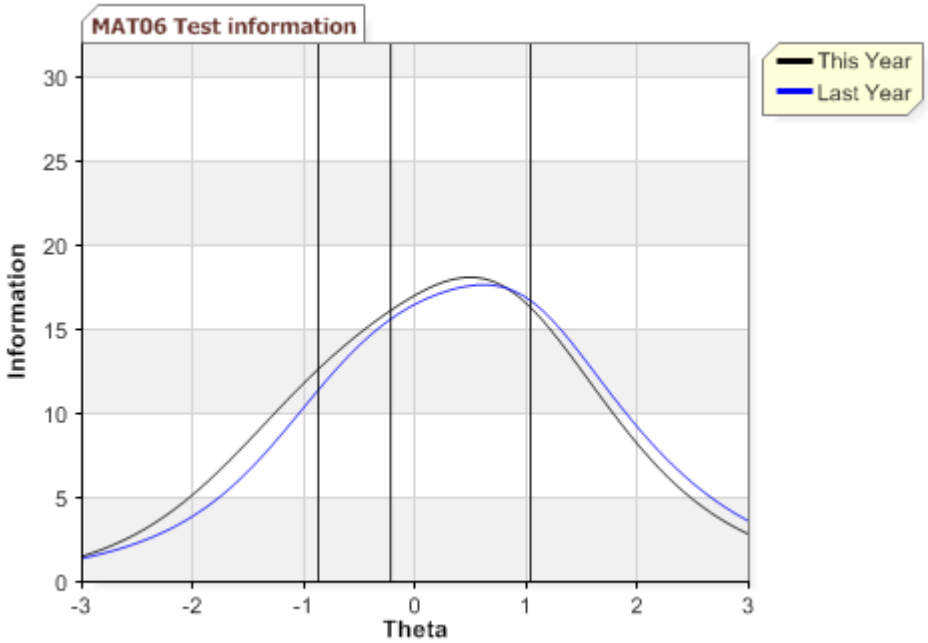

MAT04 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
60
-3
-2
-1
0
1
2
3
Theta

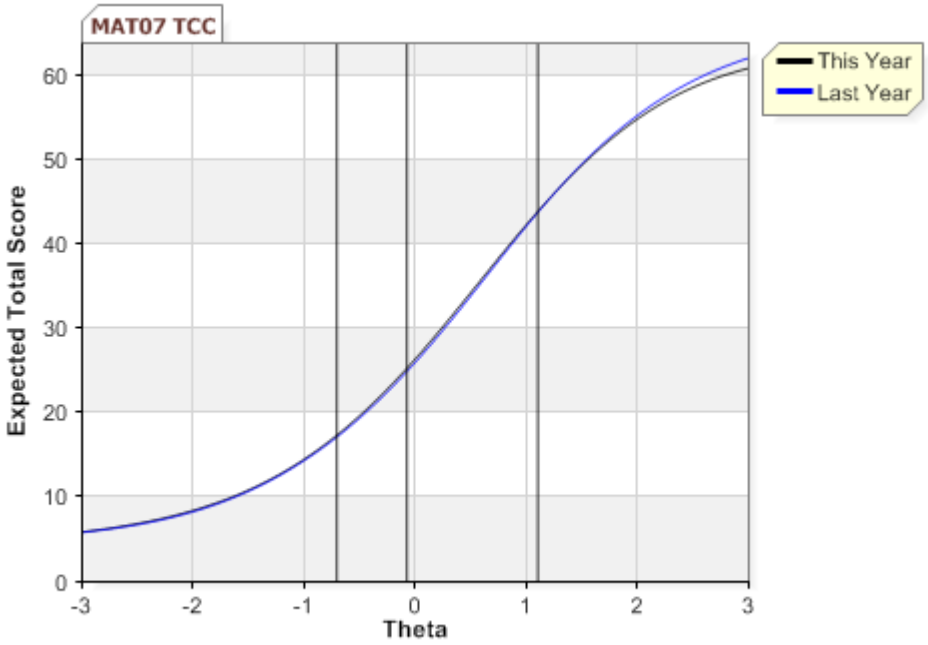

MAT04 Test information
This Year
Last Year
Information
0
5
10
15
20
25
-3
-2
-1
0
1
2
3
Theta

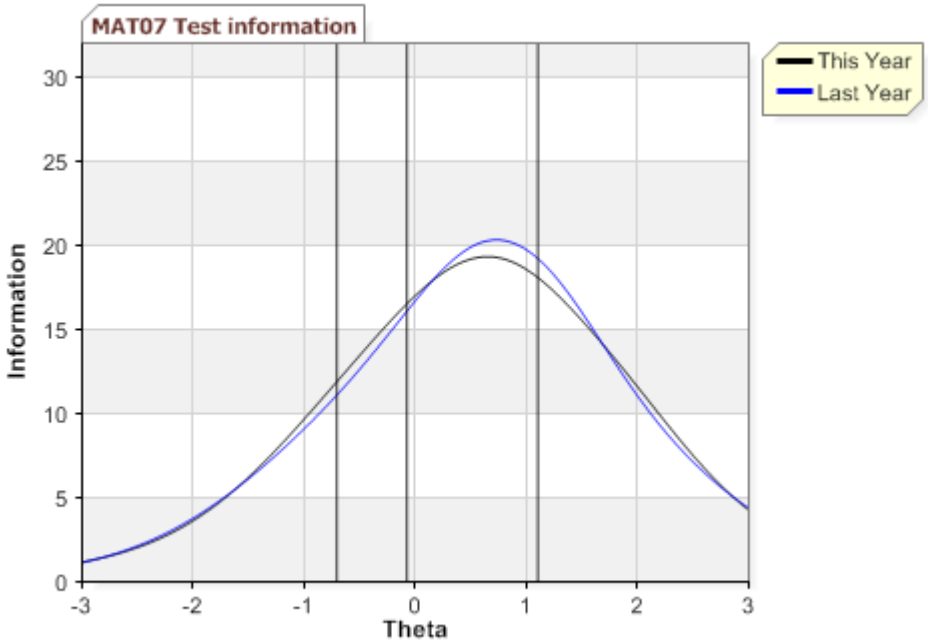

MAT05 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
60
-3
-2
-1
0
1
2
3
Theta

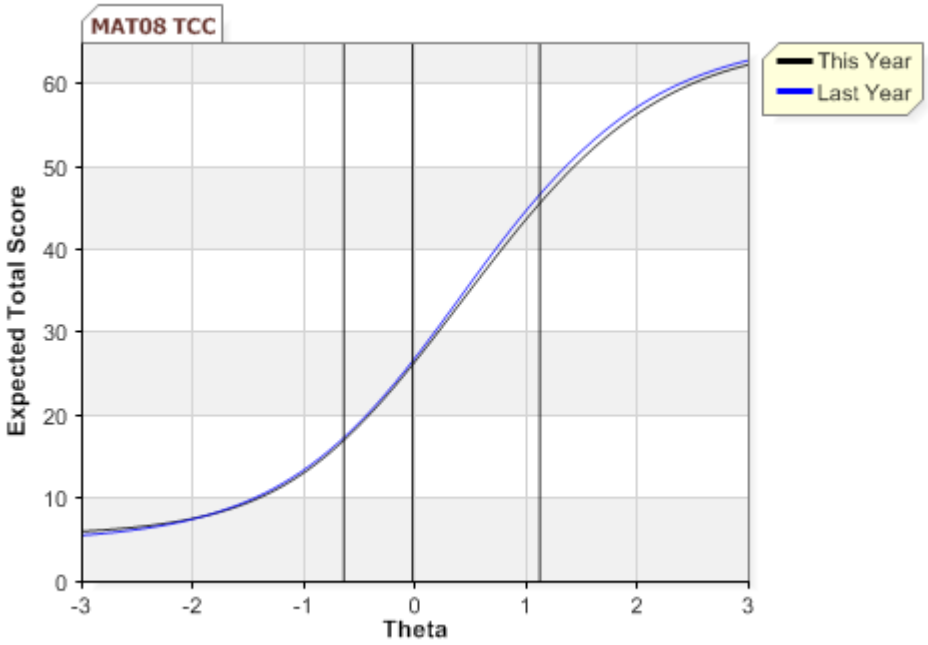

MAT05 Test information
This Year
Last Year
Information
0
5
10
15
20
25
30
-3
-2
-1
0
1
2
3
Theta

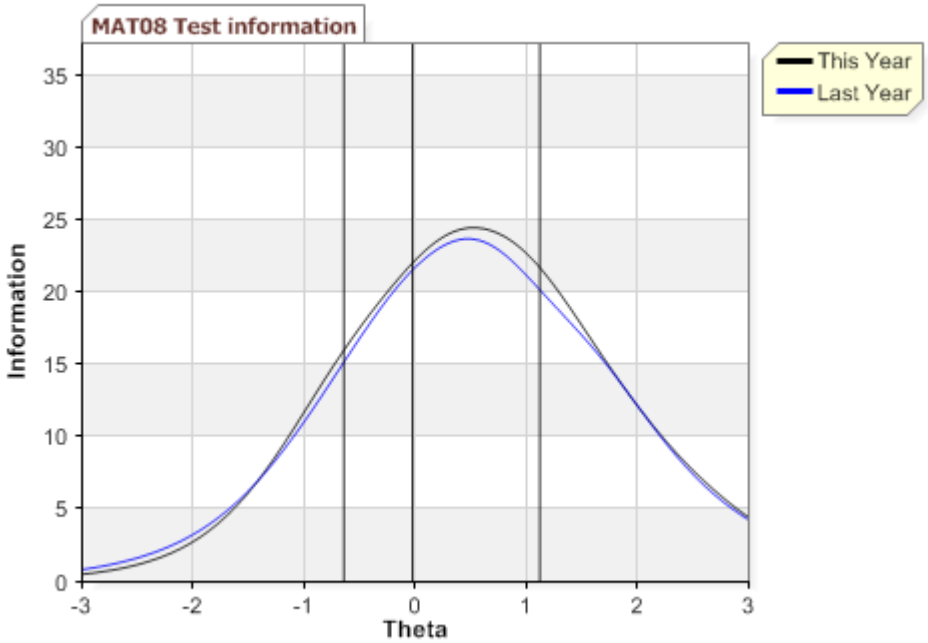

MAT06 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
60
-3
-2
-1
0
1
2
3
Theta

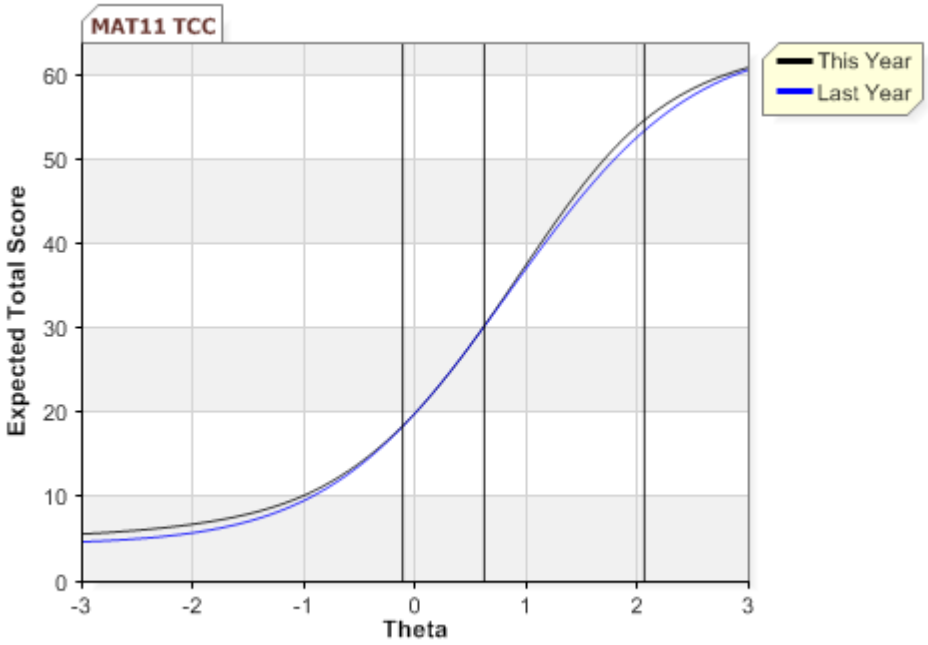

MAT06 Test information
This Year
Last Year
Information
0
5
10
15
20
25
30
-3
-2
-1
0
1
2
3
Theta

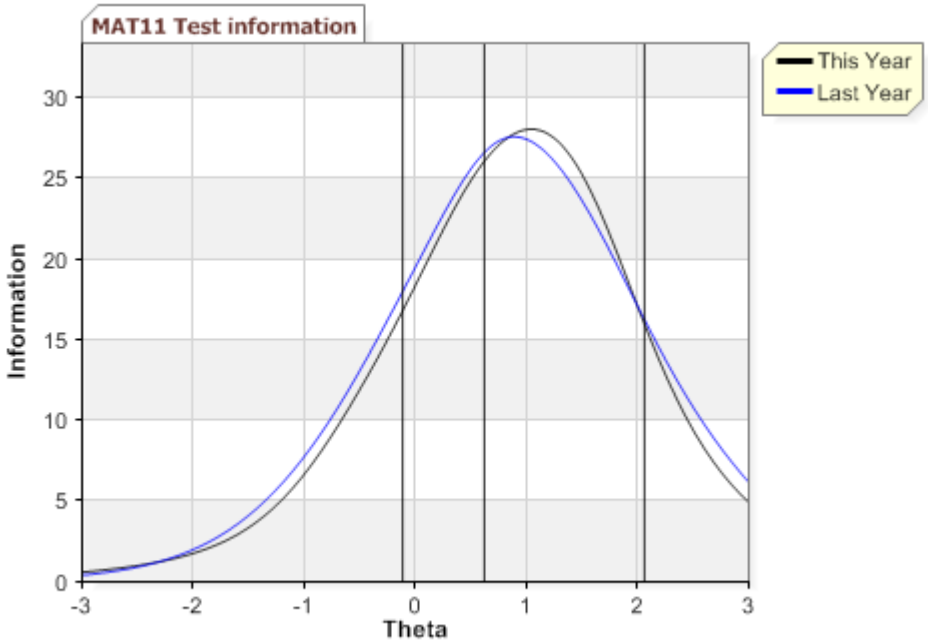

MAT07 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
60
-3
-2
-1
0
1
2
3
Theta

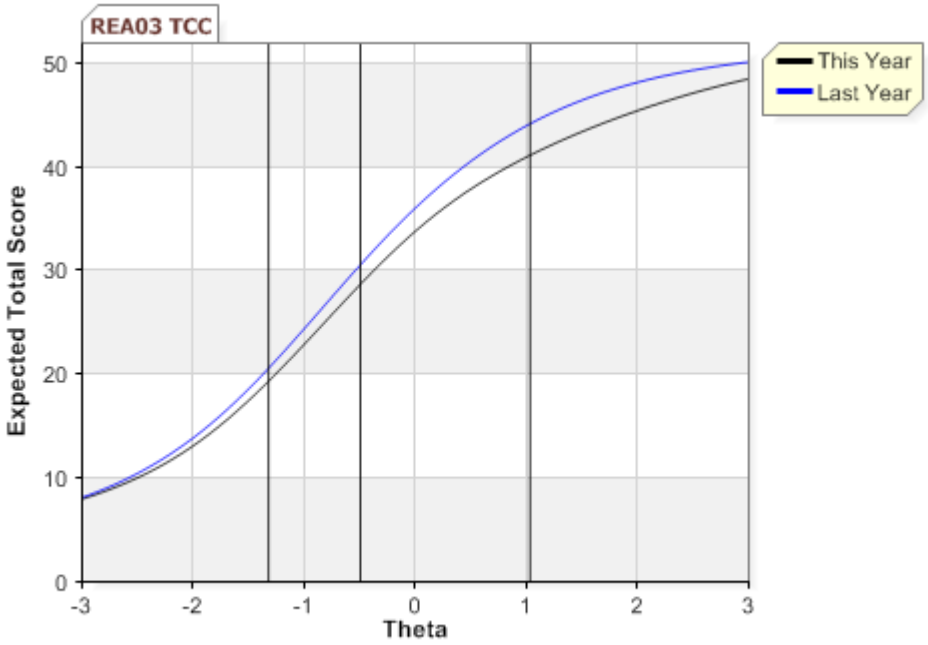

MAT07 Test information
This Year
Last Year
Information
0
5
10
15
20
25
30
-3
-2
-1
0
1
2
3
Theta

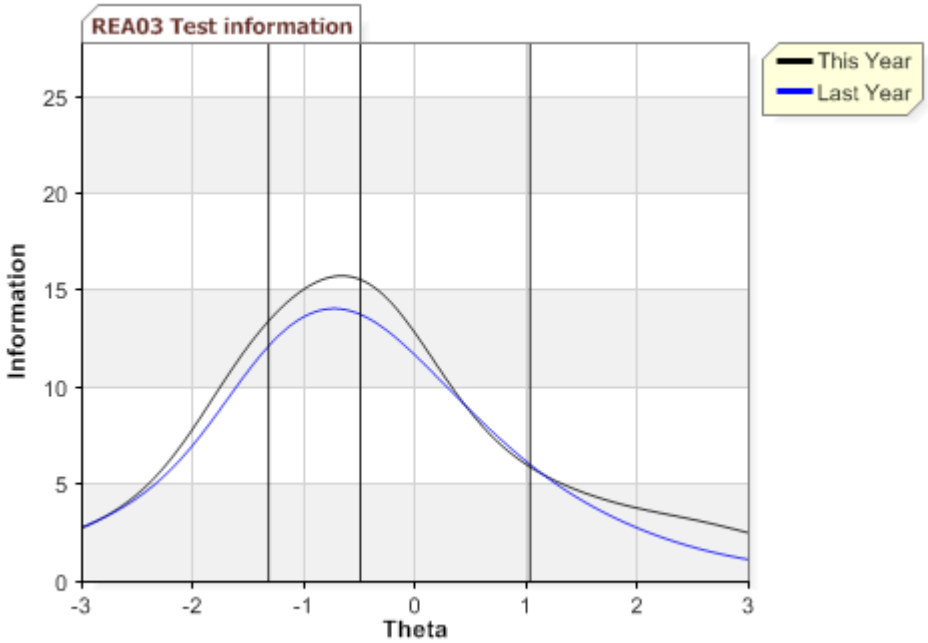

MAT08 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
60
-3
-2
-1
0
1
2
3
Theta

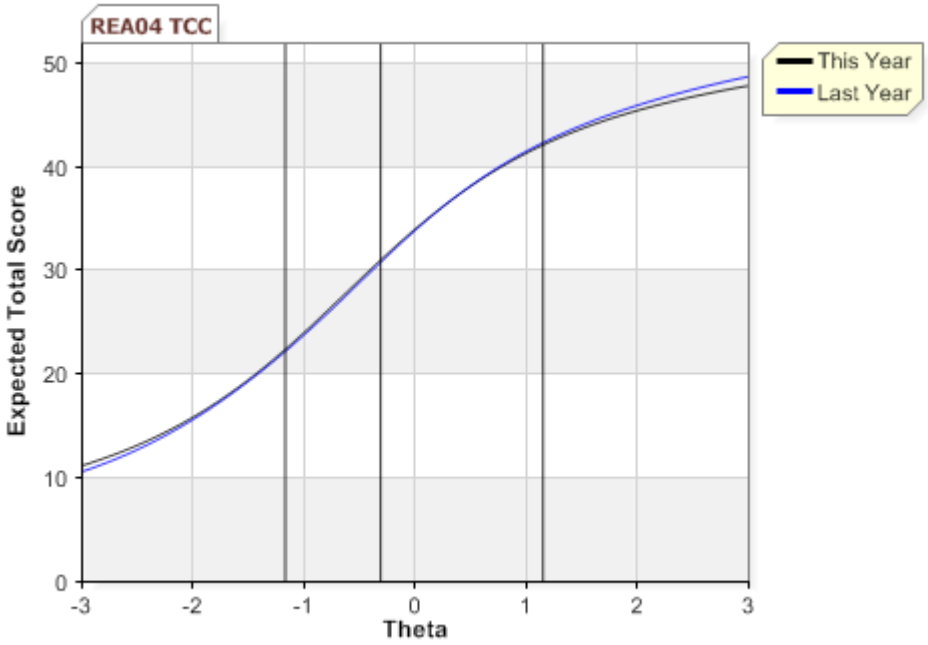

MAT08 Test information
This Year
Last Year
Information
0
5
10
15
20
25
30
35
-3
-2
-1
0
1
2
3
Theta

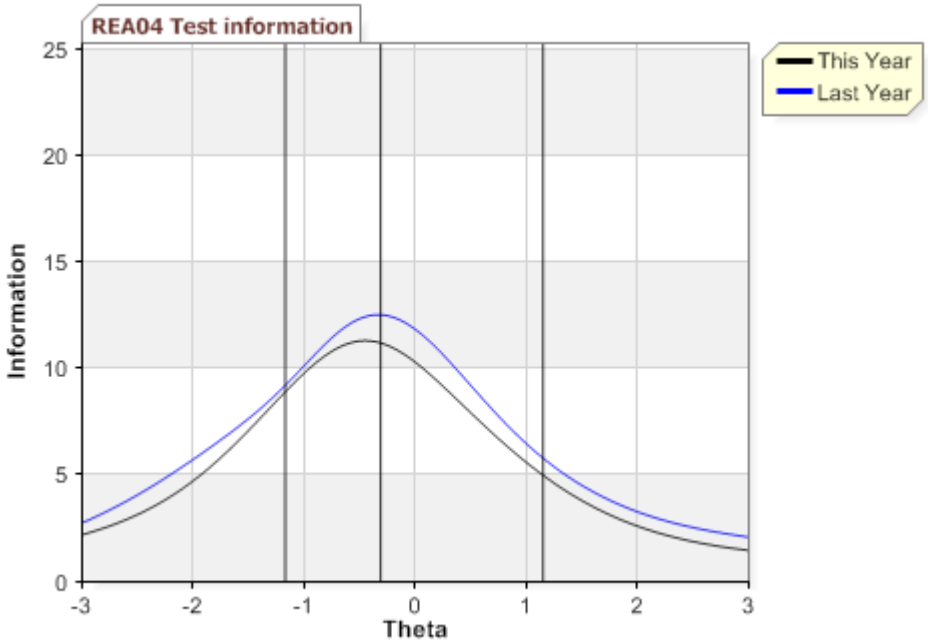

MAT11 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
60
-3
-2
-1
0
1
2
3
Theta

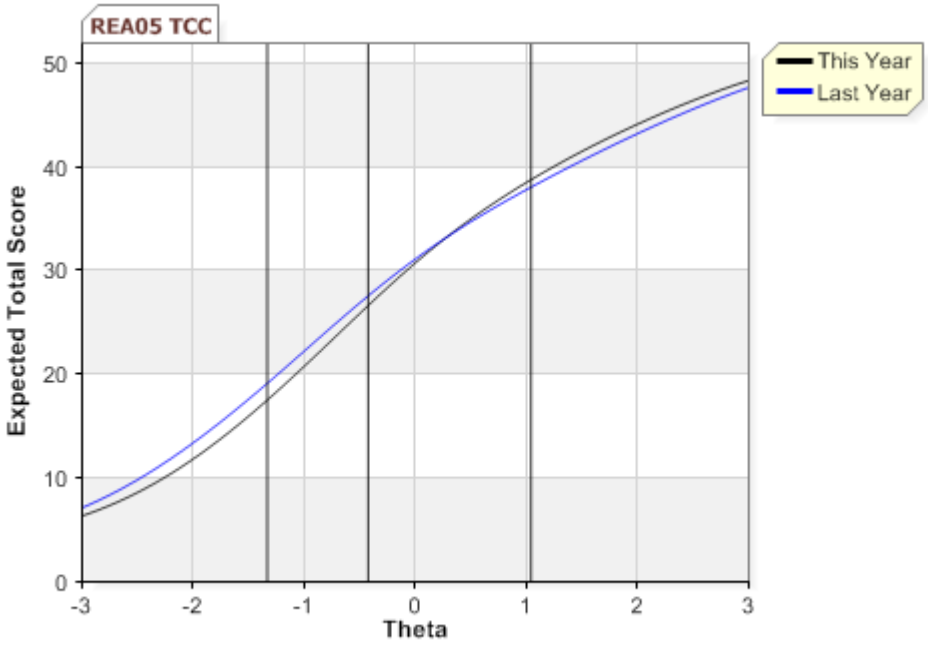

MAT11 Test information
This Year
Last Year
Information
0
5
10
15
20
25
30
-3
-2
-1
0
1
2
3
Theta

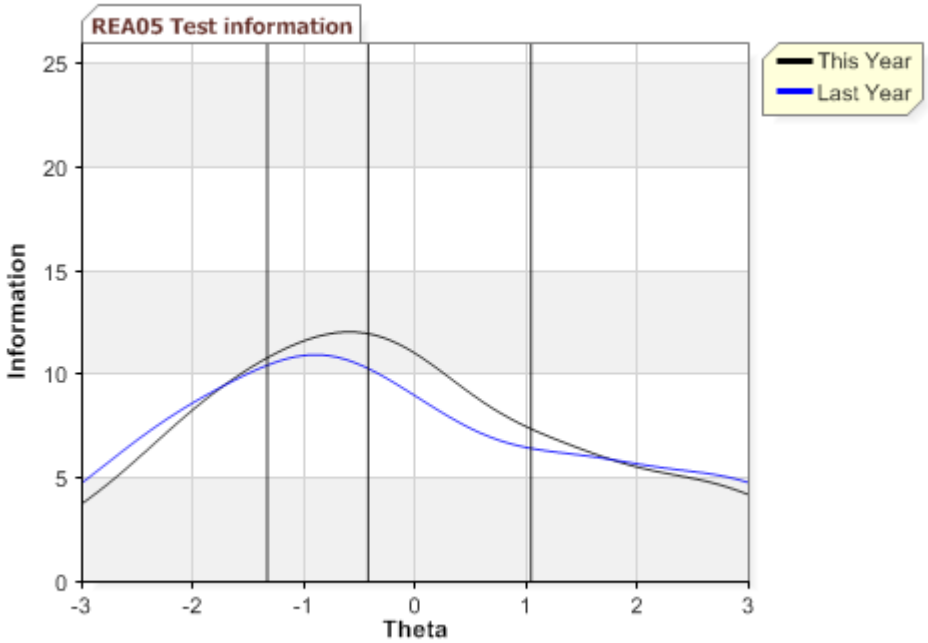

REA03 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
-3
-2
-1
0
1
2
3
Theta

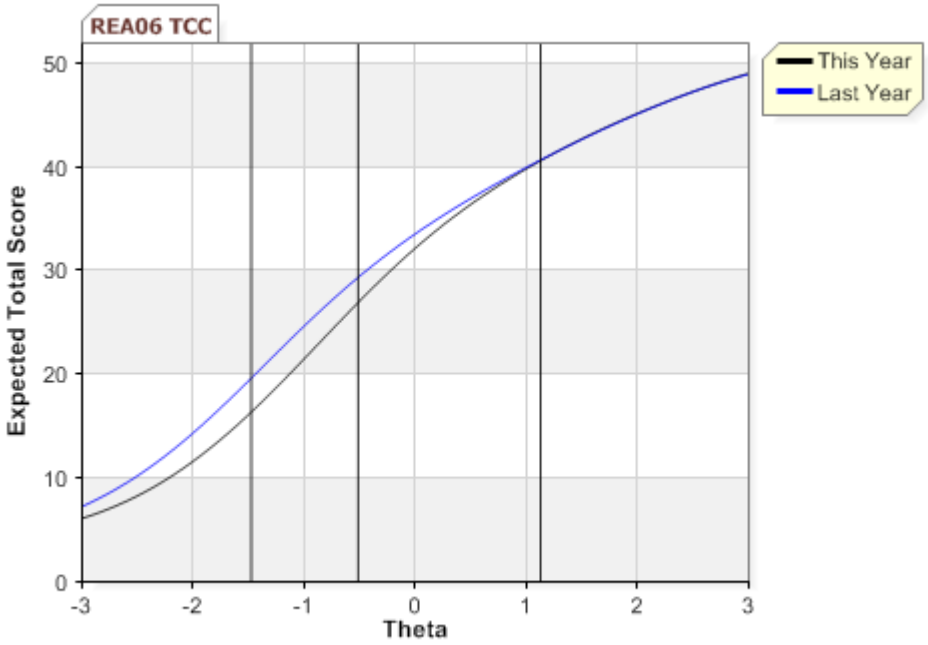

REA03 Test information
This Year
Last Year
Information
0
5
10
15
20
25
-3
-2
-1
0
1
2
3
Theta

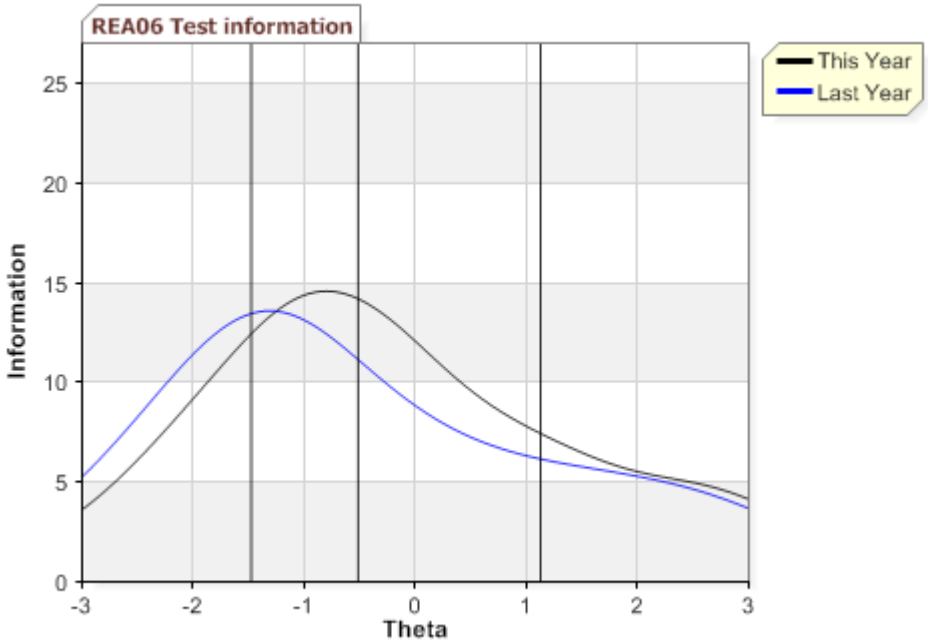

REA04 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
-3
-2
-1
0
1
2
3
Theta

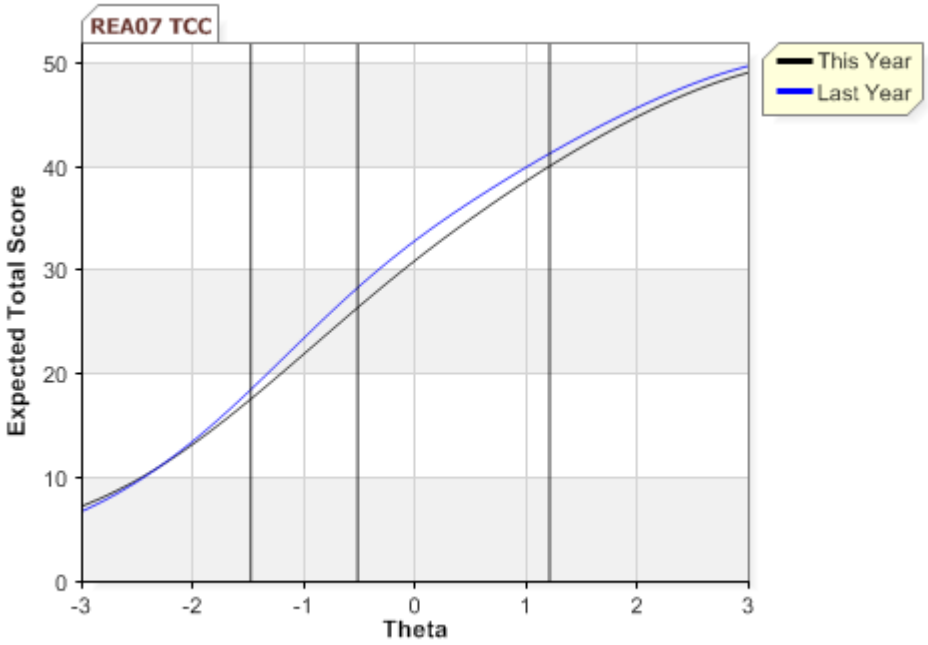

REA04 Test information
This Year
Last Year
Information
0
5
10
15
20
25
-3
-2
-1
0
1
2
3
Theta

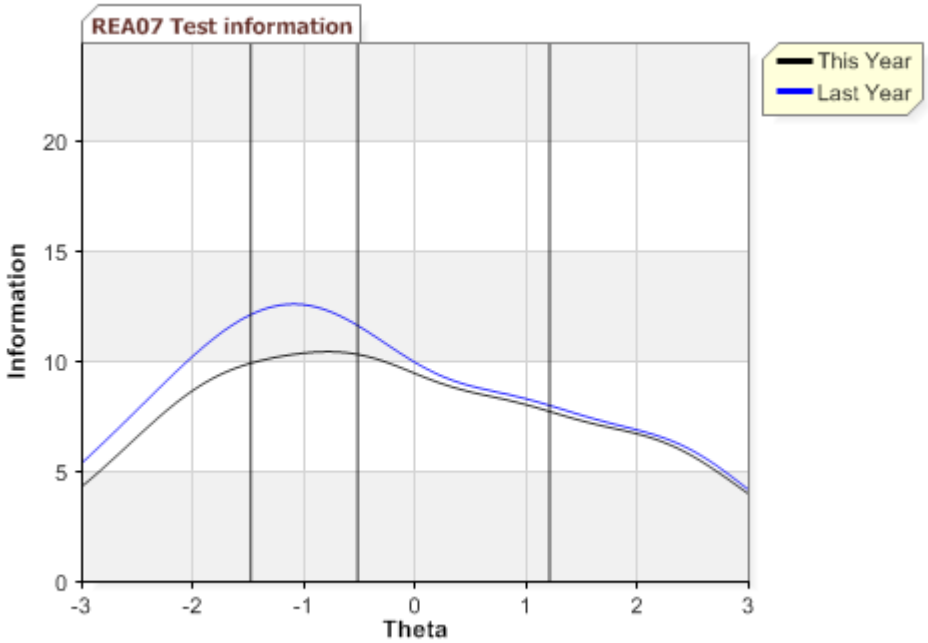

REA05 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
-3
-2
-1
0
1
2
3
Theta

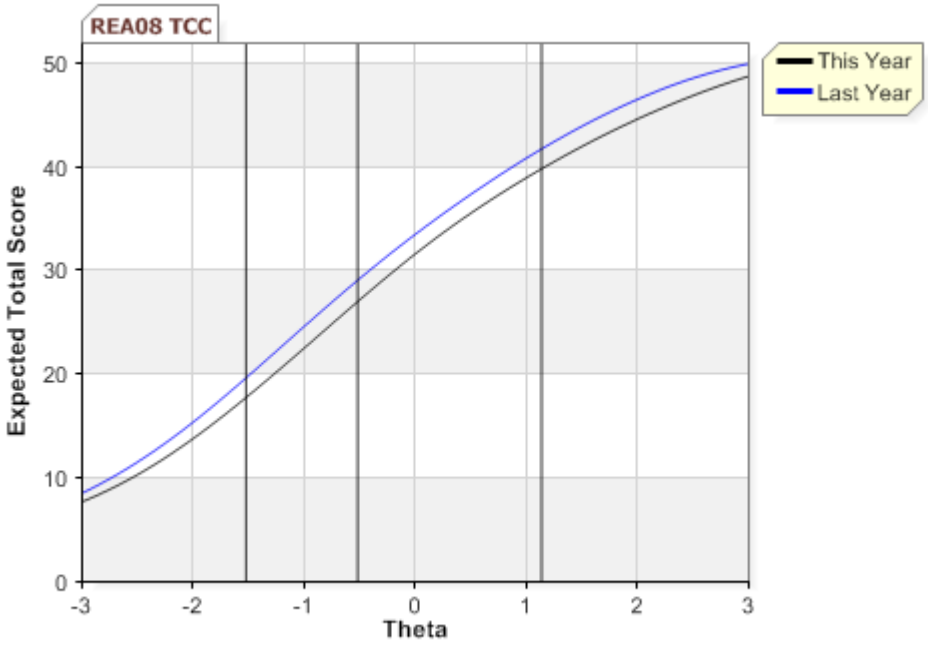

REA05 Test information
This Year
Last Year
Information
0
5
10
15
20
25
-3
-2
-1
0
1
2
3
Theta

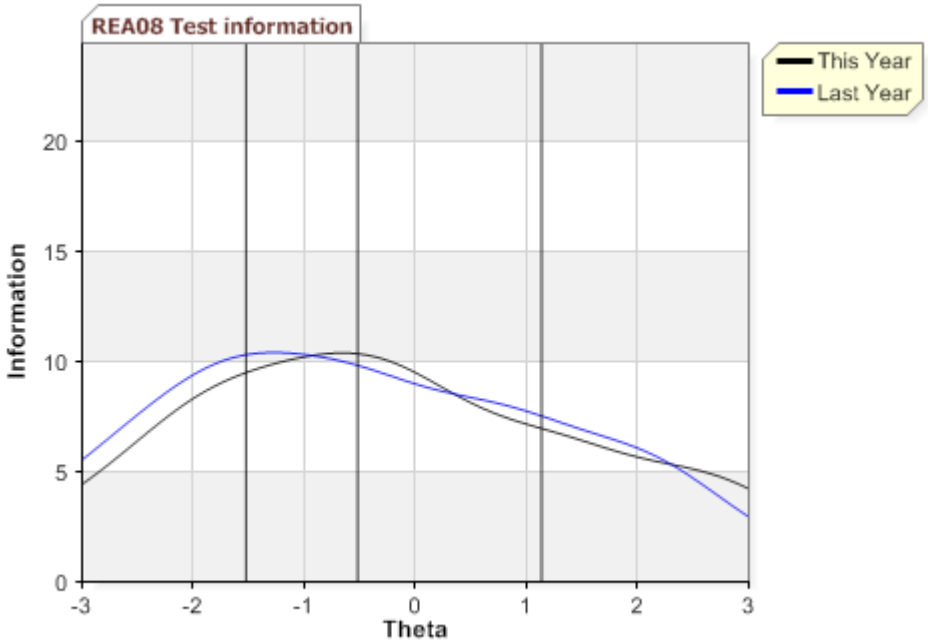

REA06 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
-3
-2
-1
0
1
2
3
Theta

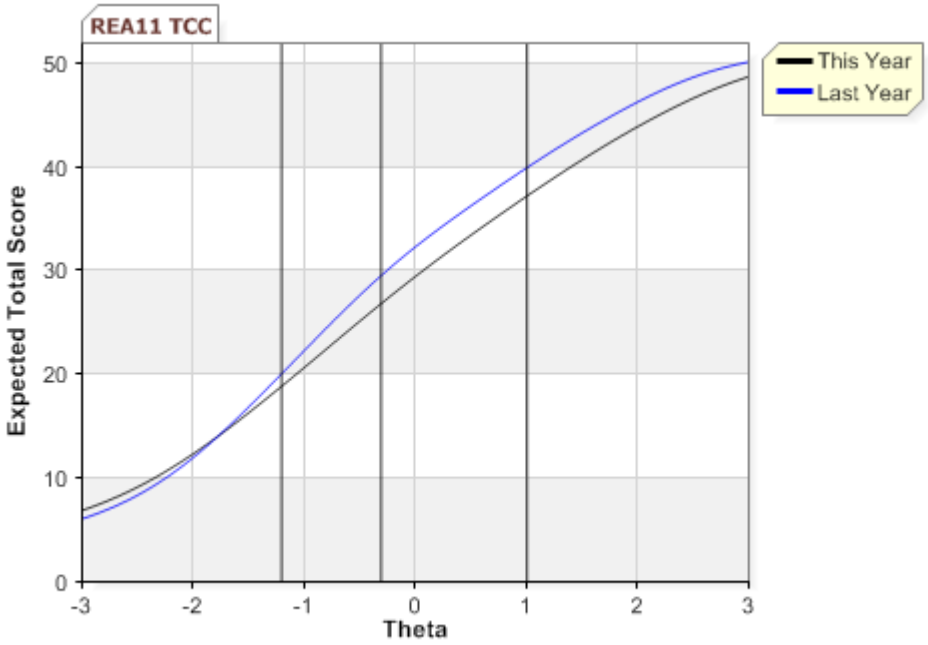

REA06 Test information
This Year
Last Year
Information
0
5
10
15
20
25
-3
-2
-1
0
1
2
3
Theta

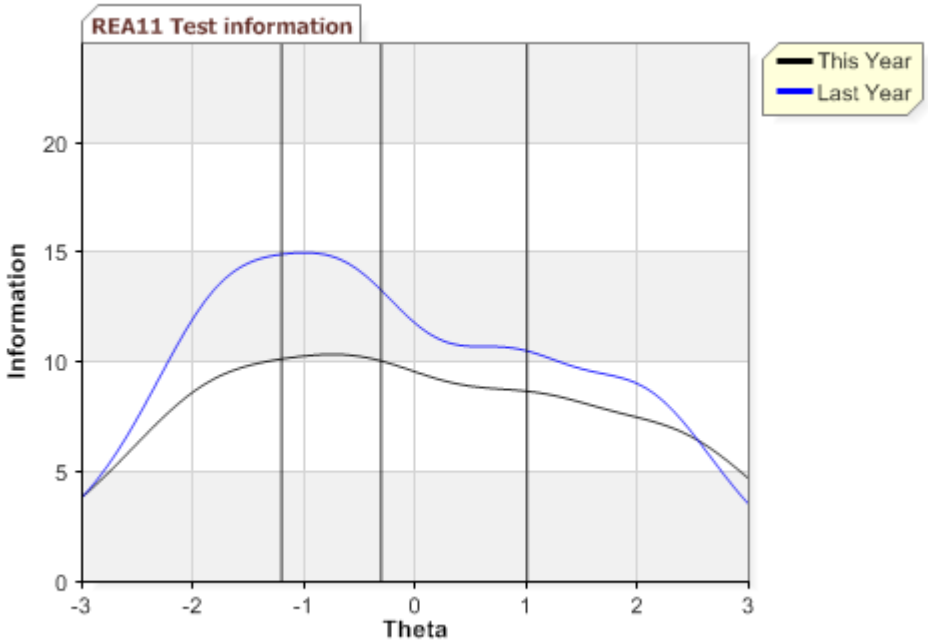

REA07 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
-3
-2
-1
0
1
2
3
Theta


REA07 Test information
This Year
Last Year
Information
0
5
10
15
20
-3
-2
-1
0
1
2
3
Theta

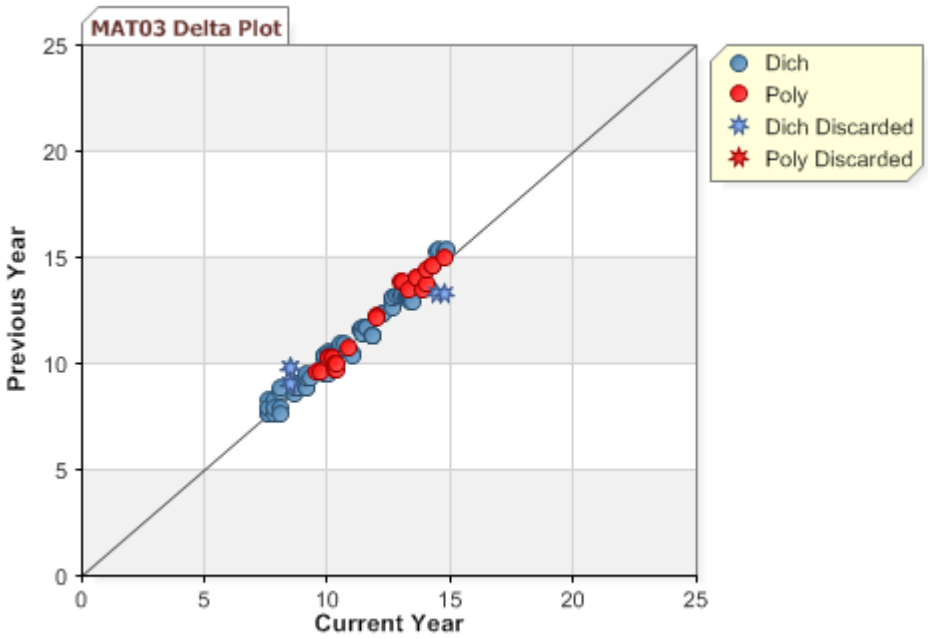

REA08 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
-3
-2
-1
0
1
2
3
Theta


REA08 Test information
This Year
Last Year
Information
0
5
10
15
20
-3
-2
-1
0
1
2
3
Theta


REA11 TCC
This Year
Last Year
Expected Total Score
0
10
20
30
40
50
-3
-2
-1
0
1
2
3
Theta


REA11 Test information
This Year
Last Year
Information
0
5
10
15
20
-3
-2
-1
0
1
2
3
Theta

Appendix K—Delta Analyses and Rescore Analyses



MAT03

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
119681	0.88	0.91	8.30	7.64	False	0.99
119699	0.90	0.91	7.87	7.64	False	-0.53
119699	0.90	0.91	7.87	7.64	False	-0.53
119699	0.91	0.91	7.64	7.64	False	-0.81
119699	0.91	0.91	7.64	7.64	False	-0.81
119701	0.81	0.78	9.49	9.91	False	0.64
119701	0.81	0.77	9.49	10.04	False	1.11
119702	0.48	0.52	13.20	12.80	False	0.17
119702	0.48	0.50	13.20	13.00	False	-0.54
119704	0.49	0.53	13.10	12.70	False	0.17
119708	0.72	0.71	10.67	10.84	False	-0.26
119715	0.81	0.83	9.49	9.18	False	-0.25
119747	0.66	0.65	11.35	11.46	False	-0.52
119752	0.85	0.86	8.85	8.68	False	-0.72
119752	0.87	0.86	8.49	8.68	False	-0.18
119761	0.36	0.40	14.43	14.07	False	0.09
119770	0.74	0.76	10.43	10.17	False	-0.42
119775	0.64	0.66	11.57	11.35	False	-0.52
119775	0.64	0.66	11.57	11.35	False	-0.52
119806	0.76	0.76	10.24	10.24	False	-0.87
119806	0.76	0.77	10.24	10.04	False	-0.63
119806	0.76	0.76	10.24	10.24	False	-0.87
119806	0.76	0.77	10.24	10.04	False	-0.63
119808	0.82	0.82	9.34	9.34	False	-0.85
119808	0.82	0.83	9.34	9.18	False	-0.79
119842	0.75	0.69	10.30	11.02	False	1.65
119842	0.74	0.69	10.43	11.02	False	1.21
119853	0.81	0.80	9.56	9.70	False	-0.35

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
119853	0.81	0.81	9.56	9.56	False	-0.86
119873	0.47	0.49	13.30	13.10	False	-0.53
119877	0.85	0.89	8.85	8.09	False	1.34
119877	0.86	0.89	8.68	8.09	False	0.72
119880	0.73	0.77	10.55	10.04	False	0.48
119886	0.54	0.53	12.60	12.70	False	-0.57
119894	0.48	0.48	13.20	13.20	False	-0.94
119894	0.49	0.48	13.10	13.20	False	-0.58
119894	0.49	0.46	13.10	13.40	False	0.13
119894	0.48	0.46	13.20	13.40	False	-0.23
119896	0.47	0.36	13.30	14.43	True	3.06
119896	0.47	0.33	13.30	14.76	True	4.21
119901	0.51	0.53	12.90	12.70	False	-0.54
119901	0.49	0.53	13.10	12.70	False	0.17
119903	0.29	0.33	15.21	14.76	False	0.41
119903	0.29	0.36	15.21	14.43	False	1.56
119909	0.43	0.40	13.71	14.01	False	0.14
119911	0.75	0.75	10.30	10.30	False	-0.88
119920	0.59	0.60	12.14	11.99	False	-0.72
119920	0.58	0.60	12.19	11.99	False	-0.54
119924	0.85	0.83	8.85	9.18	False	0.32
119924	0.82	0.83	9.34	9.18	False	-0.79
119924	0.82	0.83	9.34	9.18	False	-0.79
119924	0.85	0.83	8.85	9.18	False	0.32
119926	0.57	0.57	12.29	12.29	False	-0.92
119926	0.57	0.57	12.29	12.29	False	-0.92
119930	0.49	0.48	13.10	13.20	False	-0.58
119930	0.48	0.48	13.20	13.20	False	-0.94
119930	0.48	0.47	13.20	13.30	False	-0.59
119930	0.49	0.47	13.10	13.30	False	-0.23
124490	0.42	0.49	13.81	13.10	False	1.27
124490	0.42	0.50	13.81	13.00	False	1.62
139575	0.63	0.64	11.67	11.57	False	-0.90
139575	0.63	0.65	11.67	11.46	False	-0.52
139606	0.51	0.46	12.90	13.40	False	0.84
139606	0.51	0.45	12.90	13.50	False	1.20
139651	0.35	0.37	14.54	14.33	False	-0.46
139651	0.35	0.38	14.54	14.27	False	-0.27
139661	0.88	0.90	8.30	7.87	False	0.15
139667	0.72	0.73	10.67	10.55	False	-0.88
139667	0.72	0.73	10.67	10.55	False	-0.88
144609	0.63	0.64	11.67	11.57	False	-0.90
144609	0.63	0.64	11.67	11.57	False	-0.90
145244	0.28	0.32	15.33	14.87	False	0.43
145244	0.28	0.35	15.33	14.54	False	1.60
145259	0.31	0.33	14.98	14.76	False	-0.41
145720	0.76	0.76	10.17	10.17	False	-0.87
198283	0.84	0.85	9.02	8.85	False	-0.75
198315	0.84	0.85	9.02	8.85	False	-0.75
198507	0.80	0.75	9.70	10.30	False	1.25

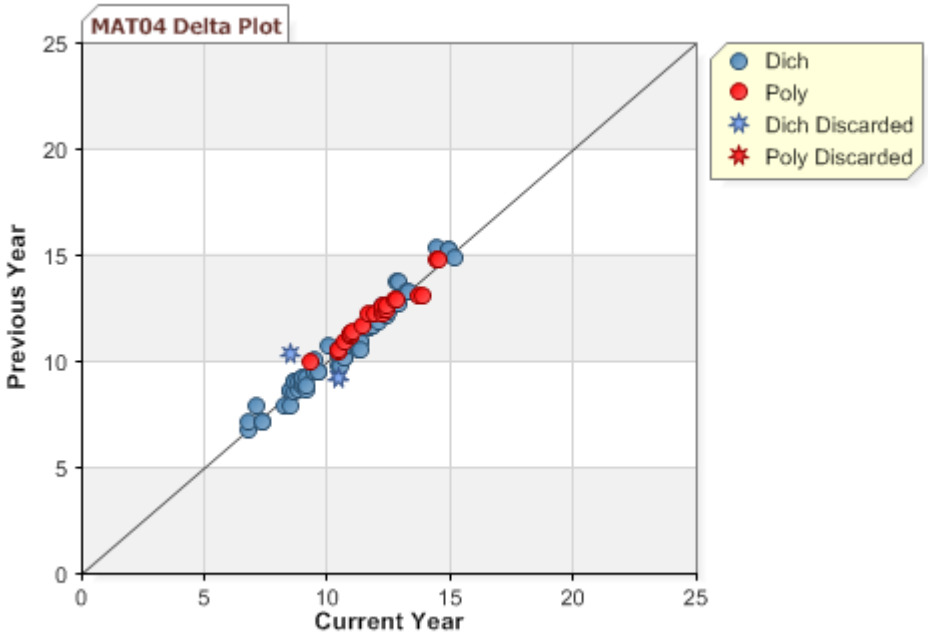
IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
198507	0.80	0.75	9.70	10.36	False	1.47
198507	0.78	0.75	9.91	10.36	False	0.74
198507	0.78	0.75	9.91	10.30	False	0.52
198521	0.46	0.47	13.45	13.30	False	-0.71
198521	0.46	0.41	13.45	13.91	False	0.67
201312	0.84	0.87	9.02	8.49	False	0.52
201312	0.84	0.87	9.02	8.49	False	0.52
201312	0.79	0.87	9.77	8.49	True	3.20
201312	0.79	0.87	9.77	8.49	True	3.20
201416	0.70	0.73	10.90	10.55	False	-0.05
201416	0.70	0.72	10.90	10.67	False	-0.47
201450	0.47	0.46	13.30	13.40	False	-0.59
201465	0.76	0.77	10.17	10.04	False	-0.85
201510	0.76	0.78	10.17	9.91	False	-0.38
201510	0.75	0.78	10.30	9.91	False	0.07
201806	0.67	0.61	11.24	11.88	False	1.37
201806	0.67	0.61	11.24	11.88	False	1.37
201806	0.67	0.61	11.24	11.88	False	1.37
201806	0.67	0.61	11.24	11.88	False	1.37
201811	0.76	0.76	10.17	10.17	False	-0.87
201811	0.76	0.75	10.17	10.30	False	-0.42
201811	0.75	0.76	10.30	10.17	False	-0.86
201811	0.75	0.75	10.30	10.30	False	-0.88
202089	0.40	0.44	14.01	13.60	False	0.22
202089	0.40	0.44	14.01	13.60	False	0.22
223883	0.75	0.76	10.30	10.17	False	-0.86
223883	0.75	0.75	10.30	10.30	False	-0.88
223883	0.75	0.76	10.30	10.17	False	-0.86
223883	0.75	0.75	10.30	10.30	False	-0.88
255679	0.85	0.85	8.85	8.85	False	-0.84
255686	0.75	0.78	10.30	9.91	False	0.07
255686	0.75	0.78	10.30	9.91	False	0.07
255900	0.40	0.41	14.01	13.91	False	-0.86
255929	0.90	0.90	7.87	7.87	False	-0.82
255929	0.90	0.89	7.87	8.09	False	-0.04
255929	0.91	0.90	7.64	7.87	False	0.02
255929	0.91	0.89	7.64	8.09	False	0.80
255932	0.42	0.45	13.81	13.50	False	-0.15
255932	0.42	0.44	13.81	13.60	False	-0.51

MAT03

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
119853	2	1.63	1.61	0.64	0.65	-0.03	No
119806	2	1.50	1.49	0.67	0.64	-0.02	No
198507	2	1.56	1.56	0.67	0.69	0.00	No
198521	2	0.90	0.95	0.86	0.83	0.05	No
124490	2	0.73	0.85	0.83	0.79	0.15	No
139651	2	0.68	0.75	0.77	0.78	0.09	No

MAT03

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
202089	2	0.81	0.79	0.93	0.92	-0.03	No
145259	2	0.58	0.57	0.78	0.78	-0.01	No
119761	2	0.69	0.76	0.78	0.77	0.09	No
119708	2	1.46	1.47	0.65	0.64	0.02	No
119909	2	0.89	0.87	0.78	0.79	-0.02	No
119920	2	1.09	1.09	0.91	0.90	-0.01	No



MAT04

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
120071	0.57	0.56	12.29	12.40	False	-0.55
120071	0.57	0.56	12.29	12.40	False	-0.55
120072	0.85	0.84	8.85	9.02	False	-0.59
120074	0.64	0.63	11.57	11.67	False	-0.59
120074	0.64	0.64	11.57	11.57	False	-0.95
120074	0.64	0.64	11.57	11.57	False	-0.95
120074	0.64	0.63	11.57	11.67	False	-0.59
120083	0.70	0.72	10.90	10.73	False	-0.34
120098	0.32	0.29	14.87	15.21	False	0.47
120102	0.84	0.84	9.02	9.02	False	-0.79
120102	0.84	0.85	9.02	8.85	False	-0.22
120108	0.71	0.72	10.79	10.67	False	-0.52
120108	0.71	0.72	10.79	10.67	False	-0.52
120126	0.84	0.86	9.02	8.68	False	0.38
120141	0.29	0.31	15.21	14.98	False	-0.48
120141	0.29	0.31	15.21	14.98	False	-0.48
120146	0.43	0.52	13.71	12.80	False	1.93
120146	0.43	0.51	13.71	12.90	False	1.59
120159	0.72	0.77	10.67	10.04	False	1.21
120173	0.56	0.58	12.40	12.19	False	-0.35
120174	0.47	0.47	13.30	13.30	False	-0.82
120174	0.47	0.48	13.30	13.20	False	-0.77
120183	0.90	0.93	7.87	7.10	False	1.94
120187	0.75	0.74	10.30	10.43	False	-0.63
120187	0.75	0.74	10.30	10.43	False	-0.63
120197	0.60	0.59	11.99	12.09	False	-0.57
120197	0.61	0.59	11.88	12.09	False	-0.23
120203	0.58	0.61	12.19	11.93	False	-0.16

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
120203	0.58	0.63	12.19	11.73	False	0.56
120214	0.85	0.83	8.85	9.18	False	-0.04
120214	0.86	0.83	8.68	9.18	False	0.54
120232	0.83	0.84	9.18	9.02	False	-0.25
120232	0.84	0.84	9.02	9.02	False	-0.79
120236	0.73	0.74	10.55	10.43	False	-0.49
120253	0.83	0.84	9.18	9.02	False	-0.25
120253	0.83	0.83	9.18	9.18	False	-0.80
120257	0.60	0.61	11.99	11.88	False	-0.66
120257	0.59	0.61	12.09	11.88	False	-0.32
120266	0.63	0.61	11.67	11.88	False	-0.23
120266	0.63	0.61	11.67	11.88	False	-0.23
120269	0.59	0.56	12.09	12.40	False	0.13
120280	0.70	0.66	10.90	11.35	False	0.52
120280	0.70	0.66	10.90	11.35	False	0.52
120290	0.66	0.69	11.35	11.07	False	-0.02
120291	0.81	0.81	9.49	9.49	False	-0.82
120299	0.73	0.74	10.55	10.43	False	-0.49
120299	0.74	0.74	10.43	10.43	False	-0.90
120301	0.52	0.53	12.85	12.75	False	-0.74
120301	0.52	0.52	12.85	12.80	False	-0.91
124526	0.72	0.71	10.67	10.79	False	-0.62
124592	0.83	0.74	9.18	10.43	True	3.09
124592	0.83	0.74	9.18	10.43	True	3.09
124700	0.58	0.58	12.19	12.24	False	-0.73
139465	0.71	0.69	10.79	11.02	False	-0.23
139482	0.63	0.66	11.67	11.40	False	-0.08
139521	0.79	0.73	9.77	10.55	False	1.54
139521	0.79	0.74	9.77	10.43	False	1.13
144645	0.33	0.35	14.76	14.54	False	-0.49
144645	0.33	0.36	14.76	14.49	False	-0.30
145079	0.73	0.67	10.55	11.24	False	1.32
145079	0.73	0.66	10.55	11.35	False	1.69
145268	0.77	0.81	10.04	9.49	False	1.02
145281	0.75	0.74	10.30	10.43	False	-0.63
145281	0.75	0.74	10.30	10.43	False	-0.63
145563	0.93	0.92	7.10	7.38	False	-0.34
145563	0.93	0.92	7.10	7.38	False	-0.34
145854	0.53	0.51	12.70	12.90	False	-0.18
145854	0.53	0.52	12.70	12.80	False	-0.52
198351	0.94	0.94	6.78	6.78	False	-0.61
198351	0.93	0.94	7.10	6.78	False	0.43
198385	0.54	0.55	12.60	12.50	False	-0.72
198385	0.54	0.57	12.60	12.29	False	-0.03
198385	0.57	0.55	12.29	12.50	False	-0.21
198385	0.57	0.57	12.29	12.29	False	-0.90
198396	0.61	0.62	11.88	11.78	False	-0.65
198431	0.50	0.43	13.05	13.71	False	1.39
198431	0.50	0.41	13.05	13.91	False	2.08
198442	0.78	0.82	9.98	9.34	False	1.31

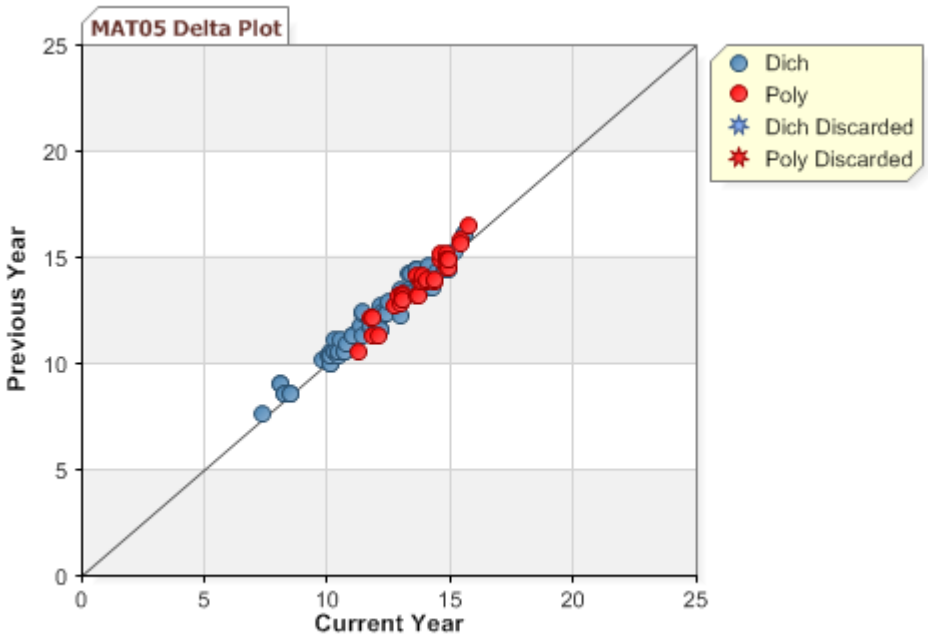
IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
202346	0.87	0.86	8.49	8.68	False	-0.56
202346	0.87	0.87	8.49	8.49	False	-0.75
202390	0.90	0.87	7.87	8.49	False	0.87
202390	0.90	0.88	7.87	8.30	False	0.21
227082	0.67	0.70	11.24	10.96	False	0.00
227082	0.67	0.70	11.24	10.96	False	0.00
227082	0.68	0.70	11.18	10.96	False	-0.19
227082	0.68	0.70	11.18	10.96	False	-0.19
227100	0.86	0.85	8.68	8.85	False	-0.58
227100	0.86	0.87	8.68	8.49	False	-0.13
227853	0.57	0.56	12.29	12.40	False	-0.55
227853	0.55	0.56	12.50	12.40	False	-0.71
232445	0.84	0.86	9.02	8.68	False	0.38
232445	0.84	0.86	9.02	8.68	False	0.38
232636	0.28	0.36	15.33	14.43	False	1.78
255664	0.75	0.87	10.30	8.49	True	5.26
255685	0.81	0.80	9.49	9.63	False	-0.62
255685	0.81	0.81	9.49	9.49	False	-0.82
255685	0.81	0.80	9.49	9.63	False	-0.62
255685	0.81	0.81	9.49	9.49	False	-0.82
255692	0.73	0.74	10.55	10.43	False	-0.49
255692	0.73	0.73	10.55	10.55	False	-0.90
255694	0.75	0.72	10.30	10.67	False	0.20
255694	0.75	0.72	10.30	10.67	False	0.20
255694	0.76	0.72	10.17	10.67	False	0.62
255694	0.76	0.72	10.17	10.67	False	0.62
270069	0.56	0.57	12.45	12.29	False	-0.53
270069	0.56	0.56	12.45	12.45	False	-0.89
270069	0.54	0.57	12.60	12.29	False	-0.03
270069	0.54	0.56	12.60	12.45	False	-0.55

MAT04

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
120203	2	1.17	1.30	0.76	0.77	0.17	No
270069	2	1.13	1.12	0.93	0.94	-0.01	No
198431	2	0.97	0.84	0.76	0.80	-0.17	No
227082	2	1.40	1.43	0.56	0.54	0.05	No
144645	2	0.67	0.73	0.88	0.90	0.06	No
124700	2	1.16	1.17	0.53	0.52	0.03	No

MAT04

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
198442	2	1.51	1.43	0.61	0.68	-0.12	No
139482	2	1.26	1.28	0.74	0.73	0.02	No
120299	2	1.49	1.48	0.70	0.71	-0.01	No
120290	2	1.44	1.43	0.77	0.77	-0.01	No
120083	2	1.44	1.43	0.69	0.69	-0.01	No



MAT05

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
120631	0.49	0.51	13.15	12.90	False	-0.56
120639	0.67	0.61	11.30	11.88	False	1.32
120639	0.67	0.59	11.30	12.09	False	2.12
120640	0.84	0.89	9.02	8.09	False	1.93
120640	0.84	0.89	9.02	8.09	False	1.93
120644	0.75	0.76	10.30	10.17	False	-1.14
120644	0.75	0.76	10.30	10.17	False	-1.14
120644	0.78	0.76	9.91	10.17	False	0.11
120644	0.78	0.76	9.91	10.17	False	0.11
120653	0.68	0.73	11.13	10.55	False	0.65
120653	0.68	0.75	11.13	10.30	False	1.60
120657	0.63	0.62	11.67	11.78	False	-0.56
120667	0.37	0.36	14.33	14.43	False	-0.64
120667	0.37	0.40	14.33	14.01	False	-0.28
120682	0.67	0.65	11.24	11.46	False	-0.11
120686	0.56	0.57	12.40	12.29	False	-1.17
120686	0.56	0.56	12.40	12.40	False	-0.99
120694	0.48	0.44	13.20	13.60	False	0.55
120694	0.48	0.43	13.20	13.71	False	0.94
120698	0.45	0.40	13.50	14.01	False	0.95
120698	0.45	0.37	13.50	14.33	False	2.17
120700	0.56	0.65	12.40	11.46	False	2.07
120700	0.57	0.65	12.29	11.46	False	1.67
120704	0.59	0.62	12.14	11.83	False	-0.37
120704	0.59	0.62	12.14	11.78	False	-0.16
120713	0.36	0.31	14.43	14.98	False	1.07
120713	0.36	0.33	14.43	14.76	False	0.21
120720	0.62	0.66	11.78	11.35	False	0.08

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
120731	0.50	0.49	12.97	13.10	False	-0.52
120731	0.50	0.49	12.97	13.08	False	-0.62
120733	0.91	0.92	7.64	7.38	False	-0.72
120762	0.60	0.59	11.99	12.09	False	-0.58
120762	0.60	0.60	11.99	11.99	False	-0.98
120762	0.61	0.60	11.88	11.99	False	-0.57
120762	0.61	0.59	11.88	12.09	False	-0.17
120764	0.73	0.72	10.55	10.67	False	-0.46
120767	0.42	0.42	13.86	13.86	False	-1.04
120769	0.51	0.50	12.90	13.00	False	-0.62
120769	0.51	0.50	12.90	13.00	False	-0.62
120793	0.36	0.32	14.46	14.90	False	0.64
120793	0.36	0.31	14.46	14.96	False	0.86
120803	0.73	0.74	10.55	10.43	False	-1.15
120803	0.73	0.74	10.55	10.43	False	-1.15
120805	0.67	0.69	11.24	11.02	False	-0.73
120805	0.67	0.69	11.24	11.02	False	-0.73
120808	0.42	0.43	13.81	13.71	False	-1.12
120810	0.35	0.39	14.54	14.12	False	0.15
120822	0.70	0.71	10.90	10.79	False	-1.16
120830	0.48	0.50	13.20	13.00	False	-0.76
120830	0.48	0.50	13.20	13.00	False	-0.76
120834	0.40	0.43	14.07	13.68	False	-0.01
120834	0.40	0.41	14.07	13.91	False	-0.91
120838	0.58	0.50	12.19	13.00	False	2.15
120850	0.63	0.60	11.67	11.99	False	0.25
120855	0.60	0.60	11.99	11.99	False	-0.98
124036	0.73	0.75	10.55	10.30	False	-0.67
124036	0.73	0.76	10.55	10.17	False	-0.17
124760	0.73	0.72	10.55	10.67	False	-0.46
124760	0.73	0.75	10.55	10.30	False	-0.67
144684	0.76	0.79	10.17	9.77	False	-0.08
198603	0.74	0.67	10.49	11.30	False	2.20
198645	0.57	0.59	12.29	12.09	False	-0.77
198645	0.57	0.56	12.29	12.40	False	-0.59
203258	0.87	0.87	8.49	8.49	False	-0.86
203258	0.87	0.88	8.49	8.30	False	-0.93
203258	0.87	0.88	8.49	8.30	False	-0.93
203258	0.87	0.87	8.49	8.49	False	-0.86
203301	0.43	0.40	13.71	14.01	False	0.16
203556	0.62	0.62	11.78	11.78	False	-0.97
203588	0.60	0.59	11.99	12.09	False	-0.58
203621	0.32	0.35	14.87	14.60	False	-0.41
203621	0.32	0.32	14.87	14.87	False	-1.07
203621	0.30	0.35	15.16	14.60	False	0.70
203621	0.30	0.32	15.16	14.87	False	-0.37
225021	0.22	0.26	16.09	15.57	False	0.56
225021	0.22	0.26	16.09	15.57	False	0.56
225316	0.53	0.58	12.70	12.19	False	0.41
225389	0.33	0.31	14.82	14.98	False	-0.42

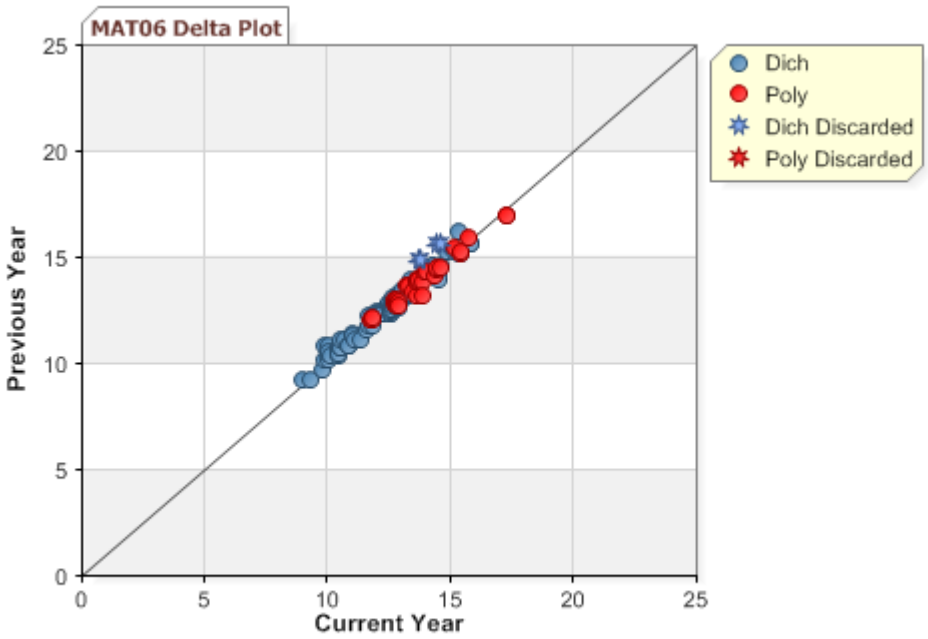
IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
225389	0.33	0.32	14.82	14.87	False	-0.85
225389	0.32	0.32	14.87	14.87	False	-1.07
225389	0.32	0.31	14.87	14.98	False	-0.63
226748	0.51	0.55	12.90	12.50	False	0.01
226810	0.29	0.29	15.21	15.21	False	-1.08
228544	0.20	0.25	16.44	15.76	False	1.20
230748	0.26	0.27	15.60	15.42	False	-0.75
230748	0.25	0.27	15.76	15.42	False	-0.13
255145	0.38	0.46	14.22	13.40	False	1.68
255145	0.38	0.47	14.22	13.30	False	2.07
255255	0.42	0.40	13.81	14.07	False	-0.04
255255	0.42	0.37	13.81	14.38	False	1.18
255255	0.41	0.40	13.91	14.07	False	-0.44
255255	0.41	0.37	13.91	14.38	False	0.78
255761	0.75	0.77	10.30	10.04	False	-0.63
255761	0.75	0.74	10.30	10.43	False	-0.44
255763	0.47	0.50	13.30	13.00	False	-0.36
255763	0.47	0.47	13.30	13.30	False	-1.02
255763	0.46	0.50	13.40	13.00	False	0.03
255763	0.46	0.47	13.40	13.30	False	-1.14
255766	0.53	0.50	12.75	13.00	False	-0.03
269314	0.42	0.42	13.78	13.78	False	-1.03
269369	0.36	0.44	14.43	13.60	False	1.72
269369	0.36	0.44	14.43	13.60	False	1.72
269369	0.36	0.44	14.43	13.60	False	1.72
269369	0.36	0.44	14.43	13.60	False	1.72
269393	0.64	0.60	11.57	11.99	False	0.67
269393	0.64	0.58	11.57	12.19	False	1.46
269393	0.64	0.60	11.57	11.99	False	0.67
269393	0.64	0.58	11.57	12.19	False	1.46
269405	0.53	0.53	12.70	12.75	False	-0.80
272113	0.47	0.50	13.30	13.05	False	-0.56
272113	0.48	0.50	13.20	13.05	False	-0.95

MAT05

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
120834	4	1.59	1.69	1.36	1.36	0.07	No
120639	2	1.35	1.35	0.88	0.89	0.00	No
255255	2	0.81	0.80	0.82	0.78	-0.01	No
120793	4	1.51	1.52	1.43	1.43	0.01	No
120704	2	1.14	1.17	0.85	0.87	0.04	No
225389	2	0.67	0.72	0.85	0.82	0.06	No

MAT05

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
120731	4	2.03	2.13	1.39	1.37	0.07	No
203621	2	0.71	0.72	0.85	0.84	0.01	No
120694	2	0.98	1.00	0.87	0.88	0.02	No
228544	2	0.40	0.50	0.72	0.74	0.14	No
255766	2	1.00	1.00	0.93	0.92	-0.01	No
120767	4	1.60	1.59	1.25	1.29	-0.01	No
198603	2	1.49	1.51	0.71	0.72	0.03	No
269405	2	1.10	1.14	0.93	0.92	0.04	No
269314	4	1.70	1.71	1.54	1.55	0.01	No
120631	2	1.10	1.08	0.92	0.95	-0.02	No
272113	2	0.94	0.95	0.79	0.79	0.01	No
230748	4	1.10	1.13	1.36	1.42	0.02	No



MAT06

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
119175	0.62	0.61	11.78	11.88	False	-0.05
119175	0.60	0.61	11.99	11.88	False	-1.01
119184	0.80	0.79	9.63	9.77	False	0.32
119186	0.83	0.84	9.18	9.02	False	-0.98
119186	0.83	0.82	9.18	9.34	False	0.43
119200	0.68	0.68	11.13	11.13	False	-0.46
119200	0.68	0.66	11.13	11.35	False	0.53
119203	0.52	0.52	12.85	12.80	False	-0.85
119203	0.54	0.52	12.65	12.80	False	0.07
119203	0.54	0.51	12.65	12.90	False	0.52
119203	0.52	0.51	12.85	12.90	False	-0.40
119219	0.64	0.64	11.57	11.57	False	-0.50
119229	0.59	0.61	12.14	11.88	False	-0.46
119232	0.75	0.74	10.30	10.43	False	0.18
119232	0.75	0.76	10.30	10.17	False	-0.94
119238	0.49	0.53	13.10	12.70	False	0.27
119238	0.49	0.51	13.10	12.90	False	-0.63
119258	0.52	0.55	12.80	12.50	False	-0.20
119258	0.52	0.55	12.80	12.50	False	-0.20
119262	0.54	0.52	12.60	12.80	False	0.30
119262	0.54	0.51	12.60	12.90	False	0.75
119277	0.71	0.70	10.79	10.90	False	0.10
119277	0.71	0.70	10.79	10.90	False	0.10
119281	0.60	0.62	11.99	11.78	False	-0.70
119281	0.60	0.62	11.99	11.83	False	-0.93
119313	0.46	0.49	13.40	13.10	False	-0.14
119316	0.58	0.63	12.19	11.67	False	0.71
119316	0.58	0.61	12.19	11.88	False	-0.23

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
119319	0.71	0.70	10.79	10.90	False	0.10
119328	0.29	0.32	15.21	14.87	False	0.22
119328	0.29	0.29	15.21	15.21	False	-0.86
119332	0.67	0.69	11.24	11.02	False	-0.71
119341	0.38	0.40	14.27	14.01	False	-0.24
119341	0.38	0.41	14.27	13.96	False	-0.01
119349	0.41	0.46	13.91	13.40	False	0.83
119351	0.66	0.69	11.35	11.02	False	-0.20
119356	0.21	0.28	16.23	15.33	False	2.79
119364	0.30	0.28	15.10	15.33	False	0.20
119364	0.30	0.27	15.10	15.45	False	0.73
119368	0.57	0.58	12.29	12.19	False	-1.03
119368	0.57	0.58	12.29	12.19	False	-1.03
119370	0.49	0.44	13.15	13.60	False	1.37
119370	0.49	0.42	13.15	13.86	False	2.51
119377	0.53	0.55	12.70	12.50	False	-0.66
119377	0.53	0.53	12.70	12.70	False	-0.61
119379	0.35	0.37	14.54	14.33	False	-0.42
119379	0.35	0.37	14.54	14.33	False	-0.42
119393	0.47	0.51	13.30	12.90	False	0.29
119393	0.48	0.51	13.20	12.90	False	-0.17
119393	0.47	0.47	13.30	13.30	False	-0.67
119393	0.48	0.47	13.20	13.30	False	-0.21
119396	0.55	0.53	12.50	12.70	False	0.31
119396	0.53	0.53	12.70	12.70	False	-0.61
119396	0.53	0.54	12.70	12.60	False	-1.06
119396	0.55	0.54	12.50	12.60	False	-0.14
122249	0.51	0.50	12.90	13.00	False	-0.18
122249	0.51	0.53	12.90	12.70	False	-0.64
123448	0.36	0.36	14.49	14.49	False	-0.79
123448	0.36	0.34	14.49	14.65	False	-0.06
123513	0.26	0.34	15.57	14.65	False	2.85
123513	0.26	0.36	15.57	14.43	True	3.82
125111	0.47	0.46	13.33	13.45	False	-0.11
139322	0.23	0.24	15.92	15.79	False	-0.66
139351	0.42	0.42	13.86	13.86	False	-0.73
198608	0.72	0.74	10.67	10.43	False	-0.68
198608	0.72	0.74	10.67	10.43	False	-0.68
198608	0.72	0.73	10.67	10.55	False	-0.95
198608	0.72	0.73	10.67	10.55	False	-0.95
198622	0.56	0.54	12.40	12.60	False	0.32
198709	0.72	0.73	10.67	10.55	False	-0.95
198715	0.35	0.34	14.54	14.65	False	-0.31
198727	0.30	0.28	15.16	15.39	False	0.20
198727	0.29	0.28	15.27	15.39	False	-0.34
203255	0.40	0.37	14.07	14.38	False	0.66
203460	0.41	0.35	13.91	14.54	False	2.09
203460	0.38	0.35	14.22	14.54	False	0.67
203483	0.76	0.78	10.17	9.91	False	-0.63
203483	0.76	0.77	10.17	10.04	False	-0.94

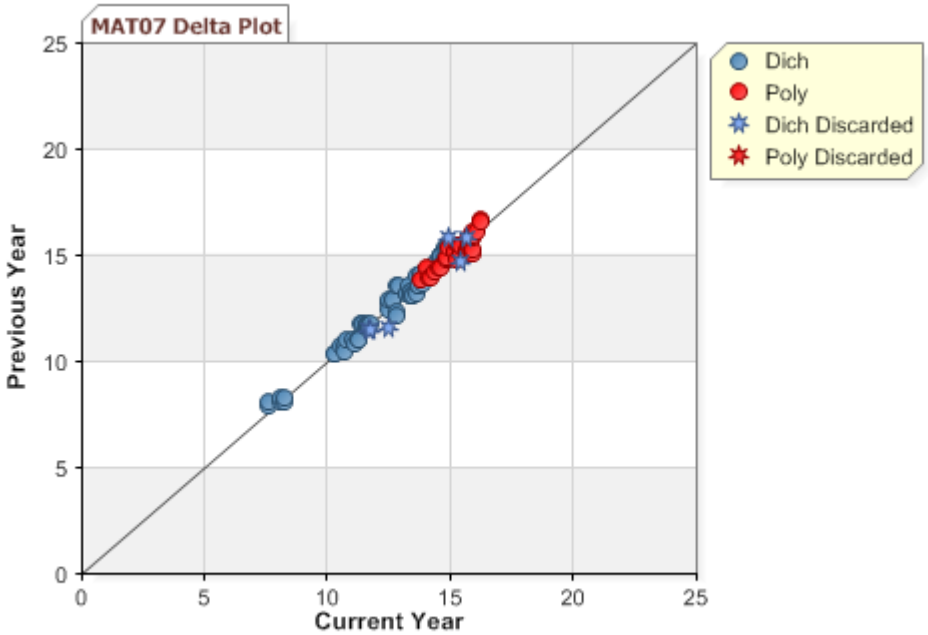
IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
203483	0.71	0.78	10.79	9.91	False	2.16
203483	0.71	0.77	10.79	10.04	False	1.57
203632	0.42	0.44	13.81	13.65	False	-0.77
203632	0.42	0.43	13.81	13.73	False	-1.06
203632	0.41	0.44	13.94	13.65	False	-0.18
203632	0.41	0.43	13.94	13.73	False	-0.52
225273	0.43	0.43	13.71	13.71	False	-0.71
225273	0.43	0.43	13.71	13.71	False	-0.71
225287	0.44	0.48	13.60	13.20	False	0.33
225287	0.44	0.47	13.60	13.35	False	-0.34
225313	0.57	0.60	12.29	11.99	False	-0.23
225313	0.57	0.58	12.29	12.19	False	-1.03
225313	0.56	0.60	12.40	11.99	False	0.24
225313	0.56	0.58	12.40	12.19	False	-0.68
225370	0.16	0.14	16.98	17.32	False	0.50
225370	0.17	0.14	16.90	17.32	False	0.87
225428	0.32	0.43	14.87	13.71	True	3.86
225428	0.32	0.42	14.87	13.81	True	3.41
228068	0.68	0.72	11.13	10.67	False	0.34
228068	0.68	0.73	11.13	10.55	False	0.88
228071	0.74	0.77	10.43	10.04	False	-0.08
228071	0.74	0.74	10.43	10.43	False	-0.39
228071	0.73	0.77	10.55	10.04	False	0.48
228071	0.73	0.74	10.55	10.43	False	-0.94
255369	0.45	0.44	13.50	13.60	False	-0.24
255371	0.57	0.57	12.29	12.29	False	-0.57
255371	0.57	0.54	12.29	12.60	False	0.79
255371	0.57	0.57	12.29	12.29	False	-0.57
255371	0.57	0.54	12.29	12.60	False	0.79
255384	0.52	0.53	12.80	12.75	False	-0.85
255384	0.52	0.52	12.80	12.85	False	-0.40
255384	0.50	0.53	13.00	12.75	False	-0.41
255384	0.50	0.52	13.00	12.85	False	-0.86
255575	0.28	0.29	15.39	15.21	False	-0.50
256902	0.37	0.36	14.38	14.43	False	-0.54
270370	0.62	0.63	11.78	11.67	False	-0.99
270584	0.26	0.24	15.57	15.83	False	0.23
270584	0.26	0.24	15.57	15.83	False	0.23

MAT06

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
119370	2	1.00	0.97	0.95	0.97	-0.04	No
119341	2	0.82	0.89	0.73	0.78	0.10	No
119203	2	0.98	0.97	0.72	0.71	-0.01	No
123448	4	1.45	1.46	1.20	1.17	0.01	No
255384	2	0.98	0.99	0.77	0.78	0.02	No
225287	2	0.92	0.93	0.92	0.93	0.01	No

MAT06

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
203632	4	1.76	1.78	1.17	1.17	0.02	No
119281	2	1.25	1.25	0.69	0.70	0.00	No
198727	2	0.63	0.61	0.80	0.80	-0.02	No
255575	2	0.65	0.69	0.78	0.81	0.05	No
139351	2	0.73	0.74	0.81	0.84	0.02	No
203255	2	0.79	0.84	0.80	0.81	0.07	No
139322	4	0.92	0.89	1.11	1.06	-0.03	No
119229	2	1.15	1.14	0.86	0.85	-0.01	No
225370	2	0.34	0.33	0.63	0.65	-0.02	No
256902	4	1.64	1.60	0.91	0.96	-0.04	No



MAT07

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
120328	0.64	0.63	11.57	11.67	False	-0.69
120329	0.65	0.62	11.46	11.78	False	0.34
120329	0.65	0.63	11.46	11.67	False	-0.17
120334	0.62	0.66	11.78	11.35	False	0.93
120355	0.41	0.39	13.88	14.14	False	-0.02

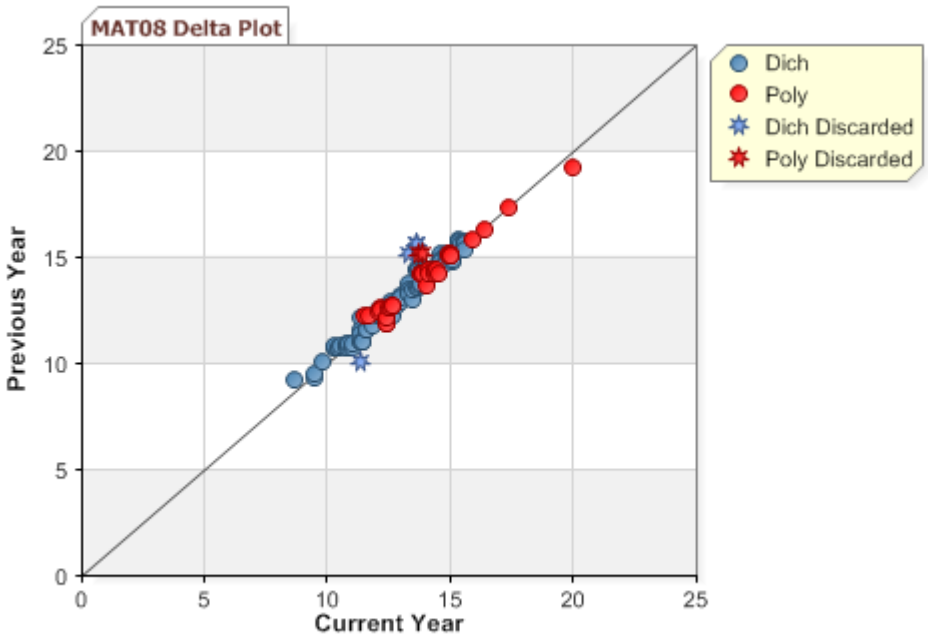
IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
120355	0.41	0.38	13.88	14.22	False	0.36
120366	0.45	0.47	13.50	13.30	False	-0.12
120380	0.48	0.44	13.20	13.60	False	0.70
120380	0.48	0.45	13.20	13.50	False	0.21
120385	0.40	0.40	14.01	14.01	False	-1.08
120385	0.39	0.42	14.12	13.81	False	0.42
120385	0.40	0.42	14.01	13.81	False	-0.08
120385	0.39	0.40	14.12	14.01	False	-0.57
120393	0.38	0.36	14.22	14.41	False	-0.39
120409	0.49	0.45	13.10	13.50	False	0.69
120411	0.89	0.91	8.09	7.64	False	0.96
120411	0.90	0.91	7.87	7.64	False	-0.11
120417	0.36	0.40	14.43	14.07	False	0.72
120434	0.59	0.52	12.09	12.80	False	2.21
120434	0.57	0.52	12.29	12.80	False	1.21
120435	0.63	0.63	11.67	11.67	False	-1.14
120444	0.62	0.65	11.78	11.46	False	0.40
120455	0.56	0.55	12.40	12.50	False	-0.74
120479	0.72	0.72	10.67	10.67	False	-1.17
120479	0.72	0.73	10.67	10.55	False	-0.59
120482	0.40	0.44	14.01	13.60	False	0.90
120487	0.63	0.64	11.67	11.57	False	-0.63
120487	0.63	0.64	11.67	11.57	False	-0.63
120492	0.36	0.35	14.43	14.54	False	-0.77
120492	0.36	0.35	14.43	14.60	False	-0.51
120501	0.74	0.72	10.43	10.67	False	0.00
120504	0.22	0.23	16.12	15.92	False	-0.05
120504	0.22	0.22	16.12	16.12	False	-1.02
120511	0.75	0.75	10.30	10.30	False	-1.17
120511	0.75	0.75	10.30	10.30	False	-1.17
120517	0.37	0.37	14.33	14.33	False	-1.07
120517	0.37	0.37	14.33	14.33	False	-1.07
120517	0.36	0.37	14.43	14.33	False	-0.55
120517	0.36	0.37	14.43	14.33	False	-0.55
120523	0.33	0.29	14.76	15.27	False	1.18
122346	0.64	0.64	11.57	11.57	False	-1.15
122346	0.64	0.63	11.57	11.67	False	-0.69
125286	0.71	0.68	10.79	11.13	False	0.47
125349	0.45	0.43	13.50	13.71	False	-0.28
125349	0.45	0.43	13.50	13.71	False	-0.28
139994	0.51	0.53	12.90	12.70	False	-0.14
139994	0.51	0.53	12.90	12.70	False	-0.14
140025	0.24	0.31	15.83	14.98	True	3.05
140025	0.24	0.25	15.83	15.70	True	-0.41
140026	0.32	0.34	14.87	14.65	False	0.02
140026	0.32	0.33	14.87	14.76	False	-0.51
145151	0.44	0.41	13.60	13.91	False	0.22
146128	0.25	0.24	15.76	15.89	False	-0.71
146128	0.25	0.24	15.76	15.83	False	-1.02
146220	0.33	0.31	14.76	14.98	False	-0.22

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
146263	0.28	0.32	15.33	14.87	False	1.19
146263	0.28	0.33	15.33	14.76	False	1.73
146266	0.53	0.55	12.70	12.50	False	-0.14
150893	0.25	0.24	15.76	15.83	False	-1.02
150893	0.25	0.25	15.76	15.76	False	-1.03
154762	0.42	0.42	13.86	13.81	False	-0.83
154775	0.64	0.55	11.57	12.50	True	3.29
199804	0.69	0.69	11.02	11.02	False	-1.16
199804	0.69	0.67	11.02	11.24	False	-0.11
199900	0.31	0.34	14.98	14.65	False	0.56
199900	0.31	0.34	14.98	14.65	False	0.56
199900	0.31	0.34	14.98	14.65	False	0.56
199900	0.31	0.34	14.98	14.65	False	0.56
199905	0.36	0.35	14.43	14.54	False	-0.77
199932	0.31	0.28	15.04	15.39	False	0.39
206092	0.45	0.52	13.50	12.80	False	2.31
206092	0.45	0.51	13.50	12.90	False	1.82
206107	0.28	0.31	15.33	14.98	False	0.64
206107	0.27	0.31	15.45	14.98	False	1.23
206112	0.34	0.36	14.65	14.43	False	-0.01
206112	0.34	0.36	14.65	14.43	False	-0.01
206144	0.62	0.64	11.78	11.57	False	-0.12
206144	0.62	0.62	11.78	11.78	False	-1.14
206198	0.19	0.21	16.55	16.23	False	0.56
206198	0.18	0.21	16.62	16.23	False	0.92
224764	0.88	0.89	8.30	8.09	False	-0.24
224764	0.88	0.88	8.30	8.30	False	-1.11
224764	0.89	0.88	8.09	8.30	False	-0.11
224764	0.89	0.89	8.09	8.09	False	-1.11
224796	0.49	0.46	13.10	13.40	False	0.21
224796	0.47	0.46	13.30	13.40	False	-0.77
224796	0.47	0.46	13.30	13.40	False	-0.77
224796	0.49	0.46	13.10	13.40	False	0.21
224856	0.31	0.25	15.04	15.70	False	1.87
224856	0.31	0.24	15.04	15.89	False	2.80
224856	0.29	0.25	15.27	15.70	False	0.74
224856	0.29	0.24	15.27	15.89	False	1.67
225135	0.26	0.25	15.57	15.76	False	-0.41
225135	0.26	0.24	15.57	15.83	False	-0.10
225135	0.28	0.25	15.33	15.76	False	0.76
225135	0.28	0.24	15.33	15.83	False	1.07
234445	0.48	0.48	13.20	13.20	False	-1.10
234445	0.48	0.48	13.20	13.20	False	-1.10
234455	0.28	0.31	15.33	14.98	False	0.64
234458	0.69	0.71	11.02	10.79	False	-0.05
234459	0.34	0.27	14.65	15.45	False	2.58
255899	0.27	0.28	15.45	15.33	False	-0.46
255899	0.27	0.27	15.45	15.45	False	-1.03
255899	0.28	0.28	15.39	15.33	False	-0.75
255899	0.28	0.27	15.39	15.45	False	-1.03

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
256095	0.32	0.32	14.87	14.87	False	-1.05
256095	0.33	0.32	14.76	14.87	False	-0.76
256118	0.29	0.29	15.18	15.18	False	-1.04
256118	0.29	0.29	15.18	15.21	False	-1.17
256141	0.51	0.55	12.90	12.50	False	0.84
256152	0.30	0.27	15.10	15.45	False	0.40
256152	0.30	0.28	15.10	15.33	False	-0.18
269013	0.68	0.67	11.13	11.24	False	-0.66
269013	0.69	0.67	11.02	11.24	False	-0.11
269083	0.24	0.23	15.89	15.96	False	-1.01
269118	0.37	0.37	14.33	14.33	False	-1.07

MAT07

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
120492	2	0.75	0.79	0.89	0.88	0.04	No
146128	2	0.56	0.52	0.72	0.65	-0.06	No
120504	4	0.99	1.00	1.10	1.07	0.00	No
199932	2	0.62	0.62	0.77	0.78	-0.01	No
234455	2	0.55	0.57	0.57	0.57	0.03	No
120393	4	1.60	1.70	1.27	1.25	0.07	No
154762	2	0.90	0.96	0.93	0.95	0.06	No
120417	2	0.75	0.77	0.82	0.82	0.02	No
206198	4	0.63	0.60	0.88	0.87	-0.04	No
255899	2	0.53	0.52	0.70	0.70	-0.01	No
120355	4	1.71	1.68	1.13	1.15	-0.03	No
224856	2	0.46	0.47	0.81	0.80	0.01	No
256095	2	0.71	0.68	0.58	0.58	-0.06	No
269083	4	0.96	0.99	1.07	1.05	0.02	No
120523	2	0.61	0.54	0.87	0.82	-0.07	No
225135	2	0.62	0.55	0.87	0.85	-0.08	No
256118	4	1.21	1.25	0.93	0.95	0.05	No
150893	2	0.42	0.45	0.77	0.80	0.03	No



MAT08

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
120876	0.56	0.59	12.40	12.09	False	-0.29
120877	0.62	0.61	11.78	11.88	False	-0.29
120886	0.77	0.79	10.04	9.77	False	-0.23
120890	0.58	0.63	12.19	11.67	False	0.39
120890	0.58	0.65	12.19	11.51	False	0.88
120895	0.72	0.71	10.67	10.79	False	-0.33
120895	0.72	0.70	10.67	10.90	False	0.03
120895	0.70	0.70	10.90	10.90	False	-0.67
120895	0.70	0.71	10.90	10.79	False	-0.77
120912	0.66	0.66	11.35	11.35	False	-0.64
120915	0.83	0.86	9.18	8.68	False	0.56
120919	0.64	0.66	11.57	11.35	False	-0.51
120921	0.82	0.81	9.34	9.49	False	-0.32
120921	0.81	0.81	9.49	9.49	False	-0.78
120931	0.46	0.47	13.40	13.30	False	-0.81
120931	0.47	0.47	13.30	13.30	False	-0.50
120931	0.47	0.47	13.30	13.30	False	-0.50
120931	0.46	0.47	13.40	13.30	False	-0.81
120932	0.77	0.66	10.04	11.35	True	3.32
120936	0.55	0.58	12.50	12.19	False	-0.30
120936	0.54	0.58	12.60	12.19	False	0.00
120974	0.53	0.54	12.70	12.60	False	-0.86
120974	0.54	0.54	12.60	12.60	False	-0.55
120976	0.57	0.58	12.29	12.19	False	-0.89
120985	0.44	0.40	13.65	14.07	False	0.80
120998	0.44	0.42	13.60	13.81	False	0.15
120998	0.43	0.42	13.71	13.81	False	-0.15
121002	0.36	0.36	14.43	14.43	False	-0.42

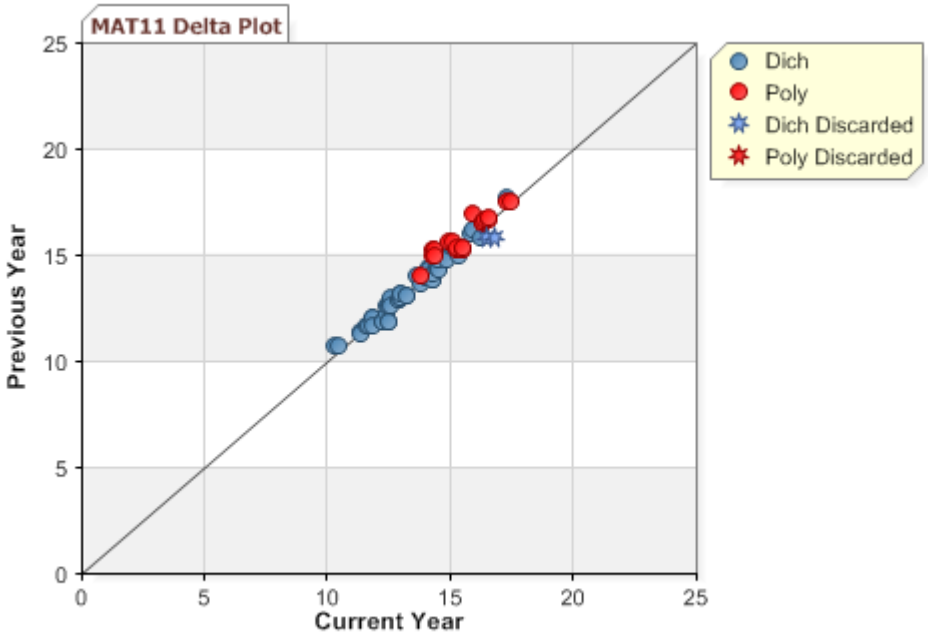
IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
121004	0.53	0.52	12.70	12.80	False	-0.23
121004	0.53	0.54	12.70	12.60	False	-0.86
121004	0.51	0.52	12.90	12.80	False	-0.84
121004	0.51	0.54	12.90	12.60	False	-0.34
121025	0.70	0.69	10.90	11.02	False	-0.32
121025	0.72	0.69	10.67	11.02	False	0.39
121027	0.58	0.59	12.19	12.09	False	-0.90
121027	0.58	0.59	12.19	12.09	False	-0.90
121048	0.25	0.28	15.70	15.33	False	-0.34
121048	0.25	0.26	15.70	15.57	False	-0.71
121071	0.21	0.20	16.30	16.44	False	0.16
121073	0.35	0.34	14.54	14.65	False	-0.07
121073	0.35	0.34	14.54	14.65	False	-0.07
121081	0.58	0.54	12.19	12.60	False	0.68
121081	0.58	0.53	12.19	12.70	False	0.99
121083	0.43	0.47	13.71	13.30	False	-0.08
121085	0.31	0.31	15.04	14.96	False	-0.64
121085	0.30	0.31	15.13	14.96	False	-0.90
121085	0.30	0.31	15.13	15.04	False	-0.63
121085	0.31	0.31	15.04	15.04	False	-0.37
121087	0.36	0.39	14.43	14.12	False	-0.41
121087	0.36	0.38	14.43	14.22	False	-0.73
121089	0.30	0.34	15.10	14.65	False	-0.05
121089	0.30	0.32	15.10	14.87	False	-0.73
121091	0.69	0.66	11.02	11.35	False	0.37
121091	0.69	0.65	11.02	11.46	False	0.71
121091	0.69	0.66	11.02	11.35	False	0.37
121091	0.69	0.65	11.02	11.46	False	0.71
121099	0.24	0.24	15.83	15.89	False	-0.12
122522	0.60	0.62	11.99	11.78	False	-0.56
122522	0.60	0.62	11.99	11.78	False	-0.56
122667	0.45	0.44	13.50	13.60	False	-0.17
122667	0.45	0.43	13.50	13.71	False	0.14
139760	0.26	0.28	15.57	15.33	False	-0.72
139760	0.26	0.28	15.57	15.33	False	-0.72
139771	0.71	0.75	10.79	10.30	False	0.38
139771	0.71	0.73	10.79	10.55	False	-0.39
139777	0.31	0.31	15.04	15.04	False	-0.37
139777	0.31	0.31	15.04	15.04	False	-0.37
139782	0.37	0.39	14.33	14.12	False	-0.73
139837	0.59	0.66	12.09	11.35	False	1.08
139837	0.59	0.64	12.09	11.57	False	0.41
139869	0.30	0.43	15.16	13.71	True	3.06
139869	0.30	0.42	15.16	13.86	True	2.58
139880	0.28	0.26	15.33	15.57	False	0.40
139885	0.55	0.55	12.55	12.50	False	-0.71
139885	0.55	0.54	12.55	12.60	False	-0.40
145640	0.26	0.26	15.57	15.57	False	-0.33
145640	0.26	0.27	15.57	15.45	False	-0.71
146287	0.34	0.33	14.65	14.76	False	-0.06

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
146287	0.34	0.35	14.65	14.54	False	-0.74
146358	0.40	0.39	14.01	14.12	False	-0.13
146455	0.50	0.45	13.00	13.50	False	1.04
199731	0.46	0.45	13.40	13.50	False	-0.18
199756	0.64	0.64	11.57	11.57	False	-0.63
199756	0.64	0.64	11.57	11.57	False	-0.63
199768	0.26	0.44	15.57	13.60	True	4.64
199768	0.26	0.44	15.57	13.60	True	4.64
206256	0.30	0.47	15.10	13.30	True	4.14
206323	0.49	0.50	13.10	13.00	False	-0.83
206323	0.48	0.50	13.20	13.00	False	-0.68
206323	0.51	0.50	12.90	13.00	False	-0.22
206324	0.61	0.56	11.88	12.45	False	1.15
206324	0.59	0.56	12.14	12.40	False	0.21
206324	0.59	0.56	12.14	12.45	False	0.36
206324	0.61	0.56	11.88	12.40	False	0.99
206331	0.53	0.54	12.67	12.62	False	-0.70
206331	0.53	0.54	12.67	12.62	False	-0.70
224873	0.34	0.34	14.65	14.65	False	-0.40
224873	0.34	0.35	14.65	14.54	False	-0.74
224873	0.34	0.34	14.65	14.65	False	-0.40
224873	0.34	0.35	14.65	14.54	False	-0.74
224878	0.36	0.43	14.43	13.71	False	0.87
224878	0.36	0.44	14.43	13.60	False	1.19
224878	0.36	0.43	14.43	13.71	False	0.87
224878	0.36	0.44	14.43	13.60	False	1.19
224936	0.37	0.35	14.33	14.54	False	0.24
224947	0.39	0.42	14.17	13.81	False	-0.25
224947	0.39	0.42	14.17	13.86	False	-0.41
224947	0.39	0.42	14.17	13.81	False	-0.25
224947	0.39	0.42	14.17	13.86	False	-0.41
225437	0.24	0.28	15.83	15.33	False	0.04
225437	0.25	0.28	15.70	15.33	False	-0.34
233602	0.57	0.59	12.29	12.09	False	-0.60
233602	0.57	0.57	12.29	12.29	False	-0.57
233602	0.55	0.59	12.50	12.09	False	0.02
233602	0.55	0.57	12.50	12.29	False	-0.62
233737	0.26	0.26	15.57	15.57	False	-0.33
242392	0.36	0.41	14.43	13.91	False	0.24
242392	0.36	0.41	14.43	13.91	False	0.24
242392	0.37	0.41	14.33	13.91	False	-0.09
242392	0.37	0.41	14.33	13.91	False	-0.09
242467	0.33	0.30	14.76	15.10	False	0.65
242467	0.33	0.34	14.76	14.65	False	-0.73
242467	0.32	0.30	14.87	15.10	False	0.32
242467	0.32	0.34	14.87	14.65	False	-0.74
256121	0.38	0.39	14.22	14.12	False	-0.76
256121	0.38	0.39	14.22	14.12	False	-0.76
256320	0.14	0.14	17.32	17.41	False	0.07
256329	0.37	0.39	14.35	14.12	False	-0.65

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
256329	0.37	0.37	14.35	14.38	False	-0.34
256329	0.38	0.39	14.20	14.12	False	-0.68
256329	0.38	0.37	14.20	14.38	False	0.14
256511	0.72	0.75	10.67	10.30	False	0.02
256511	0.72	0.74	10.67	10.43	False	-0.37
256511	0.72	0.75	10.67	10.30	False	0.02
256511	0.72	0.74	10.67	10.43	False	-0.37
269098	0.38	0.35	14.22	14.54	False	0.56
269172	0.06	0.04	19.22	20.00	False	2.36

MAT08

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
206324	2	1.34	1.27	0.80	0.80	-0.09	No
139885	2	1.13	1.13	0.85	0.87	0.00	No
256329	4	1.52	1.48	1.37	1.38	-0.03	No
120985	2	0.91	0.85	0.95	0.92	-0.06	No
120936	2	1.13	1.09	0.90	0.93	-0.05	No
121071	4	0.76	0.79	0.91	0.92	0.03	No
269098	2	0.64	0.57	0.89	0.80	-0.09	No
256320	2	0.26	0.23	0.63	0.62	-0.04	No
121099	4	1.13	1.11	0.84	0.82	-0.02	No
139777	2	0.57	0.57	0.87	0.87	0.01	No
206331	4	2.16	2.06	1.12	1.13	-0.08	No
224947	2	0.77	0.80	0.85	0.84	0.04	No
120876	2	1.11	1.13	0.72	0.71	0.03	No
121002	4	1.37	1.39	1.43	1.41	0.01	No
269172	2	0.13	0.13	0.43	0.39	0.00	No
139869	2	0.58	0.58	0.86	0.88	0.00	No
121085	4	1.40	1.33	1.49	1.49	-0.05	No
120890	2	1.15	1.33	0.82	0.77	0.22	No



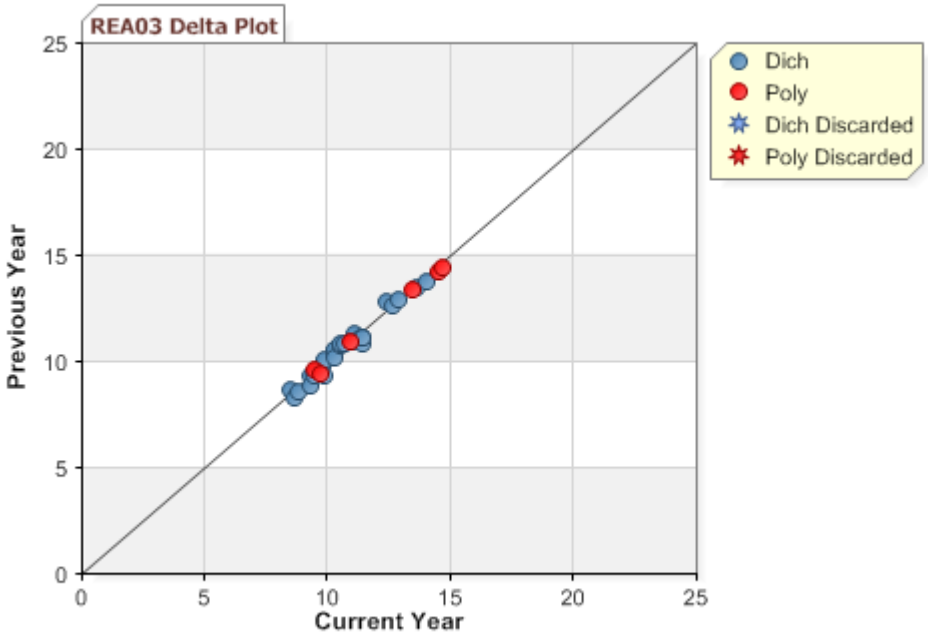
MAT11

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
119430	0.31	0.28	14.98	15.33	False	0.75
119437	0.61	0.55	11.88	12.50	False	1.51
119437	0.61	0.55	11.88	12.50	False	1.51
119437	0.61	0.57	11.88	12.29	False	0.65
119437	0.61	0.57	11.88	12.29	False	0.65
119442	0.32	0.32	14.87	14.87	False	-0.74
119450	0.40	0.40	14.01	14.01	False	-0.84
119452	0.63	0.61	11.67	11.88	False	-0.23
119460	0.19	0.21	16.47	16.30	False	-0.72
119460	0.19	0.21	16.51	16.30	False	-0.57
119464	0.42	0.37	13.81	14.33	False	1.34
119464	0.42	0.37	13.81	14.33	False	1.34
119472	0.29	0.38	15.21	14.27	False	2.65
119472	0.31	0.37	14.98	14.38	False	1.26
119472	0.31	0.38	14.98	14.27	False	1.71
119472	0.29	0.37	15.21	14.38	False	2.21
119476	0.36	0.36	14.43	14.43	False	-0.79
119484	0.50	0.54	13.00	12.60	False	0.64
119484	0.52	0.54	12.80	12.60	False	-0.19
119488	0.67	0.66	11.24	11.35	False	-0.71
119488	0.66	0.66	11.35	11.35	False	-0.87
119494	0.24	0.19	15.83	16.51	False	2.28
119494	0.24	0.17	15.83	16.82	True	3.58
119502	0.26	0.31	15.57	14.98	False	1.13
119502	0.26	0.30	15.57	15.10	False	0.65
119516	0.49	0.48	13.10	13.20	False	-0.53
119516	0.49	0.48	13.10	13.20	False	-0.53
119520	0.54	0.56	12.60	12.40	False	-0.16

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
119542	0.39	0.37	14.12	14.33	False	0.06
119546	0.40	0.42	13.99	13.78	False	-0.31
119553	0.37	0.35	14.33	14.54	False	0.10
119553	0.36	0.35	14.43	14.54	False	-0.34
119562	0.28	0.27	15.33	15.45	False	-0.18
119573	0.72	0.74	10.67	10.43	False	0.24
119573	0.72	0.75	10.67	10.30	False	0.77
119589	0.13	0.14	17.51	17.28	False	-0.63
119603	0.21	0.23	16.23	15.96	False	-0.30
119605	0.54	0.54	12.60	12.60	False	-1.01
119605	0.54	0.55	12.60	12.50	False	-0.59
119620	0.17	0.24	16.90	15.89	False	2.74
119621	0.29	0.27	15.24	15.48	False	0.32
119621	0.29	0.29	15.24	15.24	False	-0.70
119621	0.28	0.27	15.33	15.48	False	-0.05
119621	0.28	0.29	15.33	15.24	False	-0.97
119630	0.60	0.61	11.99	11.88	False	-0.50
119630	0.60	0.61	11.99	11.88	False	-0.50
125740	0.12	0.14	17.70	17.32	False	-0.02
130081	0.63	0.63	11.67	11.67	False	-0.90
130081	0.63	0.64	11.67	11.57	False	-0.45
139984	0.28	0.30	15.33	15.10	False	-0.35
140196	0.58	0.56	12.19	12.40	False	-0.20
141256	0.51	0.51	12.90	12.90	False	-0.97
141256	0.51	0.51	12.90	12.90	False	-0.97
259790	0.23	0.24	15.96	15.83	False	-0.86
259814	0.28	0.29	15.33	15.21	False	-0.84
259828	0.50	0.50	13.00	13.00	False	-0.96
259828	0.50	0.50	13.00	13.00	False	-0.96
259828	0.48	0.50	13.20	13.00	False	-0.24
259828	0.48	0.50	13.20	13.00	False	-0.24
259873	0.37	0.39	14.33	14.12	False	-0.33
259914	0.41	0.40	13.91	14.01	False	-0.42
259914	0.40	0.40	14.01	14.01	False	-0.84
259914	0.40	0.40	14.01	14.01	False	-0.84
259914	0.41	0.40	13.91	14.01	False	-0.42
259921	0.18	0.19	16.74	16.59	False	-0.86
259927	0.36	0.39	14.43	14.12	False	0.11
259947	0.40	0.44	14.01	13.60	False	0.55
259958	0.18	0.20	16.62	16.40	False	-0.56
259989	0.35	0.35	14.54	14.54	False	-0.78
259989	0.35	0.35	14.54	14.54	False	-0.78
259989	0.33	0.35	14.76	14.54	False	-0.35
259989	0.33	0.35	14.76	14.54	False	-0.35
260001	0.18	0.19	16.66	16.59	False	-0.85
260675	0.13	0.13	17.51	17.51	False	-0.42
270801	0.24	0.21	15.83	16.23	False	1.07
270853	0.39	0.39	14.12	14.12	False	-0.83
270856	0.33	0.32	14.76	14.87	False	-0.28
270868	0.44	0.42	13.60	13.81	False	-0.03

MAT11

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
259921	2	0.37	0.37	0.70	0.70	0.01	No
119546	4	1.65	1.60	1.60	1.63	-0.03	No
260675	2	0.29	0.29	0.63	0.64	0.00	No
259958	4	0.86	0.82	1.29	1.28	-0.03	No
119502	2	0.63	0.63	0.72	0.72	0.01	No
119621	4	1.19	1.19	1.19	1.24	0.00	No
260001	2	0.40	0.40	0.65	0.67	-0.01	No
119620	2	0.41	0.55	0.78	0.77	0.17	No
119589	4	0.55	0.54	0.86	0.83	-0.01	No
119472	2	0.67	0.70	0.72	0.74	0.04	No
119460	4	0.84	0.85	0.92	0.91	0.02	No



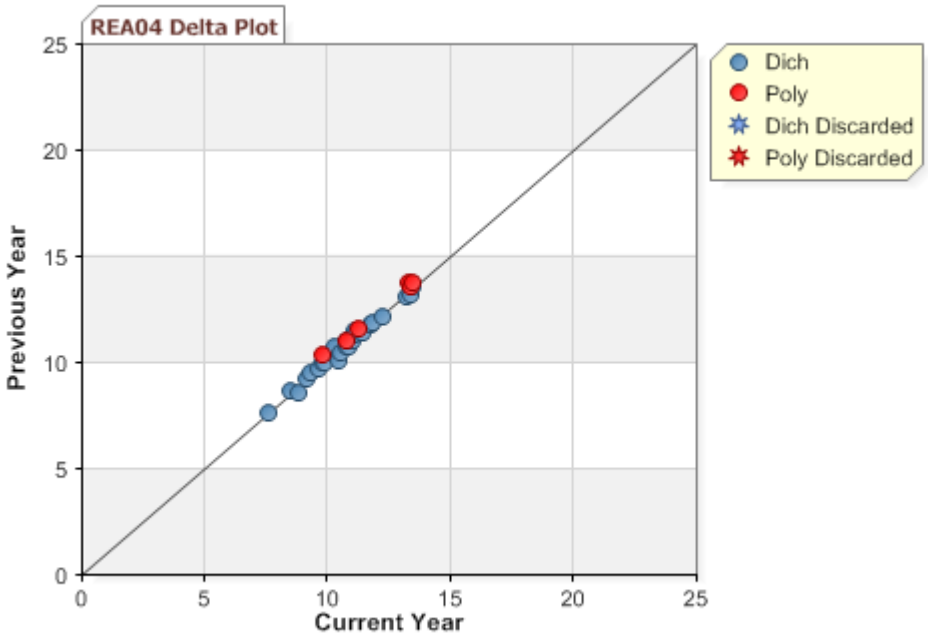
REA03

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
117787	0.68	0.68	11.13	11.13	False	-0.72
117789	0.72	0.73	10.67	10.55	False	0.10
117790	0.74	0.75	10.43	10.30	False	0.12
117793	0.71	0.72	10.79	10.67	False	0.09
117794	0.68	0.66	11.13	11.35	False	-0.57
117796	0.46	0.44	13.40	13.60	False	-0.86
117797	0.54	0.53	12.60	12.70	False	-1.34
117798	0.77	0.78	10.04	9.91	False	0.15
117801	0.70	0.70	10.90	10.96	False	-1.14
117802	0.80	0.81	9.60	9.45	False	0.21
117808	0.75	0.75	10.30	10.30	False	-0.78
147819	0.82	0.78	9.34	9.91	False	2.05
147835	0.67	0.68	11.24	11.13	False	0.07
147847	0.70	0.69	10.90	11.02	False	-1.31
147849	0.43	0.40	13.71	14.01	False	-0.13
147851	0.86	0.87	8.68	8.49	False	0.42
147856	0.71	0.65	10.79	11.46	False	2.65
147865	0.68	0.65	11.13	11.46	False	0.20
147865	0.68	0.65	11.13	11.46	False	0.20
147870	0.74	0.75	10.43	10.30	False	0.12
147870	0.76	0.75	10.17	10.30	False	-1.17
147875	0.82	0.79	9.41	9.74	False	0.29
147889	0.36	0.33	14.41	14.73	False	-0.05
148178	0.79	0.79	9.77	9.77	False	-0.81
148219	0.85	0.82	8.85	9.34	False	1.45
148224	0.88	0.86	8.30	8.68	False	0.74
202178	0.78	0.78	9.91	9.91	False	-0.80
202179	0.87	0.85	8.49	8.85	False	0.59

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
202180	0.73	0.75	10.55	10.30	False	0.99
202183	0.82	0.82	9.34	9.34	False	-0.84
230973	0.47	0.46	13.35	13.45	False	-1.29
270034	0.52	0.56	12.80	12.40	False	2.25
270042	0.51	0.51	12.90	12.90	False	-0.60
270043	0.82	0.81	9.34	9.49	False	-0.95
270062	0.71	0.73	10.79	10.55	False	0.94
270066	0.38	0.35	14.22	14.51	False	-0.27

REA03

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
270066	4	1.52	1.49	0.87	0.84	-0.04	No
117802	4	3.28	3.29	0.75	0.78	0.01	No
117801	4	2.70	2.71	1.06	1.01	0.01	No
230973	4	1.84	1.94	1.12	1.02	0.08	No
147875	4	3.31	3.32	0.85	0.85	0.01	No
147889	4	1.54	1.57	0.96	0.94	0.04	No



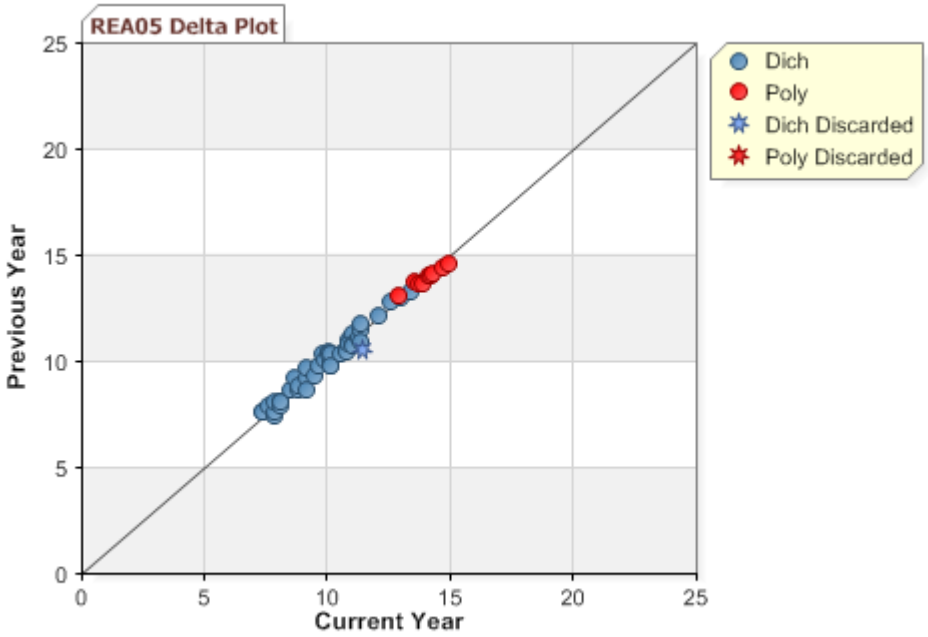
REA04

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
117924	0.69	0.69	11.02	11.02	False	-0.97
117925	0.86	0.87	8.68	8.49	False	-0.05
117926	0.73	0.75	10.55	10.30	False	0.24
117927	0.66	0.65	11.35	11.46	False	-0.04
117931	0.43	0.47	13.73	13.30	False	1.35
148096	0.91	0.91	7.64	7.64	False	-1.38
148097	0.72	0.75	10.67	10.30	False	1.20
148099	0.83	0.83	9.18	9.18	False	-1.19
148109	0.62	0.62	11.78	11.78	False	-0.88
148116	0.64	0.67	11.57	11.27	False	0.53
148398	0.77	0.74	10.04	10.43	False	2.04
148403	0.69	0.71	11.02	10.79	False	0.04
148411	0.81	0.82	9.49	9.34	False	-0.43
203668	0.45	0.45	13.50	13.50	False	-0.67
203670	0.49	0.48	13.10	13.20	False	0.11
203673	0.67	0.68	11.24	11.13	False	-0.96
203675	0.72	0.70	10.67	10.90	False	0.90
203678	0.48	0.46	13.20	13.40	False	0.94
203684	0.44	0.46	13.58	13.43	False	-0.91
205951	0.75	0.79	10.33	9.81	False	2.54
232576	0.87	0.85	8.49	8.85	False	1.67
232579	0.70	0.71	10.90	10.79	False	-0.88
232585	0.74	0.73	10.43	10.55	False	-0.04
270467	0.80	0.80	9.63	9.63	False	-1.14
270468	0.78	0.79	9.91	9.77	False	-0.58
270473	0.61	0.61	11.88	11.88	False	-0.86
270477	0.59	0.57	12.09	12.29	False	0.84

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
270479	0.72	0.71	10.67	10.79	False	-0.05
270485	0.71	0.70	10.79	10.90	False	-0.05
270502	0.78	0.78	9.91	9.91	False	-1.10
270506	0.66	0.66	11.35	11.35	False	-0.93
270511	0.70	0.71	10.96	10.79	False	-0.41
270516	0.43	0.46	13.76	13.45	False	0.32
270923	0.65	0.68	11.46	11.13	False	0.80

REA04

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
117931	4	1.74	1.77	0.85	0.98	0.03	No
205951	4	3.09	3.27	1.09	0.97	0.17	No
203684	4	1.76	1.75	0.99	0.93	-0.01	No
148116	4	2.44	2.71	1.19	1.19	0.23	No
270511	4	2.90	2.89	1.08	1.08	-0.01	No
270516	4	1.62	1.59	0.80	0.85	-0.04	No



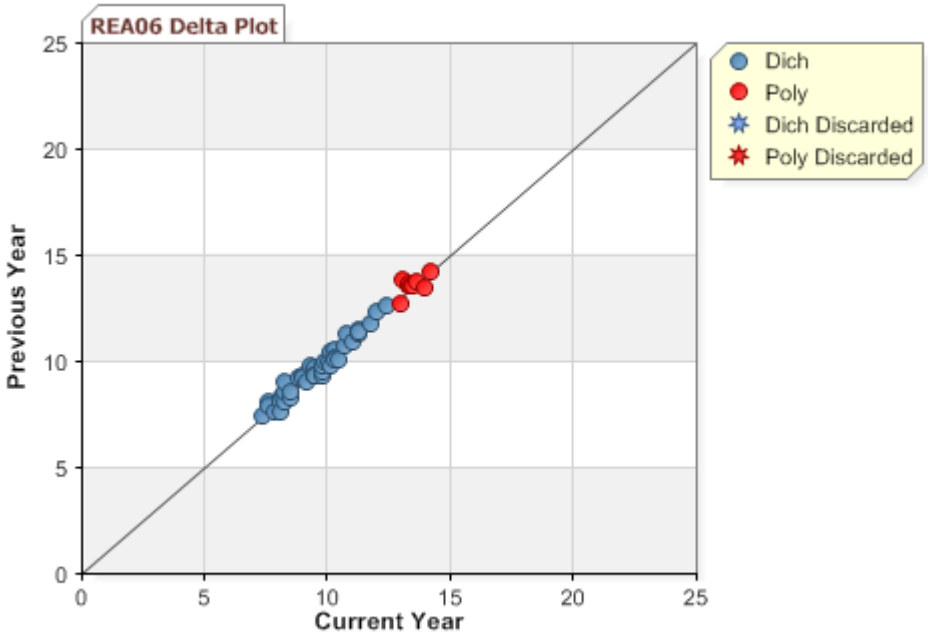
REA05

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
118142	0.75	0.79	10.30	9.77	False	1.93
118144	0.89	0.89	8.09	8.09	False	-1.25
118145	0.77	0.78	10.04	9.91	False	-0.39
118148	0.86	0.87	8.68	8.49	False	-0.31
118151	0.91	0.92	7.64	7.38	False	-0.06
118152	0.91	0.92	7.64	7.38	False	-0.06
118154	0.83	0.86	9.18	8.68	False	1.62
118155	0.79	0.80	9.77	9.63	False	-0.39
118156	0.43	0.44	13.71	13.58	False	0.16
118158	0.40	0.38	13.99	14.20	False	-0.99
118159	0.90	0.91	7.87	7.64	False	-0.14
118160	0.91	0.90	7.64	7.87	False	0.19
118163	0.73	0.71	10.55	10.79	False	-0.27
118165	0.73	0.65	10.55	11.46	True	3.62
118167	0.44	0.43	13.65	13.73	False	-1.02
118170	0.89	0.89	8.09	8.09	False	-1.25
118172	0.59	0.59	12.09	12.09	False	-0.83
118175	0.89	0.90	8.09	7.87	False	-0.20
118176	0.71	0.69	10.79	11.02	False	-0.35
118178	0.40	0.39	14.01	14.09	False	-0.97
118199	0.52	0.54	12.80	12.60	False	0.44
118205	0.69	0.70	11.02	10.90	False	-0.34
147820	0.82	0.81	9.34	9.49	False	-0.59
147825	0.75	0.77	10.30	10.04	False	0.37
147825	0.74	0.77	10.43	10.04	False	1.11
147986	0.83	0.83	9.18	9.18	False	-1.30
147991	0.75	0.73	10.30	10.55	False	-0.18
147992	0.80	0.83	9.63	9.18	False	1.38

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
147997	0.71	0.70	10.79	10.90	False	-1.01
147999	0.79	0.76	9.77	10.17	False	0.79
147999	0.79	0.76	9.77	10.17	False	0.79
148008	0.62	0.66	11.78	11.35	False	1.59
148137	0.35	0.31	14.54	14.98	False	0.27
148158	0.37	0.34	14.35	14.70	False	-0.23
148595	0.78	0.76	9.91	10.17	False	-0.02
148614	0.74	0.71	10.43	10.79	False	0.46
148616	0.86	0.83	8.68	9.18	False	1.57
148617	0.70	0.66	10.90	11.35	False	0.89
148618	0.86	0.85	8.68	8.85	False	-0.33
148639	0.68	0.67	11.13	11.24	False	-1.10
148667	0.72	0.69	10.67	11.02	False	0.35
148675	0.90	0.89	7.87	8.09	False	0.06
148680	0.40	0.37	14.07	14.33	False	-0.69
148698	0.44	0.41	13.60	13.88	False	-0.51
149086	0.85	0.85	8.85	8.85	False	-1.35
149116	0.89	0.89	8.09	8.09	False	-1.25
149127	0.75	0.76	10.30	10.17	False	-0.38
201746	0.92	0.90	7.38	7.87	False	1.72
201752	0.50	0.50	13.00	13.00	False	-0.69
201757	0.65	0.66	11.46	11.35	False	-0.31
201760	0.47	0.46	13.30	13.40	False	-1.22
201769	0.49	0.51	13.08	12.90	False	0.34
269423	0.67	0.69	11.24	11.02	False	0.33

REA05

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
118167	4	1.67	1.71	0.76	0.75	0.05	No
148680	4	1.54	1.56	0.89	0.92	0.02	No
148698	4	1.84	1.81	0.91	0.89	-0.04	No
201769	4	1.84	2.06	1.06	1.14	0.21	No
148158	4	1.50	1.38	0.89	0.98	-0.14	No
148137	4	1.39	1.35	0.86	0.88	-0.04	No
118178	4	1.69	1.63	0.94	0.97	-0.07	No
118156	4	1.80	1.81	1.12	1.13	0.02	No
118158	4	1.59	1.53	1.03	0.95	-0.06	No



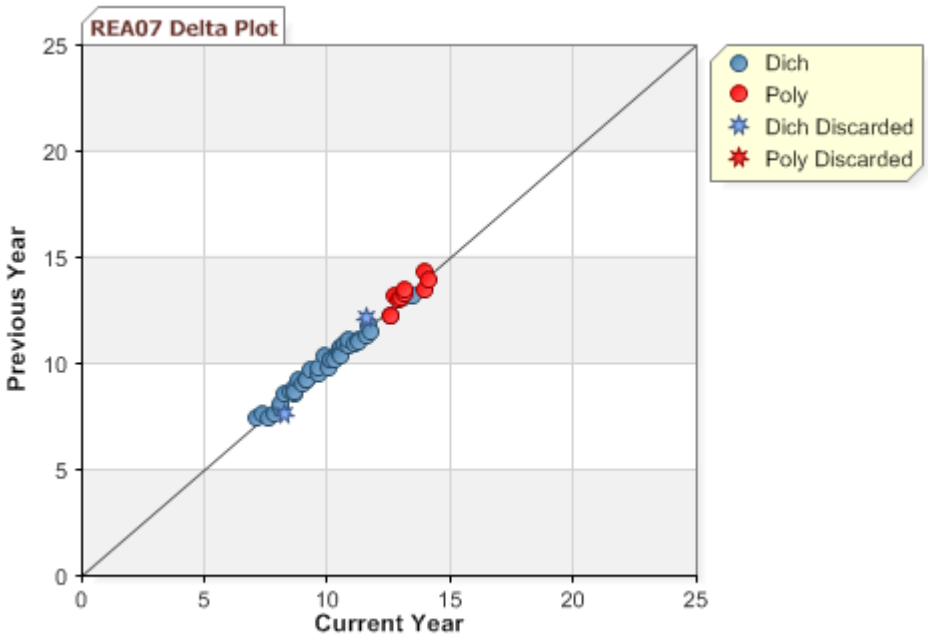
REA06

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
118282	0.79	0.79	9.77	9.77	False	-1.11
118284	0.70	0.69	10.90	11.02	False	-0.34
118285	0.84	0.83	9.02	9.18	False	-0.25
118287	0.81	0.79	9.49	9.77	False	0.52
118289	0.38	0.39	14.20	14.17	False	-0.84
118327	0.72	0.72	10.67	10.67	False	-1.02
118328	0.82	0.81	9.34	9.49	False	-0.28
118330	0.91	0.90	7.64	7.87	False	0.06
118333	0.66	0.67	11.35	11.24	False	-0.91
118335	0.89	0.88	8.09	8.30	False	-0.07
118338	0.79	0.76	9.77	10.17	False	1.21
118339	0.87	0.87	8.49	8.49	False	-1.23
118341	0.82	0.81	9.34	9.49	False	-0.28
118343	0.44	0.47	13.60	13.33	False	-0.15
118344	0.43	0.44	13.68	13.60	False	-1.18
118384	0.74	0.76	10.43	10.17	False	0.00
118387	0.67	0.71	11.24	10.79	False	1.09
118393	0.54	0.56	12.60	12.40	False	-0.50
129078	0.80	0.81	9.63	9.49	False	-0.55
129081	0.77	0.74	10.04	10.43	False	1.13
129082	0.82	0.81	9.34	9.49	False	-0.28
129083	0.88	0.89	8.30	8.09	False	-0.07
129084	0.45	0.46	13.50	13.43	False	-1.20
148388	0.91	0.89	7.64	8.09	False	1.33
148393	0.65	0.67	11.46	11.24	False	-0.29
148399	0.88	0.87	8.30	8.49	False	-0.12
148401	0.76	0.75	10.17	10.30	False	-0.33
148401	0.77	0.75	10.04	10.30	False	0.41

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
148410	0.82	0.84	9.34	9.02	False	0.47
148416	0.57	0.60	12.29	11.99	False	0.14
148425	0.87	0.88	8.49	8.30	False	-0.16
148428	0.82	0.79	9.34	9.77	False	1.37
148440	0.46	0.41	13.45	13.94	False	2.03
148455	0.45	0.45	13.53	13.50	False	-0.90
148549	0.82	0.81	9.34	9.49	False	-0.28
148563	0.90	0.91	7.87	7.64	False	0.15
148566	0.62	0.62	11.78	11.78	False	-0.92
148573	0.78	0.78	9.91	9.91	False	-1.09
148579	0.42	0.50	13.83	13.05	False	2.75
148673	0.89	0.89	8.09	8.09	False	-1.24
148684	0.84	0.88	9.02	8.30	False	2.85
148692	0.83	0.84	9.18	9.02	False	-0.41
148695	0.79	0.82	9.77	9.34	False	1.12
148699	0.73	0.75	10.55	10.30	False	-0.05
148703	0.89	0.91	8.09	7.64	False	1.40
148709	0.83	0.85	9.18	8.85	False	0.56
148711	0.92	0.92	7.38	7.38	False	-1.18
148714	0.44	0.47	13.58	13.28	False	-0.01
148718	0.54	0.50	12.65	13.00	False	1.19
148951	0.78	0.77	9.91	10.04	False	-0.32
148984	0.89	0.89	8.09	8.09	False	-1.24
256674	0.67	0.67	11.24	11.24	False	-0.97

REA06

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
129084	4	1.80	1.83	0.94	0.98	0.04	No
118343	4	1.79	1.92	0.94	1.01	0.14	No
118344	4	1.76	1.73	0.82	0.91	-0.04	No
118289	4	1.49	1.60	0.95	1.00	0.12	No
148455	4	1.78	1.71	0.81	0.88	-0.09	No
148440	4	1.94	1.85	0.87	0.91	-0.10	No
148579	4	1.70	1.83	0.85	0.90	0.15	No
148714	4	1.76	1.67	0.90	0.85	-0.10	No
148718	4	2.08	2.08	0.86	0.88	0.00	No



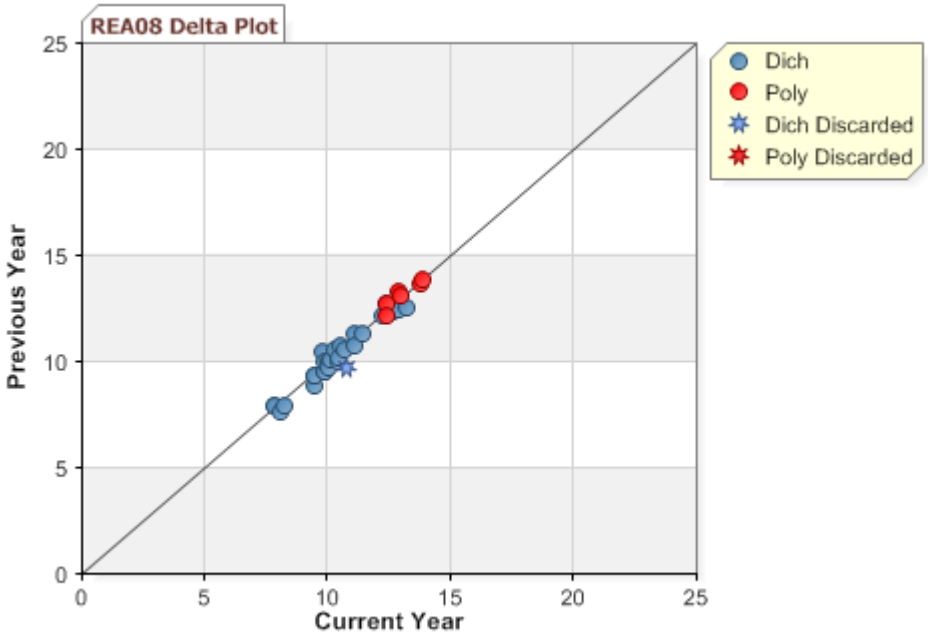
REA07

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
118526	0.75	0.78	10.30	9.91	False	1.31
118528	0.74	0.73	10.43	10.55	False	-0.63
118532	0.46	0.41	13.40	13.96	False	2.32
118572	0.84	0.84	9.02	9.02	False	-1.34
118573	0.79	0.77	9.77	10.04	False	0.38
129210	0.67	0.64	11.24	11.57	False	0.75
129211	0.89	0.89	8.09	8.09	False	-1.34
129212	0.83	0.83	9.18	9.18	False	-1.34
129213	0.71	0.70	10.79	10.90	False	-0.67
129214	0.81	0.80	9.49	9.63	False	-0.47
129215	0.70	0.68	10.90	11.13	False	0.08
129216	0.49	0.49	13.10	13.10	False	-1.32
129217	0.72	0.73	10.67	10.55	False	-0.52
129218	0.41	0.39	13.91	14.09	False	-0.25
129219	0.49	0.50	13.10	12.97	False	-0.47
147154	0.92	0.93	7.38	7.10	False	0.57
147157	0.87	0.88	8.49	8.30	False	-0.03
147161	0.86	0.86	8.68	8.68	False	-1.34
147162	0.72	0.73	10.67	10.55	False	-0.52
147164	0.86	0.87	8.68	8.49	False	-0.09
147171	0.83	0.85	9.18	8.85	False	0.89
147172	0.80	0.82	9.63	9.34	False	0.66
147174	0.65	0.62	11.46	11.78	False	0.70
147179	0.48	0.49	13.23	13.13	False	-0.64
147181	0.37	0.41	14.33	13.94	False	1.34
147526	0.90	0.89	7.87	8.09	False	0.05
147539	0.92	0.91	7.38	7.64	False	0.30
147546	0.59	0.64	12.09	11.57	False	2.22

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
147549	0.91	0.88	7.64	8.30	True	3.05
147555	0.58	0.55	12.24	12.55	False	0.60
147641	0.79	0.80	9.77	9.63	False	-0.38
147643	0.68	0.70	11.13	10.90	False	0.21
147646	0.68	0.67	11.13	11.24	False	-0.71
147646	0.69	0.67	11.02	11.24	False	0.06
147647	0.75	0.73	10.30	10.55	False	0.21
147649	0.50	0.51	13.00	12.90	False	-0.64
147666	0.87	0.86	8.49	8.68	False	-0.20
147668	0.73	0.74	10.55	10.43	False	-0.51
147677	0.70	0.72	10.90	10.67	False	0.25
147681	0.48	0.45	13.20	13.50	False	0.58
147686	0.58	0.54	12.19	12.60	False	1.28
147690	0.76	0.75	10.17	10.30	False	-0.59
201466	0.76	0.76	10.17	10.17	False	-1.34
201468	0.83	0.83	9.18	9.18	False	-1.34
201470	0.91	0.92	7.64	7.38	False	0.40
201472	0.83	0.83	9.18	9.18	False	-1.34
201476	0.85	0.86	8.85	8.68	False	-0.15
201479	0.76	0.76	10.17	10.17	False	-1.34
201482	0.62	0.63	11.78	11.67	False	-0.61
201487	0.91	0.90	7.64	7.87	False	0.16
201490	0.49	0.52	13.15	12.77	False	1.23
201492	0.46	0.49	13.43	13.15	False	0.55

REA07

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
118532	4	2.02	1.88	0.92	0.97	-0.15	No
201492	4	1.81	1.78	0.99	1.04	-0.02	No
201490	4	1.98	2.06	0.89	0.92	0.10	No
147555	4	2.27	2.17	0.85	0.97	-0.12	No
129219	4	1.97	1.99	1.06	1.15	0.02	No
129218	4	1.63	1.45	1.04	1.15	-0.17	No
147649	4	2.11	2.05	0.98	0.98	-0.06	No
147179	4	1.82	1.81	1.11	1.17	-0.01	No
147181	4	1.44	1.49	1.03	1.12	0.04	No



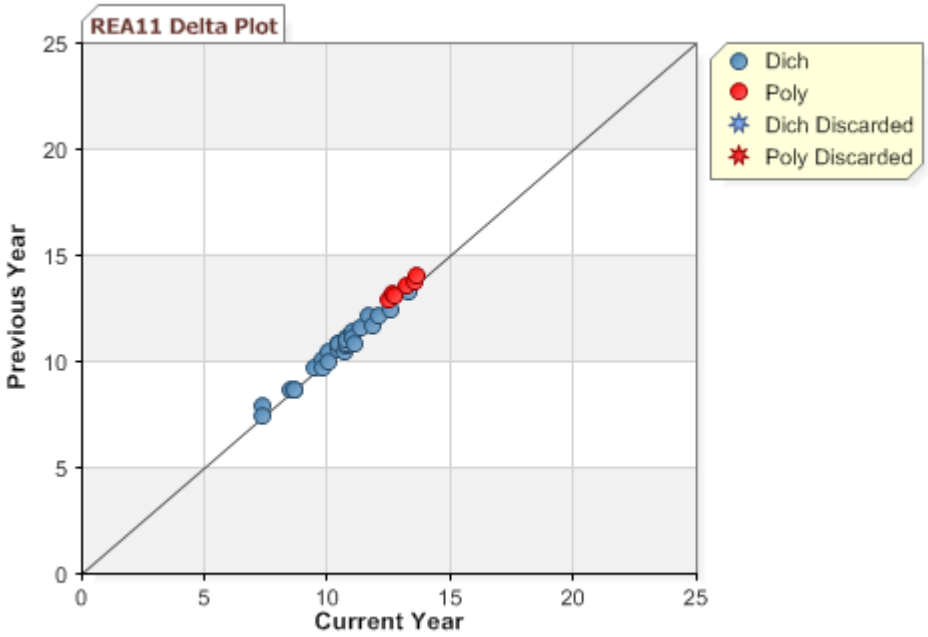
REA08

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
118704	0.77	0.76	10.04	10.17	False	-0.89
118706	0.81	0.78	9.49	9.91	False	-0.27
118707	0.59	0.57	12.09	12.29	False	-0.88
118708	0.55	0.55	12.50	12.50	False	-0.67
118711	0.59	0.56	12.12	12.40	False	-0.52
118742	0.67	0.68	11.24	11.13	False	0.07
147318	0.73	0.72	10.55	10.67	False	-0.92
147319	0.53	0.56	12.70	12.40	False	0.75
147322	0.67	0.65	11.24	11.46	False	-0.96
147324	0.78	0.78	9.91	9.91	False	-0.24
147332	0.76	0.74	10.17	10.43	False	-0.98
147333	0.82	0.81	9.34	9.49	False	-0.86
147334	0.81	0.78	9.49	9.91	False	-0.27
147340	0.55	0.48	12.50	13.20	False	1.57
147346	0.48	0.51	13.25	12.92	False	0.77
147351	0.42	0.42	13.81	13.86	False	-1.13
147492	0.78	0.77	9.91	10.04	False	-0.88
147496	0.90	0.88	7.87	8.30	False	-0.52
147498	0.57	0.53	12.29	12.70	False	0.11
147505	0.74	0.79	10.43	9.77	False	2.80
147509	0.44	0.42	13.65	13.83	False	-0.75
147516	0.72	0.68	10.67	11.13	False	0.10
147521	0.90	0.90	7.87	7.87	False	0.10
147523	0.56	0.51	12.40	12.90	False	0.60
147531	0.78	0.74	9.91	10.43	False	0.24
147537	0.82	0.81	9.34	9.49	False	-0.86
147540	0.85	0.81	8.85	9.49	False	0.63
147545	0.80	0.77	9.63	10.04	False	-0.30

IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
147547	0.91	0.89	7.64	8.09	False	-0.42
147557	0.53	0.56	12.67	12.40	False	0.63
147562	0.49	0.50	13.10	13.03	False	-0.41
147600	0.73	0.75	10.55	10.30	False	0.83
147611	0.80	0.71	9.63	10.79	True	3.25
147616	0.72	0.73	10.67	10.55	False	0.21
147621	0.90	0.90	7.87	7.87	False	0.10

REA08

IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
147509	4	1.81	1.79	1.15	1.03	-0.03	No
147351	4	1.55	1.60	1.04	1.11	0.05	No
147346	4	1.86	2.03	1.10	1.07	0.16	No
118711	4	2.45	2.26	0.88	0.93	-0.21	No
147557	4	2.25	2.36	1.16	1.08	0.10	No
147562	4	2.00	1.93	1.14	1.10	-0.06	No



REA11

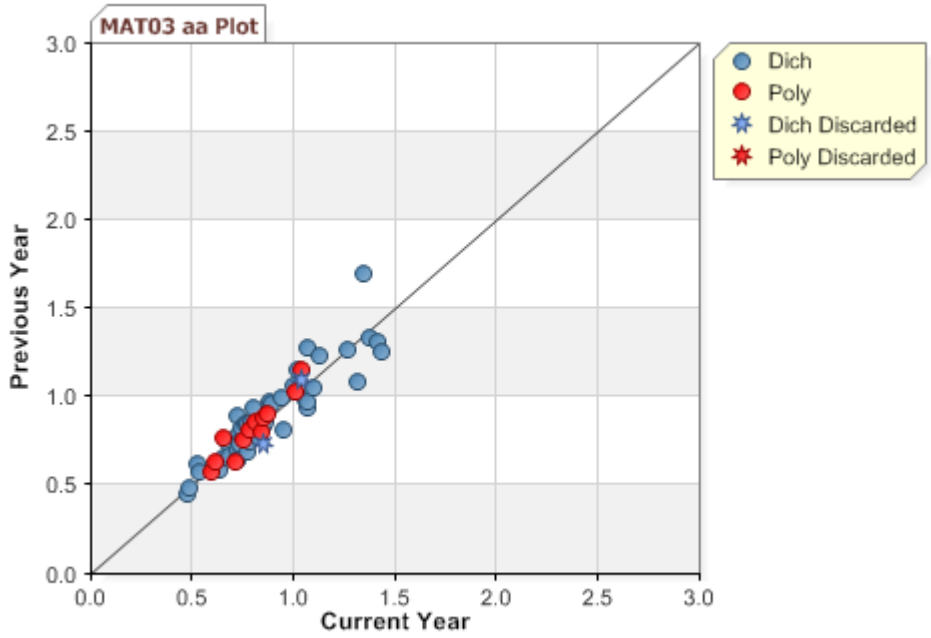
IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
118874	0.86	0.87	8.68	8.49	False	-0.78
118876	0.73	0.74	10.55	10.43	False	-1.46
118878	0.67	0.69	11.24	11.02	False	-0.85
118880	0.72	0.71	10.67	10.79	False	0.56
118883	0.43	0.44	13.76	13.58	False	-1.43
147381	0.59	0.63	12.09	11.67	False	0.64
147382	0.71	0.68	10.79	11.13	False	2.46
147383	0.68	0.71	11.13	10.79	False	0.17
147385	0.71	0.71	10.79	10.79	False	-0.41
147388	0.40	0.44	14.01	13.63	False	0.06
147423	0.86	0.86	8.68	8.68	False	-0.73
147433	0.80	0.81	9.63	9.49	False	-1.25
147435	0.68	0.69	11.13	11.02	False	-1.29
147435	0.66	0.69	11.35	11.02	False	0.05
147439	0.86	0.86	8.68	8.68	False	-0.73
147450	0.47	0.47	13.30	13.30	False	-0.01
147450	0.46	0.47	13.40	13.30	False	-0.84
147456	0.68	0.69	11.13	11.02	False	-1.29
147463	0.77	0.79	10.04	9.77	False	-0.27
147473	0.69	0.71	11.02	10.79	False	-0.76
147473	0.72	0.71	10.67	10.79	False	0.56
147484	0.50	0.52	13.05	12.77	False	-0.69
147488	0.48	0.54	13.18	12.65	False	1.39
147735	0.92	0.92	7.38	7.38	False	-0.94
147736	0.90	0.92	7.87	7.38	False	1.94
147747	0.71	0.74	10.79	10.43	False	0.36
147747	0.71	0.74	10.79	10.43	False	0.36
147749	0.59	0.59	12.09	12.09	False	-0.20

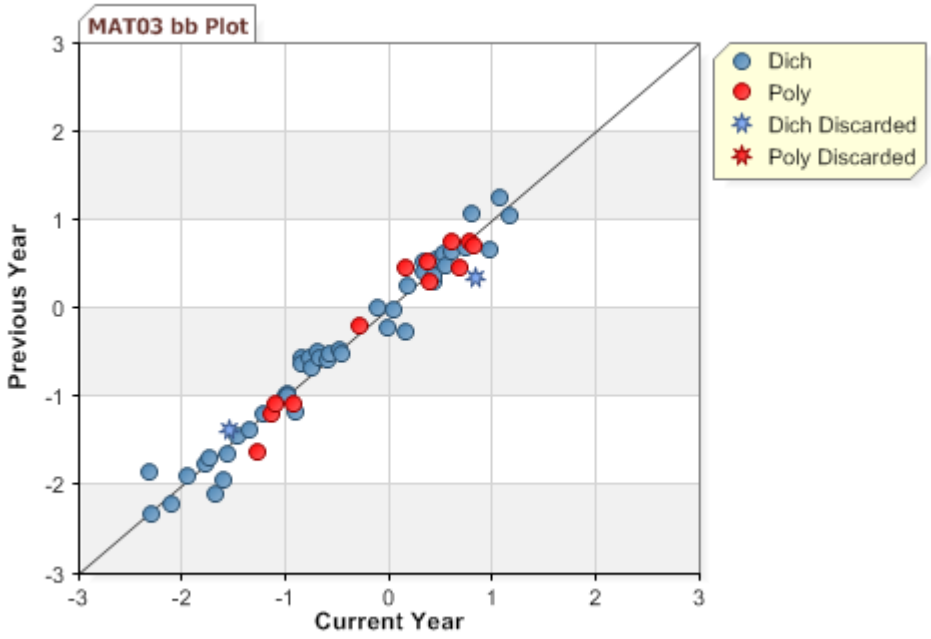
IREF	OldP	NewP	OldDelta	NewDelta	Discard	Std
147753	0.74	0.77	10.43	10.04	False	0.60
147764	0.64	0.66	11.57	11.35	False	-0.96
147766	0.80	0.79	9.63	9.77	False	0.59
147767	0.78	0.77	9.91	10.04	False	0.57
147774	0.51	0.55	12.90	12.52	False	0.18
147775	0.45	0.48	13.55	13.25	False	-0.55
147900	0.56	0.54	12.40	12.60	False	1.54
147908	0.63	0.61	11.67	11.88	False	1.49
147926	0.71	0.74	10.79	10.43	False	0.36
147929	0.74	0.72	10.43	10.67	False	1.56

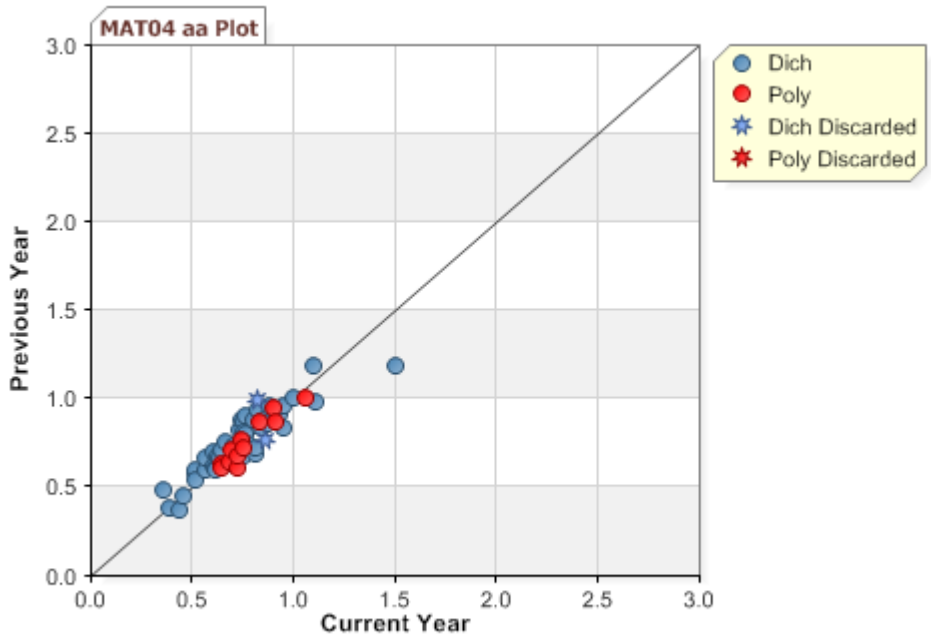
REA11

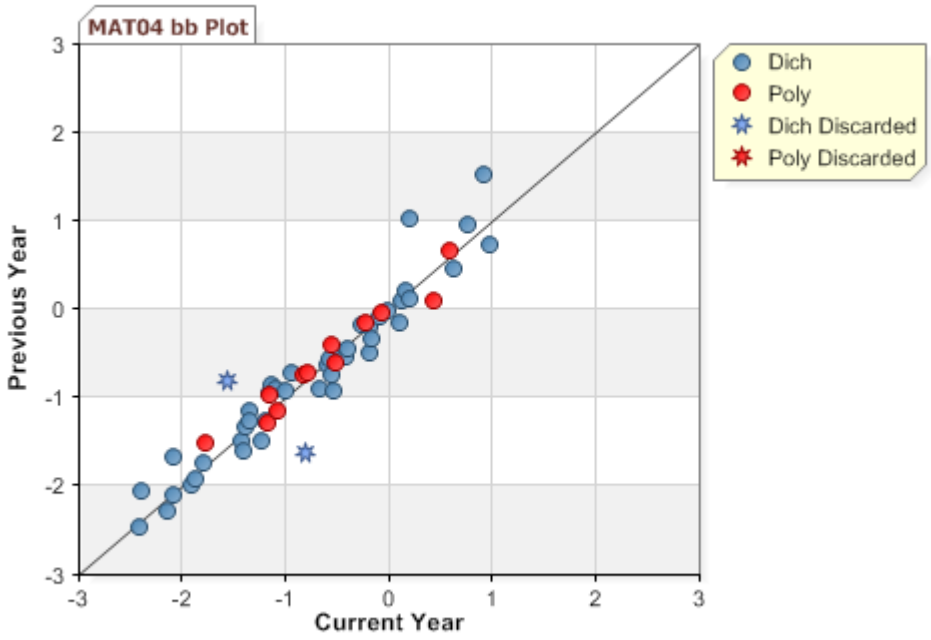
IREF	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
147774	4	2.20	2.14	1.09	1.02	-0.05	No
147775	4	1.75	1.86	1.02	1.08	0.11	No
147388	4	1.68	1.63	0.98	0.95	-0.06	No
147484	4	1.94	1.97	1.02	1.08	0.03	No
147488	4	1.86	1.99	0.95	1.02	0.13	No
118883	4	1.83	1.71	0.94	0.96	-0.13	No

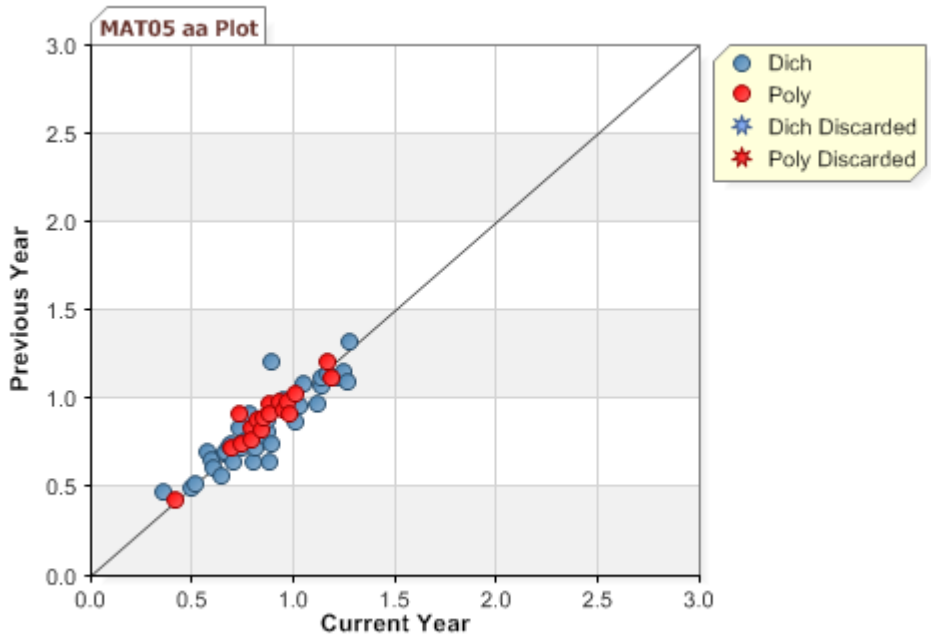
Appendix L—a-Plots and b-Plots

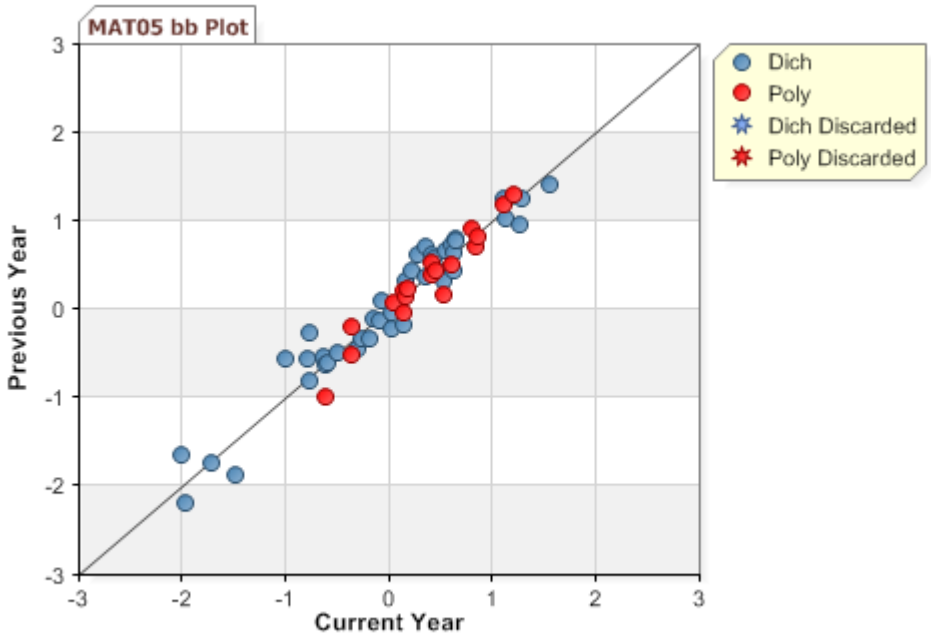

MAT03 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

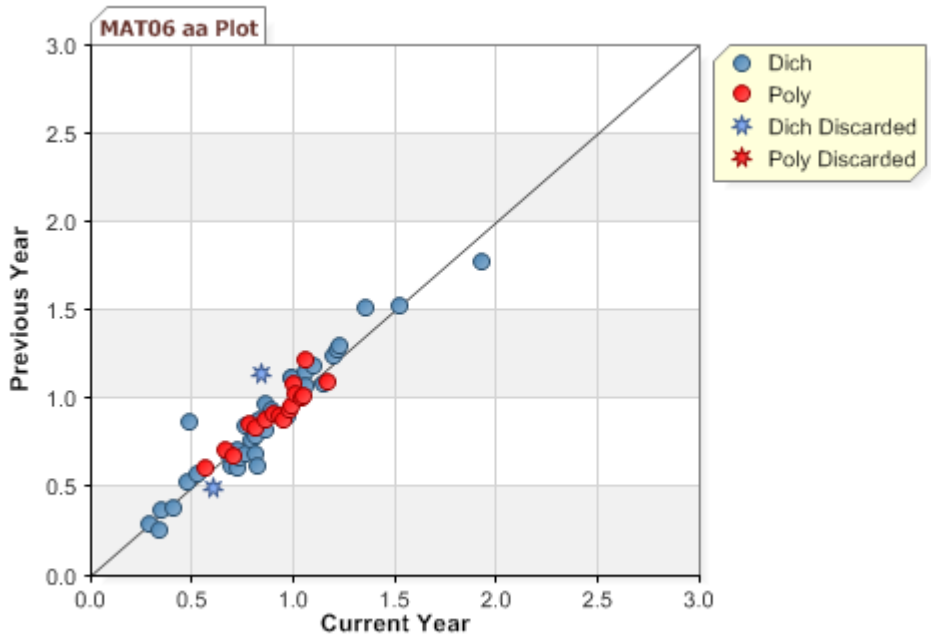

MAT03 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

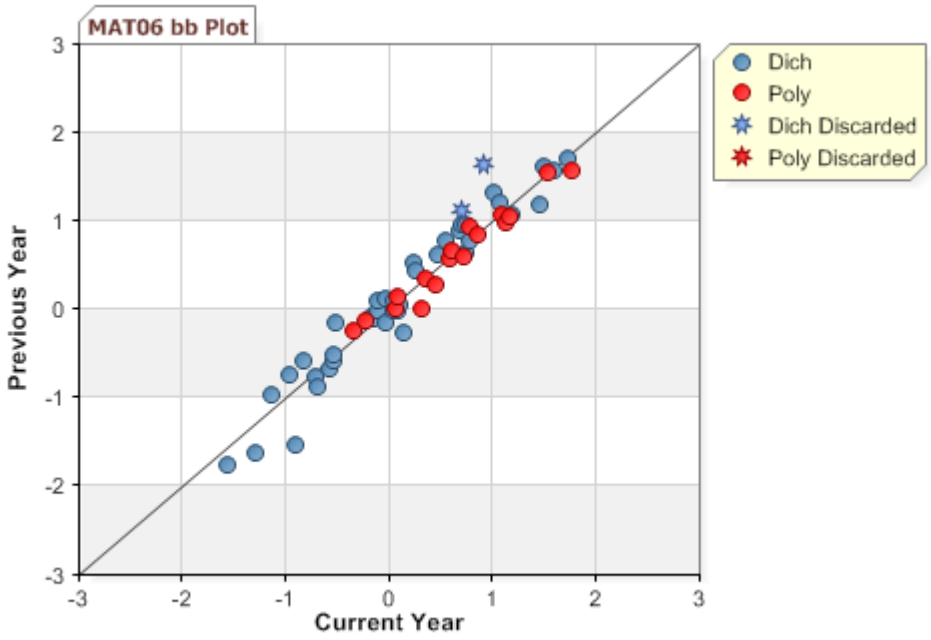

MAT04 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

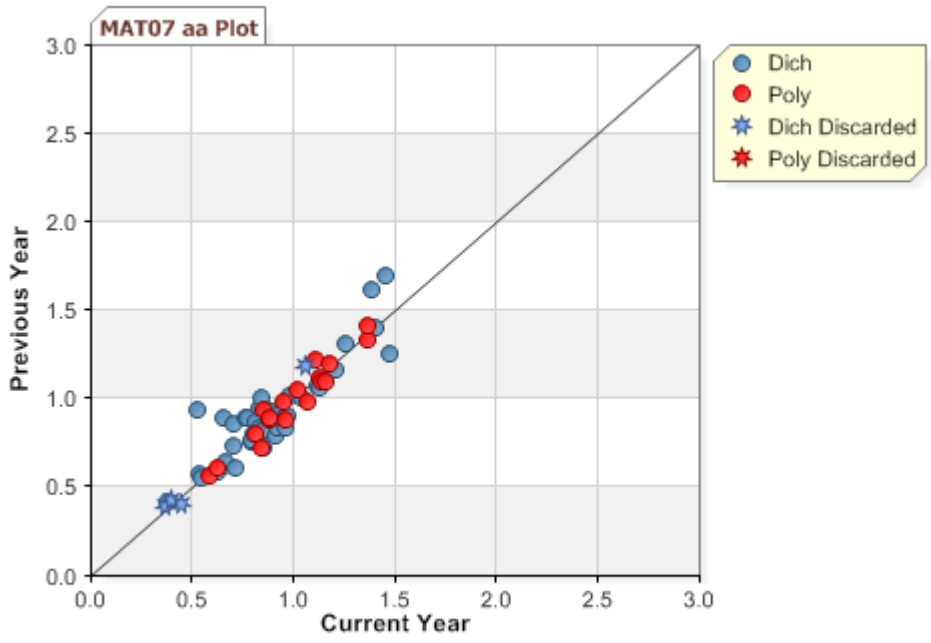

MAT04 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

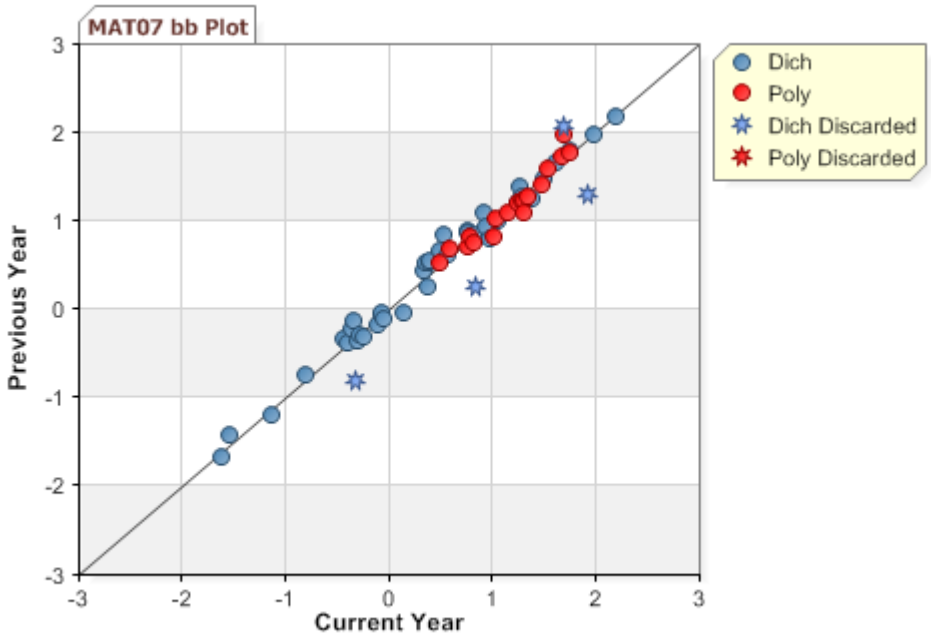

MAT05 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

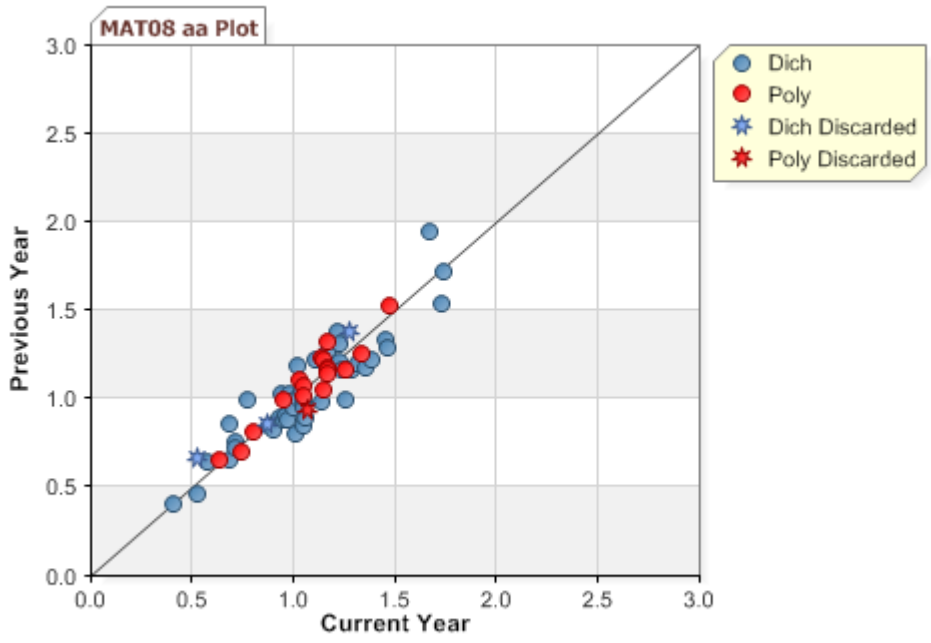

MAT05 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

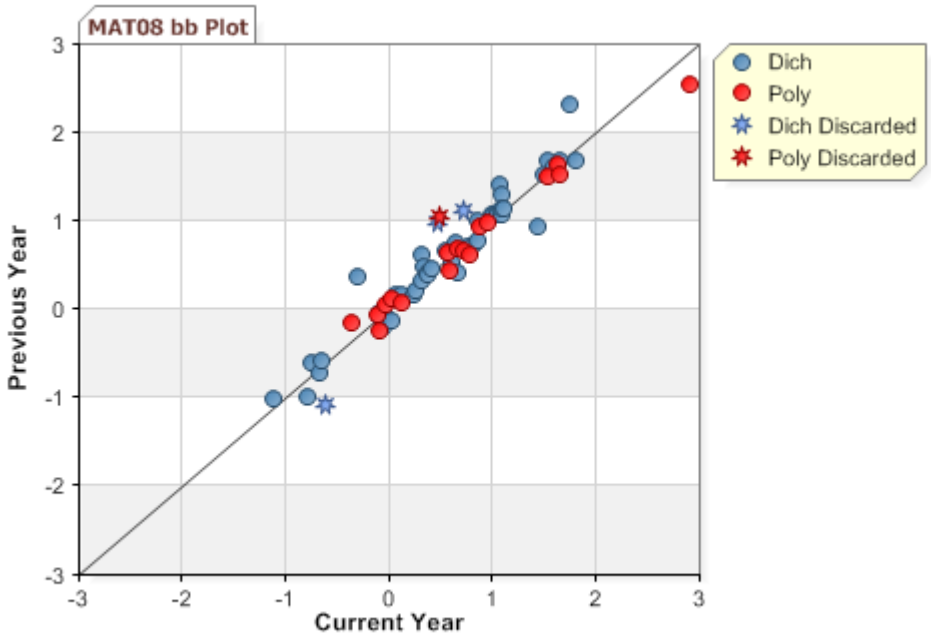

MAT06 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

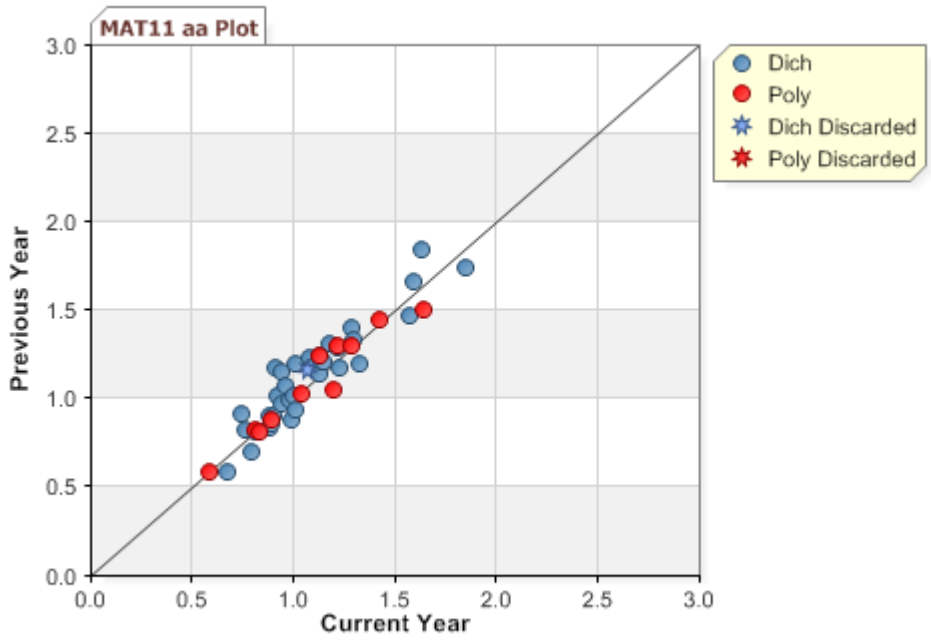

MAT06 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

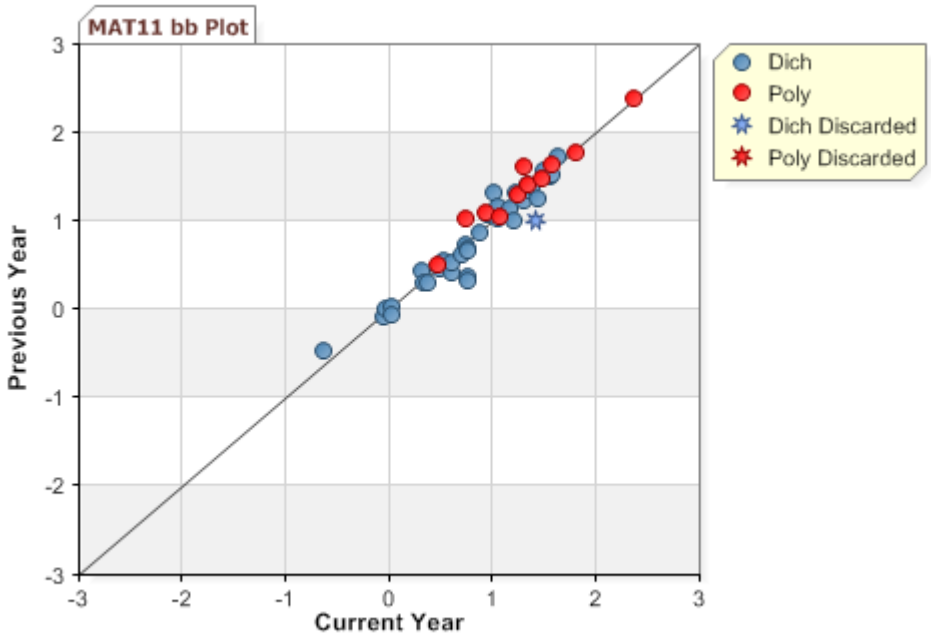

MAT07 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

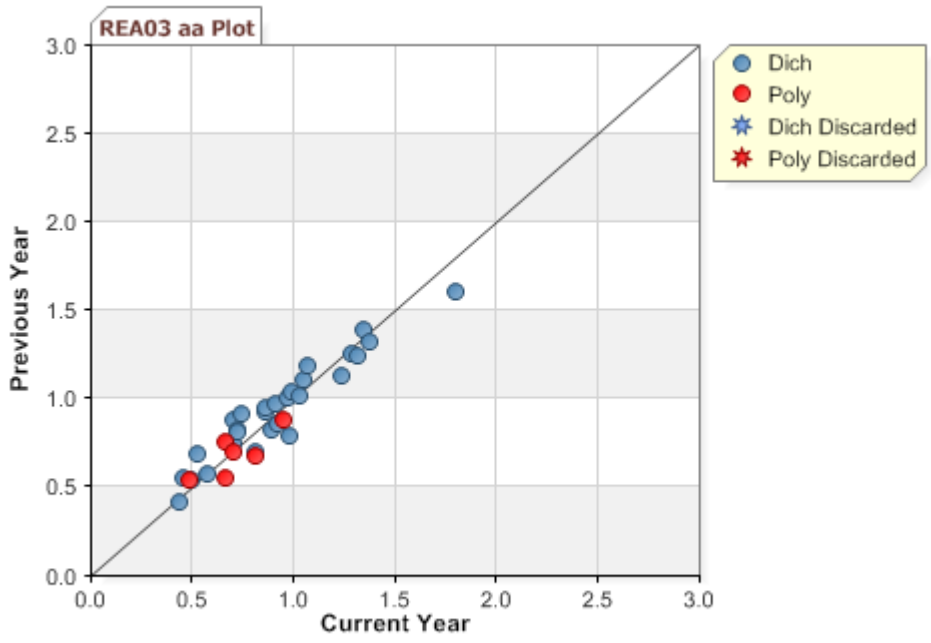

MAT07 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

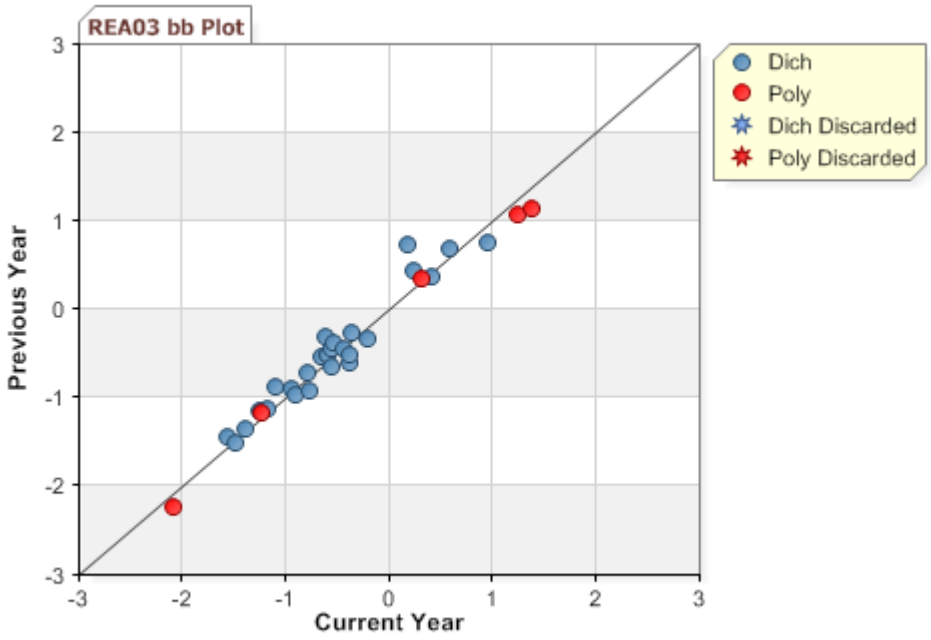

MAT08 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

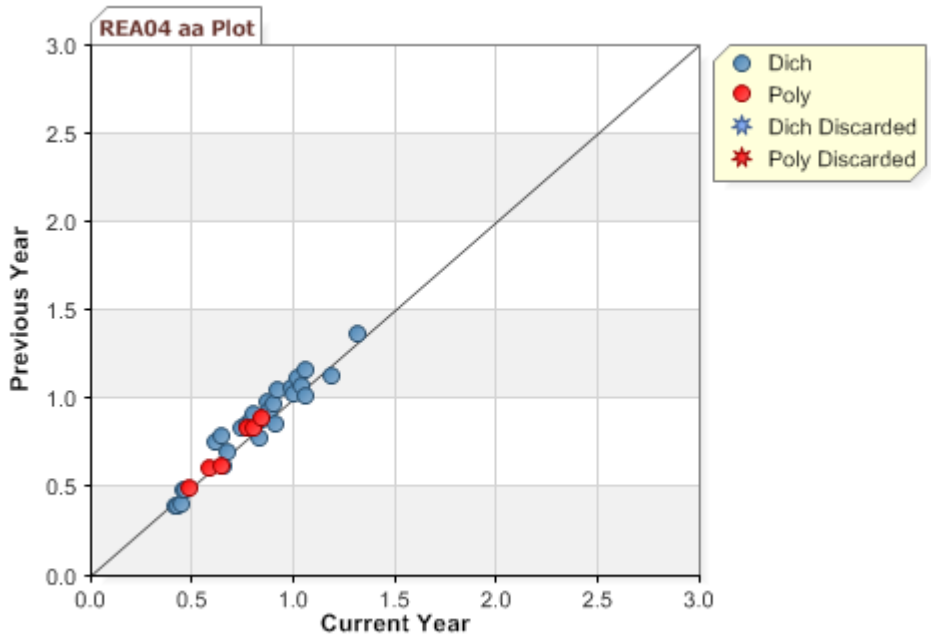

MAT08 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

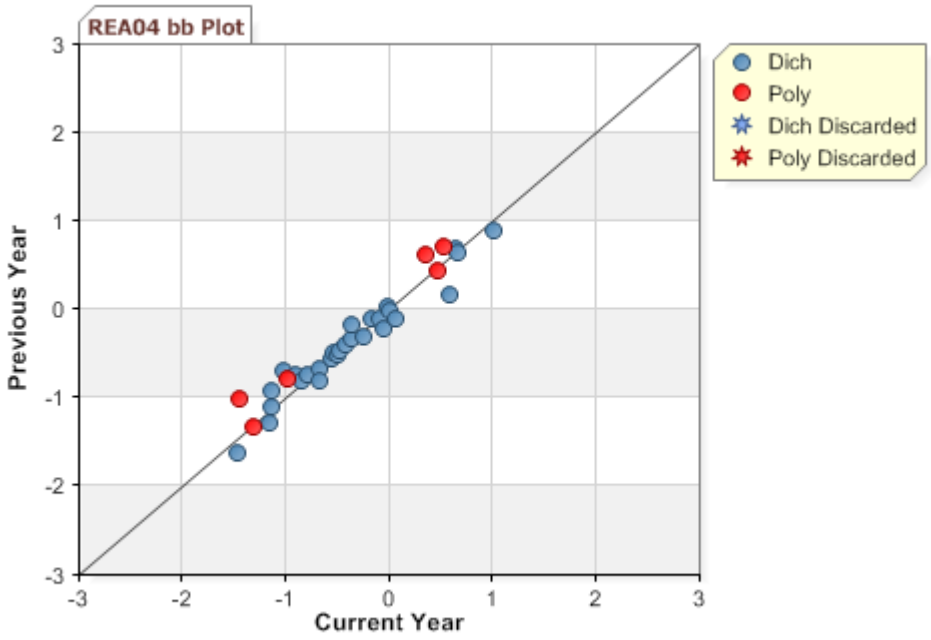

MAT11 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

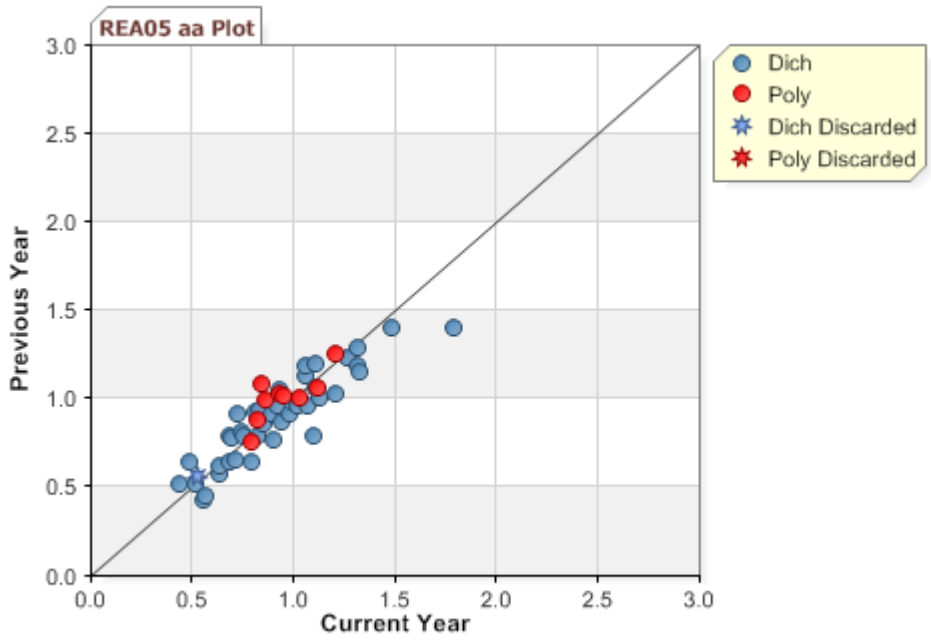

MAT11 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

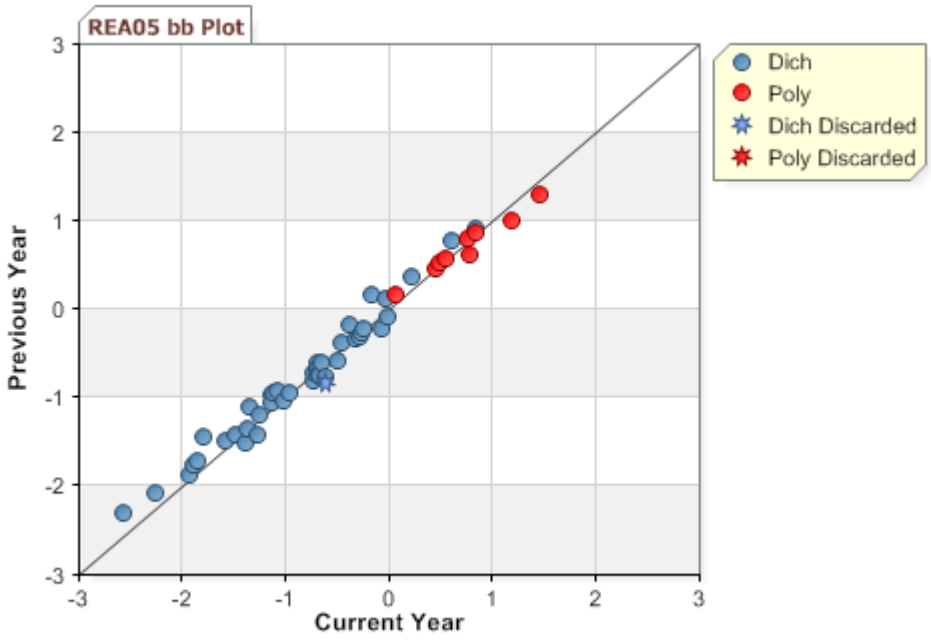

REA03 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

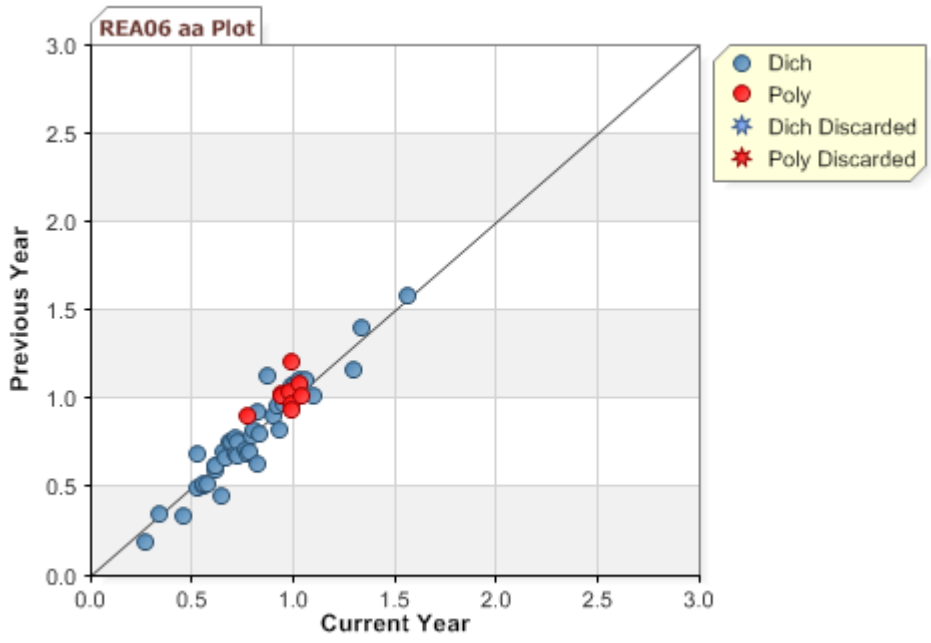

REA03 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

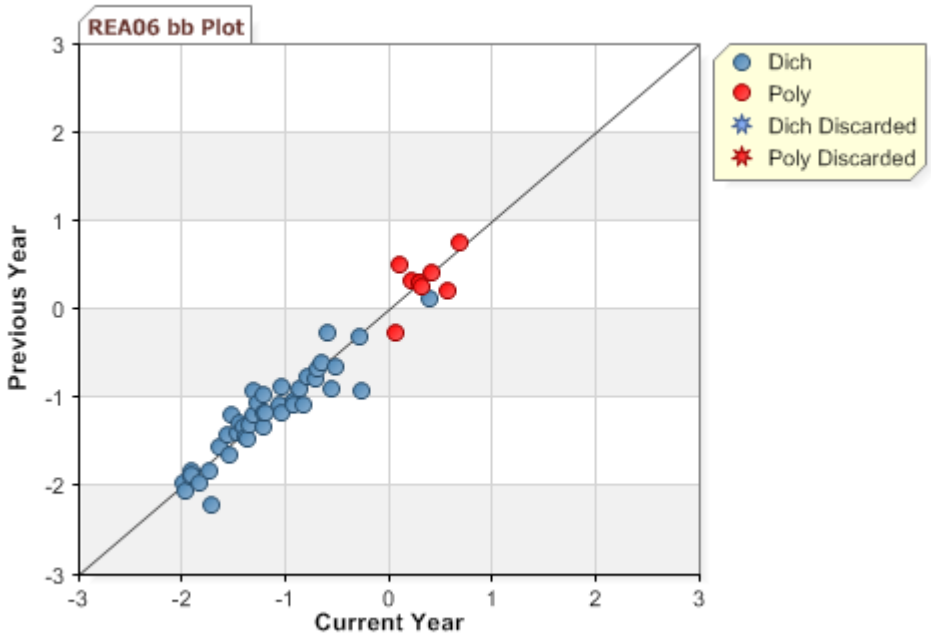

REA04 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

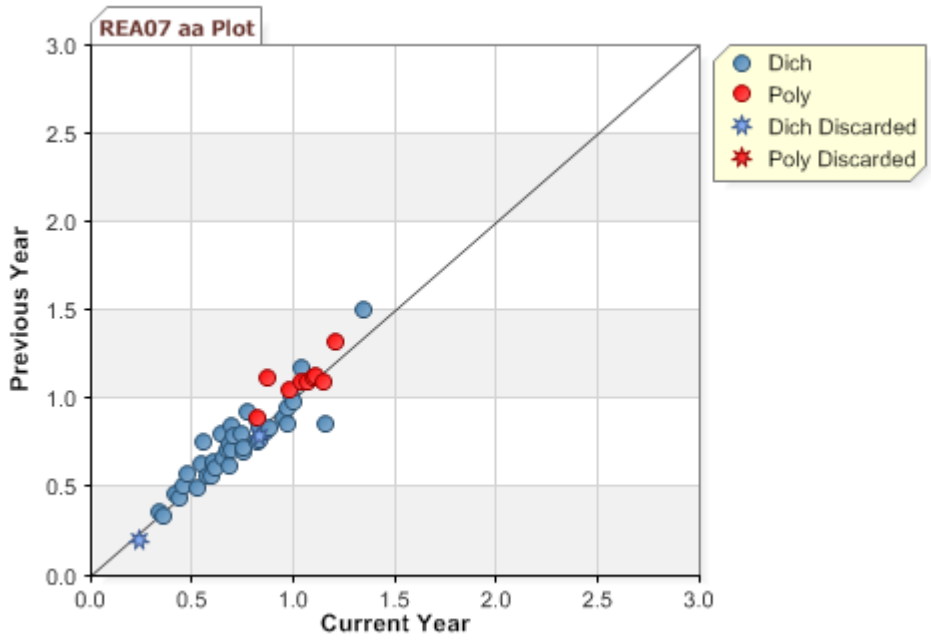

REA04 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

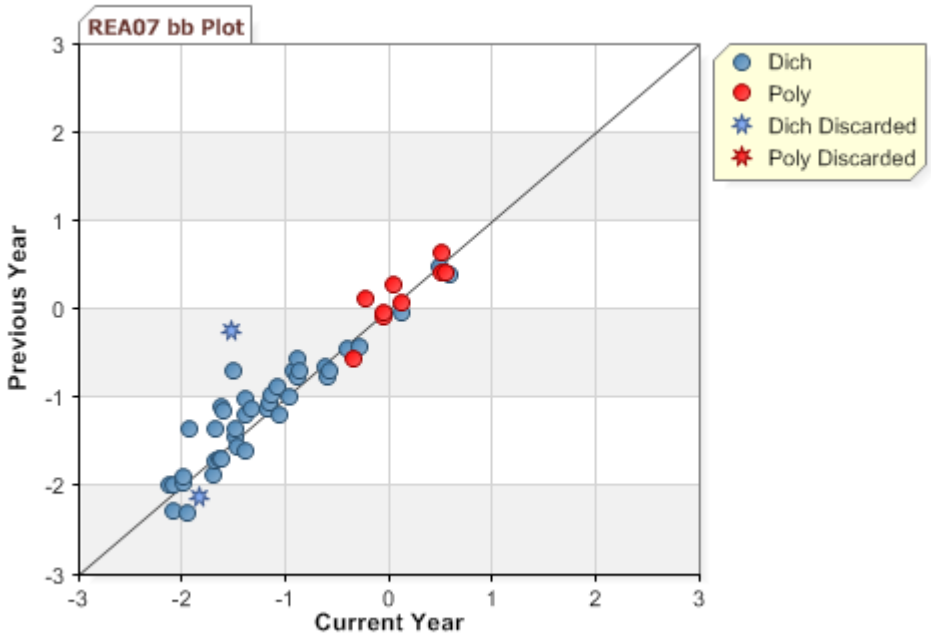

REA05 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

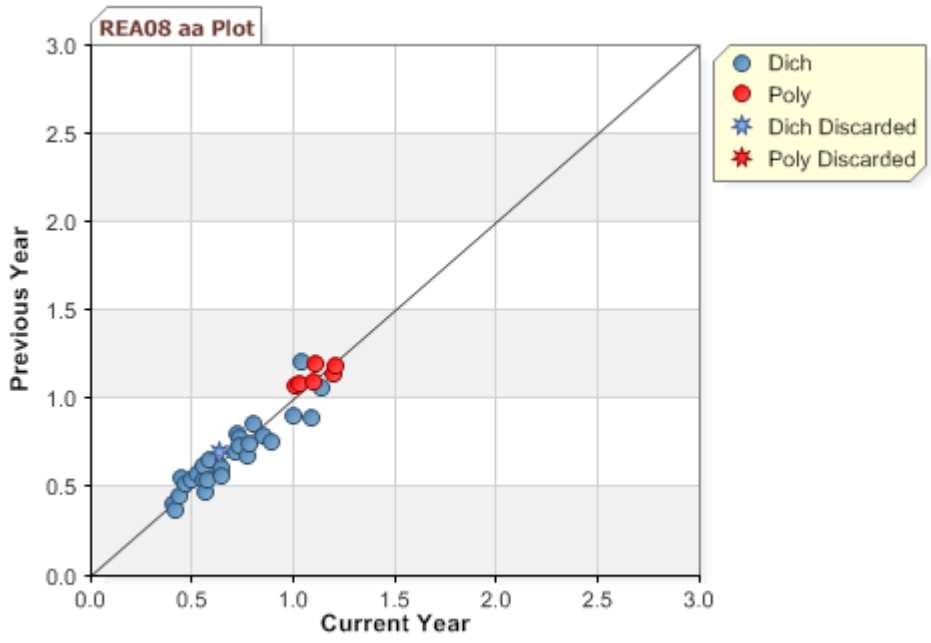

REA05 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

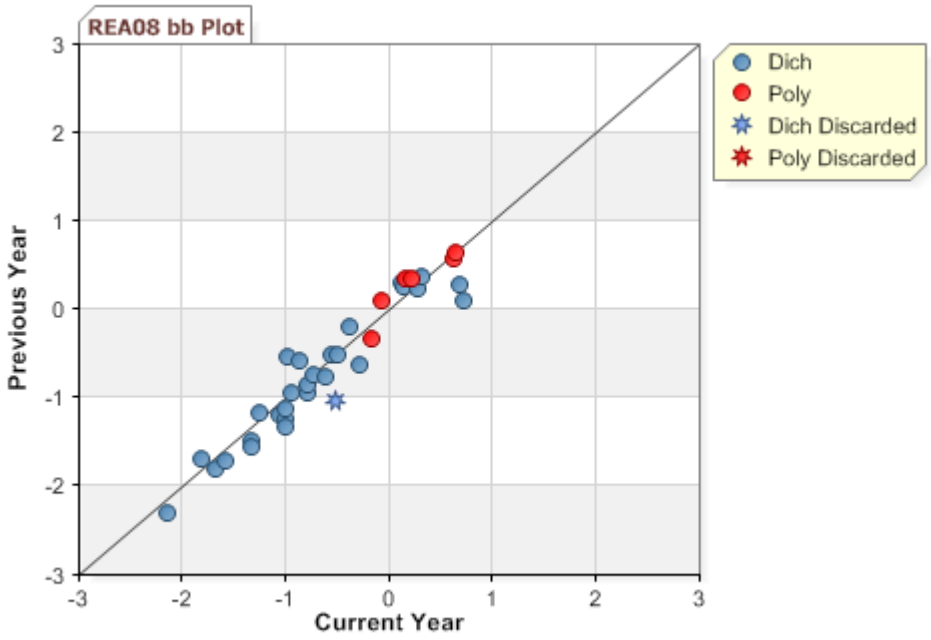

REA06 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded

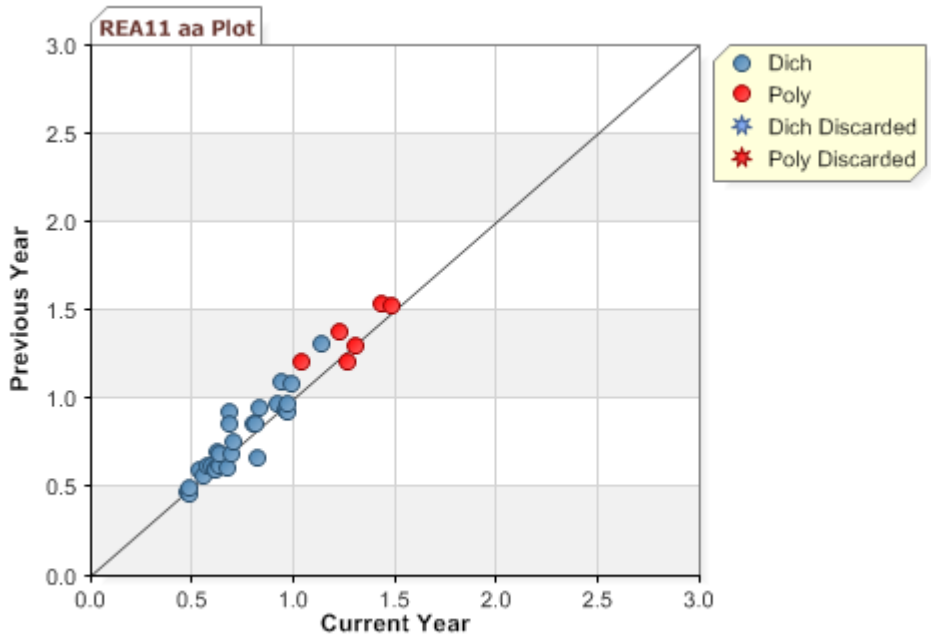

REA06 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

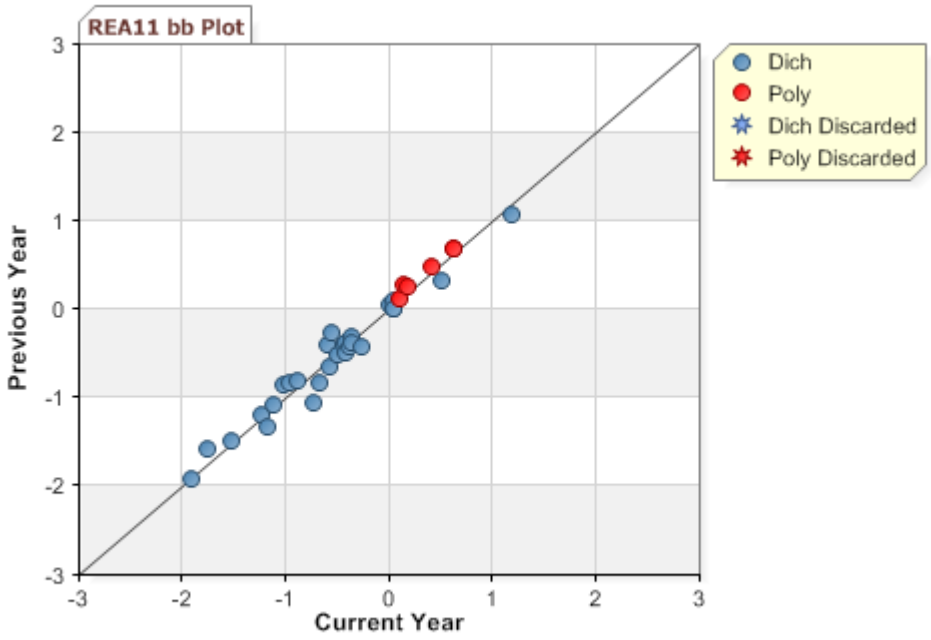

REA07 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded


REA07 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded


REA08 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded


REA08 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded


REA11 aa Plot
Previous Year
Current Year
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dich
Poly
Dich Discarded
Poly Discarded


REA11 bb Plot
Previous Year
Current Year
-3
-2
-1
0
1
2
3
Dich
Poly
Dich Discarded
Poly Discarded

Appendix M—2010-11 NECAP STANDARD SETTING REPORT

New England Common Assessment Program

Standard Setting Report

DECEMBER 8–9, 2010

PORTSMOUTH, NH

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

TABLE OF CONTENTS

CHAPTER 1. OVERVIEW OF STANDARD SETTING 2
CHAPTER 2. TASKS COMPLETED PRIOR TO STANDARD SETTING 3
2.1 Creation of Achievement Level Descriptions 3
2.2 Selection of Student Bodies of Work 3
2.3 Preparation of Materials for Panelists 3
2.4 Preparation of Presentation Materials 4
2.5 Preparation of Instructions for Facilitators Document 4
2.6 Preparation of Systems and Materials for Analysis During the Meeting 4
2.7 Selection of Panelists 4
CHAPTER 3. TASKS COMPLETED DURING STANDARD SETTING 6
3.1 Overview of Body of Work Method 6
3.2 Orientation 6
3.3 Review of Assessment Materials 6
3.4 Review of Achievement Level Descriptions 6
3.5 Training Round 7
3.6 Training Evaluation 7
3.7 Round 1 Judgments 7
3.8 Tabulation of Round 1 Results 7
3.9 Round 2 Judgments 8
3.10 Tabulation of Round 2 Results 8
3.11 Round 3 Judgments 9
3.12 Tabulation of Round 3 Results 9
3.13 Evaluation 11
CHAPTER 4. TASKS COMPLETED AFTER STANDARD SETTING 12
4.1 Analysis and Review of Panelists' Feedback 12
4.2 Final Cutpoints 12
4.3 Preparation of this Standard Setting Report 12

REFERENCES 13

APPENDICES 14

APPENDIX A. ACHIEVEMENT LEVEL DESCRIPTIONS
APPENDIX B. AGENDA
APPENDIX C. NONDISCLOSURE FORM
APPENDIX D. RATING FORMS
APPENDIX E. EVALUATION FORMS
APPENDIX F. OPENING SESSION POWERPOINT
APPENDIX G. FACILITATOR SCRIPT
APPENDIX H. PANELIST AFFILIATIONS
APPENDIX I. EVALUATION RESULTS

Chapter 1. OVERVIEW OF STANDARD SETTING

The purpose of this report is to summarize the activities of the standard setting meeting for the New England Common Assessment Program (NECAP) writing assessment (grades 5 and 8). The NECAP standard setting meeting was held December 8–9, 2010. In all, 29 panelists participated in the process. Standard setting was necessary because a new set of writing prompts was developed for use starting in fall 2010. Because the content standards, achievement level descriptions, and test design remained the same, it was anticipated that the cuts established on the 2010 test with the new prompts should yield very similar results to those obtained in past administrations of the test. Therefore, the standard setting meeting was conceptualized as a standards transfer process, i.e., to find the cuts on the tests containing the new prompts that are equivalent to the currently existing cuts on the earlier versions of the assessment. To address this goal: 1) exemplar student samples from a previous operational administration for each of the achievement levels were introduced into the process and 2) panelists were provided with the percent of students who scored at or above the proficiency level on all previous operational administrations of the assessment.

A modified version of the body of work method was used for setting standards for NECAP writing. The body of work standard setting method was developed specifically for use with assessments that are designed to allow for a range of student responses, such as performance-based assessments (Kingston, Kahl, Sweeney, & Bay, 2001). The modified version of the method, which has been in use for a number of years, substantially reduces the logistical burden of the procedure and has been found to yield reasonable and defensible cutpoints. In the body of work method, panelists are presented with samples of actual student work (bodies of work) and make their judgments based on those samples. Specifically, panelists examine each student body of work and determine which achievement level best matches the particular knowledge, skills, and abilities the student exhibits through his or her performance on the work sample.

This report is organized into three major sections, describing tasks completed prior to, during, and after the standard setting meeting.

Chapter 2. TASKS COMPLETED PRIOR TO STANDARD SETTING

2.1 Creation of Achievement Level Descriptions

The Achievement Level Descriptions (ALDs) describe the set of knowledge, skills, and abilities that students in each achievement level are expected to display. The NECAP writing ALDs, in operational use since 2006, were presented to the panelists. The ALDs are provided in Appendix A.

2.2 Selection of Student Bodies of Work

In selecting student bodies of work to use for the standard setting, the goal was to choose a total of approximately 50 work samples, spread as evenly as possible across the range of obtainable total raw scores. A list of bodies of work was generated with approximately five times the target number of bodies of work at each score point. Each body of work was reviewed by Measured Progress scoring and program management staff to ensure that they were suitable for use during the standard setting process. The surplus was identified so that a replacement body of work was available should scoring or program management staff deem a work sample to be inappropriate.

In addition to the set of 50 bodies of work used for setting standards, a set of 12 exemplar bodies of work was selected for each grade level. The purpose of the exemplar bodies of work, which were selected from the 2007–08 administration of the NECAP writing assessment, was to provide panelists with concrete examples of the knowledge, skills, and abilities demonstrated by students who scored in each achievement level based on the old prompts. Because the content standards and achievement level descriptions were unchanged, panelists could then use this information to help them identify equivalent performance on the 2010 test with the new prompts. The exemplars were chosen such that:

- each of the four achievement levels was represented;
- there were exemplars that fell just above and below each of the three cuts; and
- the majority of the exemplars (nine of the twelve) represented the Partially Proficient and Proficient achievement levels.

Finally, a set of 5 training bodies of work was selected, also from the 2007-08 administration. The training bodies of work were used for a practice round completed by the panelists prior to starting Round 1 of the standard setting (see below).

2.3 Preparation of Materials for Panelists

The following materials were assembled for presentation to the panelists at the standard setting meeting:

- Meeting agenda

- Nondisclosure agreement
- Achievement Level Descriptions (ALDs)
- Student bodies of work (training, exemplar, and standard setting)
- Rating forms
- Evaluation forms

Copies of the ALDs, meeting agenda, nondisclosure form, rating forms, and evaluation forms are included in Appendices A through E.

2.4 Preparation of Presentation Materials

The PowerPoint presentation used in the opening session was prepared prior to the meeting. The presentation was designed to give panelists an overview of the assessment and how it is scored, as well as a preview of what to expect throughout the standard setting process. A copy of the presentation is included in Appendix F.

2.5 Preparation of Instructions for Facilitators Document

A script was created for the group facilitators to refer to while working through each step of the standard setting process. This document is included in Appendix G. The facilitators also attended a training session, led by a Measured Progress psychometrician, approximately one week before the standard setting meeting. The purpose of the training was to prepare the facilitators for the panel activities and to ensure consistency in the implementation of procedures.

2.6 Preparation of Systems and Materials for Analysis During the Meeting

The computational programming used to carry out all analyses during the standard setting meeting was completed and thoroughly tested in advance.

2.7 Selection of Panelists

As was emphasized in Cizek and Bunch (2007), regardless of the method used, the selection of panelists is an important factor in determining standard setting outcomes and maximizing the validity of the standard setting process. The guidance provided by *Standards for Educational and Psychological Testing* (AERA, APA & NCME, 1999) states “a sufficiently large and representative group of judges should be involved to provide reasonable assurance that results would not vary greatly if the process were repeated.” Consistent with the above guidance, as well as practical considerations regarding the maximum size of a group that can be successfully managed by group facilitators, the goal was to recruit standard setting panels of approximately 16 members representing different stakeholder groups to set standards for each grade. Targets

for the size and composition of the panels were also consistent with federal guidelines as described in *Standards and Assessment Peer Review Guidance: Information and Examples for Meeting Requirements of the No Child Left Behind Act of 2001* (U.S. Department of Education, 2009).

Panelists were selected by the member departments and Measured Progress prior to the standard setting meeting. The goal was for each panel to consist of approximately 16 participants and to include general education teachers, literacy specialists, and teachers with expertise with English language learners, special education students, and gifted and talented students. In addition, to the extent possible, panels were assembled so as to reflect a balance of gender, race/ethnicity, and geographic location. The final number of panelists who participated was 15 for grade 5 and 14 for grade 8. A list of the panelists and their affiliations is included in Appendix H.

Chapter 3. TASKS COMPLETED DURING STANDARD SETTING

3.1 Overview of Body of Work Method

The body of work standard setting method was developed specifically for use with assessments that are designed to allow for a range of student responses, such as assessments that consist largely of open-ended items. For a number of years, a modified version of the method has been in use that substantially reduces the logistical burden of the procedure and has been found to yield reasonable and defensible cutpoints. Panelists were asked to evaluate each work sample from a holistic perspective before classifying it into a single achievement level.

3.2 Orientation

With regard to panelist training, *Standards for Educational and Psychological Testing* states the following:

> Care must be taken to assure that judges understand what they are to do. The process must be such that well-qualified judges can apply their knowledge and experience to reach meaningful and relevant judgments that accurately reflect their understanding and intentions. (AERA, APA, & NCME, 1999, p. 54)

The training of the panelists began with a general orientation at the start of the standard setting meeting. The purpose of the orientation was to ensure that all panelists received the same information about the need for and goals of standard setting and about their part in the process. First, staff from the member departments provided some pertinent context about the NECAP program and an introduction to the issues of standard setting. Next, a Measured Progress psychometrician presented a brief overview of the body of work procedure and the activities that would occur during the standard setting meeting. Once the general orientation was complete, each panel convened in a breakout room, where the panelists received more detailed training and completed the standard setting activities.

3.3 Review of Assessment Materials

The first step after the opening session was for the panelists to take the NECAP writing test. The purpose of this step was to make sure the panelists were thoroughly familiar with the test items and to gain an understanding of the experience of the students who take the test.

3.4 Review of Achievement Level Descriptions

The next step in the process was to discuss the Achievement Level Descriptions (ALDs). This important step was designed to ensure that panelists thoroughly understood the knowledge, skills, and abilities needed for students to be classified into achievement levels (Substantially Below Proficient, Partially

Proficient, Proficient, and Proficient with Distinction). In addition, at this time, panelists reviewed the 12 exemplar bodies of work and used those to help identify the knowledge, skills, and abilities demonstrated by students in each achievement level. Panelists first reviewed the ALDs and exemplar papers on their own and then participated in group discussion of the ALDs, clarifying the description for each achievement level. The discussions focused on the knowledge, skills, and abilities that differentiated adjacent achievement levels. Bulleted lists of characteristics for each level were generated based on the group discussion and were posted in the room for panelists to refer to during the rounds of ratings.

3.5 Training Round

Before beginning the rating rounds, the panelists completed a training round using the set of five training bodies of work. The training set was chosen to include straightforward papers that would fit well into one of the achievement levels and papers that, based on the ALDs, would likely lead to discrepant placement by the panelists. The facilitators walked the panelists through the five BOWs, engaged the panelists in a discussion about how to categorize into achievement levels, and checked for understanding before proceeding.

3.6 Training Evaluation

Prior to beginning the Round 1 ratings, the panelists anonymously completed a training evaluation form. The purpose of the evaluation was to ensure that panelists were comfortable with the process and ready to move on to the rating task. Any issues or problems that came up in the training evaluations were addressed before the facilitator proceeded to Round 1. Results of the evaluation are included in Appendix I.

3.7 Round 1 Judgments

In the first round, panelists worked individually with the ALDs, the student bodies of work, and the rating form. The work samples consisted of 50 bodies of work, covering the full range of possible total scores. For each body of work, the panelists considered the skills and abilities demonstrated in the work sample and decided which achievement level was the best match. The panelists worked their way through the bodies of work, decided on a rating for each one, and recorded their ratings on the rating form. While the bodies of work were presented in order of total score, panelists were not required to rate them strictly in increasing order. Instead, panelists were encouraged to take a holistic look at the knowledge, skills, and abilities demonstrated in the body of work, rather than making a judgment based primarily on the total raw score.

3.8 Tabulation of Round 1 Results

After all panelists had completed their individual ratings, the Measured Progress data analysis team calculated the average cutpoints for the group based on the Round 1 ratings. Cutpoints were calculated using

SAS statistical software. Logistic regression was used to determine each panelist's individual cutpoints, and then the cutpoints were averaged across the group. In addition, impact data were calculated, which reflected the percentage of students who would fall into each achievement level based on the group average Round 1 ratings. The Round 1 results are outlined in Table 3-1.

Table 3-1. 2010 NECAP Standard Setting: Round 1

Grade	*Achievement level*	*Average theta cut*	*Standard error*	*Raw score*		*Percent of students*
				Min	*Max*	
5	Proficient with Distinction	1.75	0.10	27	34	5.13
	Proficient	-0.89	0.08	16	26	75.89
	Partially Proficient	-2.50	0.12	8	15	17.47
	Substantially Below Proficient	NA	NA	0	7	1.51
8	Proficient with Distinction	1.08	0.10	27	34	16.27
	Proficient	-0.60	0.07	18	26	61.68
	Partially Proficient	-1.68	0.08	12	17	16.79
	Substantially Below Proficient	NA	NA	0	11	5.26

3.9 Round 2 Judgments

The purpose of Round 2 was for panelists to discuss their Round 1 judgments as a group and determine whether any revisions were necessary. A psychometrician shared the average cutpoint locations with the panelists to help inform their group discussion and Round 2 ratings. It is important to note that the impact data and raw score ranges are presented in Table 3-1, they were not shared with the panelists. Prior to the group discussion, the facilitator asked for a show of hands to determine the number of panelists who had placed each body of work into each achievement level; the facilitator then recorded the results on chart paper. Starting with the first body of work they disagreed on, the panelists began discussing the categorization of the bodies of work according to their initial ratings in the context of the classifications made by other members of the group. Panelists were encouraged to share their own points of view as well as listen to the comments of their colleagues. Facilitators made sure the panelists knew that the purpose of the discussion was not to reach consensus; at every point throughout the standard setting process, panelists were asked to provide their own best judgment. Once the discussions were complete, the panelists completed the Round 2 rating form.

3.10 Tabulation of Round 2 Results

When Round 2 ratings were complete, the Measured Progress data analysis team calculated the average cutpoints for the room and associated impact data. The results of the panelists' Round 2 ratings are outlined in Table 3-2.

Table 3-2. 2010 NECAP Standard Setting: Round 2

Grade	*Achievement level*	*Average theta cut*	*Standard error*	*Raw score*		*Percent of students*
				Min	*Max*	
5	Proficient with Distinction	1.83	0.06	27	34	5.13
	Proficient	-0.87	0.05	16	26	75.89
	Partially Proficient	-2.41	0.07	9	15	16.83
	Substantially Below Proficient	NA	NA	0	8	2.15
8	Proficient with Distinction	1.06	0.04	27	34	16.27
	Proficient	-0.56	0.04	19	26	56.48
	Partially Proficient	-1.81	0.02	11	18	23.17
	Substantially Below Proficient	NA	NA	0	10	4.07

3.11 Round 3 Judgments

The purpose of Round 3 was for panelists to discuss their Round 2 ratings as a whole group and, if necessary, revise their judgments. Prior to the group discussion, the facilitator once again asked for a show of hands to determine the number of panelists who had placed each body of work into each achievement level; the facilitator recorded the results on chart paper. The group average cutpoints based on the Round 2 results were presented. In addition, in this round, the group was presented with the impact data. The psychometrician presented the group average cuts and impact data to the group and explained how to use the information as they completed their Round 3 discussions. Panelists were encouraged to discuss whether the percentage of students classified in each performance level seemed reasonable, given their perceptions of the students and the knowledge, skills, and abilities demonstrated in the bodies of work. Panelists were also encouraged to refer to the exemplars and the relationship between the knowledge, skills, and abilities demonstrated in the exemplars and the BOWs that bordered the cutpoints. As in Round 2, for each body of work for which there was disagreement, the panelists discussed their ratings, and considered the impact data as additional context for the discussion. After the discussions were complete, panelists were given a final opportunity to revise their ratings. Once again, the facilitator reminded the panelists that they should use their individual best judgment and that it was not necessary for them to reach consensus.

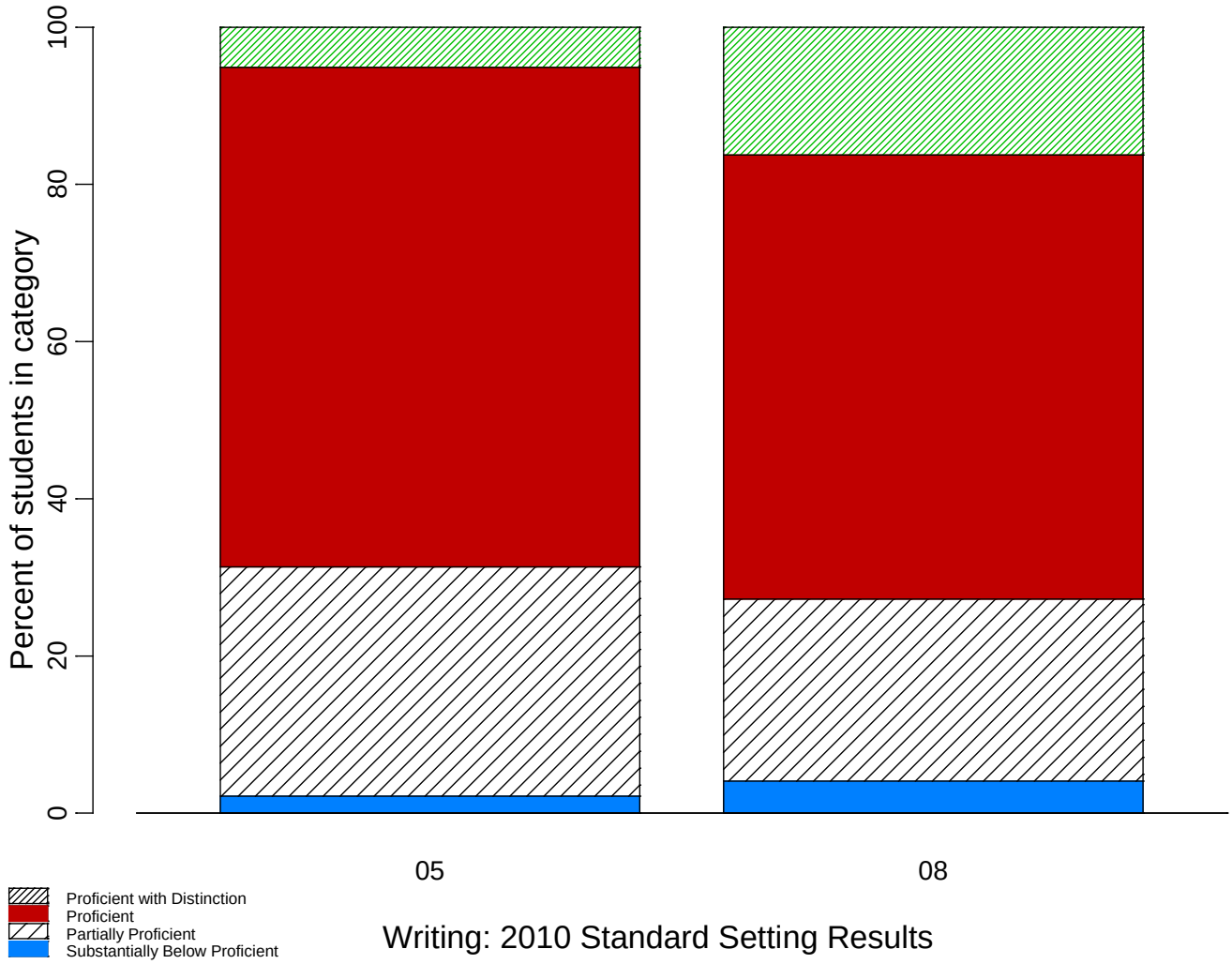
3.12 Tabulation of Round 3 Results

When Round 3 ratings were complete, the Measured Progress data analysis team once again calculated the average cutpoints for the room and associated impact data. The results of the panelists' Round 3 ratings are outlined in Table 3-3 and Figure 3-1.

Table 3-3. 2010 NECAP Standard Setting: Round 3

Grade	*Achievement level*	*Average theta cut*	*Standard error*	*Raw score*		*Percent of students*
				Min	*Max*	
5	Proficient with Distinction	1.91	0.04	27	34	5.13
	Proficient	-0.47	0.05	18	26	63.53
	Partially Proficient	-2.41	0.08	9	17	29.19
	Substantially Below Proficient	NA	NA	0	8	2.15
8	Proficient with Distinction	1.07	0.04	27	34	16.27
	Proficient	-0.39	0.04	19	26	56.48
	Partially Proficient	-1.80	0.02	11	18	23.17
	Substantially Below Proficient	NA	NA	0	10	4.07

Figure 3-1. Round 3 Results: Percent of Students in each Achievement Level by Grade



3.13 Evaluation

The measurement literature sometimes considers the evaluation process to be another product of the standard setting process (e.g., Reckase, 2001). To provide evidence of the participants' views of the standard setting process, panelists were asked to complete questionnaires prior to beginning Round 1 and again at the end of the standard setting process. The results of the evaluations are presented in Appendix I.

Chapter 4. TASKS COMPLETED AFTER STANDARD SETTING

Upon conclusion of the standard setting meeting, several important tasks were completed. These tasks centered on reviewing the standard setting process and addressing anomalies that may have occurred in the process or in the outcomes, presenting the results to the member departments, and making any final revisions or adjustments.

4.1 Analysis and Review of Panelists' Feedback

Upon completion of the evaluation forms, panelists' responses were reviewed. In general, this review did not reveal any anomalies in the standard setting process or indicate any reason that a particular panelist's data should not be included when the final cutpoints were calculated.

4.2 Final Cutpoints

At the conclusion of the standard setting meeting, the Round 3 cutpoints were presented to the member departments of education. In addition to these cutpoints, the member departments of education also considered other relevant information when determining the final cutpoints. For additional information on the standard setting process and the final cutpoints used for the reporting of grade 5 & 8 writing results, contact your state department of education.

4.3 Preparation of this Standard Setting Report

Following final compilation of standard setting results, Measured Progress prepared this report, which documents the procedures and results of the 2010 standard setting meeting in order to establish performance standards for the NECAP writing assessments.

REFERENCES

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

Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: Establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.

Kingston, N., Kahl, S., Sweeney, K., & Bay, L. (2001). Setting performance standards using the body of work method. In G. J. Cizek (Ed.) *Setting performance standards: Concepts, methods, and perspectives* (pp. 219–248). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Reckase, M.D. (2001). Innovative methods for helping standard-setting participants to perform their task: The role of feedback regarding consistency, accuracy, and impact. In G. J. Cizek (Ed.) *Setting performance standards: Concepts, methods, and perspectives* (pp. 159-173). Mahwah, NJ: Lawrence Erlbaum.

U.S. Department of Education. (2009). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. (Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education), www2.ed.gov/policy/elsec/guid/saaprguidance.pdf. (accessed June 10, 2010).

APPENDICES

Appendix A—Writing Achievement Level Descriptions

NECAP Writing Achievement Level Descriptions

Proficient with Distinction	Student's writing demonstrates an ability to respond to prompt/task with clarity and insight. Focus is well developed and maintained throughout response. Response demonstrates use of strong organizational structures. A variety of elaboration strategies is evident. Sentence structures and language choices are varied and used effectively. Response demonstrates control of conventions; minor errors may occur.
Proficient	Student's writing demonstrates an ability to respond to prompt/task. Focus is clear and maintained throughout the response. Response is organized with a beginning, middle and end with appropriate transitions. Details are sufficiently elaborated to support focus. Sentence structures and language use are varied. Response demonstrates control of conventions; errors may occur but do not interfere with meaning.
Partially Proficient	Student's writing demonstrates an attempt to respond to prompt/task. Focus may be present but not maintained. Organizational structure is inconsistent with limited use of transitions. Details may be listed and lack elaboration. Sentence structures and language use are unsophisticated and may be repetitive. Response demonstrates inconsistent control of conventions.
Substantially Below Proficient	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.

Appendix B—AGENDA

New England Common Assessment Program

Standard Setting: Grades 5 & 8 Writing

Sheraton Harborside – Portsmouth, NH

AGENDA

Wednesday, December 8, 2010

Breakfast will be available in the Langdon Room from 7:30 – 8:30

8:00 – 8:30	Registration
8:30 – Noon	Welcome and Introductions Overview of the Standard Setting Process Committee Work Session – *Writing Committee Work will be in the Whipple & Roberts Rooms*
	10:15 Morning Beverage Break – Ballroom Lobby
Noon – 1:00	*Lunch – Langdon Room*
1:00 – End	Committee Work Session
	2:30 Afternoon Snack Break – Ballroom Lobby

Thursday, December 9, 2010

Breakfast will be available in the Langdon Room from 7:30 – 8:30

8:30 – Noon	Committee Work Session
	10:15 Morning Beverage Break-Ballroom Lobby
Noon – 1:00	*Lunch – Langdon Room*
1:00 – Conclusion	Committee Work Session
	2:00 Afternoon Snack Break – Ballroom Lobby

Appendix C—NONDISCLOSURE FORM

New England Common Assessment Program

NONDISCLOSURE AGREEMENT

The New England Common Assessment Program is a partnership among the Maine, New Hampshire, Rhode Island, and Vermont Departments of Education to develop and administer a shared statewide assessment system. The design of the program requires that the test questions remain secure. To maintain the security of the tests, only authorized persons are permitted to view the test questions and reading passages. With the exception of questions and reading passages released by the Maine, New Hampshire, Rhode Island, and Vermont Departments of Education with official reports and on their state websites, all test questions, draft or final, and reading passages are to be regarded as secure instruments.

I understand that it is my professional responsibility to maintain the security of the tests. I will not reproduce, discuss, or in any way release, share, or distribute the test questions or reading passages to unauthorized personnel.

The undersigned is an employee, contractor, consultant, or committee member for the New England Common Assessment Program, or person otherwise authorized to view secure New England Common Assessment Program materials and hereby agrees to be bound to the terms of this agreement restricting the disclosure of said materials.

_____
Name (please print)

_____
School

_____
Position

_____
State

_____
Signature

_____
Date

Appendix D—Rating Forms

ID_____

NECAP Writing Grade 5
Rating Form

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

BOW #	Round 1	Round 2	Round 3
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			

Enter your ratings as follows:

1 = Substantially Below Proficient
2 = Partially Proficient
3 = Proficient
4 = Proficient with Distinction

ID_____

NECAP Writing Grade 8
Rating Form

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

BOW #	Round 1	Round 2	Round 3
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			

Enter your ratings as follows:

1 = Substantially Below Proficient
2 = Partially Proficient
3 = Proficient
4 = Proficient with Distinction

Appendix E—EVALUATION FORMS

Content Area: _____
Grade: _____

Standards Transfer Training Evaluation

The purpose of this evaluation form is to obtain your feedback about the training you have received. Please complete the information below. **Do not put your name on the form.** We want your feedback to be anonymous.

Please mark the appropriate box for each statement.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
I understand the goals of the standards transfer meeting.	☐	☐	☐	☐	☐
I understand the procedures we are using to transfer standards.	☐	☐	☐	☐	☐
I understand how to use the standard transferring materials.	☐	☐	☐	☐	☐
I understand the differences between the achievement levels.	☐	☐	☐	☐	☐
I understand how to make the cut score judgment.	☐	☐	☐	☐	☐
I know what tasks to expect for the remainder of the meeting.	☐	☐	☐	☐	☐
I am confident in my understanding of the standards transferring task.	☐	☐	☐	☐	☐

Please indicate any areas in which you would like more information before you continue.

Please indicate any questions you may have about the remainder of the standards transfer meeting.

Content Area: _____
Grade: _____

Standards Transfer Final Evaluation

Please complete the information below. Your feedback will provide a basis for evaluating the training, methods, and materials. **Do not put your name on the form.** We want your feedback to be anonymous.

Gender: Male ☐ Female ☐

Race/ethnicity: White ☐ Black ☐ Hispanic ☐ Asian ☐ Pacific Islander ☐ American Indian ☐

Years of experience in education: 0-5 ☐ 5-10 ☐ 10-15 ☐ More than 15 ☐

Area of Expertise (Check all that apply):

- Students with Disabilities ☐
- Students with Limited English Proficiency ☐
- Economically Disadvantaged Students ☐
- Gifted and Talented Students ☐
- General Education ☐

Please mark the appropriate box for each statement.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
I understood the goals of the standards transfer meeting.	☐	☐	☐	☐	☐
I understood the procedures we used to transfer standards.	☐	☐	☐	☐	☐
The facilitator helped me understand the process.	☐	☐	☐	☐	☐
The materials contained the information needed to transfer standards.	☐	☐	☐	☐	☐
I understood how to use the materials provided.	☐	☐	☐	☐	☐
The achievement level definitions were clear.	☐	☐	☐	☐	☐
I understood how to make the cut score judgments.	☐	☐	☐	☐	☐
I understood how to use the feedback provided after each round.	☐	☐	☐	☐	☐
I understood how to use the impact data.	☐	☐	☐	☐	☐
I understood how the cut scores were calculated.	☐	☐	☐	☐	☐
The facilitator was able to get answers to my questions.	☐	☐	☐	☐	☐
Sufficient time was allotted for training on the standards transfer tasks.	☐	☐	☐	☐	☐
Sufficient time was allotted to complete the standards transfer tasks.	☐	☐	☐	☐	☐
The facilitator helped the standards transfer process run smoothly.	☐	☐	☐	☐	☐

Content Area: _____
Grade: _____
Please rate the usefulness of each of the following:

	Not at all useful				Extremely useful
The opening session.	☐	☐	☐	☐	☐
The small group activities.	☐	☐	☐	☐	☐
Becoming familiar with the assessment.	☐	☐	☐	☐	☐
Articulating the differences between the achievement levels.	☐	☐	☐	☐	☐
Discussions with other participants.	☐	☐	☐	☐	☐

Please rate the influence of the following when transferring standards.

	Not at all influential				Extremely influential
The achievement level definitions.	☐	☐	☐	☐	☐
My expectations of students.	☐	☐	☐	☐	☐
The difficulty of the test materials.	☐	☐	☐	☐	☐
The student responses.	☐	☐	☐	☐	☐
My experience in the field.	☐	☐	☐	☐	☐
Discussions with other participants.	☐	☐	☐	☐	☐
Cut scores of other participants.	☐	☐	☐	☐	☐
Impact data.	☐	☐	☐	☐	☐

Content Area: _____
Grade: _____

Do you believe the final recommended cut score for each of the achievement levels is too low, about right, or too high?

	Too Low	About Right	Too High
Proficient With Distinction/Proficient	☐	☐	☐
Proficient/Partially Proficient	☐	☐	☐
Partially Proficient/Substantially Below Proficient	☐	☐	☐

Please provide any additional comments about this standard transferring process or suggestions as to how the training and process could be improved.

Appendix F—Opening Session PowerPoint

New England Common Assessment Program (NECAP)

Transferring Achievement Standards

Writing Grades 5 and 8

Measured Progress

© 2010 Measured Progress. All rights reserved.

Welcome and Meeting Overview

Measured Progress

Overview of the Standards Transfer Process

Measured Progress

Content Standards

vs.

Achievement Standards

- Content standards = “What”
- Describe the knowledge and skills students are expected to demonstrate by grade level

- Achievement standards = “How well”
- Describe attributes of student performance on the assessment with respect to the achievement level definitions

Measured Progress

What is your job?

- To recommend cut scores for each of the achievement levels that will be used to report results:

  Substantially Below Proficient

  Partially Proficient

  Proficient

  Proficient with Distinction

Measured Progress

We are trying to determine:

- What knowledge, skills and abilities (KSAs) need to be demonstrated to be classified in each achievement level?
- How much is enough?
- What test performance corresponds to each achievement level?

Substantially Below Proficient

Partially Proficient

Proficient

Proficient with Distinction

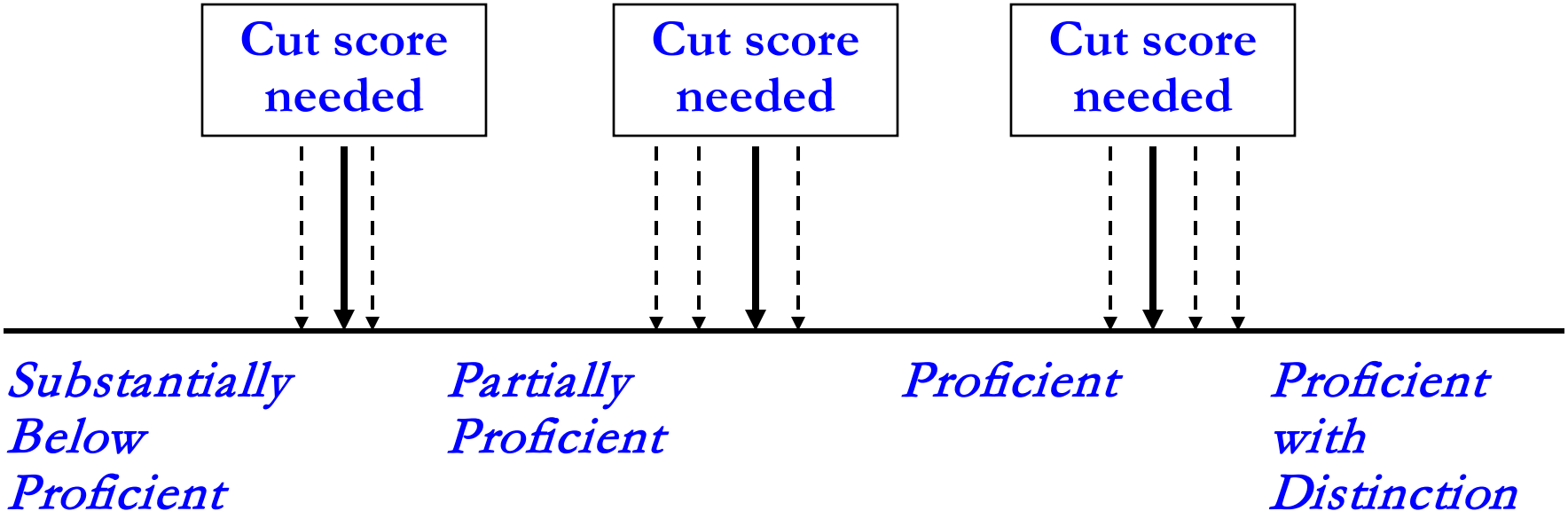
Measured Progress

Achievement Continuum

Substantially Below Proficient	Partially Proficient	Proficient	Proficient with Distinction

Measured Progress

Based on achievement level definitions, you will recommend cut scores…



Achievement Continuum

Measured Progress

Two Phases of the Standards Transfer

- Data Collection
- Policy-making/Decision-making

Measured Progress

Final Recommendations

- Your recommendations will be reviewed and presented to the policy-makers (e.g., technical advisory committee, state boards of education, state commissioners, etc.)
- The policy-makers will be responsible for final adoption of the cut scores.
- The recommendations may be accepted or modified by policy-makers

Measured Progress

Overview of the Standards Transfer Method

Measured Progress

Today's Training

- We will cover: implementation of the body of work procedure
- Note: this session is intended to be an overview
- Your facilitator will give you more details and guide you through the process step by step

Measured Progress

Cut Score Recommendations

Substantially Below Proficient

← Cut Score

Partially Proficient

← Cut Score

Proficient

← Cut Score

Proficient with Distinction

Measured Progress

Factors that Influence Selection of Standard Setting/Transfer Method

- Prior usage/history
- Recommendation/requirement by policy-making authority
- Type of assessment

Body of Work method chosen

Measured Progress

What is the Body of Work Method and How Does it Work?

- The BOW method uses approximately 50 complete sets of student work, including their answers to the writing prompt and a display of their responses to the multiple-choice items
- Bodies of work cover the range of possible scores and are presented in order from lowest to highest total score
- Panelists will classify each BOW into the achievement level in which they feel it belongs:

  Substantially Below Proficient

  Partially Proficient

  Proficient

  Proficient with Distinction

Measured Progress

Important Concepts and Resources

- Achievement Level Definitions
- Test items
- Knowledge, skills and abilities (KSAs) needed to answer each question
- Cut scores

Measured Progress

Steps for Body of Work Method

Measured Progress

Before You Do Your Ratings:

- **Take the test** to familiarize yourself with the test items and the test-taking experience
- Review and discuss the **achievement level definitions** and **exemplar papers**
- Develop **lists of characteristics** (KSAs) that define performance in each achievement level

Measured Progress

Exemplar Papers

- Bodies of work selected from the most recent administration of the same prompt type
- Set of approximately 12 exemplars, chosen to represent the four achievement levels
- You will use these exemplars to help “flesh out” your definitions of the four achievement levels

Measured Progress

You must reach consensus as a group about the knowledge, skills and abilities that define student performance at each achievement level.

Measured Progress

Practice Round / Check for Understanding

- You will have an opportunity to familiarize yourself with the rating process using 5 practice BOWs
- Your facilitator will check with you for understanding and answer any questions you may have during and after the practice round
- You will then complete a training evaluation form

Measured Progress

Actual Ratings: Three Rounds

- Round 1 (without discussion)

  Work through the BOWs and make your initial judgments as to how each should be categorized

- Round 2 (with group discussion)

  Discuss the Round 1 ratings – focus on the KSAs

  Examine your ratings in relation to the rest of the group

  Review and revise ratings as appropriate

- Round 3 (with group discussion)

  Discuss the Round 2 ratings – focus on the KSAs

  Examine your ratings in relation to the rest of the group and impact data

  Review and revise ratings as appropriate

Measured Progress

A Few Reminders:

- It is **not** necessary for panelists to reach consensus as to how to rate each BOW.
- You should be open-minded when listening to your colleagues' rationales for their ratings.
- You may or may not change your mind as a result of the discussions.
- We want each panelist to use his or her own **best judgment** in each round of rating.

Measured Progress

Questions about the Body of Work Method?

Measured Progress

After the ratings…

- The final step in the process will be for you to complete an evaluation of the standards transfer process.
- Your honest feedback is important!

Measured Progress

What’s next?

- Take the test
- Discuss the Achievement Level Definitions
- Practice Round
- Round 1
- Round 2
- Round 3
- Evaluation

Measured Progress

And that’s it…

Please be sure to ask your facilitator any questions you may have about the Body of Work procedure.

Measured Progress

Before you break into groups…

Measured Progress

Top 10 Misconceptions about Standard Setting:

10. Standard setting is a great opportunity to review and revise the test.
9. Standard setting is the same thing as scoring.
8. This is a good time to discuss the test administration policy.
7. This is a good time to revise the content standards.
6. This is a good time to revise the Achievement Level Definitions.

Measured Progress

Top 10 Misconceptions about Standard Setting:

5. This is a good time to discuss effective teaching strategies.
4. Only scholars and researchers are qualified to do this work.
3. Only educators are qualified to do this work.
2. The process is rigged.
1. Disagreement is bad.

Measured Progress

Role of the Facilitator

- Lead and keep the group on track
- Ensure that all panelists clearly understand the procedures
- Ensure that the rating forms are completed
- Ensure that the evaluation forms are completed.

Measured Progress

Ground Rules

- Process is focused solely on recommending achievement standards (cut scores).
- The achievement levels and their definitions are not open for debate.
- Panelists' recommendations are crucial, but final cut score decisions will be made by policy-makers of the NECAP states.
- Each panelist is expected to participate in the entire process or his/her judgments will be discounted.
- Please: no cell phone use except during breaks.
- Please be sure to return promptly after breaks and in the morning.

Measured Progress

Good Luck!

Measured Progress

Appendix G—FACILITATOR SCRIPT

GENERAL INSTRUCTIONS FOR

NEW ENGLAND COMMON ASSESSMENT PROGRAM (NECAP)

STANDARD TRANSER GROUP FACILITATORS

WRITING GRADES 5 AND 8

DECEMBER 8-9, 2010

Preliminaries

Introductions:

- Welcome group; introduce yourself (name, affiliation, a little background information).
- Have each participant introduce him/herself.
- Ask each participant to sign a nondisclosure form. Do not proceed until a signed nondisclosure form has been collected from each participant.

Take the Test

Overview: In order to establish an understanding of the test items and for panelists to gain an understanding of the experience of the students who take the test, each participant will take the test for their grade level and content area. Panelists may wish to discuss or take issue with the items in the test. Tell them we will gladly take their feedback to the NECAP state representatives. However, this is the actual assessment that students took and it is the set of items on which we must set standards.

Activities:

- Introduce the assessment and convey/do each of the following:
    - Tell panelists that they are about to take the actual NECAP assessment.

  - The purpose of the exercise is to help them establish a good understanding of the test items and to gain an understanding of the experience of the students who take the assessment. Let panelists know they do not need to completely answer the constructed-response questions; they can just jot down a few notes.
- Give each panelist a test booklet.
- Tell panelists to try to take on the perspective of a student as they complete the test.
- When the majority of the panelists have finished, pass out the answer key/scoring guides.

Discuss Achievement Level Definitions (ALDs)

Overview: In order to establish an understanding of the expected performance of students on the test, panelists must have a clear understanding of:

- the definition of the four achievement levels, and
- what the key characteristics are that distinguish students in adjacent achievement level categories.

The purpose of this activity is for the panelists to come to consensus about what characterizes students in each achievement level category, starting with the Proficient achievement level. This activity is critical since the ratings panelists will be making will be based on these understandings.

Activities:

- Introduce the task. In this activity they will:
  - individually review the Achievement Level Definitions;
  - individually review the exemplar papers;
  - discuss the definitions and exemplars as a group, starting with the Proficient level; and
  - develop lists of the knowledge, skills and abilities (KSAs) that characterize each achievement level.
- The goal here is for the panelists to come to a common understanding of what it means to be in each achievement level. It is not unusual for panelists to disagree with the definitions they will see; almost certainly there will be some panelists who will want to change them. However, the task at hand is for panelists to have a common understanding of what knowledge, skills, and abilities (KSAs) are described by each of the Achievement Level Definitions.
- Pass out the Achievement Level Definitions and exemplar papers
- Orient panelists to the exemplar papers. These papers have been sampled from the 2007 operational administration of the Writing test. The 2007 administration was selected because it was the last time the item prompt genre in the 2010 test was administered. Walk them through the first exemplar, pointing out that it

consists of the full set of student responses to the writing items, including a display of the multiple-choice items which shows the text of the item, either an indicator showing that the student got the item correct (+) or the distractor they chose, and the percent of students who got the item correct. Point out that the exemplar papers are presented in order, from lowest scoring to highest scoring.

- Have panelists review the ALDs and exemplars individually. Panelists can make notes if they like.
- After individually reviewing the definitions and exemplars, have panelists discuss each definition as a group, starting with Proficient. Referring to the exemplars categorized as Proficient, the panelists will discuss the KSAs a student must demonstrate in order to be classified in the Proficient category. Or, put another way, have panelists identify the most important KSAs that distinguish a Partially Proficient student from a student in the Proficient category.
- After completing Proficient, the panelists will repeat this process for the Partially Proficient and Proficient with Distinction categories, using the exemplars corresponding to each achievement level.
- Have the panelists identify the most important KSAs describing students at each achievement level and record those as bulleted lists on chart paper. These should be posted on the walls for panelists to refer to as they complete the three rounds of rating.
- The purpose of this step is to have a collegial discussion in which to bring up/clarify any issues or questions that any individual may have and to reach consensus on an understanding of the ALDs.

Practice Round

Overview of Practice Round: The primary purpose of the practice round is for panelists to become familiar with the rating task with five practice BOWs, sampled from the 2007 administration. Panelists will review the BOWs with the facilitator and discuss them as a group. The facilitator will walk the panelists through the practice ratings, engage the panelists in a readiness discussion and check for understanding. If any of the panelists indicate an incomplete understanding of the practice rating task, then the facilitator will continue to work with the panelists to clarify any misconceptions before proceeding to Round 1.

Activities:

- Make sure the panelists have the following materials:
    - Set of practice BOWs.
    - Scoring guides.
    - Achievement Level Definitions.
    - Practice rating form.

- Give the panelists a few minutes to briefly review the set of BOWs. As they are reviewing the BOWs, they should be thinking about the knowledge, skills, and abilities demonstrated in each BOW and how they relate to the ALDs.
- After the panelists have had a few minutes to review the BOWs, begin discussing them as a group:
    - For the first BOW walk through it as a group and, referring to the scoring guides, discuss the KSAs represented by each item. Discuss the ALDs and determine which achievement level best fits the KSAs represented in this BOW. This BOW has been selected to be clearly in Substantially Below Proficient.
    - For the second BOW, again, walk through it as a group and discuss the KSAs represented by each item. This BOW is one that could be considered to be either in Substantially Below Proficient or Partially Proficient. Lead them through a discussion of their rationale. Remind panelists that for a BOW to fall into Partially Proficient they are looking for the BOW to be **just over the line** between Substantially Below Proficient and Partially Proficient – they are not looking for a solid Partially Proficient performance.
    - For the third BOW, have the panelists rate it on their own and then lead them through a discussion of their rationale. Ask for a show of hands indicating which panelists placed the BOW in each of the achievement levels. Have the panelists discuss the rationale behind their placements. Repeat this same process two more times for the remaining BOWs. Note that the fifth BOW will be chosen to represent the upper end of the score scale.
    - Make sure panelists understand that they do not need to agree about the placement of the practice BOWs; the purpose here is to understand the process of rating the BOWs based on their consensus definitions of the ALDs and the KSAs represented by the items.
- Panelists may want to take notes as they work if there are particular points they would like to discuss with their colleagues.
- Go over the practice rating form with panelists:
    - Have panelists write their ID number on the rating form. The ID number is on their name tags.
    - Lead panelists through a step-by-step demonstration of how to fill in the rating form.
    - Collect the rating forms and check to make sure they have been completed accurately.
    - NOTE: you do not need to return the practice rating forms to the data analysis work room.

Readiness Discussion

After the panelists have completed the practice rating form, lead a readiness discussion by posing the questions below.

The purpose of this discussion is to determine how well each panelist understands the rating task, to correct any misunderstandings, and if necessary, to identify panelists whose ratings should be excluded from the standard transfer if their understanding doesn't improve.

The "correct" answers for each of the questions are listed directly under each question. Please guide the discussion to ensure each panelist understands the rating task. **Make sure any questions or concerns are resolved prior to moving on.**

- What type of information should I be considering when making my classification?
    - The ALDs.
    - The holistic set of knowledge, skills and abilities represented in each BOW.
    - My understanding of who these students are.

- Why is it important to refer to the scoring guides and the ALDs?
    - The scoring guides specify the knowledge, skills and abilities demonstrated by the student who obtained each score on a given item. It is these KSAs that need to correspond to the KSAs outlined in each achievement level.

- If a student profile reflects KSAs that "just barely" correspond to Partially Proficient, what achievement level should be assigned to the BOW?
    - Partially Proficient: A BOW does not need to represent solid Partially Proficient performance to be classified in that level, it just needs to make it into the category.

Training Evaluation

Have panelists fill out the training evaluation form before proceeding to Round 1. Before you start the Round 1 activities, scan the completed evaluations to see if there are any problems or concerns that need to be addressed before proceeding. Return the completed evaluations to the data analysis work room at the next convenient opportunity.

Round 1 Ratings

Overview of Round 1: The primary purpose of Round 1 is to ask the panelists to make their initial determination as to which achievement level category each BOW should be classified into. In this round, panelists will be working individually, without discussion with their colleagues.

Activities:

- Make sure panelists have the following materials:
  - Rating form
  - Set of BOWs
  - Achievement Level Definitions
- Orient panelists to the set of BOWs. Point out that the BOWs are presented in order, from lowest scoring to highest. Make sure panelists understand that, even though the BOWs are presented from lowest- to highest-scoring, their own ratings do not need to be in strictly increasing order.
- Starting with the first BOW, the panelists will review each BOW in turn. As they are reviewing the BOWs, the panelists should keep in mind the Achievement Level Definitions. They should consider the knowledge, skills, and abilities demonstrated by each and how they relate to the definitions of the achievement levels. The purpose of this step is for panelists to make their initial determinations as to how the BOWs should be categorized into the four achievement levels. The panelists are free to make notes on the BOWs, sort them into piles, use sticky notes, or use whatever system helps them to keep track of their categorizations.
- Panelists may want to take notes as they work if there are particular points they would like to discuss with their colleagues in Round 2.
  - Have panelists write their ID and round number on the rating form. The ID number is on their name tags.
  - Briefly remind them how to fill in the rating form.
  - Answer questions the panelists may have about the work in Round 1.
  - Once everyone understands what they are to do in Round 1, tell them to begin.
- As panelists complete the task, ask them to carefully inspect their rating forms to ensure they are filled out properly.
  - The ID and round number must be filled in.
  - Each BOW must be assigned to one and only one achievement level.
  - Reiterate that although the BOWs are presented in order from lowest- to highest-scoring, the panelists' category assignments do not need to be in strictly increasing order.
- Facilitators should bring all the completed rating forms together to the data analysis work room for tabulation. Prior to submitting them, however, using a show of hands, indicate on a piece of chart paper how many panelists assigned

each BOW to each achievement level category. This chart will be used for the Round 2 discussions.

Tabulation of Round 1 Results

Tabulation of Round 1 results will be completed as quickly as possible after receipt of the rating forms. While the tabulation occurs, the group may take a break.

Round 2 Ratings

Overview of Round 2: In Round 2, the panelists will have an opportunity to discuss their Round 1 placements and to revise their ratings on the basis of that discussion. Prior to beginning the Round 2 discussions, the psychometrician will share the group's average cut points based on the Round 1 ratings.

Focusing on any BOWs for which there was disagreement as to how they should be categorized, the panelists will discuss why they categorized each BOW as they did, making sure that all different points of view are included in the discussion.

Once panelists have reviewed and discussed the Round 1 categorizations, the panelists will make their Round 2 ratings.

Activities:

- Make sure panelists have the following materials:
    a. Rating form
    b. Set of BOWs
    c. Achievement Level Definitions
- Have panelists write their ID number and Round 2 on the rating form.
- Provide an overview of Round 2. Paraphrase the following:
    a. As in Round 1, the primary purpose is to categorize each BOW into the achievement level category where you believe it belongs.
    b. Each panelist needs to base his/her judgments on his/her experience with the content area, understanding of students, the definition of each achievement level category, discussions with other panelists, and the knowledge, skills, and abilities required to answer each item.
- Show the panelists how the BOWs would be categorized based on the room average Round 1 cut point placements.
- Remind panelists that they will be discussing each BOW with their colleagues, but that they will be categorizing the BOWs individually. It is not necessary for the panelists to reach consensus about how to categorize each BOW.
- Give panelists an opportunity to ask questions about the task for Round 2.
- Beginning with the first BOW for which there is disagreement as to how it should be categorized, the panelists should begin discussing the categorization of the BOWs according to the Round 1 ratings.
    a. Panelists only need to discuss those BOWs for which there was disagreement as to how they should be categorized.
    b. Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
    c. If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.

d. On the basis of the discussions and the feedback presented, panelists should make their Round 2 ratings.
e. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating with which they disagree.

Encourage the panelists to use the discussion and feedback to assess how stringent or lenient a judge they are. If a panelist is categorizing BOWs consistently higher or lower than the group, he or she may have a different understanding of the Achievement Level Definitions than the rest of the group. **It is acceptable for panelists to disagree, but that disagreement should be based on a common understanding of the Achievement Level Definitions.**

- When the group has completed their Round 2 ratings, collect the rating forms. When you collect the rating forms, carefully inspect them to ensure they are filled out properly.
  a. The ID and round number must be filled in.
  b. Each BOW must have one (and only one) rating.
- Facilitators should bring all the completed rating forms together to the data analysis work room for tabulation. Prior to submitting them, however, using a show of hands, indicate on a piece of chart paper how many panelists assigned each BOW to each achievement level category. This chart will be used for the Round 3 discussions.

Tabulation of Round 2 Results

Tabulation of Round 2 results will be completed as quickly as possible after receipt of the rating forms. While the tabulation occurs, the group may take a break.

Round 3 Ratings

Overview of Round 3: In Round 3, the panelists will have a final opportunity to discuss their Round 2 placements and to revise their ratings on the basis of that discussion. Prior to beginning the Round 3 discussions, the psychometrician will share the Round 2 results with the group, including the group's average cut points and associated impact data (i.e., the approximate percentage of students who would be classified into each achievement level category based on the room average cut points from Round 2). In addition, the psychometrician will present the percentages of students classified as Proficient or above for previous administrations of the Writing test. This information will provide the panelists with some context within which to evaluate their cuts and the associated impact data.

Once panelists have reviewed and discussed the Round 2 categorizations, they will make their final ratings.

Activities:

- Make sure panelists have the following materials:
    - Rating form
    - Set of BOWs
    - Achievement Level Definitions
- Have panelists write their ID number and Round 3 on the rating form.
- Provide an overview of Round 3. Paraphrase the following:
    - As in Round 2, the primary purpose is to categorize each BOW into the achievement level category where you believe it belongs.
    - Each panelist needs to base his/her judgments on his/her experience with the content area, understanding of students, the definition of each achievement level category, discussions with other panelists, and the knowledge, skills, and abilities required to answer each item.
- Review the feedback information with the panelists.
    - Show the panelists how the BOWs would be categorized based on the room average Round 2 cut point placements.
    - Go over the impact data, explaining that if the Round 2 ratings were to be used to set the final cut points, these are the approximate percentages of students who would be classified into each achievement level category.
    - Panelists will use the impact data as a reasonableness check. Specifically, panelists will discuss whether the percentages of students in each achievement level make sense and, if not, which cut or cuts may need to be adjusted.
    - In addition to the Round 2 impact data, panelists will also have the percentages of students that were classified as Proficient or above based on previous administrations of the Writing test. The panelists will use that information to evaluate the impact data based on their Round 2 cuts. Specifically, panelists should discuss whether the impact data seem

reasonable, given previous performance on the Writing test and the expected amount of growth.

- Remind panelists that they will be discussing each BOW with their colleagues, but that they will be categorizing the BOWs individually. It is not necessary for the panelists to reach consensus about how to categorize each BOW.
- Give panelists an opportunity to ask questions about the feedback information or about the task for Round 3.
- Beginning with the first BOW for which there is disagreement as to how it should be categorized, the panelists should begin discussing the categorization of the BOWs according to the Round 2 ratings.
    - Panelists should discuss the BOWs for which there was disagreement as to how they should be categorized, focusing in particular on the BOWs around the cuts.
    - Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
    - If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
    - On the basis of the discussions and the feedback presented, panelists should make their final ratings.
    - The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating with which they disagree.

  Encourage the panelists to use the discussion and feedback to assess how stringent or lenient a judge they are. If a panelist is categorizing BOWs consistently higher or lower than the group, he or she may have a different understanding of the Achievement Level Definitions than the rest of the group. **It is acceptable for panelists to disagree, but that disagreement should be based on a common understanding of the Achievement Level Definitions.**

- When the group has completed their final ratings, collect the rating forms. When you collect the rating forms, carefully inspect them to ensure they are filled out properly.
    - The ID and round number must be filled in.
    - Each BOW for Round 3 must have one (and only one) rating.
    - Return the completed rating forms to the data analysis work room.

Complete Final Evaluation Form

After the rounds of ratings are completed, have the panelists fill out the final evaluation form. Emphasize that their honest feedback is important.

Appendix H—PANELIST AFFILIATIONS

State	First Name	Last Name	Grade	Association/Affiliation	Position
VT	Janice	Garrow	5	Rutland Int. School	Gr. 6 Teacher
VT	Elaine	Collins	5	Browington Central School	Grs. 6-8 Humanities Teacher
VT	Carol	Owen	5	Orange East S.U.	Literacy Teacher Leader
VT	Kathryn	Grace	5	Founders School	Literacy Teacher Leader
VT	Geoff	Lawrence	8	Proctor Jr./Sr. High	MS English Teacher
VT	Maggie	Eaton	8	U-32 Middle/ High	Teacher/ Curriculum Leader
VT	Kimberly	Means	8	Stern Center	Special Educator
NH	Michelle	Trafton	5	Milton Elementary	Gr. 5 Teacher
NH	Cynthia	Douglass	5	Oyster River Middle	Reading Specialist
NH	Jill	Duffield	5	New Durham School	Literacy Teacher
NH	Kathleen	Woodbury	5	Gilsum School	Reading Teacher, Special Education Care Manager, Assessment Coordinator
NH	Bonnie	Robinson	8	Newport Middle High	Reading Specialist
NH	Alison	Bryant	8	Belmont High School	English Teacher
RI	Alison	Santerre	5	Kingston Hill Academy	Gr. 5 Teacher
RI	Patricia	Vecchione	5	Winsor Hill School	Literacy Coach/ Reading Specialist
RI	Kathleen	Desrosiers	5	Warwick S.D.	Coordinator/ ELA/ K-12
RI	Kevin	Montoya	5	Segue Institute for Learning	Gr. 6 ELA
RI	Melissa	Lourenco	8	Segue Institute for Learning	Literacy Coach/ Title 1
RI	Elizabeth	Ferguson	8	Curtis Corner Middle & South Kingstown	Teacher/ Curriculum Coordinator
RI	Janice	Place	8	Pilgrim High & Aldrich Jr. High Schools	English Dept. Chairperson/ Teacher
RI	Brenda	Aspelund	8	Aldrich Jr. High School	ELA Teacher
RI	Sally	Caruso	8	Kickemuit Middle	Gr. 8 ELA Teacher
ME	Gloria	Westhrin	5	Glenburn School	Gr. 4 Teacher
ME	Jo Ellen	Merry	5	Houlton Southside	Gr. 5 Teacher
ME	Jan	Obery	5	Gilbert Elementary School	Literacy Specialist
ME	Patricia	Friesland	8	Camden Rockport Middle	Teacher
ME	Kathy	Kauffman	8	Oak Hill Middle School	Gr. 8 ELA
ME	Beth	Ahlholm	8	Medomak Middle	7-8 LA/SS Teacher
ME	Wendy	Dunbar	8	Mt. Jefferson Jr. High	Literacy Specialist/ ELA Teacher

Appendix I—EVALUATION RESULTS

2010 NECAP Writing Standard Setting Report: Grade 5 Training Evaluation

	N	Mean	% SD	% D	% U	% A	% SA
I understand the goals of the standard setting meeting.	15	4.53	0	0	7	33	60
I understand the procedures we are using to set standards.	15	4.2	0	0	7	67	27
I understand how to use the standard setting materials.	15	4.27	0	0	0	73	27
I understand the differences between the achievement levels.	15	4.53	0	0	0	47	53
I understand how to make the cut score judgment.	15	3.87	0	0	33	47	20
I know what tasks to expect for the remainder of the meeting.	15	4.2	0	7	0	60	33
I am confident in my understanding of the standard setting task.	14	4.21	0	0	7	64	29

Please indicate any areas in which you would like more information before you continue.
Review the use of multiple choice responses in the decision making process. Do the multiple choice responses in the decision making process. Do the multiple choice items provide/prevent a pass-through to the next level?
The second day's procedures will be helpful in a more thorough understanding of the cut score.
I'm all set
Not 100% sure about the 3 rounds tomorrow - will be scoring the same 50 samples 3 times or given 50 different samples each time?
What happens if two scorers have a huge discrepancy in their scores? Example someone gives score of substantially below and another gives a score of proficient?
I need more direct and explicit instructions BEFORE the tasks begin. I feel as though we are asked to start and then new points are given which often skew the work that was already completed
I am very knowledgeable of the standards so my background knowledge was more helpful than the facilitator. The facilitator was not always clear about the task at hand which often led to confusions in the group. Her answers changed from minute to minute. Each task took much longer than necessary because of this.
Discussing scaling of grades and determining cut scores. I am a bit confused as to how our work tomorrow will assist the NECAP process and long term implications.

Please indicate any questions you may have about the remainder of the standard setting meeting.

No question. Great experience. Well picked team.

No questions

This is not a question but a comment - once it was clear that we were looking at each student's body of work holistically the standards transfer task became clearer. Thank you, Luz!

No questions at this time

Once the "metric folks" have computed the cut scores base on our work; does that get discussed further by policy makers in the second phase of standards transfer?

I believe if someone had discussed the standard vocabulary in terms of school data teams at the outset of the training, some of the confusion could have been avoided. The language of the facilitator (re: language of psychometric standards) did not seem to match the language of "teacher standards info".

What is it that we are truly providing Measured Progress with?

My question is actually a comment about the process. This particular protocol of looking at a body of work, including the multiple-choice items, muddies the water a bit. If we are determining proficiency for a writing piece, and the characteristics of why it is proficient, it would be helpful to have a proficient piece of writing to look at (rather than a proficient BOW). Also, not understanding how the multiple-choice answers affect the scores makes the determination more difficult.

2010 NECAP Writing Standard Setting Report: Grade 5 Final Evaluation

Panelist Demographics	N = 15
Gender:	
Male	1
Female	14
Race/Ethnicity:	
White	14
Black	0
Hispanic	1
Asian	0
Pacific Islander	0
American Indian	0
Years of Educational Experience:	
0-5	1
5-10	0
10-15	3
More than 15	11
Profesional Experience (check all that apply):	
Students with Disabilities	8
Students with Limited English Proficiency	0
Economically Disadvantaged Students	7
Gifted and Talented Students	3
General Education	15

	N	Mean	% SD	% D	% U	% A	% SA
I understood the goals of the standard setting meeting.	15	4.47	0	0	0	53	47
I understood the procedures we used to set standards.	15	4.27	0	0	7	60	33
The facilitator helped me understand the process.	15	4	0	0	13	73	13
The materials contained the information needed to set standards.	15	4.33	0	0	7	53	40
I understood how to use the materials provided.	15	4.13	0	0	7	73	20
The achievement level descriptions were clear.	15	4.27	0	0	7	60	33
I understood how to make the cut score judgments.	15	4.27	0	0	0	73	27
I understood how to use the feedback provided after each round.	15	4.33	0	0	7	53	40
I understood how to use the impact data.	15	4.2	0	0	7	67	27

	N	Mean	% SD	% D	% U	% A	% SA
I understood how the cut scores were calculated.	14	4.07	0	0	14	64	21
The facilitator was able to get answers to my questions.	15	4.2	0	0	7	67	27
Sufficient time was allotted for training on the standard setting tasks.	15	4.27	0	0	0	73	27
Sufficient time was allotted to complete the standard setting tasks.	15	4.33	0	0	0	67	33
The facilitator helped the standard setting process run smoothly.	15	3.87	0	7	20	53	20

Usefulness of	N	Mean	%Low				%High
The opening session.	15	3.93	0	7	20	47	27
The small group activities.	14	4.21	0	0	14	50	36
Becoming familiar with the assessment.	15	4.27	0	0	13	47	40
Articulating the differences between the achievement levels.	15	4.4	0	0	7	47	47
Discussions with other participants.	15	4.6	0	0	7	27	67

Influence of	N	Mean	%Low				%High
The performance level descriptors.	15	4.2	0	0	20	40	40
My expectations of students.	15	4.13	0	0	20	47	33
The difficulty of the test materials.	15	4.07	0	7	13	47	33
The student responses.	15	4.6	0	0	7	27	67
My experience in the field.	15	4.67	0	0	0	33	67
Discussions with other participants.	15	4.6	0	0	7	27	67
Cut scores of other participants.	15	4.13	0	0	20	47	33
Impact data.	15	4.27	0	0	7	60	33

Please provide any additional comments about the standard setting process or suggestions as to how the training and process could be improved.

#11 should have had modification of testing administration. New to school? Unable to have an accommodations meeting? This student had a message. A word processor would have been a big bonus.

It was not made clear that we should have been using the exemplar book from the earlier test when scoring until the 3rd round. This changed a lot of opinions on the student samples. I also believe that our scoring would have been different had the pieces not been ranked. I found that this strongly influenced my scoring of pieces.

While it was helpful to have a facilitator who understands the "number crunching" it was often hard to discern what was being asked of us. Furthermore, clear group "rules" should be set at the forefront.

Let the panel use the exemplar book to refer to as benchmark papers when scoring 1st and 2nd round.

I believe facilitators need to set the tone as to how meetings will function. I know how difficult it is sometimes to guide professionals but setting norms and guidelines prior to group discussions is crucial. Much of what people said was lost because there were times that many took the spotlight.

bit of confusion on how #'s impacted ea. other. Wish we compared with exemplars earlier. Great experience though. Thx =-)

The idea of comparing the exemplars to the 2010 writing, I believe, should have happened earlier in the process. I also will repeat my earlier comment. It would be easier to not include multiple choice questions in the holistic scoring, as it is unclear to me how much to count them for.

I believe facilitators should continually remind folks, if needed, to follow discussion protocol. Perhaps "rules" about one speaker at a time, raising hands, no side conversations, should be articulated @ the outset. Most people intuitively know and follow these rules. However, it takes just one participant (w/ ADHD??) to ruin the experience for everyone.

#34, #11 seemed out of place in the rank ordering despite lots of discussion.

Establish some procedural conventions at sharing/discussion time would help with giving everyone an equal voice. Thank you for an extremely rewarding experience.

2010 NECAP Writing Standard Setting Report: Grade 8 Training Evaluation

	N	Mean	% SD	% D	% U	% A	% SA
I understand the goals of the standard setting meeting.	14	4.79	0	0	0	21	79
I understand the procedures we are using to set standards.	14	4.64	0	0	0	36	64
I understand how to use the standard setting materials.	14	4.71	0	0	0	29	71
I understand the differences between the achievement levels.	14	4.64	0	0	0	36	64
I understand how to make the cut score judgment.	14	4.14	0	7	7	50	36
I know what tasks to expect for the remainder of the meeting.	14	4.43	0	0	7	43	50
I am confident in my understanding of the standard setting task.	14	4.43	0	0	21	14	64

Please indicate any areas in which you would like more information before you continue.
I feel confident that I understand the purpose and process
If two pieces are adjacent and there are just a few (1 or 2) elements that are proficient, should it still be scored as proficient?
N/A
All Set
I would like to clarify making the cut score judgment....one more time. I just tried to study the achievement levels to "fully" understand the differences.
More training on what (other than a blank) constitutes a zero.
As we go along, everything I need to know has been revealed in a timely manner. I'm fine.

Please indicate any questions you may have about the remainder of the standard setting meeting.
I don't have any at this time
Very Good facilitation
N/A
Can we keep materials (from 2007 prompts/tests) for our own training purposes?
None
I'm confident that any questions I have will be answered in session 2.

2010 NECAP Writing Standard Setting Report: Grade 8 Final Evaluation

Panelist Demographics	N = 14
Gender:	
Male	1
Female	13
Race/Ethnicity:	
White	14
Black	0
Hispanic	0
Asian	0
Pacific Islander	0
American Indian	0
Years of Educational Experience:	
0-5	1
5-10	2
10-15	2
More than 15	9
Profesional Experience (check all that apply):	
Students with Disabilities	2
Students with Limited English Proficiency	1
Economically Disadvantaged Students	2
Gifted and Talented Students	2
General Education	13

	N	Mean	% SD	% D	% U	% A	% SA
I understood the goals of the standard setting meeting.	14	4.64	0	0	0	36	64
I understood the procedures we used to set standards.	14	4.57	0	0	0	43	57
The facilitator helped me understand the process.	14	4.57	0	0	0	43	57
The materials contained the information needed to set standards.	14	4.29	0	0	7	57	36
I understood how to use the materials provided.	14	4.5	0	0	0	50	50
The achievement level descriptions were clear.	14	4.5	0	0	0	50	50
I understood how to make the cut score judgments.	14	4.57	0	0	0	43	57
I understood how to use the feedback provided after each round.	14	4.64	0	0	0	36	64
I understood how to use the impact data.	14	4.5	0	0	7	36	57

	N	Mean	% SD	% D	% U	% A	% SA
I understood how the cut scores were calculated.	14	4.64	0	0	0	36	64
The facilitator was able to get answers to my questions.	14	4.71	0	0	0	29	71
Sufficient time was allotted for training on the standard setting tasks.	14	4.64	0	0	0	36	64
Sufficient time was allotted to complete the standard setting tasks.	14	4.71	0	0	0	29	71
The facilitator helped the standard setting process run smoothly.	14	4.64	0	0	0	36	64

Usefulness of	N	Mean	%Low				%High
The opening session.	14	4.07	0	0	14	64	21
The small group activities.	13	4.38	0	0	8	46	46
Becoming familiar with the assessment.	14	4.43	0	7	0	36	57
Articulating the differences between the achievement levels.	14	4.36	0	0	14	36	50
Discussions with other participants.	14	4.86	0	0	0	14	86

Influence of	N	Mean	%Low				%High
The achievement level descriptions.	13	4.54	0	0	0	46	54
My expectations of students.	13	3.77	0	8	31	38	23
The difficulty of the test materials.	13	3.46	8	8	23	54	8
The student responses.	13	4.54	0	0	0	46	54
My experience in the field.	13	4.38	0	0	8	46	46
Discussions with other participants.	13	4.46	0	0	15	23	62
Cut scores of other participants.	13	3.62	0	15	31	31	23
Impact data.	13	3.69	0	8	23	62	8

Please provide any additional comments about the standard setting process or suggestions as to how the training and process could be improved.

One member of our group is a 12th grade teacher, and while she taught 8th grade 4 years ago, she has lost touch with the expectations. She is a nice person, but her selection for this co. was not advised. I commend all other members of the group. A new standard setting seems to be needed since many of us did not agree with the scoring of the exemplars. I would love to work on that committee.

The exemplars for "lowest" proficient and "highest" proficient were seemingly too low - which caused confusion and some discord.

I wish individual scores appeared on each exemplar response; it would have helped me when weighing the various pieces in the binder and better link them to exemplar b.o.w.

The scores of the exemplars would have helped.

I feel that we transferred the standards from the exemplars to the test material, appropriately - though I am quite concerned about the influence of the impact data -

Students know the test and a lot of emphasis has been put on writing in schools.

I think that the achievement level definitions should have guided our discussion more. We could have defined and fleshed out each level more effectively by posting each level on the wall (w/ its definition from the rubric) and going from there - mapping or creating a web around it. I am grateful that I've had the opportunity to work on this committee. I would be interested to be involved again - the accommodations were great, also. Thanks!

Appendix N—Performance Level Score Distributions

Table N-1. 2010–11 NECAP: Performance Level Distributions by Grade – Mathematics

Grade	Performance Level	Percent in Level		
		2010-11	2009-10	2008-09
3	4	21.48	19.29	19.42
	3	45.51	47.66	47.27
	2	20.09	18.98	19.18
	1	12.92	14.06	14.13
4	4	20.62	18.02	22.25
	3	45.29	48.82	46.90
	2	20.72	19.12	17.74
	1	13.37	14.03	13.12
5	4	18.93	20.31	19.60
	3	46.74	46.59	48.24
	2	16.95	15.86	14.92
	1	17.38	17.25	17.24
6	4	23.77	22.55	20.08
	3	41.99	42.75	43.44
	2	17.15	17.62	17.80
	1	17.08	17.08	18.68
7	4	18.89	20.24	18.91
	3	41.44	40.61	41.44
	2	19.66	19.17	18.40
	1	20.01	19.98	21.25
8	4	18.33	18.64	16.93
	3	42.15	42.39	43.39
	2	20.33	19.56	20.18
	1	19.19	19.42	19.50
11	4	2.67	2.23	1.54
	3	32.30	29.70	29.45
	2	28.30	28.14	29.04
	1	36.73	39.93	39.98

Table N-2. 2010–11 NECAP: Performance Level Distributions by Grade – Reading

Grade	Performance Level	Percent in Level		
		2010-11	2009-10	2008-09
3	4	14.45	16.82	20.91
	3	59.19	58.31	52.96
	2	16.83	15.76	15.26
	1	9.53	9.10	10.87
4	4	21.03	17.85	21.50
	3	50.05	52.42	50.20
	2	19.16	18.40	18.05
	1	9.76	11.33	10.25
5	4	19.67	17.82	15.86
	3	54.22	56.50	55.82
	2	18.45	17.81	19.86
	1	7.66	7.87	8.47
6	4	16.62	15.19	13.68
	3	57.27	55.82	57.32
	2	18.31	20.60	19.74
	1	7.80	8.38	9.26

Grade	Performance Level	Percent in Level		
		2010-11	*2009-10*	*2008-09*
7	4	13.80	13.50	16.90
	3	54.43	58.54	57.91
	2	22.16	19.99	17.71
	1	9.61	7.96	7.48
8	4	21.99	19.72	14.58
	3	53.34	52.40	53.75
	2	19.31	20.78	22.37
	1	5.36	7.10	9.31
11	4	26.50	22.97	18.94
	3	47.43	49.80	52.37
	2	17.04	17.50	18.56
	1	9.04	9.72	10.13

Table N-3. 2010–11 NECAP: Performance Level Distributions by Grade – Writing

Grade	Performance Level	Percent in Level		
		2010-11	*2009-10*	*2008-09*
5	4	12.49		
	3	39.62		
	2	39.81		
	1	8.08		
8	4	10.91		
	3	48.94		
	2	31.83		
	1	8.32		
11	4	1.06	6.95	3.62
	3	46.96	44.82	37.19
	2	45.58	41.23	49.15
	1	6.40	6.99	10.05

Appendix O—Raw to Scaled Score Look-up Tables

Table O-1. 2010–11 NECAP: Raw to Scaled Score Correspondence – Mathematics Grade 3

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	300	10.00	1	300	10.00	1
1	300	10.00	1	300	10.00	1
2	300	10.00	1	300	10.00	1
3	300	10.00	1	300	10.00	1
4	300	10.00	1	300	10.00	1
5	300	10.00	1	300	10.00	1
6	300	9.60	1	300	10.00	1
7	303	7.60	1	303	8.80	1
8	307	6.30	1	308	6.90	1
9	310	5.50	1	311	5.90	1
10	312	4.90	1	314	5.10	1
11	314	4.50	1	316	4.60	1
12	316	4.20	1	318	4.20	1
13	318	3.90	1	320	3.90	1
14	319	3.70	1	321	3.70	1
15	320	3.60	1	322	3.50	1
16	322	3.40	1	323	3.30	1
17	323	3.30	1	324	3.20	1
18	324	3.20	1	325	3.10	1
19	325	3.10	1	326	3.00	1
20	326	3.00	1	327	2.90	1
21	327	2.90	1	328	2.80	1
22	328	2.90	1	329	2.70	1
23	328	2.80	1	330	2.70	1
24	329	2.80	1	331	2.60	1
25	330	2.70	1	331	2.60	1
26	331	2.70	1	332	2.50	2
27	331	2.60	1	333	2.50	2
28	332	2.60	2	334	2.50	2
29	333	2.60	2	334	2.50	2
30	334	2.60	2	335	2.40	2
31	335	2.50	2	336	2.40	2
32	335	2.50	2	336	2.40	2
33	336	2.50	2	337	2.40	2
34	337	2.50	2	338	2.40	2
35	337	2.50	2	338	2.40	2
36	338	2.50	2	339	2.40	2
37	339	2.50	2	339	2.40	2
38	339	2.50	2	341	2.40	3
39	340	2.50	3	341	2.40	3
40	341	2.50	3	342	2.40	3
41	342	2.50	3	343	2.40	3
42	342	2.50	3	343	2.50	3
43	343	2.60	3	344	2.50	3
44	344	2.60	3	345	2.50	3
45	345	2.60	3	346	2.50	3
46	345	2.60	3	346	2.60	3
47	346	2.70	3	347	2.60	3
48	347	2.70	3	348	2.60	3

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
49	348	2.80	3	349	2.70	3
50	349	2.80	3	350	2.70	3
51	350	2.90	3	351	2.70	3
52	351	2.90	3	351	2.80	3
53	352	3.00	3	352	2.90	3
54	352	3.10	3	353	3.00	4
55	354	3.20	4	355	3.10	4
56	355	3.30	4	356	3.20	4
57	357	3.50	4	357	3.30	4
58	358	3.70	4	358	3.50	4
59	360	3.90	4	360	3.80	4
60	362	4.20	4	362	4.10	4
61	364	4.70	4	364	4.60	4
62	367	5.40	4	367	5.40	4
63	371	6.70	4	371	6.70	4
64	377	9.80	4	377	9.70	4
65	380	10.00	4	380	10.00	4

Table O-2. 2010–11 NECAP: Raw to Scaled Score Correspondence – Mathematics Grade 4

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	400	10.00	1	400	10.00	1
1	400	10.00	1	400	10.00	1
2	400	10.00	1	400	10.00	1
3	400	10.00	1	400	10.00	1
4	400	10.00	1	400	10.00	1
5	400	10.00	1	400	10.00	1
6	400	8.80	1	400	9.40	1
7	401	7.50	1	400	8.40	1
8	405	6.30	1	402	7.30	1
9	408	5.60	1	406	6.40	1
10	410	5.10	1	408	5.80	1
11	412	4.70	1	411	5.30	1
12	414	4.40	1	413	4.90	1
13	415	4.10	1	415	4.60	1
14	416	4.00	1	416	4.30	1
15	418	3.80	1	418	4.10	1
16	419	3.70	1	419	4.00	1
17	420	3.60	1	420	3.80	1
18	421	3.50	1	422	3.70	1
19	422	3.40	1	423	3.60	1
20	423	3.30	1	424	3.50	1
21	424	3.30	1	425	3.40	1
22	425	3.20	1	426	3.30	1
23	426	3.20	1	427	3.30	1
24	427	3.10	1	428	3.20	1
25	428	3.10	1	429	3.20	1
26	429	3.00	1	430	3.10	1

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
27	430	3.00	1	430	3.10	1
28	430	3.00	1	432	3.10	2
29	430	3.00	1	433	3.10	2
30	432	3.00	2	434	3.00	2
31	433	2.90	2	435	3.00	2
32	434	2.90	2	436	3.00	2
33	434	2.90	2	436	3.00	2
34	435	2.90	2	437	3.00	2
35	436	2.90	2	438	3.00	2
36	437	2.90	2	439	3.00	2
37	438	2.90	2	439	3.00	2
38	439	2.90	2	441	3.00	3
39	439	2.90	2	442	2.90	3
40	440	2.90	3	443	2.90	3
41	441	3.00	3	443	2.90	3
42	442	3.00	3	444	2.90	3
43	443	3.00	3	445	2.90	3
44	444	3.00	3	446	2.90	3
45	445	3.00	3	447	2.90	3
46	446	3.10	3	448	2.90	3
47	447	3.10	3	449	3.00	3
48	448	3.10	3	450	3.00	3
49	449	3.20	3	451	3.00	3
50	450	3.20	3	452	3.00	3
51	451	3.20	3	453	3.00	3
52	452	3.30	3	454	3.10	3
53	453	3.30	3	455	3.10	4
54	454	3.40	3	456	3.20	4
55	456	3.50	4	457	3.20	4
56	457	3.60	4	459	3.30	4
57	458	3.70	4	460	3.50	4
58	460	3.90	4	462	3.70	4
59	462	4.20	4	463	3.90	4
60	464	4.60	4	465	4.30	4
61	467	5.20	4	468	4.80	4
62	470	6.20	4	471	5.70	4
63	475	8.10	4	476	7.30	4
64	480	10.00	4	480	10.00	4
65	480	10.00	4	480	10.00	4

Table O-3. 2010–11 NECAP: Raw to Scaled Score Correspondence – Mathematics Grade 5

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	500	10.00	1	500	10.00	1
1	500	10.00	1	500	10.00	1
2	500	10.00	1	500	10.00	1
3	500	10.00	1	500	10.00	1
4	500	10.00	1	500	10.00	1

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
5	500	10.00	1	500	10.00	1
6	502	10.00	1	500	10.00	1
7	509	8.30	1	506	9.10	1
8	513	6.60	1	511	7.20	1
9	517	5.60	1	514	6.10	1
10	519	5.00	1	517	5.40	1
11	521	4.50	1	519	4.90	1
12	523	4.20	1	521	4.50	1
13	524	3.90	1	523	4.20	1
14	526	3.70	1	524	4.00	1
15	527	3.50	1	526	3.80	1
16	528	3.40	1	527	3.60	1
17	529	3.30	1	528	3.50	1
18	530	3.20	1	529	3.40	1
19	531	3.10	1	530	3.30	1
20	532	3.00	1	531	3.20	1
21	532	3.00	1	532	3.10	1
22	534	2.90	2	532	3.10	1
23	535	2.90	2	534	3.00	2
24	536	2.80	2	535	2.90	2
25	536	2.80	2	535	2.90	2
26	537	2.80	2	536	2.80	2
27	538	2.70	2	537	2.80	2
28	539	2.70	2	538	2.80	2
29	539	2.70	2	538	2.70	2
30	540	2.60	3	539	2.70	2
31	541	2.60	3	539	2.70	2
32	541	2.60	3	541	2.60	3
33	542	2.60	3	541	2.60	3
34	543	2.50	3	542	2.60	3
35	544	2.50	3	543	2.60	3
36	544	2.50	3	543	2.60	3
37	545	2.50	3	544	2.50	3
38	546	2.50	3	545	2.50	3
39	546	2.50	3	545	2.50	3
40	547	2.50	3	546	2.50	3
41	548	2.50	3	547	2.50	3
42	548	2.50	3	547	2.50	3
43	549	2.50	3	548	2.50	3
44	550	2.50	3	549	2.50	3
45	550	2.50	3	549	2.50	3
46	551	2.50	3	550	2.50	3
47	552	2.60	3	551	2.50	3
48	553	2.60	3	552	2.50	3
49	553	2.70	3	552	2.60	3
50	554	2.70	4	553	2.60	3
51	555	2.80	4	553	2.60	3
52	556	2.80	4	555	2.70	4
53	557	2.90	4	556	2.80	4
54	558	3.00	4	556	2.80	4

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
55	559	3.10	4	557	2.90	4
56	560	3.20	4	558	3.00	4
57	561	3.40	4	560	3.20	4
58	562	3.50	4	561	3.30	4
59	564	3.70	4	562	3.50	4
60	565	3.90	4	564	3.80	4
61	567	4.30	4	565	4.10	4
62	570	4.70	4	567	4.50	4
63	573	5.50	4	570	5.20	4
64	577	7.00	4	574	6.30	4
65	580	10.00	4	580	9.00	4
66	580	10.00	4	580	10.00	4

Table O-4. 2010–11 NECAP: Raw to Scaled Score Correspondence – Mathematics Grade 6

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	600	10.00	1	600	10.00	1
1	600	10.00	1	600	10.00	1
2	600	10.00	1	600	10.00	1
3	600	10.00	1	600	10.00	1
4	600	10.00	1	600	10.00	1
5	600	10.00	1	600	10.00	1
6	600	10.00	1	606	10.00	1
7	608	9.90	1	613	7.60	1
8	614	7.60	1	616	6.00	1
9	618	6.40	1	619	5.20	1
10	620	5.50	1	621	4.60	1
11	623	4.90	1	623	4.20	1
12	625	4.50	1	625	3.90	1
13	626	4.10	1	626	3.70	1
14	628	3.90	1	627	3.50	1
15	629	3.60	1	628	3.40	1
16	630	3.50	1	630	3.30	1
17	631	3.30	1	631	3.20	1
18	632	3.20	1	632	3.10	1
19	633	3.10	2	632	3.00	1
20	634	3.00	2	633	3.00	2
21	635	2.90	2	634	2.90	2
22	636	2.90	2	635	2.80	2
23	637	2.80	2	636	2.80	2
24	638	2.80	2	637	2.80	2
25	638	2.70	2	637	2.70	2
26	639	2.70	2	638	2.70	2
27	639	2.70	2	639	2.70	2
28	641	2.70	3	639	2.60	2
29	641	2.60	3	640	2.60	3
30	642	2.60	3	641	2.60	3
31	643	2.60	3	642	2.60	3

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
32	643	2.60	3	642	2.60	3
33	644	2.60	3	643	2.50	3
34	645	2.60	3	644	2.50	3
35	645	2.50	3	644	2.50	3
36	646	2.50	3	645	2.50	3
37	647	2.50	3	646	2.50	3
38	647	2.50	3	646	2.50	3
39	648	2.50	3	647	2.50	3
40	649	2.50	3	648	2.50	3
41	649	2.50	3	648	2.50	3
42	650	2.50	3	649	2.50	3
43	651	2.50	3	650	2.50	3
44	652	2.50	3	650	2.50	3
45	652	2.60	3	651	2.50	3
46	652	2.60	3	652	2.60	3
47	654	2.60	4	652	2.60	3
48	655	2.60	4	653	2.60	4
49	655	2.70	4	654	2.70	4
50	656	2.70	4	655	2.70	4
51	657	2.80	4	656	2.80	4
52	658	2.90	4	657	2.80	4
53	659	2.90	4	658	2.90	4
54	660	3.10	4	658	3.00	4
55	661	3.20	4	659	3.10	4
56	662	3.30	4	661	3.30	4
57	664	3.50	4	662	3.40	4
58	665	3.70	4	663	3.60	4
59	667	4.00	4	665	3.90	4
60	668	4.30	4	666	4.20	4
61	671	4.70	4	668	4.70	4
62	673	5.30	4	671	5.30	4
63	677	6.20	4	674	6.30	4
64	680	7.70	4	679	7.90	4
65	680	9.10	4	680	10.00	4
66	680	9.10	4	680	10.00	4

Table O-5. 2010–11 NECAP: Raw to Scaled Score Correspondence – Mathematics Grade 7

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	700	10.00	1	700	10.00	1
1	700	10.00	1	700	10.00	1
2	700	10.00	1	700	10.00	1
3	700	10.00	1	700	10.00	1
4	700	10.00	1	700	10.00	1
5	700	10.00	1	702	10.00	1
6	711	9.10	1	711	9.20	1
7	716	6.70	1	716	7.00	1
8	719	5.50	1	719	5.80	1

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
9	722	4.80	1	722	5.00	1
10	724	4.40	1	724	4.40	1
11	726	4.10	1	726	4.10	1
12	727	3.80	1	727	3.80	1
13	729	3.60	1	729	3.50	1
14	730	3.40	1	730	3.40	1
15	731	3.30	1	731	3.20	1
16	732	3.20	1	732	3.10	1
17	733	3.10	1	733	3.00	1
18	734	3.00	2	734	2.90	2
19	735	2.90	2	735	2.80	2
20	736	2.80	2	736	2.80	2
21	737	2.80	2	737	2.70	2
22	738	2.70	2	737	2.60	2
23	739	2.60	2	738	2.60	2
24	739	2.60	2	739	2.60	2
25	739	2.50	2	739	2.50	2
26	741	2.50	3	740	2.50	3
27	741	2.50	3	741	2.50	3
28	742	2.40	3	742	2.40	3
29	743	2.40	3	742	2.40	3
30	743	2.40	3	743	2.40	3
31	744	2.30	3	744	2.40	3
32	745	2.30	3	744	2.40	3
33	745	2.30	3	745	2.30	3
34	746	2.30	3	746	2.30	3
35	746	2.30	3	746	2.30	3
36	747	2.30	3	747	2.30	3
37	748	2.30	3	747	2.30	3
38	748	2.30	3	748	2.30	3
39	749	2.30	3	749	2.30	3
40	750	2.30	3	749	2.30	3
41	750	2.30	3	750	2.30	3
42	751	2.30	3	751	2.40	3
43	751	2.30	3	751	2.40	3
44	752	2.30	4	751	2.40	3
45	753	2.40	4	753	2.40	4
46	753	2.40	4	753	2.40	4
47	754	2.40	4	754	2.50	4
48	755	2.50	4	755	2.50	4
49	756	2.50	4	756	2.60	4
50	756	2.60	4	756	2.60	4
51	757	2.60	4	757	2.70	4
52	758	2.70	4	758	2.70	4
53	759	2.80	4	759	2.80	4
54	760	2.90	4	760	2.90	4
55	761	3.00	4	761	3.00	4
56	762	3.10	4	762	3.10	4
57	763	3.30	4	764	3.30	4
58	764	3.50	4	765	3.50	4

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
59	765	3.70	4	767	3.80	4
60	767	3.90	4	769	4.20	4
61	769	4.30	4	771	4.90	4
62	771	4.70	4	775	6.10	4
63	773	5.40	4	780	9.30	4
64	777	6.50	4	780	9.40	4
65	780	8.60	4			
66	780	8.60	4			

Table O-6. 2010–11 NECAP: Raw to Scaled Score Correspondence – Mathematics Grade 8

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	800	10.00	1	800	10.00	1
1	800	10.00	1	800	10.00	1
2	800	10.00	1	800	10.00	1
3	800	10.00	1	800	10.00	1
4	800	10.00	1	800	10.00	1
5	800	10.00	1	800	10.00	1
6	813	9.00	1	808	10.00	1
7	818	6.40	1	817	7.70	1
8	821	5.20	1	821	5.50	1
9	824	4.40	1	824	4.50	1
10	825	3.90	1	826	3.80	1
11	827	3.60	1	827	3.50	1
12	828	3.30	1	829	3.20	1
13	830	3.10	1	830	3.00	1
14	831	3.00	1	831	2.80	1
15	832	2.80	1	832	2.70	1
16	833	2.70	1	833	2.60	1
17	833	2.60	1	833	2.50	1
18	834	2.50	2	834	2.50	2
19	835	2.50	2	835	2.40	2
20	836	2.40	2	836	2.30	2
21	836	2.40	2	837	2.30	2
22	837	2.30	2	837	2.30	2
23	838	2.30	2	838	2.20	2
24	838	2.20	2	839	2.20	2
25	839	2.20	2	839	2.20	2
26	839	2.20	2	839	2.10	2
27	840	2.20	3	840	2.10	3
28	841	2.10	3	841	2.10	3
29	841	2.10	3	842	2.10	3
30	842	2.10	3	842	2.10	3
31	842	2.10	3	843	2.10	3
32	843	2.10	3	843	2.00	3
33	844	2.10	3	844	2.00	3
34	844	2.10	3	845	2.00	3
35	845	2.10	3	845	2.00	3

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
36	845	2.10	3	846	2.00	3
37	846	2.10	3	846	2.00	3
38	846	2.10	3	847	2.00	3
39	847	2.10	3	847	2.00	3
40	847	2.10	3	848	2.10	3
41	848	2.10	3	849	2.10	3
42	849	2.10	3	849	2.10	3
43	849	2.10	3	850	2.10	3
44	850	2.20	3	850	2.10	3
45	850	2.20	3	851	2.10	3
46	851	2.20	3	852	2.20	4
47	852	2.20	4	852	2.20	4
48	852	2.30	4	853	2.20	4
49	853	2.30	4	854	2.30	4
50	854	2.30	4	855	2.30	4
51	855	2.40	4	855	2.40	4
52	855	2.40	4	856	2.40	4
53	856	2.50	4	857	2.50	4
54	857	2.50	4	858	2.60	4
55	858	2.60	4	859	2.70	4
56	859	2.70	4	860	2.80	4
57	860	2.80	4	861	3.00	4
58	861	3.00	4	862	3.10	4
59	863	3.20	4	864	3.30	4
60	864	3.40	4	865	3.60	4
61	866	3.80	4	867	3.90	4
62	868	4.20	4	869	4.40	4
63	871	4.90	4	872	5.40	4
64	874	6.00	4	878	7.70	4
65	880	8.30	4	880	9.50	4
66	880	8.50	4			

Table O-7. 2010–11 NECAP: Raw to Scaled Score Correspondence – Mathematics Grade 11

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	1100	10.00	1	1100	10.00	1
1	1100	10.00	1	1100	10.00	1
2	1100	10.00	1	1100	10.00	1
3	1100	10.00	1	1100	10.00	1
4	1100	10.00	1	1100	10.00	1
5	1112	10.00	1	1100	10.00	1
6	1118	5.80	1	1112	9.20	1
7	1121	4.40	1	1118	6.20	1
8	1123	3.80	1	1121	4.80	1
9	1125	3.30	1	1124	4.00	1
10	1126	3.00	1	1125	3.50	1
11	1128	2.80	1	1127	3.10	1
12	1129	2.60	1	1128	2.90	1

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
13	1130	2.50	1	1129	2.70	1
14	1130	2.40	1	1130	2.50	1
15	1131	2.30	1	1131	2.40	1
16	1132	2.20	1	1132	2.30	1
17	1133	2.10	1	1133	2.20	1
18	1133	2.10	1	1133	2.10	1
19	1134	2.00	2	1134	2.10	2
20	1135	2.00	2	1135	2.00	2
21	1135	1.90	2	1135	2.00	2
22	1136	1.90	2	1136	1.90	2
23	1136	1.80	2	1136	1.90	2
24	1137	1.80	2	1137	1.90	2
25	1137	1.80	2	1137	1.80	2
26	1138	1.80	2	1138	1.80	2
27	1138	1.70	2	1138	1.80	2
28	1139	1.70	2	1139	1.70	2
29	1139	1.70	2	1139	1.70	2
30	1139	1.70	2	1139	1.70	2
31	1140	1.70	3	1140	1.70	3
32	1141	1.70	3	1141	1.70	3
33	1141	1.70	3	1141	1.70	3
34	1142	1.60	3	1142	1.70	3
35	1142	1.60	3	1142	1.60	3
36	1143	1.60	3	1142	1.60	3
37	1143	1.70	3	1143	1.60	3
38	1144	1.70	3	1143	1.60	3
39	1144	1.70	3	1144	1.60	3
40	1145	1.70	3	1144	1.60	3
41	1145	1.70	3	1145	1.60	3
42	1146	1.70	3	1145	1.60	3
43	1146	1.70	3	1146	1.70	3
44	1147	1.70	3	1146	1.70	3
45	1147	1.80	3	1147	1.70	3
46	1148	1.80	3	1147	1.70	3
47	1148	1.80	3	1148	1.70	3
48	1149	1.80	3	1148	1.70	3
49	1149	1.90	3	1149	1.80	3
50	1150	1.90	3	1149	1.80	3
51	1151	2.00	3	1150	1.90	3
52	1151	2.00	3	1150	1.90	3
53	1151	2.10	3	1151	2.00	3
54	1153	2.20	4	1151	2.10	3
55	1154	2.30	4	1153	2.20	4
56	1154	2.40	4	1153	2.30	4
57	1155	2.50	4	1154	2.50	4
58	1156	2.70	4	1155	2.70	4
59	1158	2.90	4	1157	2.90	4
60	1159	3.10	4	1158	3.30	4
61	1161	3.50	4	1160	3.80	4
62	1163	4.30	4	1163	4.60	4

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
63	1167	6.10	4	1168	6.60	4
64	1180	7.30	4	1180	7.60	4

Table O-8. 2010–11 NECAP: Raw to Scaled Score Correspondence – Reading Grade 3

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	300	10.00	1	300	10.00	1
1	300	10.00	1	300	10.00	1
2	300	10.00	1	300	10.00	1
3	300	10.00	1	300	10.00	1
4	300	10.00	1	300	10.00	1
5	300	10.00	1	300	10.00	1
6	303	9.20	1	302	9.60	1
7	307	7.80	1	307	7.90	1
8	311	6.90	1	311	6.80	1
9	314	6.20	1	314	6.00	1
10	316	5.70	1	317	5.40	1
11	318	5.30	1	319	4.90	1
12	320	4.90	1	321	4.40	1
13	321	4.50	1	323	4.10	1
14	323	4.30	1	324	3.80	1
15	324	4.00	1	325	3.60	1
16	326	3.80	1	327	3.50	1
17	327	3.70	1	328	3.30	1
18	328	3.50	1	329	3.20	1
19	329	3.40	1	330	3.10	1
20	330	3.30	1	331	3.10	2
21	331	3.20	2	332	3.00	2
22	332	3.20	2	333	3.00	2
23	333	3.10	2	334	2.90	2
24	334	3.10	2	335	2.90	2
25	335	3.10	2	336	2.90	2
26	336	3.10	2	337	2.90	2
27	337	3.00	2	338	2.90	2
28	338	3.00	2	339	2.90	2
29	338	3.00	2	340	2.90	3
30	339	3.10	2	341	2.90	3
31	340	3.10	3	342	3.00	3
32	341	3.10	3	343	3.00	3
33	342	3.20	3	345	3.10	3
34	343	3.20	3	346	3.20	3
35	344	3.30	3	347	3.30	3
36	346	3.30	3	348	3.50	3
37	347	3.40	3	350	3.70	3
38	348	3.50	3	351	3.90	3
39	349	3.60	3	353	4.10	3
40	350	3.70	3	355	4.40	3
41	352	3.90	3	356	4.60	3

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
42	353	4.10	3	359	4.90	4
43	355	4.30	3	362	5.20	4
44	356	4.60	3	364	5.40	4
45	359	4.90	4	367	5.70	4
46	361	5.30	4	370	6.00	4
47	364	5.80	4	373	6.30	4
48	368	6.60	4	377	6.70	4
49	372	7.80	4	380	7.60	4
50	378	9.90	4	380	9.30	4
51	380	10.00	4	380	10.00	4
52	380	10.00	4	380	10.00	4

Table O-9. 2010–11 NECAP: Raw to Scaled Score Correspondence – Reading Grade 4

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	400	10.00	1	400	10.00	1
1	400	10.00	1	400	10.00	1
2	400	10.00	1	400	10.00	1
3	400	10.00	1	400	10.00	1
4	400	10.00	1	400	10.00	1
5	400	10.00	1	400	10.00	1
6	400	10.00	1	400	10.00	1
7	400	10.00	1	400	10.00	1
8	400	10.00	1	400	10.00	1
9	405	8.30	1	401	9.40	1
10	409	7.10	1	406	8.30	1
11	412	6.30	1	410	7.40	1
12	414	5.70	1	413	6.80	1
13	417	5.30	1	416	6.20	1
14	419	4.90	1	418	5.70	1
15	421	4.70	1	420	5.30	1
16	422	4.50	1	422	4.90	1
17	424	4.30	1	423	4.60	1
18	425	4.10	1	425	4.40	1
19	427	4.00	1	426	4.20	1
20	428	3.80	1	428	4.00	1
21	429	3.70	1	429	3.80	1
22	430	3.60	1	430	3.70	1
23	432	3.50	2	431	3.60	2
24	433	3.40	2	432	3.50	2
25	434	3.30	2	434	3.40	2
26	435	3.20	2	435	3.30	2
27	436	3.20	2	436	3.30	2
28	437	3.10	2	437	3.20	2
29	438	3.10	2	438	3.20	2
30	439	3.10	2	439	3.20	2
31	440	3.10	3	439	3.20	2
32	441	3.10	3	441	3.30	3

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
33	442	3.10	3	442	3.30	3
34	443	3.20	3	443	3.40	3
35	445	3.20	3	444	3.40	3
36	446	3.30	3	446	3.50	3
37	447	3.40	3	447	3.70	3
38	448	3.50	3	448	3.80	3
39	450	3.70	3	450	4.00	3
40	451	3.90	3	452	4.20	3
41	453	4.10	3	453	4.50	3
42	455	4.40	3	455	4.80	3
43	457	4.70	4	458	5.20	4
44	459	5.10	4	460	5.70	4
45	462	5.50	4	463	6.30	4
46	465	6.00	4	467	7.10	4
47	468	6.50	4	471	8.00	4
48	472	7.00	4	476	9.00	4
49	477	7.60	4	480	9.80	4
50	480	8.50	4	480	10.00	4
51	480	9.50	4	480	10.00	4
52	480	9.50	4	480	10.00	4

Table O-10. 2010–11 NECAP: Raw to Scaled Score Correspondence – Reading Grade 5

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	500	10.00	1	500	10.00	1
1	500	10.00	1	500	10.00	1
2	500	10.00	1	500	10.00	1
3	500	10.00	1	500	10.00	1
4	500	8.80	1	500	9.80	1
5	504	6.90	1	506	7.30	1
6	508	5.80	1	510	6.10	1
7	511	5.20	1	513	5.30	1
8	513	4.80	1	515	4.80	1
9	515	4.50	1	517	4.50	1
10	517	4.20	1	519	4.20	1
11	519	4.10	1	521	4.00	1
12	520	3.90	1	523	3.90	1
13	522	3.80	1	524	3.70	1
14	523	3.80	1	525	3.60	1
15	525	3.70	1	527	3.60	1
16	526	3.60	1	528	3.50	1
17	527	3.60	1	529	3.40	1
18	528	3.50	1	530	3.40	2
19	529	3.50	1	532	3.30	2
20	531	3.40	2	533	3.30	2
21	532	3.40	2	534	3.30	2
22	533	3.40	2	535	3.30	2
23	534	3.40	2	536	3.20	2

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
24	536	3.40	2	537	3.20	2
25	537	3.40	2	538	3.20	2
26	538	3.40	2	539	3.20	2
27	539	3.50	2	540	3.20	3
28	540	3.50	3	541	3.30	3
29	542	3.60	3	543	3.30	3
30	543	3.60	3	544	3.30	3
31	545	3.70	3	545	3.40	3
32	546	3.80	3	546	3.40	3
33	547	3.90	3	548	3.50	3
34	549	4.00	3	549	3.60	3
35	551	4.10	3	550	3.70	3
36	552	4.20	3	552	3.80	3
37	554	4.30	3	553	3.90	3
38	555	4.40	3	555	4.00	3
39	558	4.40	4	557	4.10	4
40	560	4.50	4	559	4.20	4
41	562	4.50	4	560	4.40	4
42	564	4.60	4	562	4.50	4
43	567	4.70	4	564	4.60	4
44	569	4.70	4	567	4.70	4
45	571	4.80	4	569	4.80	4
46	574	4.90	4	571	4.90	4
47	576	5.00	4	574	5.10	4
48	579	5.20	4	577	5.30	4
49	580	5.60	4	580	5.70	4
50	580	6.60	4	580	6.70	4
51	580	7.50	4	580	8.30	4
52	580	7.50	4	580	8.30	4

Table O-11. 2010–11 NECAP: Raw to Scaled Score Correspondence – Reading Grade 6

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	600	10.00	1	600	10.00	1
1	600	10.00	1	600	10.00	1
2	600	10.00	1	600	10.00	1
3	600	10.00	1	600	10.00	1
4	600	9.60	1	600	10.00	1
5	604	7.20	1	607	7.80	1
6	608	5.90	1	611	6.20	1
7	611	5.20	1	614	5.30	1
8	613	4.70	1	616	4.80	1
9	615	4.30	1	619	4.40	1
10	617	4.00	1	620	4.10	1
11	618	3.80	1	622	3.90	1
12	620	3.70	1	623	3.70	1
13	621	3.50	1	625	3.60	1
14	622	3.40	1	626	3.50	1

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
15	624	3.30	1	627	3.40	1
16	625	3.30	1	628	3.30	1
17	626	3.20	1	630	3.20	2
18	627	3.20	1	631	3.10	2
19	628	3.10	1	632	3.10	2
20	629	3.10	2	633	3.10	2
21	630	3.10	2	634	3.00	2
22	631	3.10	2	635	3.00	2
23	633	3.10	2	636	3.00	2
24	634	3.10	2	637	3.00	2
25	635	3.20	2	638	3.00	2
26	636	3.20	2	639	3.00	2
27	637	3.30	2	640	3.00	3
28	638	3.30	2	641	3.10	3
29	639	3.40	2	642	3.10	3
30	641	3.50	3	643	3.20	3
31	642	3.60	3	644	3.20	3
32	644	3.70	3	646	3.30	3
33	645	3.80	3	647	3.40	3
34	647	3.90	3	648	3.40	3
35	648	4.00	3	650	3.50	3
36	650	4.10	3	651	3.60	3
37	652	4.30	3	652	3.80	3
38	653	4.40	3	654	3.90	3
39	655	4.50	3	656	4.00	3
40	657	4.60	3	658	4.10	3
41	659	4.60	4	659	4.20	4
42	662	4.70	4	661	4.40	4
43	664	4.80	4	664	4.50	4
44	666	4.90	4	666	4.70	4
45	668	5.00	4	668	4.80	4
46	671	5.10	4	671	5.00	4
47	674	5.20	4	674	5.10	4
48	677	5.50	4	677	5.30	4
49	680	5.90	4	680	5.60	4
50	680	6.90	4	680	6.30	4
51	680	9.50	4	680	8.30	4
52	680	9.50	4	680	8.90	4

Table O-12. 2010–11 NECAP: Raw to Scaled Score Correspondence – Reading Grade 7

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	700	10.00	1	700	10.00	1
1	700	10.00	1	700	10.00	1
2	700	10.00	1	700	10.00	1
3	700	9.70	1	700	10.00	1
4	702	7.70	1	700	10.00	1
5	706	6.20	1	703	8.10	1

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
6	709	5.40	1	707	6.60	1
7	712	4.90	1	711	5.70	1
8	714	4.50	1	713	5.10	1
9	716	4.30	1	715	4.70	1
10	718	4.10	1	717	4.40	1
11	719	3.90	1	719	4.20	1
12	721	3.80	1	721	4.10	1
13	722	3.60	1	722	3.90	1
14	723	3.60	1	724	3.80	1
15	725	3.50	1	725	3.80	1
16	726	3.40	1	727	3.70	1
17	727	3.40	1	728	3.70	1
18	728	3.30	1	729	3.60	2
19	729	3.30	2	731	3.60	2
20	731	3.30	2	732	3.60	2
21	732	3.20	2	733	3.60	2
22	733	3.20	2	734	3.60	2
23	734	3.20	2	736	3.60	2
24	735	3.20	2	737	3.60	2
25	736	3.30	2	738	3.60	2
26	737	3.30	2	739	3.60	2
27	738	3.30	2	741	3.60	3
28	739	3.40	2	742	3.60	3
29	741	3.40	3	743	3.60	3
30	742	3.50	3	745	3.70	3
31	743	3.50	3	746	3.70	3
32	745	3.60	3	747	3.80	3
33	746	3.60	3	749	3.80	3
34	747	3.70	3	750	3.90	3
35	749	3.80	3	752	3.90	3
36	751	3.80	3	753	3.90	3
37	752	3.90	3	755	4.00	3
38	754	3.90	3	756	4.00	3
39	756	3.90	3	758	4.10	3
40	757	4.00	3	759	4.10	3
41	759	4.00	3	761	4.20	4
42	761	4.10	4	763	4.20	4
43	763	4.20	4	765	4.30	4
44	765	4.20	4	767	4.40	4
45	767	4.30	4	769	4.40	4
46	769	4.40	4	771	4.50	4
47	772	4.50	4	774	4.70	4
48	774	4.70	4	777	5.00	4
49	777	5.00	4	780	5.60	4
50	780	5.80	4	780	6.80	4
51	780	7.90	4	780	10.00	4
52	780	10.00	4	780	10.00	4

Table O-13. 2010–11 NECAP: Raw to Scaled Score Correspondence – Reading Grade 8

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	800	10.00	1	800	10.00	1
1	800	10.00	1	800	10.00	1
2	800	10.00	1	800	10.00	1
3	800	10.00	1	800	10.00	1
4	800	8.80	1	800	9.50	1
5	800	7.60	1	802	7.90	1
6	804	6.30	1	806	6.60	1
7	808	5.60	1	809	5.90	1
8	810	5.10	1	812	5.30	1
9	812	4.70	1	814	4.90	1
10	814	4.50	1	817	4.60	1
11	816	4.30	1	818	4.40	1
12	818	4.10	1	820	4.20	1
13	819	4.00	1	822	4.10	1
14	821	3.90	1	823	4.00	1
15	822	3.80	1	825	3.90	1
16	824	3.70	1	826	3.80	1
17	825	3.70	1	827	3.80	1
18	826	3.60	1	829	3.70	2
19	827	3.60	1	830	3.70	2
20	829	3.60	2	831	3.70	2
21	830	3.60	2	833	3.60	2
22	831	3.60	2	834	3.60	2
23	832	3.60	2	835	3.60	2
24	834	3.60	2	836	3.60	2
25	835	3.60	2	837	3.60	2
26	836	3.60	2	839	3.60	2
27	837	3.60	2	839	3.60	2
28	839	3.60	2	841	3.60	3
29	839	3.70	2	842	3.60	3
30	841	3.70	3	844	3.60	3
31	842	3.70	3	845	3.70	3
32	844	3.80	3	846	3.70	3
33	845	3.80	3	848	3.80	3
34	847	3.80	3	849	3.90	3
35	848	3.90	3	851	4.00	3
36	850	3.90	3	852	4.10	3
37	851	4.00	3	854	4.10	3
38	853	4.00	3	856	4.20	3
39	854	4.00	3	857	4.30	3
40	856	4.10	3	859	4.40	4
41	858	4.10	3	861	4.40	4
42	859	4.20	4	863	4.50	4
43	861	4.30	4	865	4.60	4
44	863	4.40	4	867	4.70	4
45	865	4.50	4	870	4.90	4
46	868	4.60	4	872	5.00	4
47	870	4.70	4	875	5.10	4
48	873	5.00	4	878	5.30	4

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
49	876	5.50	4	880	5.70	4
50	880	6.60	4	880	6.60	4
51	880	10.00	4	880	9.10	4
52	880	10.00	4	880	9.10	4

Table O-14. 2010–11 NECAP: Raw to Scaled Score Correspondence – Reading Grade 11

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	1100	10.00	1	1100	10.00	1
1	1100	10.00	1	1100	10.00	1
2	1100	10.00	1	1100	10.00	1
3	1100	10.00	1	1100	10.00	1
4	1101	10.00	1	1100	10.00	1
5	1107	7.20	1	1103	8.40	1
6	1110	5.60	1	1108	6.50	1
7	1113	4.70	1	1111	5.40	1
8	1115	4.10	1	1114	4.80	1
9	1117	3.70	1	1116	4.40	1
10	1119	3.50	1	1118	4.10	1
11	1120	3.30	1	1120	3.90	1
12	1122	3.10	1	1121	3.70	1
13	1123	3.00	1	1123	3.60	1
14	1124	3.00	1	1124	3.50	1
15	1125	2.90	1	1125	3.50	1
16	1126	2.90	1	1127	3.50	1
17	1127	2.80	1	1128	3.40	1
18	1128	2.80	1	1129	3.40	1
19	1129	2.80	1	1130	3.40	2
20	1129	2.80	1	1132	3.40	2
21	1131	2.80	2	1133	3.40	2
22	1132	2.80	2	1134	3.40	2
23	1133	2.80	2	1135	3.40	2
24	1134	2.80	2	1136	3.40	2
25	1135	2.80	2	1138	3.40	2
26	1136	2.80	2	1139	3.40	2
27	1137	2.90	2	1140	3.40	3
28	1138	2.90	2	1141	3.40	3
29	1139	2.90	2	1143	3.50	3
30	1140	3.00	3	1144	3.50	3
31	1142	3.10	3	1145	3.60	3
32	1143	3.10	3	1147	3.60	3
33	1144	3.20	3	1148	3.60	3
34	1146	3.20	3	1149	3.60	3
35	1147	3.30	3	1151	3.60	3
36	1148	3.30	3	1152	3.70	3
37	1150	3.30	3	1153	3.70	3
38	1151	3.30	3	1155	3.70	4
39	1153	3.30	3	1157	3.70	4

Raw Score	2010			2011		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
40	1153	3.30	3	1158	3.80	4
41	1156	3.40	4	1160	3.80	4
42	1157	3.40	4	1162	3.90	4
43	1159	3.50	4	1163	3.90	4
44	1161	3.50	4	1165	4.00	4
45	1162	3.50	4	1167	4.00	4
46	1164	3.60	4	1169	4.10	4
47	1166	3.70	4	1171	4.30	4
48	1169	3.90	4	1174	4.60	4
49	1171	4.30	4	1177	5.20	4
50	1175	5.40	4	1180	6.40	4
51	1180	9.20	4	1180	9.30	4
52	1180	10.00	4	1180	9.30	4

Table O-15. 2010–11 NECAP: Raw to Scaled Score Correspondence – Writing Grade 5

Raw Score	2011		
	Scaled Score	Standard Error	Performance Level
0	500	10.00	1
1	500	10.00	1
2	500	9.80	1
3	505	7.00	1
4	509	5.80	1
5	512	5.20	1
6	514	4.80	1
7	516	4.60	1
8	518	4.50	1
9	520	4.40	1
10	522	4.30	1
11	524	4.30	1
12	525	4.30	1
13	527	4.30	2
14	529	4.30	2
15	531	4.40	2
16	533	4.40	2
17	535	4.50	2
18	537	4.60	2
19	539	4.80	2
20	542	4.90	3
21	545	5.00	3
22	547	5.10	3
23	550	5.20	3
24	553	5.20	3
25	556	5.30	4
26	559	5.40	4
27	562	5.50	4
28	566	5.60	4
29	569	5.80	4
30	574	6.30	4

Raw Score	2011		
	Scaled Score	Standard Error	Performance Level
31	579	7.50	4
32	580	7.80	4
33	580	7.80	4
34	580	7.80	4

Table O-16. 2010–11 NECAP: Raw to Scaled Score Correspondence – Writing Grade 8

Raw Score	2011		
	Scaled Score	Standard Error	Performance Level
0	800	10.00	1
1	800	10.00	1
2	800	10.00	1
3	800	10.00	1
4	807	7.30	1
5	810	5.60	1
6	813	4.90	1
7	815	4.50	1
8	817	4.30	1
9	819	4.20	1
10	821	4.10	1
11	822	4.10	1
12	824	4.10	1
13	826	4.00	1
14	827	4.00	2
15	829	3.90	2
16	831	3.90	2
17	832	4.00	2
18	834	4.00	2
19	836	4.10	2
20	838	4.10	2
21	840	4.00	3
22	842	4.00	3
23	844	3.90	3
24	846	4.00	3
25	849	4.10	3
26	851	4.20	3
27	853	4.30	3
28	856	4.30	4
29	859	4.50	4
30	863	5.10	4
31	868	6.30	4
32	875	7.90	4
33	880	8.10	4
34	880	8.10	4

Appendix P—Scaled Score Distributions

Figure P-1. 2010–11 NECAP: Cumulative Score Distribution—Mathematics Grade 3

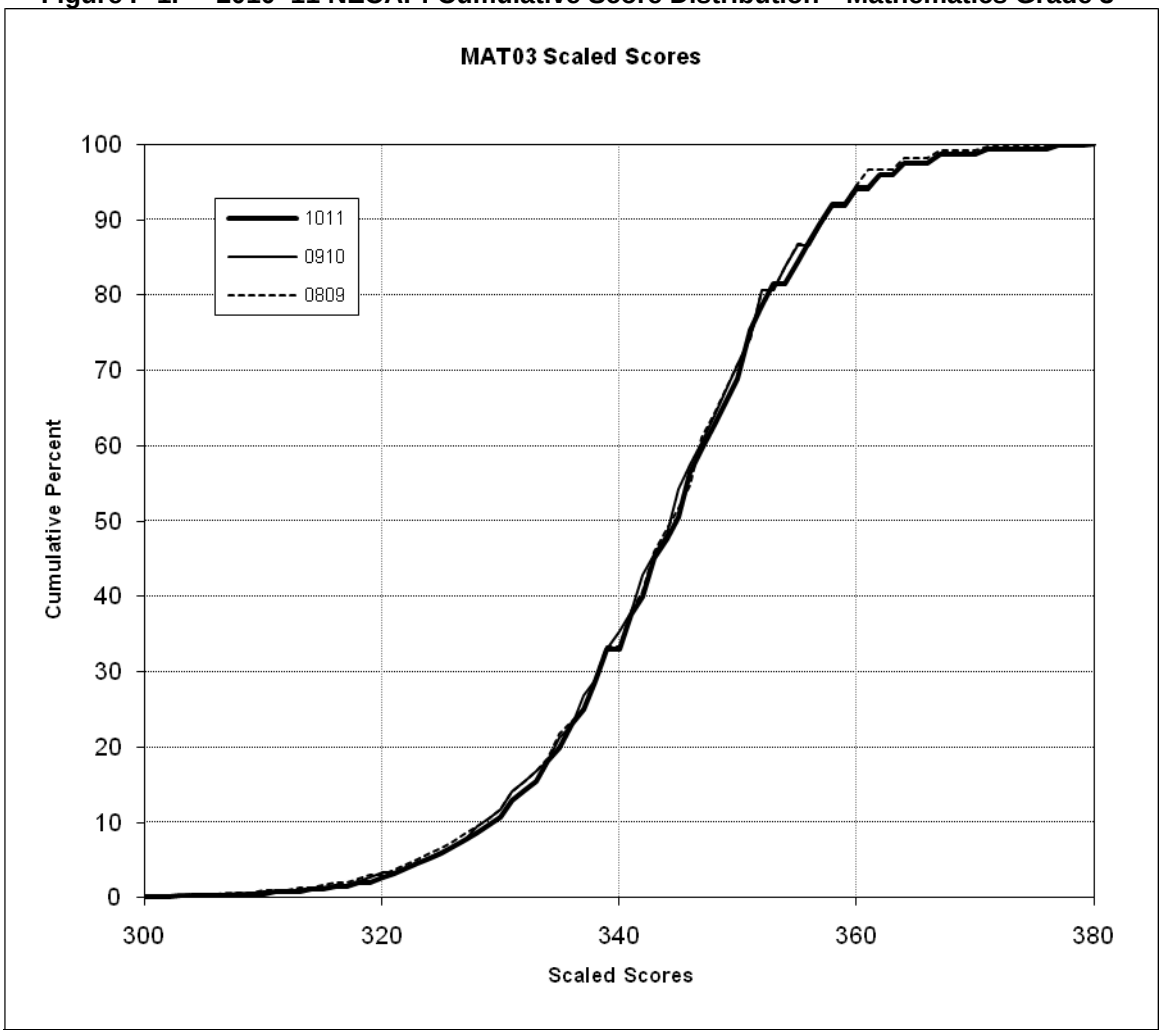


Figure P-2. 2010–11 NECAP: Cumulative Score Distribution—Mathematics Grade 4

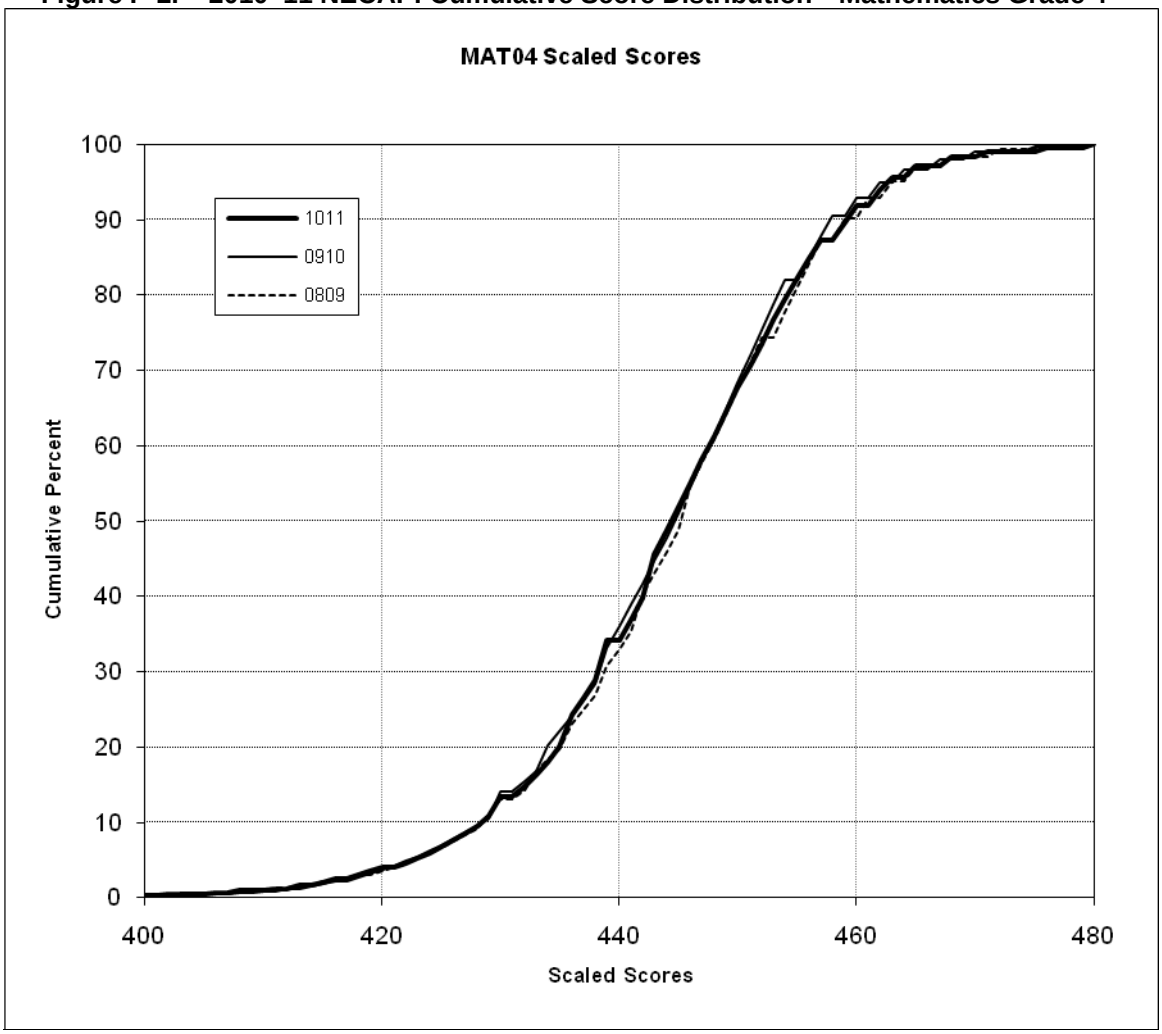


Figure P-3. 2010–11 NECAP: Cumulative Score Distribution—Mathematics Grade 5

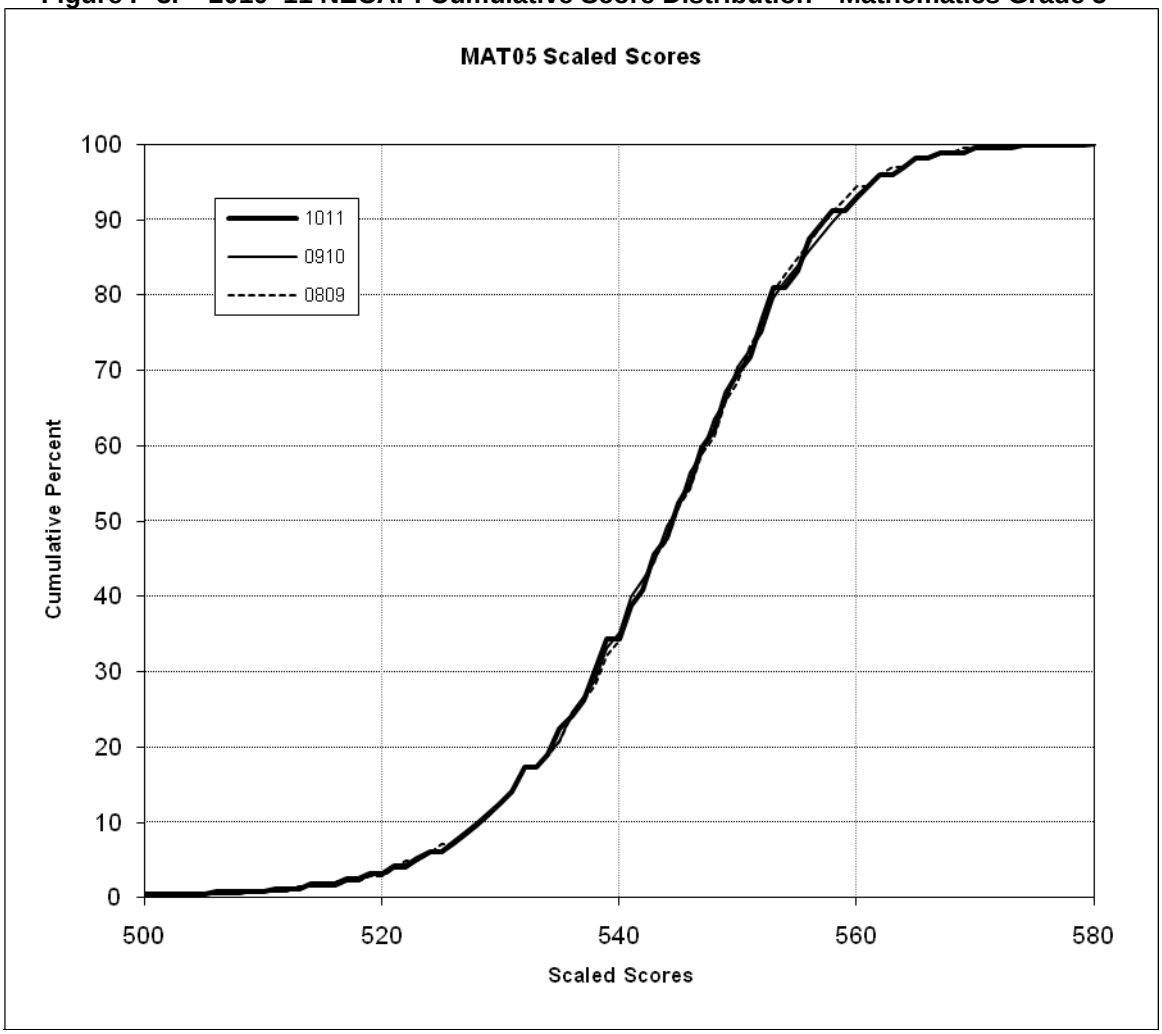


Figure P-4. 2010–11 NECAP: Cumulative Score Distribution—Mathematics Grade 6

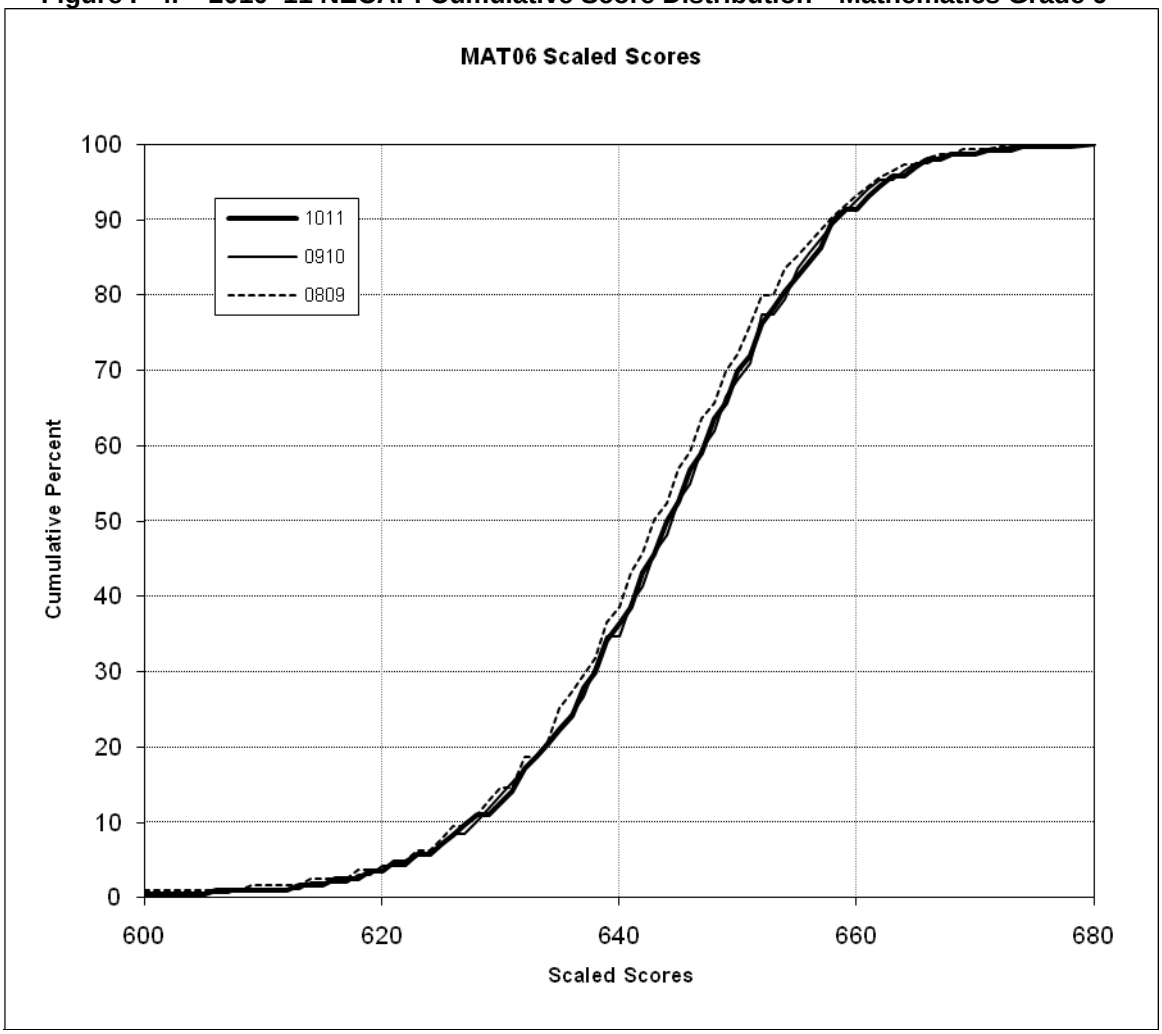


Figure P-5. 2010–11 NECAP: Cumulative Score Distribution—Mathematics Grade 7

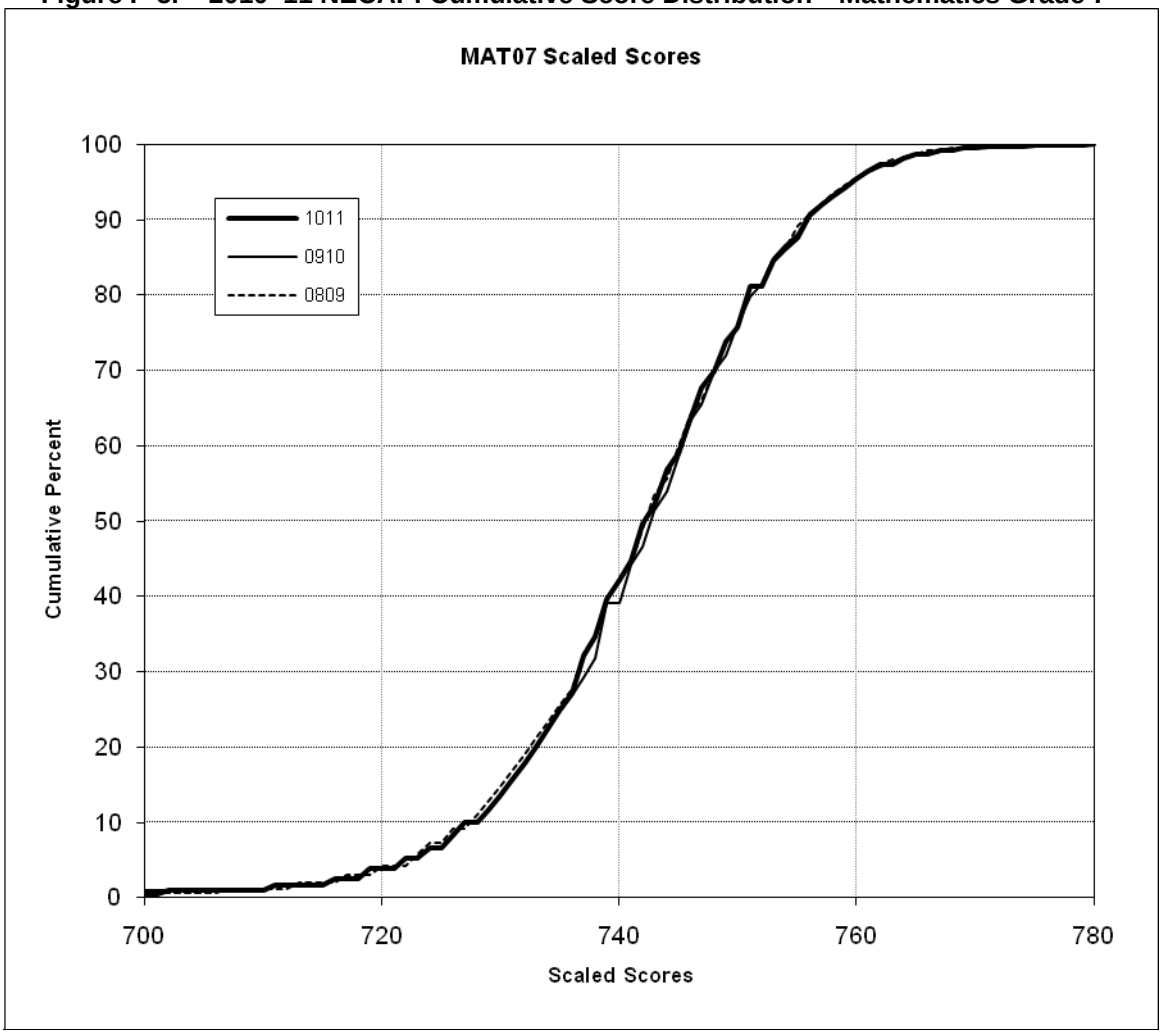


Figure P-6. 2010–11 NECAP: Cumulative Score Distribution—Mathematics Grade 8

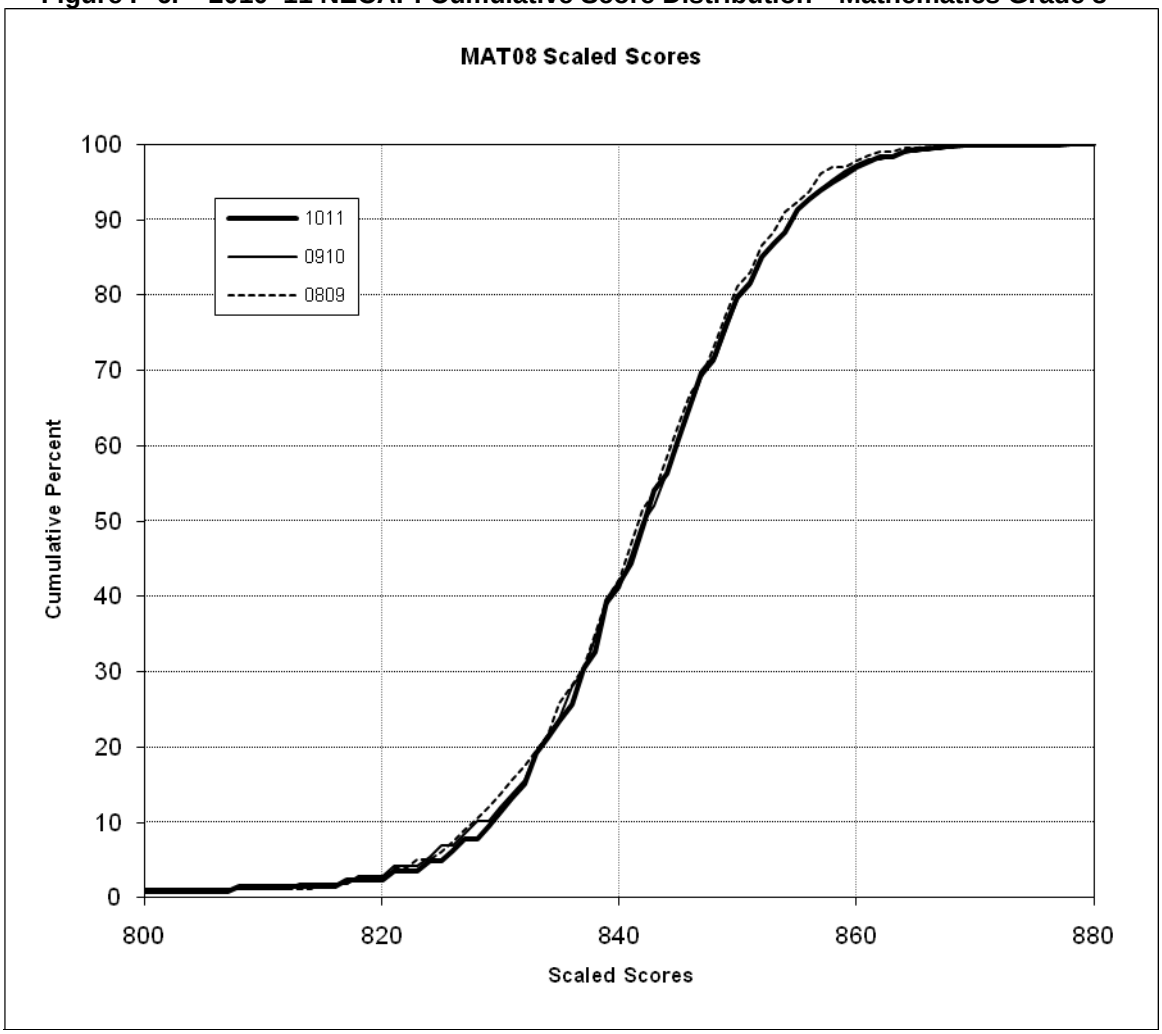


Figure P-7. 2010–11 NECAP: Cumulative Score Distribution—Mathematics Grade 11

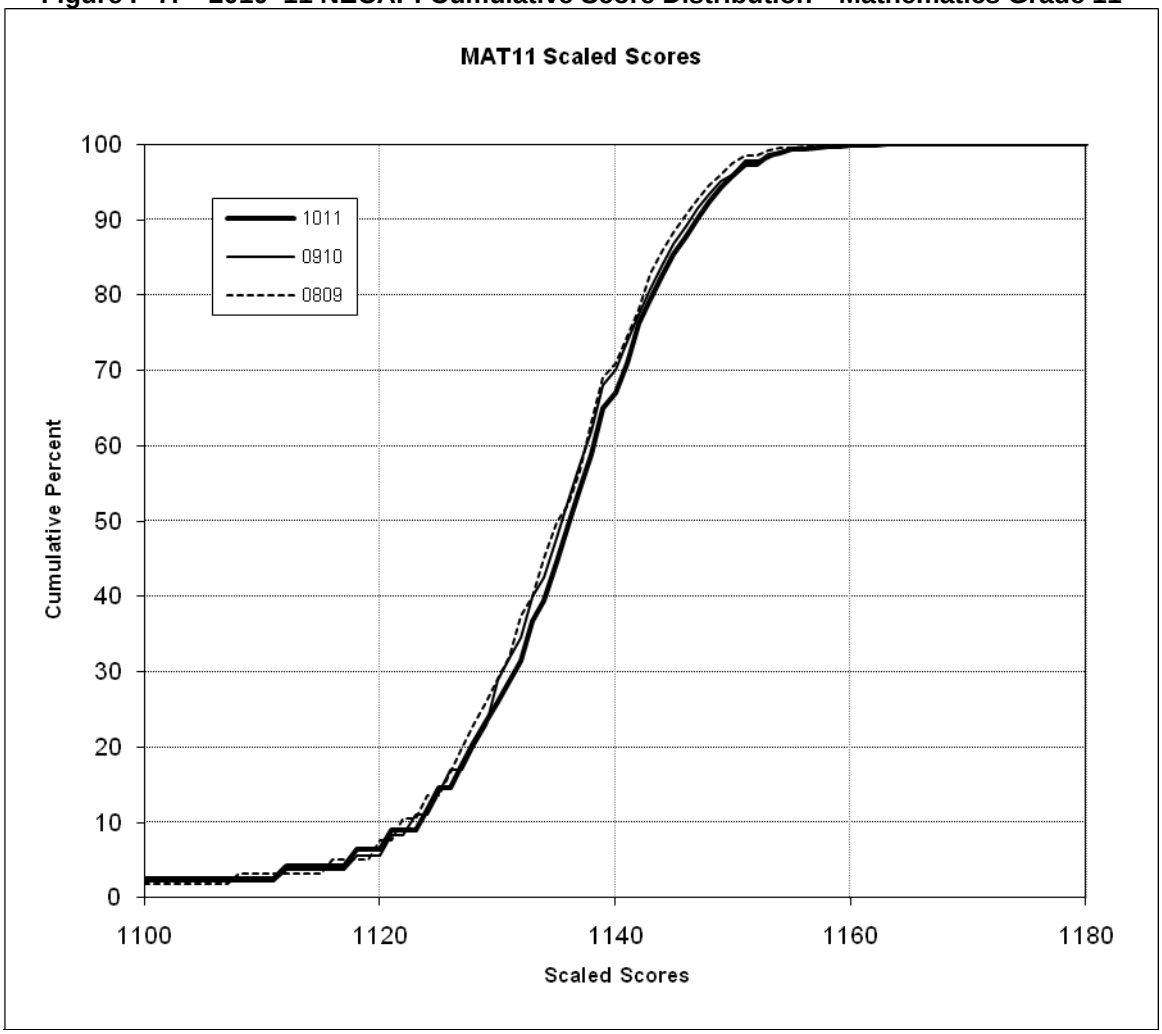


Figure P-8. 2010–11 NECAP: Cumulative Score Distribution—Reading Grade 3

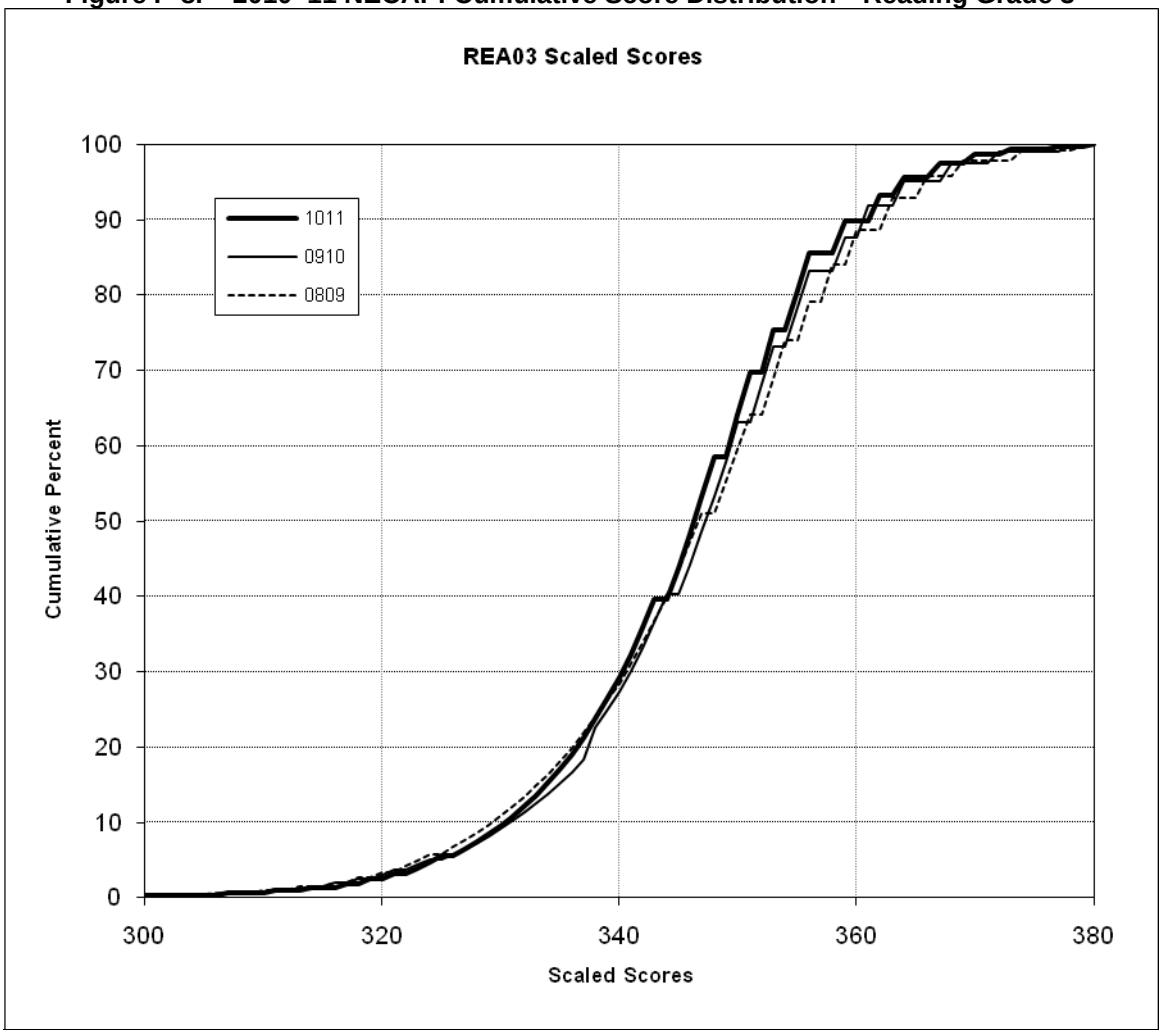


Figure P-9. 2010–11 NECAP: Cumulative Score Distribution—Reading Grade 4

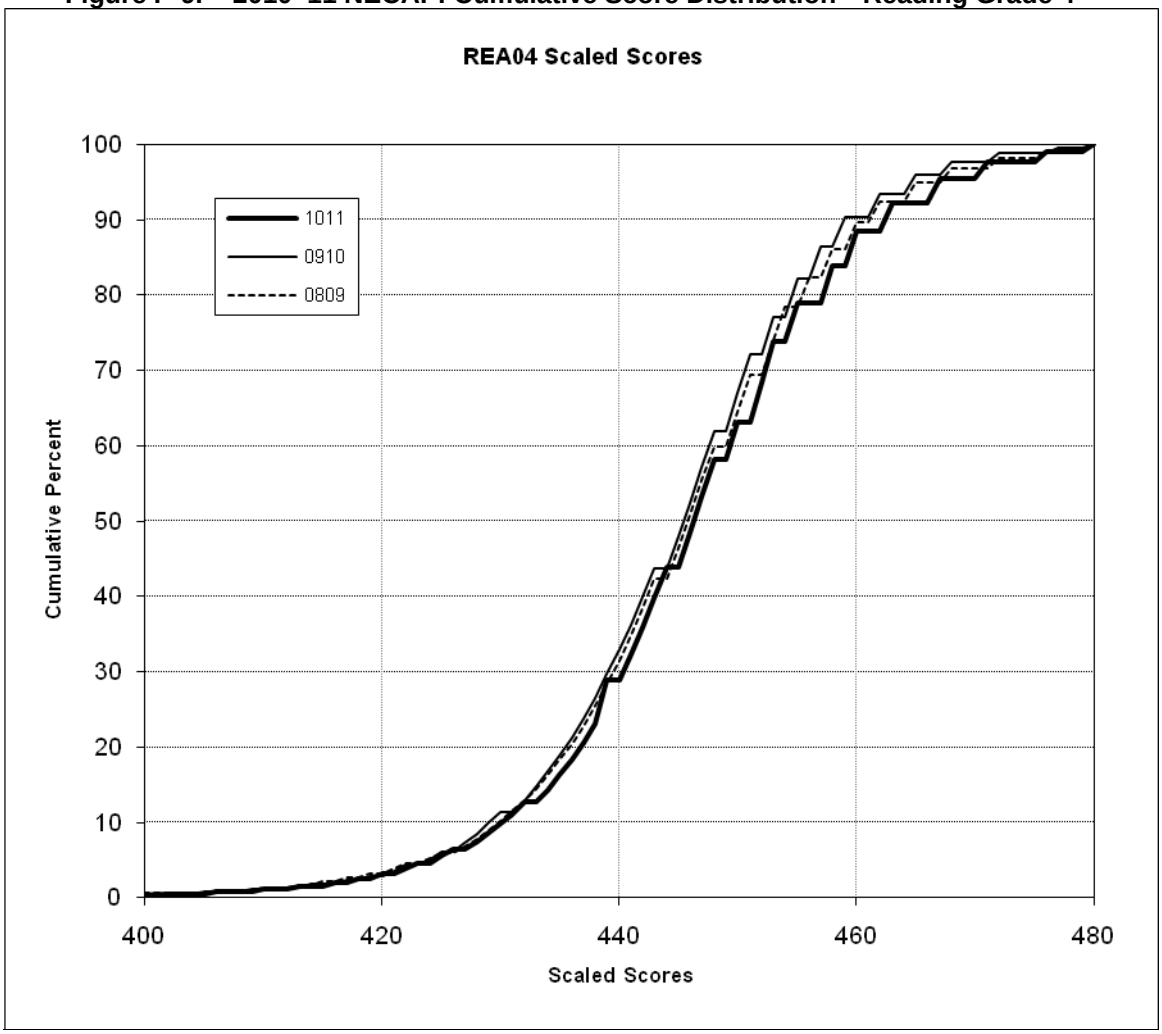


Figure P-10. 2010–11 NECAP: Cumulative Score Distribution—Reading Grade 5

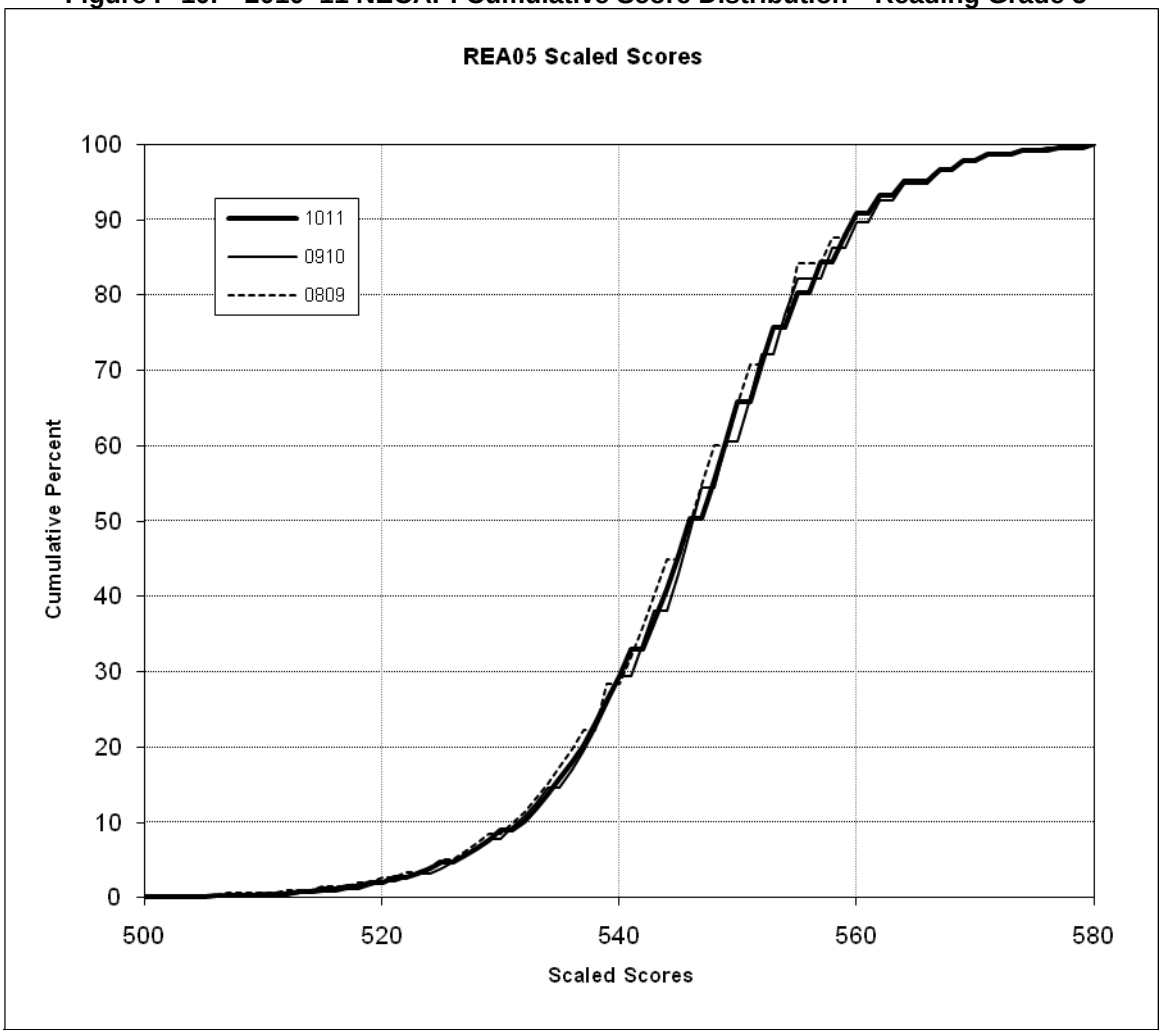


Figure P-11. 2010–11 NECAP: Cumulative Score Distribution—Reading Grade 6

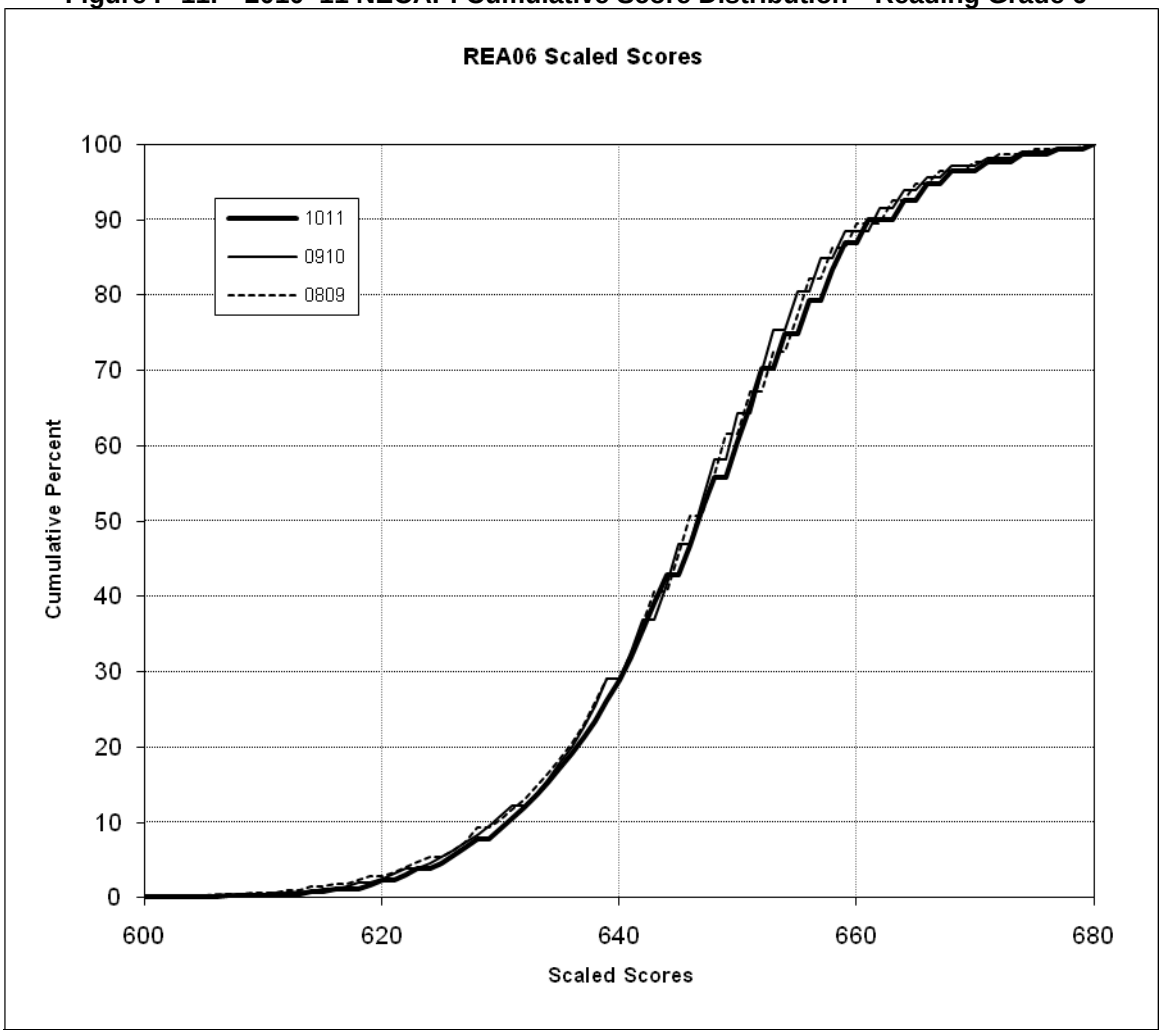


Figure P-12. 2010–11 NECAP: Cumulative Score Distribution—Reading Grade 7

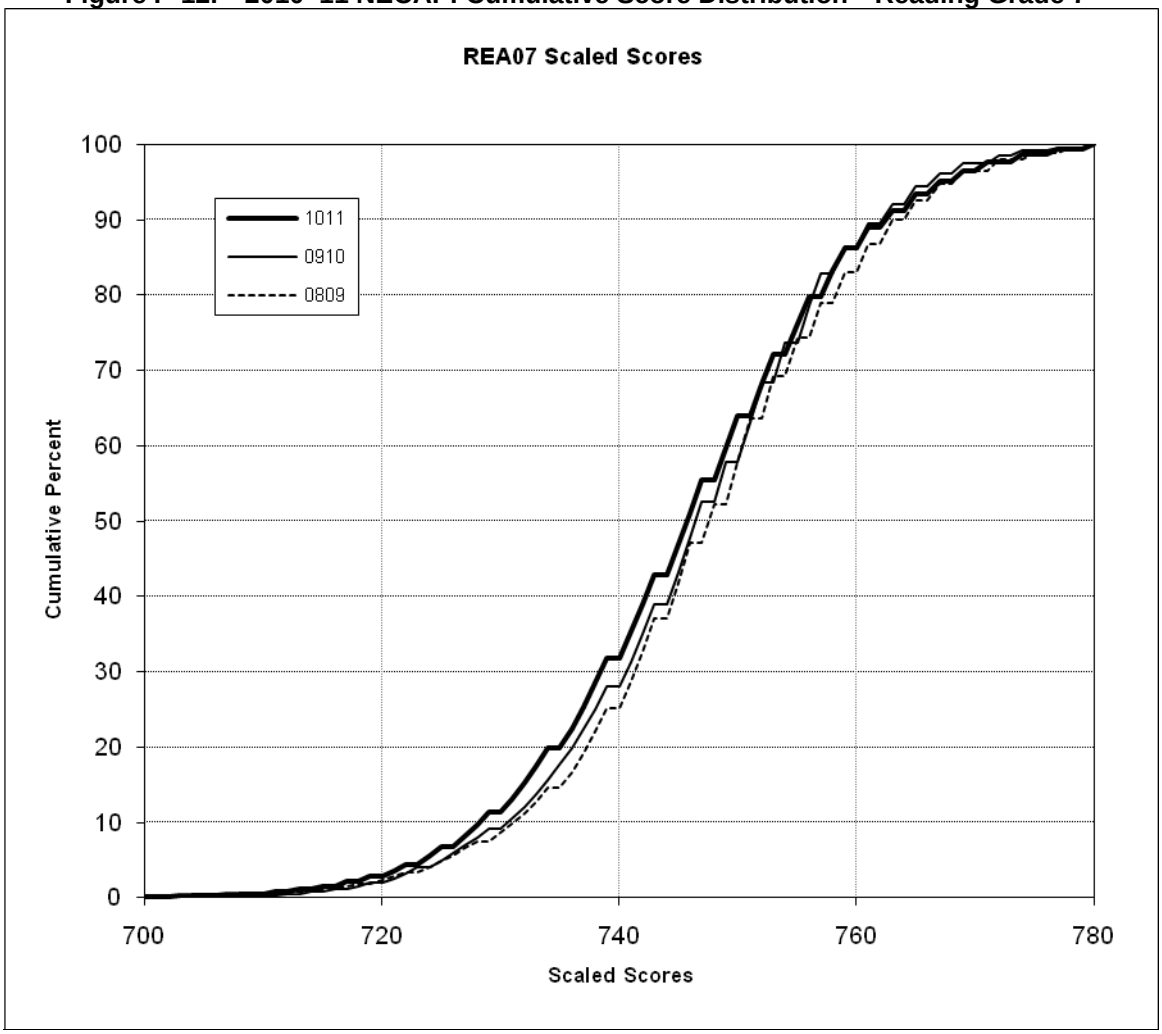


Figure P-13. 2010–11 NECAP: Cumulative Score Distribution—Reading Grade 8

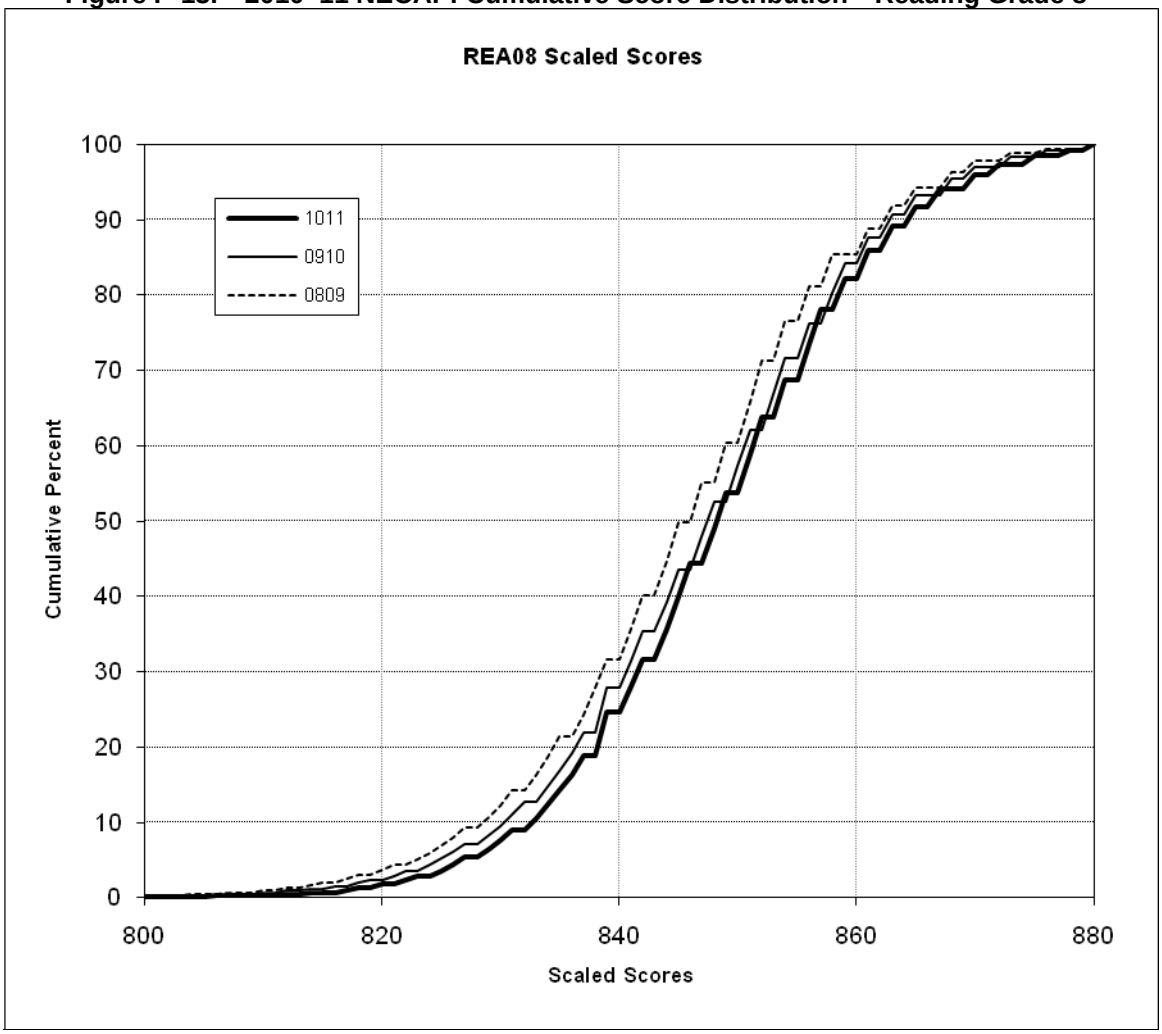


Figure P-14. 2010–11 NECAP: Cumulative Score Distribution—Reading Grade 11

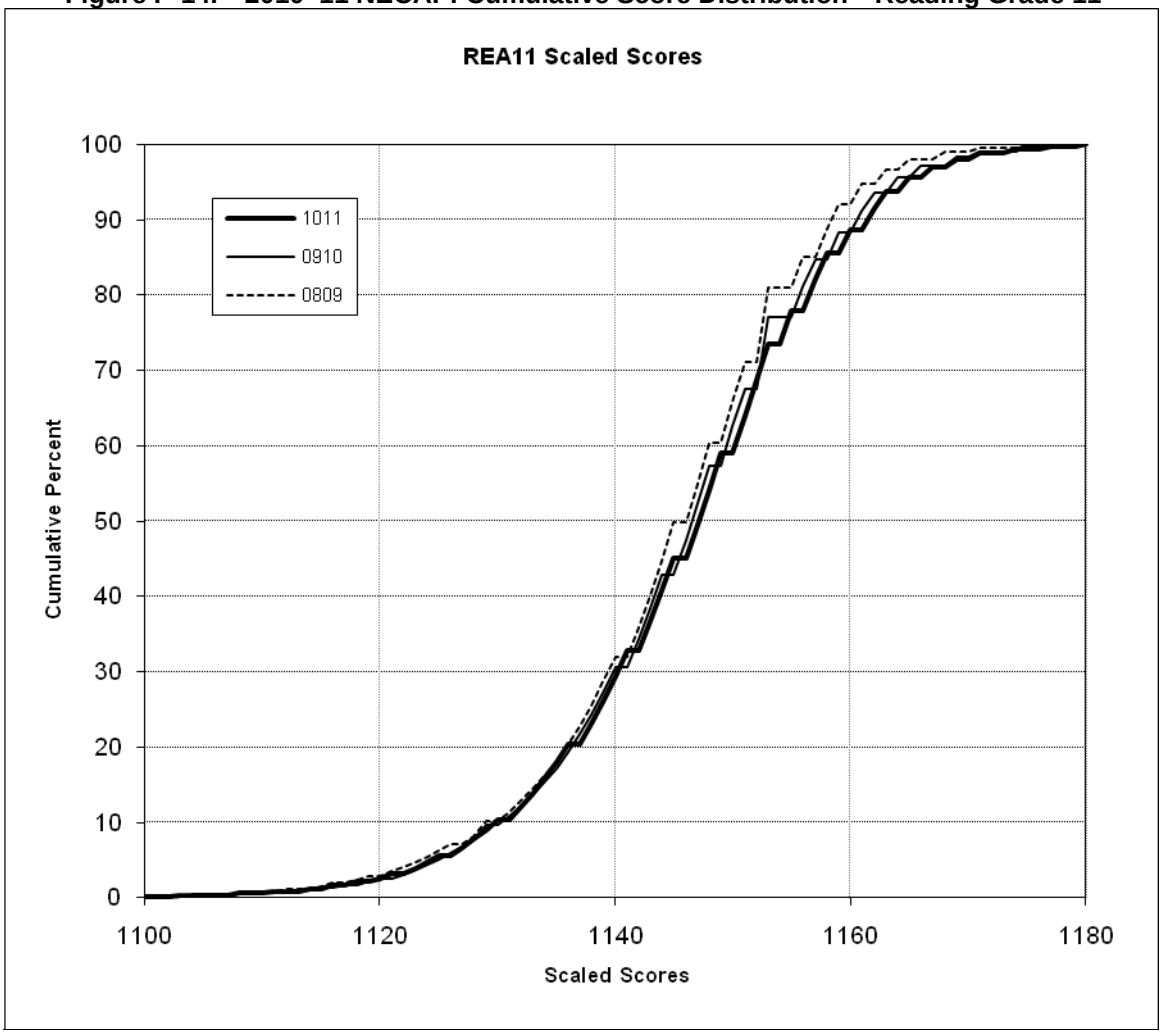


Figure P-15. 2010–11 NECAP: Cumulative Score Distribution—Writing Grade 5

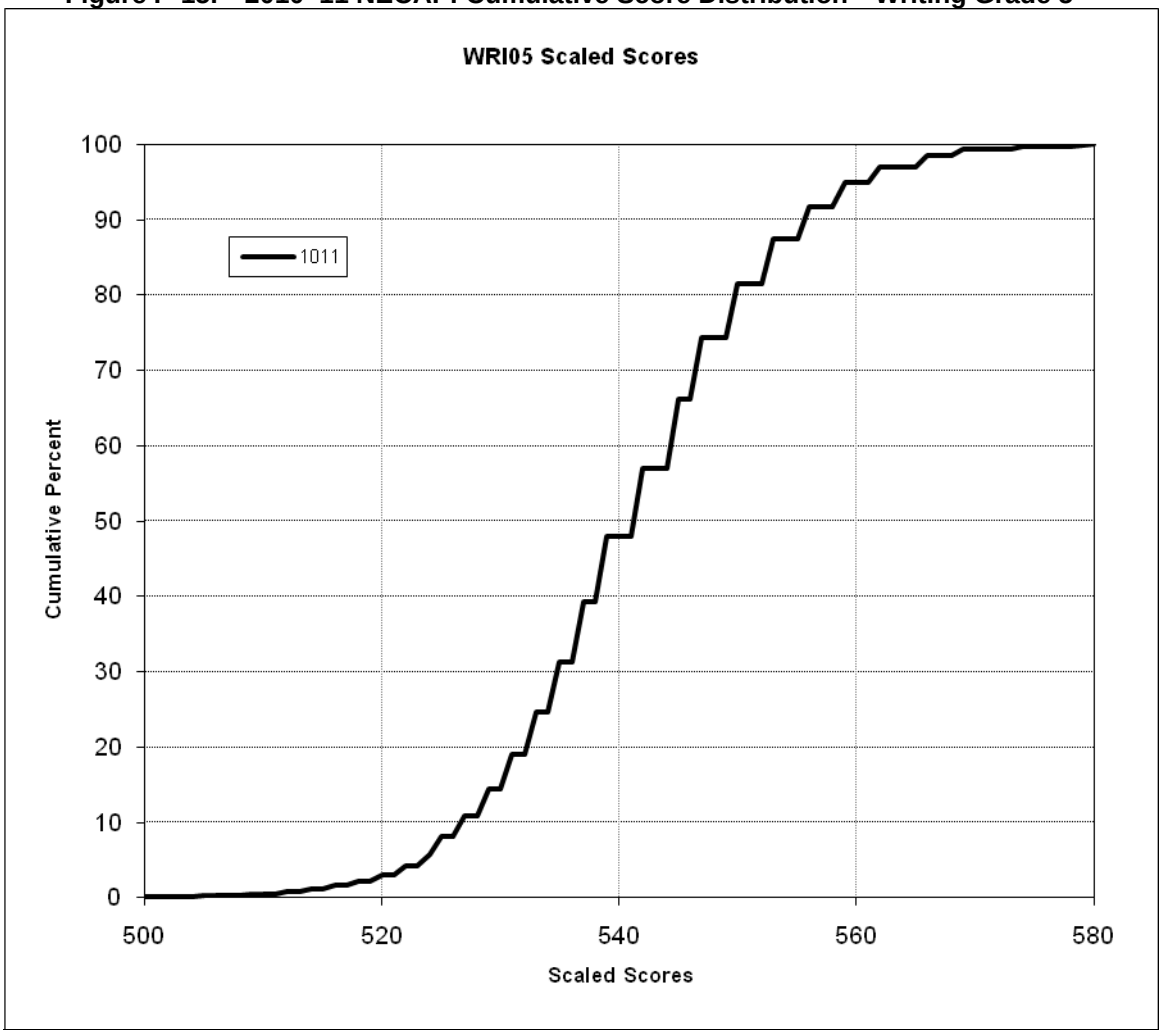


Figure P-16. 2010–11 NECAP: Cumulative Score Distribution—Writing Grade 8

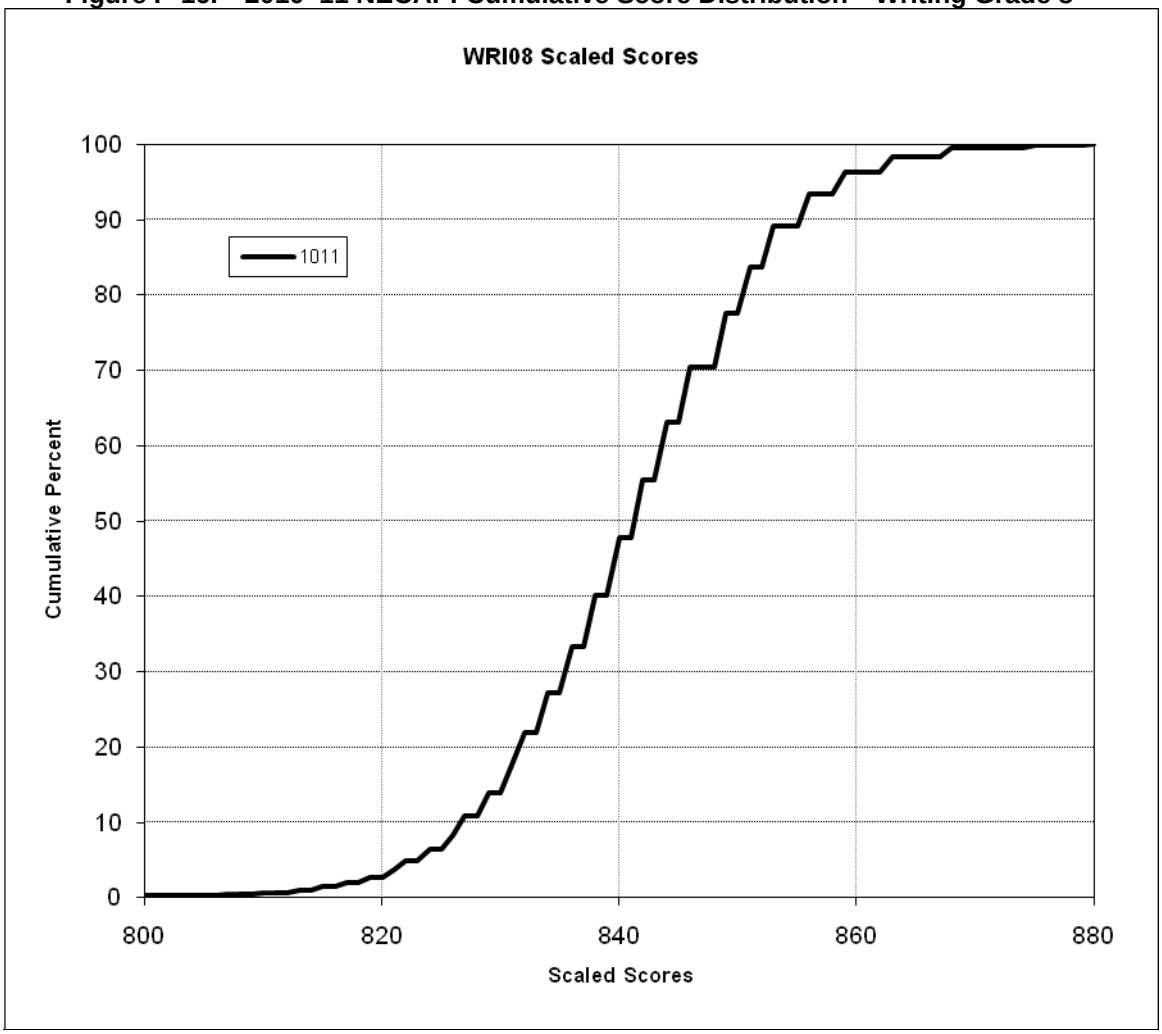


Figure P-17. 2010–11 NECAP: Cumulative Score Distribution—Writing Grade 11

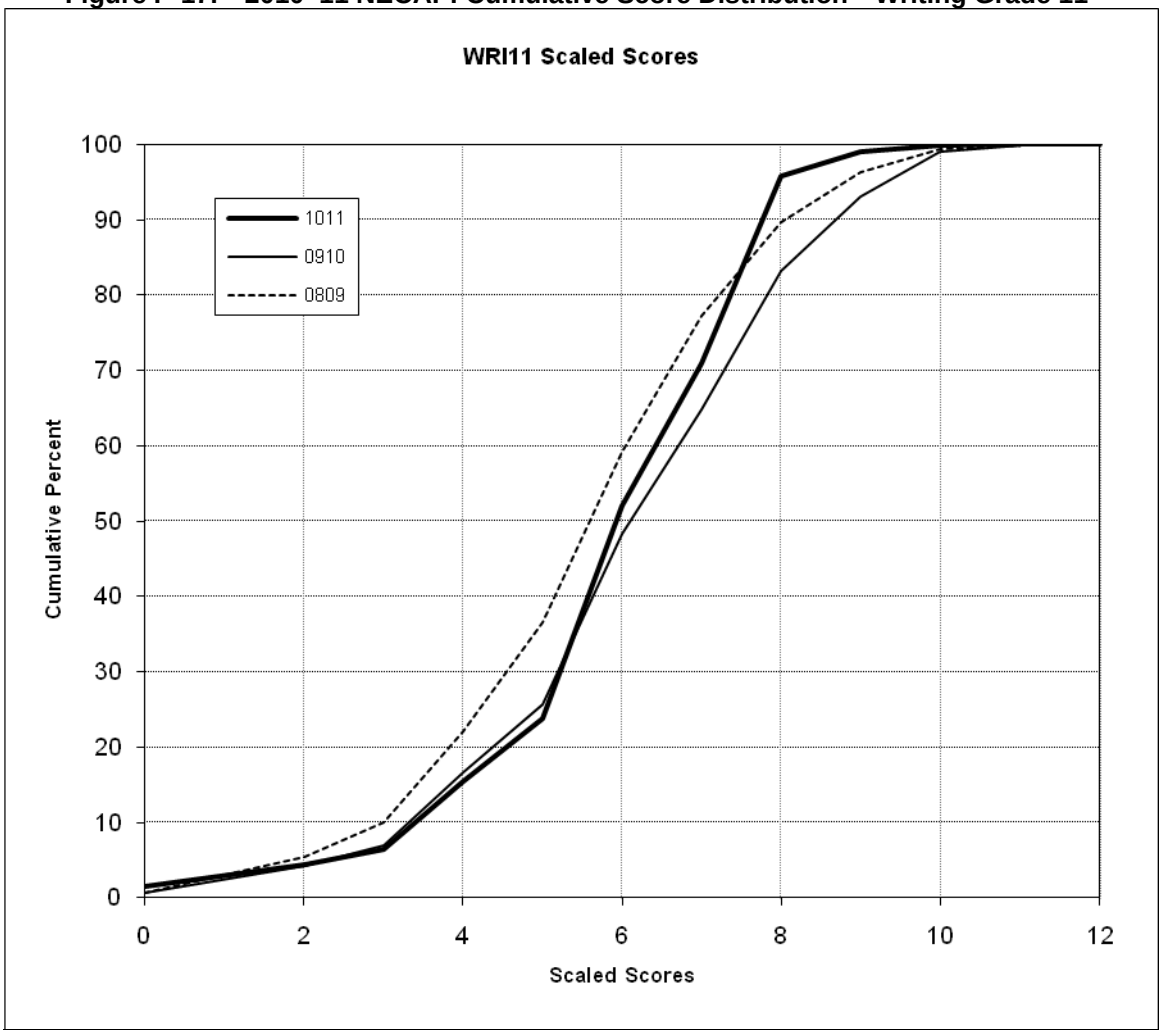


Table P-1. 2010–11 NECAP: Scaled Score Distribution—Mathematics Grade 3

Scaled score	*Percentage*	*Cumulative percentage*
300	0.2	0.2
303	0.1	0.3
308	0.2	0.5
311	0.3	0.8
314	0.3	1.2
316	0.4	1.5
318	0.5	2.0
320	0.6	2.6
321	0.6	3.2
322	0.6	3.8
323	0.7	4.5
324	0.7	5.2
325	0.7	5.9
326	0.9	6.8
327	0.9	7.7
328	0.9	8.6
329	1.0	9.6
330	1.1	10.7
331	2.2	12.9
332	1.3	14.2
333	1.3	15.5
334	2.7	18.2
335	1.6	19.8
336	3.4	23.2
337	1.8	25.0
338	3.8	28.8
339	4.2	33.0
341	4.6	37.6
342	2.4	40.0
343	5.2	45.1
344	2.6	47.7
345	2.8	50.4
346	6.1	56.5
347	3.1	59.6
348	3.2	62.7
349	3.2	65.9
350	3.0	68.9
351	6.5	75.3
352	3.2	78.5
353	3.0	81.6
355	2.8	84.4
356	2.8	87.2
357	2.6	89.7
358	2.4	92.1
360	2.1	94.3
362	1.8	96.0
364	1.5	97.5
367	1.1	98.6
371	0.8	99.4
377	0.4	99.9
380	0.2	100.0

Table P-2. 2010–11 NECAP—Scaled Score Distribution—Mathematics Grade 4

Scaled score	*Percentage*	*Cumulative percentage*
400	0.2	0.2
402	0.2	0.4
406	0.2	0.6
408	0.3	0.9
411	0.3	1.2
413	0.4	1.6
415	0.5	2.1
416	0.5	2.5
418	0.5	3.0
419	0.5	3.5
420	0.6	4.0
422	0.6	4.6
423	0.6	5.3
424	0.7	5.9
425	0.8	6.7
426	0.9	7.6
427	1.0	8.5
428	1.1	9.6
429	1.2	10.8
430	2.6	13.4
432	1.5	14.8
433	1.5	16.4
434	1.7	18.1
435	2.0	20.0
436	4.2	24.3
437	2.3	26.6
438	2.3	28.9
439	5.2	34.1
441	2.9	36.9
442	2.8	39.7
443	5.9	45.6
444	3.2	48.8
445	3.1	51.9
446	3.0	54.9
447	3.2	58.1
448	3.1	61.2
449	3.2	64.5
450	3.2	67.6
451	3.0	70.6
452	2.9	73.5
453	3.1	76.6
454	2.7	79.4
455	2.9	82.3
456	2.6	84.9
457	2.5	87.4
459	2.4	89.7
460	2.2	91.9
462	2.1	94.0
463	1.7	95.7
465	1.4	97.1
468	1.2	98.3
471	0.8	99.1
476	0.5	99.6
480	0.4	100.0

Table P-3. 2010–11 NECAP: Scaled Score Distribution—Mathematics Grade 5

Scaled score	*Percentage*	*Cumulative percentage*
500	0.5	0.5
506	0.3	0.8
511	0.4	1.2
514	0.6	1.8
517	0.7	2.4
519	0.8	3.2
521	0.9	4.1
523	1.0	5.2
524	1.0	6.2
526	1.2	7.3
527	1.2	8.6
528	1.3	9.8
529	1.4	11.2
530	1.4	12.6
531	1.5	14.1
532	3.2	17.4
534	1.7	19.1
535	3.3	22.4
536	1.8	24.2
537	2.0	26.1
538	4.1	30.2
539	4.1	34.3
541	4.4	38.7
542	2.2	40.9
543	4.6	45.5
544	2.3	47.8
545	4.6	52.4
546	2.5	54.9
547	4.8	59.7
548	2.5	62.1
549	4.9	67.0
550	2.5	69.5
551	2.4	71.9
552	4.8	76.6
553	4.5	81.1
555	2.1	83.2
556	4.2	87.4
557	1.9	89.3
558	1.8	91.2
560	1.7	92.9
561	1.6	94.4
562	1.4	95.9
564	1.2	97.1
565	1.1	98.1
567	0.8	98.9
570	0.6	99.5
574	0.3	99.8
580	0.2	100.0

Table P-4. 2010–11 NECAP: Scaled Score Distribution—Mathematics Grade 6

Scaled score	*Percentage*	*Cumulative percentage*
600	0.5	0.5
606	0.5	1.0
613	0.7	1.7
616	0.8	2.6
619	1.0	3.5
621	1.1	4.6
623	1.2	5.8
625	1.3	7.1
626	1.3	8.3
627	1.3	9.7
628	1.4	11.1
630	1.5	12.6
631	1.5	14.1
632	3.0	17.1
633	1.5	18.6
634	1.7	20.3
635	1.9	22.2
636	1.9	24.1
637	3.9	27.9
638	1.9	29.8
639	4.4	34.2
640	2.2	36.4
641	2.3	38.6
642	4.5	43.2
643	2.4	45.6
644	4.5	50.1
645	2.4	52.5
646	4.4	56.9
647	2.2	59.1
648	4.5	63.6
649	2.1	65.7
650	4.3	69.9
651	2.1	72.0
652	4.2	76.2
653	2.1	78.3
654	2.1	80.4
655	2.0	82.4
656	1.9	84.4
657	1.9	86.3
658	3.5	89.8
659	1.6	91.4
661	1.7	93.0
662	1.5	94.5
663	1.3	95.8
665	1.2	97.0
666	1.0	97.9
668	0.7	98.7
671	0.6	99.2
674	0.4	99.6
679	0.2	99.9
680	0.2	100.0

Table P-5. 2010–11 NECAP: Scaled Score Distribution—Mathematics Grade 7

Scaled score	*Percentage*	*Cumulative percentage*
700	0.5	0.5
702	0.5	1.0
711	0.7	1.7
716	0.9	2.6
719	1.2	3.8
722	1.4	5.2
724	1.5	6.6
726	1.7	8.3
727	1.7	10.0
729	1.7	11.7
730	1.8	13.6
731	2.1	15.7
732	2.1	17.8
733	2.2	20.0
734	2.4	22.4
735	2.4	24.8
736	2.4	27.2
737	5.0	32.2
738	2.5	34.7
739	5.0	39.7
740	2.5	42.2
741	2.6	44.7
742	5.0	49.7
743	2.4	52.1
744	4.7	56.8
745	2.2	59.0
746	4.4	63.4
747	4.3	67.7
748	2.1	69.8
749	3.9	73.8
750	2.0	75.8
751	5.4	81.1
753	3.4	84.5
754	1.6	86.1
755	1.6	87.7
756	2.8	90.5
757	1.4	91.9
758	1.2	93.1
759	1.2	94.3
760	1.1	95.4
761	1.0	96.4
762	1.0	97.4
764	0.7	98.1
765	0.6	98.7
767	0.5	99.2
769	0.3	99.5
771	0.2	99.7
775	0.2	99.9
780	0.1	100.0

Table P-6. 2010–11 NECAP: Scaled Score Distribution—Mathematics Grade 8

Scaled score	*Percentage*	*Cumulative percentage*
800	0.8	0.8
808	0.6	1.4
817	1.0	2.4
821	1.1	3.5
824	1.3	4.9
826	1.5	6.3
827	1.5	7.9
829	1.7	9.5
830	1.9	11.4
831	1.8	13.2
832	1.9	15.1
833	4.1	19.2
834	2.1	21.3
835	2.2	23.5
836	2.2	25.7
837	4.7	30.4
838	2.2	32.6
839	7.0	39.5
840	2.4	41.9
841	2.5	44.3
842	4.8	49.2
843	4.9	54.1
844	2.3	56.3
845	4.6	60.9
846	4.5	65.4
847	4.4	69.8
848	2.1	71.9
849	4.1	76.0
850	3.8	79.8
851	1.8	81.7
852	3.5	85.1
853	1.6	86.7
854	1.6	88.3
855	3.0	91.3
856	1.4	92.7
857	1.2	93.9
858	1.1	95.0
859	1.1	96.0
860	0.9	96.9
861	0.8	97.7
862	0.7	98.4
864	0.5	98.9
865	0.4	99.3
867	0.3	99.7
869	0.2	99.9
872	0.1	99.9
878	0.1	100.0
880	0.0	100.0

Table P-7. 2010–11 NECAP: Scaled Score Distribution—Mathematics Grade 11

Scaled score	*Percentage*	*Cumulative percentage*
1100	2.5	2.5
1112	1.8	4.3
1118	2.2	6.5
1121	2.6	9.0
1124	2.7	11.7
1125	2.9	14.7
1127	3.0	17.7
1128	2.9	20.6
1129	2.8	23.4
1130	2.6	26.0
1131	2.8	28.8
1132	2.7	31.5
1133	5.3	36.7
1134	2.7	39.4
1135	5.2	44.6
1136	4.9	49.5
1137	4.7	54.2
1138	4.6	58.8
1139	6.2	65.0
1140	2.1	67.1
1141	3.9	71.0
1142	5.3	76.3
1143	3.3	79.5
1144	3.0	82.5
1145	2.8	85.3
1146	2.5	87.8
1147	2.4	90.2
1148	2.2	92.4
1149	1.8	94.3
1150	1.6	95.8
1151	1.5	97.3
1153	1.1	98.5
1154	0.4	98.9
1155	0.4	99.3
1157	0.3	99.6
1158	0.2	99.8
1160	0.1	99.9
1163	0.1	100.0
1168	0.0	100.0
1180	0.0	100.0

Table P-8. 2010–11 NECAP: Scaled Score Distribution—Reading Grade 3

Scaled score	*Percentage*	*Cumulative percentage*
300	0.2	0.2
302	0.2	0.4
307	0.3	0.6
311	0.3	1.0
314	0.4	1.4
317	0.5	1.9
319	0.6	2.5
321	0.7	3.1
323	0.7	3.8
324	0.8	4.6
325	0.9	5.5
327	0.9	6.5
328	1.0	7.4
329	1.0	8.4
330	1.1	9.5
331	1.2	10.7
332	1.4	12.1
333	1.4	13.6
334	1.7	15.3
335	1.8	17.0
336	2.1	19.1
337	2.2	21.3
338	2.5	23.8
339	2.6	26.4
340	2.9	29.3
341	3.0	32.3
342	3.6	35.9
343	3.8	39.7
345	4.0	43.7
346	4.5	48.2
347	5.1	53.3
348	5.3	58.6
350	5.6	64.1
351	5.6	69.7
353	5.6	75.3
355	5.3	80.6
356	4.9	85.6
359	4.3	89.9
362	3.4	93.2
364	2.5	95.7
367	1.8	97.5
370	1.2	98.7
373	0.7	99.4
377	0.3	99.7
380	0.3	100.0

Table P-9. 2010–11 NECAP: Scaled Score Distribution—Reading Grade 4

Scaled score	*Percentage*	*Cumulative percentage*
400	0.3	0.3
401	0.2	0.5
406	0.3	0.8
410	0.3	1.1
413	0.4	1.5
416	0.5	2.0
418	0.5	2.5
420	0.6	3.1
422	0.7	3.8
423	0.8	4.6
425	0.9	5.5
426	0.9	6.4
428	1.0	7.4
429	1.2	8.6
430	1.2	9.8
431	1.3	11.1
432	1.6	12.7
434	1.6	14.3
435	2.0	16.3
436	2.1	18.3
437	2.2	20.5
438	2.5	23.0
439	5.9	28.9
441	3.4	32.3
442	3.6	35.9
443	3.9	39.8
444	4.1	43.9
446	4.6	48.5
447	4.7	53.2
448	5.0	58.1
450	5.0	63.1
452	5.3	68.4
453	5.4	73.8
455	5.2	79.0
458	4.9	83.9
460	4.5	88.4
463	3.7	92.2
467	3.2	95.4
471	2.3	97.6
476	1.4	99.0
480	1.0	100.0

Table P-10. 2010–11 NECAP: Scaled Score Distribution—Reading Grade 5

Scaled score	*Percentage*	*Cumulative percentage*
500	0.2	0.2
506	0.2	0.3
510	0.2	0.5
513	0.3	0.8
515	0.4	1.2
517	0.4	1.6
519	0.5	2.0
521	0.6	2.6
523	0.6	3.2
524	0.7	3.9
525	0.8	4.7
527	0.9	5.6
528	1.0	6.6
529	1.1	7.7
530	1.3	8.9
532	1.3	10.3
533	1.6	11.9
534	1.9	13.7
535	2.1	15.9
536	2.1	17.9
537	2.3	20.3
538	2.9	23.2
539	2.9	26.1
540	3.3	29.4
541	3.6	33.0
543	3.8	36.8
544	4.2	41.0
545	4.5	45.4
546	4.9	50.3
548	5.1	55.4
549	5.2	60.6
550	5.3	65.9
552	5.1	71.0
553	4.8	75.7
555	4.6	80.3
557	4.1	84.5
559	3.5	87.9
560	3.0	90.9
562	2.4	93.3
564	1.9	95.2
567	1.5	96.7
569	1.1	97.8
571	0.8	98.6
574	0.5	99.1
577	0.4	99.5
580	0.5	100.0

Table P-11. 2010–11 NECAP: Scaled Score Distribution—Reading Grade 6

Scaled score	*Percentage*	*Cumulative percentage*
600	0.1	0.1
607	0.1	0.3
611	0.2	0.5
614	0.3	0.8
616	0.4	1.2
619	0.5	1.7
620	0.6	2.3
622	0.7	3.0
623	0.8	3.8
625	0.8	4.6
626	0.9	5.5
627	1.1	6.6
628	1.2	7.8
630	1.3	9.1
631	1.4	10.5
632	1.6	12.0
633	1.6	13.6
634	1.6	15.3
635	1.8	17.1
636	2.1	19.2
637	2.1	21.3
638	2.2	23.5
639	2.6	26.1
640	2.8	28.9
641	3.1	32.0
642	3.5	35.5
643	3.6	39.1
644	3.8	42.9
646	4.0	46.9
647	4.5	51.3
648	4.5	55.8
650	4.7	60.5
651	4.9	65.4
652	4.9	70.3
654	4.5	74.8
656	4.6	79.3
658	4.1	83.4
659	3.6	87.0
661	3.0	90.0
664	2.7	92.6
666	2.1	94.7
668	1.8	96.5
671	1.2	97.7
674	1.0	98.6
677	0.7	99.3
680	0.7	100.0

Table P-12. 2010–11 NECAP: Scaled Score Distribution—Reading Grade 7

Scaled score	*Percentage*	*Cumulative percentage*
700	0.2	0.2
703	0.1	0.3
707	0.2	0.5
711	0.3	0.8
713	0.4	1.1
715	0.4	1.6
717	0.6	2.1
719	0.7	2.8
721	0.7	3.5
722	0.9	4.4
724	1.1	5.5
725	1.3	6.8
727	1.3	8.1
728	1.5	9.6
729	1.7	11.3
731	1.8	13.1
732	2.1	15.2
733	2.2	17.4
734	2.4	19.8
736	2.7	22.5
737	2.8	25.3
738	3.2	28.5
739	3.3	31.8
741	3.6	35.3
742	3.6	39.0
743	3.9	42.9
745	4.1	47.0
746	4.1	51.1
747	4.4	55.5
749	4.3	59.8
750	4.3	64.0
752	4.1	68.2
753	4.1	72.2
755	4.0	76.2
756	3.6	79.9
758	3.3	83.1
759	3.1	86.2
761	2.7	88.9
763	2.3	91.2
765	2.1	93.3
767	1.8	95.1
769	1.4	96.5
771	1.2	97.7
774	1.0	98.6
777	0.7	99.3
780	0.7	100.0

Table P-13. 2010–11 NECAP: Scaled Score Distribution—Reading Grade 8

Scaled score	*Percentage*	*Cumulative percentage*
800	0.1	0.1
802	0.1	0.2
806	0.1	0.3
809	0.1	0.4
812	0.1	0.5
814	0.2	0.7
817	0.3	1.1
818	0.3	1.4
820	0.4	1.8
822	0.5	2.3
823	0.6	2.9
825	0.7	3.6
826	0.8	4.4
827	0.9	5.4
829	1.0	6.4
830	1.2	7.6
831	1.3	9.0
833	1.5	10.5
834	1.8	12.3
835	1.9	14.2
836	2.2	16.4
837	2.4	18.8
839	5.9	24.7
841	3.4	28.0
842	3.7	31.7
844	4.0	35.7
845	4.3	39.9
846	4.4	44.3
848	4.7	49.0
849	4.8	53.8
851	5.0	58.8
852	5.0	63.8
854	5.0	68.8
856	4.7	73.5
857	4.5	78.0
859	4.1	82.1
861	3.8	85.9
863	3.2	89.1
865	2.7	91.8
867	2.3	94.0
870	1.9	95.9
872	1.4	97.3
875	1.2	98.5
878	0.7	99.2
880	0.8	100.0

Table P-14. 2010–11 NECAP: Scaled Score Distribution—Reading Grade 11

Scaled score	*Percentage*	*Cumulative percentage*
1100	0.2	0.2
1103	0.1	0.4
1108	0.2	0.6
1111	0.3	0.8
1114	0.4	1.2
1116	0.4	1.6
1118	0.5	2.1
1120	0.5	2.6
1121	0.6	3.2
1123	0.7	3.9
1124	0.8	4.7
1125	0.9	5.6
1127	1.0	6.6
1128	1.2	7.7
1129	1.3	9.0
1130	1.4	10.4
1132	1.7	12.1
1133	1.7	13.8
1134	2.0	15.8
1135	2.2	18.0
1136	2.5	20.5
1138	2.6	23.1
1139	3.0	26.1
1140	3.3	29.3
1141	3.5	32.9
1143	3.9	36.7
1144	4.1	40.8
1145	4.3	45.1
1147	4.4	49.5
1148	4.6	54.1
1149	4.9	59.1
1151	4.9	63.9
1152	4.8	68.8
1153	4.8	73.5
1155	4.4	77.9
1157	4.2	82.1
1158	3.5	85.6
1160	3.1	88.7
1162	2.7	91.5
1163	2.3	93.7
1165	1.8	95.5
1167	1.4	96.9
1169	1.1	98.1
1171	0.8	98.9
1174	0.5	99.4
1177	0.3	99.7
1180	0.3	100.0

Table P-15. 2010–11 NECAP: Scaled Score Distribution—Writing Grade 5

Scaled score	*Percentage*	*Cumulative percentage*
500	0.2	0.2
505	0.1	0.3
509	0.2	0.5
512	0.3	0.7
514	0.4	1.1
516	0.5	1.6
518	0.6	2.2
520	0.8	3.0
522	1.2	4.2
524	1.7	5.8
525	2.2	8.1
527	2.8	10.8
529	3.5	14.4
531	4.6	19.0
533	5.6	24.6
535	6.7	31.4
537	7.9	39.3
539	8.6	47.9
542	9.1	57.0
545	9.2	66.2
547	8.2	74.3
550	7.2	81.5
553	6.0	87.5
556	4.3	91.8
559	3.1	94.9
562	2.2	97.1
566	1.4	98.5
569	0.8	99.3
574	0.5	99.7
579	0.2	99.9
580	0.1	100.0

Table P-16. 2010–11 NECAP: Scaled Score Distribution—Writing Grade 8

Scaled score	*Percentage*	*Cumulative percentage*
800	0.3	0.3
807	0.2	0.5
810	0.2	0.7
813	0.4	1.0
815	0.5	1.5
817	0.6	2.0
819	0.7	2.8
821	0.9	3.7
822	1.3	4.9
824	1.5	6.4
826	1.9	8.3
827	2.5	10.8
829	3.1	13.8
831	3.8	17.6
832	4.3	22.0
834	5.2	27.2
836	6.1	33.3
838	6.9	40.1
840	7.6	47.7
842	7.8	55.5
844	7.7	63.2
846	7.3	70.4
849	7.1	77.5
851	6.2	83.7
853	5.4	89.1
856	4.2	93.3
859	3.1	96.4
863	2.0	98.4
868	1.1	99.5
875	0.4	99.9
880	0.1	100.0

Table P-17. 2010–11 NECAP: Scaled Score Distribution—Writing Grade 11

Scaled score	*Percentage*	*Cumulative percentage*
0	1.5	1.5
2	2.8	4.3
3	2.1	6.4
4	9.1	15.5
5	8.3	23.8
6	28.2	52.0
7	18.9	70.9
8	24.8	95.7
9	3.2	98.9
10	0.9	99.9
11	0.1	100.0
12	0.0	100.0

Appendix Q—CLASSICAL RELIABILITIES

Table Q-1. 2010–11 NECAP: Subgroup Reliabilities – Mathematics

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
3	All Students	43893	65	42.14	13.02	0.93	3.43
	Male	22551	65	42.47	13.01	0.93	3.42
	Female	21335	65	41.79	13.02	0.93	3.43
	Gender Not Reported	7	65				
	Hispanic or Latino	3245	65	34.59	13.31	0.93	3.51
	American Indian or Alaskan Native	284	65	38.38	13.16	0.93	3.53
	Asian	1014	65	43.74	13.25	0.94	3.36
	Black or African American	1573	65	32.74	13.42	0.93	3.52
	Native Hawaiian or Pacific Islander	38	65	41.29	13.44	0.94	3.40
	White (Non-Hispanic)	36911	65	43.26	12.55	0.93	3.41
	Two or More Races (non-Hispanic)	556	65	38.15	13.53	0.93	3.48
	No Primary Race/Ethnicity Reported	272	65	40.92	13.83	0.94	3.42
	Currently Receiving LEP services	1843	65	31.03	13.59	0.93	3.55
	Former LEP Student - Monitoring Year 1	165	65	43.01	12.14	0.92	3.39
	Former LEP Student - Monitoring Year 2	25	65	44.00	13.16	0.94	3.29
	LEP: All Other Students	41860	65	42.63	12.78	0.93	3.42
	Students with an IEP	5775	65	31.61	14.08	0.94	3.56
	IEP: All Other Students	38118	65	43.74	12.08	0.92	3.40
	Economically Disadvantaged Students	17148	65	36.76	13.14	0.93	3.52
	SES: All Other Students	26745	65	45.59	11.71	0.92	3.35
	Migrant Students	17	65	34.29	16.50	0.96	3.48
	Migrant: All Other Students	43876	65	42.14	13.02	0.93	3.43
	Students Receiving Title 1 Services	9731	65	35.41	13.12	0.93	3.53
	Title 1: All Other Students	34162	65	44.06	12.33	0.92	3.39
	Plan 504	265	65	40.26	12.44	0.92	3.51
	Plan 504: All Other Students	43628	65	42.15	13.02	0.93	3.43
4	All Students	44350	65	41.67	12.08	0.92	3.39
	Male	22829	65	41.69	12.37	0.92	3.39
	Female	21518	65	41.65	11.75	0.92	3.38
	Gender Not Reported	3	65				
	Hispanic or Latino	3178	65	35.56	12.54	0.92	3.48
	American Indian or Alaskan Native	287	65	37.76	13.55	0.93	3.48
	Asian	1069	65	44.21	12.58	0.93	3.30
	Black or African American	1610	65	33.93	13.04	0.93	3.49
	Native Hawaiian or Pacific Islander	37	65	40.81	11.93	0.91	3.50
	White (Non-Hispanic)	37395	65	42.51	11.66	0.92	3.37
	Two or More Races (non-Hispanic)	531	65	39.65	12.66	0.93	3.44
	No Primary Race/Ethnicity Reported	243	65	41.62	12.77	0.93	3.36
	Currently Receiving LEP services	1648	65	30.95	12.92	0.93	3.53
	Former LEP Student - Monitoring Year 1	370	65	41.95	10.88	0.90	3.36
	Former LEP Student - Monitoring Year 2	60	65	44.83	12.14	0.93	3.28
	LEP: All Other Students	42272	65	42.08	11.86	0.92	3.38
	Students with an IEP	6158	65	30.54	12.63	0.92	3.55
	IEP: All Other Students	38192	65	43.46	10.97	0.91	3.35
	Economically Disadvantaged Students	17258	65	37.10	12.18	0.92	3.49
	SES: All Other Students	27092	65	44.58	11.07	0.91	3.31
	Migrant Students	16	65	31.31	13.11	0.93	3.50
	Migrant: All Other Students	44334	65	41.67	12.07	0.92	3.39
	Students Receiving Title 1 Services	9240	65	35.94	11.95	0.91	3.50

	Title 1: All Other Students	35110	65	43.18	11.65	0.92	3.35
	Plan 504	338	65	39.21	10.95	0.90	3.48
	Plan 504: All Other Students	44012	65	41.69	12.08	0.92	3.39
5	All Students	44207	66	37.54	14.05	0.92	3.97
	Male	22989	66	37.72	14.11	0.92	3.95
	Female	21214	66	37.35	13.99	0.92	3.98
	Gender Not Reported	4	66				
	Hispanic or Latino	2961	66	29.70	13.62	0.91	3.99
	American Indian or Alaskan Native	266	66	32.89	15.16	0.93	4.01
	Asian	952	66	41.25	14.81	0.93	3.94
	Black or African American	1563	66	26.90	13.42	0.91	3.96
	Native Hawaiian or Pacific Islander	37	66	33.35	14.50	0.93	3.91
	White (Non-Hispanic)	37705	66	38.62	13.66	0.92	3.95
	Two or More Races (non-Hispanic)	510	66	32.66	14.32	0.92	4.00
	No Primary Race/Ethnicity Reported	213	66	35.65	15.58	0.93	4.01
	Currently Receiving LEP services	1399	66	23.44	13.14	0.91	3.88
	Former LEP Student - Monitoring Year 1	254	66	36.49	12.69	0.90	3.99
	Former LEP Student - Monitoring Year 2	121	66	40.96	13.12	0.91	3.98
	LEP: All Other Students	42433	66	38.01	13.85	0.92	3.96
	Students with an IEP	6677	66	25.23	12.53	0.90	3.94
	IEP: All Other Students	37530	66	39.73	13.15	0.91	3.94
	Economically Disadvantaged Students	16824	66	31.67	13.46	0.91	3.99
	SES: All Other Students	27383	66	41.15	13.16	0.91	3.91
	Migrant Students	16	66	25.44	10.07	0.83	4.12
	Migrant: All Other Students	44191	66	37.55	14.05	0.92	3.97
	Students Receiving Title 1 Services	8295	66	30.24	13.16	0.91	4.00
	Title 1: All Other Students	35912	66	39.23	13.71	0.92	3.94
	Plan 504	402	66	36.00	12.93	0.90	4.00
	Plan 504: All Other Students	43805	66	37.56	14.06	0.92	3.97
6	All Students	44477	66	35.16	14.52	0.93	3.94
	Male	23010	66	35.28	14.62	0.93	3.92
	Female	21465	66	35.04	14.42	0.93	3.95
	Gender Not Reported	2	66				
	Hispanic or Latino	2787	66	26.39	13.59	0.92	3.81
	American Indian or Alaskan Native	297	66	29.85	13.65	0.92	3.85
	Asian	943	66	39.41	15.35	0.93	3.93
	Black or African American	1554	66	25.38	13.89	0.93	3.78
	Native Hawaiian or Pacific Islander	33	66	35.85	16.04	0.94	3.85
	White (Non-Hispanic)	38181	66	36.19	14.20	0.92	3.94
	Two or More Races (non-Hispanic)	483	66	31.91	14.27	0.92	3.93
	No Primary Race/Ethnicity Reported	199	66	33.13	15.73	0.94	3.84
	Currently Receiving LEP services	1166	66	20.80	13.00	0.92	3.59
	Former LEP Student - Monitoring Year 1	206	66	31.62	13.90	0.92	3.91
	Former LEP Student - Monitoring Year 2	101	66	37.50	13.85	0.92	3.91
	LEP: All Other Students	43004	66	35.57	14.37	0.92	3.94
	Students with an IEP	6829	66	21.22	12.11	0.91	3.60
	IEP: All Other Students	37648	66	37.69	13.45	0.91	3.94
	Economically Disadvantaged Students	16564	66	28.81	13.55	0.92	3.85
	SES: All Other Students	27913	66	38.93	13.75	0.92	3.94
	Migrant Students	16	66	22.19	13.60	0.93	3.48
	Migrant: All Other Students	44461	66	35.17	14.52	0.93	3.94
	Students Receiving Title 1 Services	6643	66	26.92	13.08	0.92	3.80
	Title 1: All Other Students	37834	66	36.61	14.28	0.92	3.94

	Plan 504	499	66	33.36	13.84	0.92	3.95
	Plan 504: All Other Students	43978	66	35.19	14.53	0.93	3.94
7	All Students	46536	64	30.56	13.79	0.92	3.90
	Male	23776	64	30.65	14.02	0.92	3.88
	Female	22756	64	30.47	13.54	0.92	3.90
	Gender Not Reported	4	64				
	Hispanic or Latino	3190	64	21.93	12.19	0.91	3.59
	American Indian or Alaskan Native	350	64	26.73	13.81	0.92	3.81
	Asian	957	64	33.63	15.26	0.93	3.97
	Black or African American	1644	64	21.23	12.06	0.91	3.53
	Native Hawaiian or Pacific Islander	29	64	34.00	14.28	0.92	4.05
	White (Non-Hispanic)	39694	64	31.64	13.53	0.92	3.92
	Two or More Races (non-Hispanic)	457	64	26.45	13.27	0.92	3.77
	No Primary Race/Ethnicity Reported	215	64	30.63	14.03	0.92	3.92
	Currently Receiving LEP services	1139	64	16.43	10.38	0.90	3.27
	Former LEP Student - Monitoring Year 1	110	64	28.31	13.63	0.92	3.83
	Former LEP Student - Monitoring Year 2	102	64	32.04	13.95	0.92	3.96
	LEP: All Other Students	45185	64	30.92	13.67	0.92	3.90
	Students with an IEP	7184	64	17.66	10.05	0.89	3.33
	IEP: All Other Students	39352	64	32.91	13.05	0.91	3.94
	Economically Disadvantaged Students	17138	64	24.38	12.18	0.91	3.69
	SES: All Other Students	29398	64	34.16	13.39	0.91	3.95
	Migrant Students	15	64	20.07	9.65	0.87	3.50
	Migrant: All Other Students	46521	64	30.56	13.79	0.92	3.90
	Students Receiving Title 1 Services	5120	64	22.23	11.55	0.90	3.59
	Title 1: All Other Students	41416	64	31.59	13.69	0.92	3.92
	Plan 504	530	64	29.05	12.62	0.91	3.87
	Plan 504: All Other Students	46006	64	30.58	13.80	0.92	3.90
8	All Students	46567	65	31.29	13.88	0.93	3.72
	Male	24076	65	31.42	14.07	0.93	3.70
	Female	22486	65	31.17	13.67	0.93	3.74
	Gender Not Reported	5	65				
	Hispanic or Latino	2989	65	22.96	12.03	0.91	3.54
	American Indian or Alaskan Native	310	65	27.17	13.71	0.93	3.65
	Asian	961	65	35.53	14.98	0.94	3.79
	Black or African American	1637	65	22.14	12.14	0.92	3.52
	Native Hawaiian or Pacific Islander	37	65	29.03	11.83	0.91	3.63
	White (Non-Hispanic)	39959	65	32.28	13.68	0.93	3.73
	Two or More Races (non-Hispanic)	483	65	28.71	13.65	0.93	3.70
	No Primary Race/Ethnicity Reported	191	65	27.38	13.94	0.93	3.64
	Currently Receiving LEP services	1044	65	17.69	10.65	0.90	3.32
	Former LEP Student - Monitoring Year 1	89	65	28.57	13.39	0.92	3.70
	Former LEP Student - Monitoring Year 2	103	65	30.31	13.09	0.92	3.76
	LEP: All Other Students	45331	65	31.62	13.79	0.93	3.73
	Students with an IEP	7307	65	18.05	10.47	0.90	3.33
	IEP: All Other Students	39260	65	33.76	13.01	0.92	3.75
	Economically Disadvantaged Students	16294	65	25.11	12.34	0.91	3.60
	SES: All Other Students	30273	65	34.62	13.52	0.92	3.75
	Migrant Students	13	65	20.15	8.82	0.84	3.54
	Migrant: All Other Students	46554	65	31.30	13.88	0.93	3.72
	Students Receiving Title 1 Services	4417	65	22.61	11.58	0.91	3.53
	Title 1: All Other Students	42150	65	32.20	13.79	0.93	3.73
	Plan 504	570	65	30.02	12.59	0.91	3.71

	Plan 504: All Other Students	45997	65	31.31	13.89	0.93	3.72
11	All Students	32526	64	25.72	13.91	0.93	3.78
	Male	16473	64	26.17	14.42	0.93	3.80
	Female	16044	64	25.26	13.36	0.92	3.76
	Gender Not Reported	9	64				
	Hispanic or Latino	2364	64	17.96	11.03	0.90	3.44
	American Indian or Alaskan Native	135	64	21.67	13.02	0.92	3.63
	Asian	765	64	30.20	15.38	0.94	3.90
	Black or African American	1242	64	17.25	10.86	0.90	3.40
	Native Hawaiian or Pacific Islander	41	64	20.00	11.45	0.90	3.63
	White (Non-Hispanic)	27512	64	26.74	13.87	0.92	3.81
	Two or More Races (non-Hispanic)	271	64	22.69	12.40	0.91	3.66
	No Primary Race/Ethnicity Reported	196	64	19.80	12.20	0.92	3.51
	Currently Receiving LEP services	616	64	12.49	9.16	0.90	2.94
	Former LEP Student - Monitoring Year 1	74	64	18.64	8.79	0.84	3.49
	Former LEP Student - Monitoring Year 2	69	64	20.28	10.47	0.88	3.62
	LEP: All Other Students	31767	64	26.00	13.87	0.93	3.79
	Students with an IEP	4602	64	12.85	8.34	0.87	3.00
	IEP: All Other Students	27924	64	27.84	13.50	0.92	3.83
	Economically Disadvantaged Students	8348	64	19.19	11.48	0.91	3.51
	SES: All Other Students	24178	64	27.97	13.97	0.92	3.84
	Migrant Students	4	64				
	Migrant: All Other Students	32522	64	25.72	13.91	0.93	3.78
	Students Receiving Title 1 Services	2758	64	18.27	11.24	0.90	3.47
	Title 1: All Other Students	29768	64	26.41	13.94	0.93	3.80
	Plan 504	250	64	27.82	13.73	0.92	3.78
	Plan 504: All Other Students	32276	64	25.70	13.91	0.93	3.78

Table Q-2. 2010–11 NECAP: Subgroup Reliabilities – Reading

Grade	*Description*	*Number of Students*	*Raw Score*			*Alpha*	*Standard Error*
			Maximum	*Mean*	*Standard Deviation*		
3	All Students	43736	52	32.93	8.82	0.89	2.96
	Male	22457	52	31.86	9.02	0.89	2.95
	Female	21272	52	34.07	8.46	0.88	2.94
	Gender Not Reported	7	52				
	Hispanic or Latino	3180	52	28.51	9.34	0.89	3.10
	American Indian or Alaskan Native	282	52	30.90	9.15	0.89	3.10
	Asian	985	52	34.31	8.79	0.89	2.97
	Black or African American	1544	52	28.13	9.67	0.90	3.07
	Native Hawaiian or Pacific Islander	38	52	32.95	8.09	0.86	3.00
	White (Non-Hispanic)	36883	52	33.53	8.55	0.88	2.93
	Two or More Races (non-Hispanic)	555	52	31.06	9.30	0.89	3.05
	No Primary Race/Ethnicity Reported	269	52	32.10	8.90	0.89	2.95
	Currently Receiving LEP services	1699	52	25.38	9.60	0.89	3.17
	Former LEP Student - Monitoring Year 1	165	52	33.91	7.66	0.85	3.00
	Former LEP Student - Monitoring Year 2	25	52	36.20	8.00	0.87	2.92
	LEP: All Other Students	41847	52	33.24	8.65	0.88	2.94
	Students with an IEP	5756	52	23.64	9.82	0.90	3.14
	IEP: All Other Students	37980	52	34.34	7.74	0.86	2.91
	Economically Disadvantaged Students	17049	52	29.48	9.31	0.89	3.05
	SES: All Other Students	26687	52	35.14	7.72	0.86	2.86

	Migrant Students	17	52	26.65	10.64	0.91	3.14
	Migrant: All Other Students	43719	52	32.94	8.82	0.89	2.96
	Students Receiving Title 1 Services	10309	52	28.70	9.01	0.88	3.08
	Title 1: All Other Students	33427	52	34.24	8.34	0.88	2.90
	Plan 504	265	52	32.77	8.23	0.87	2.98
	Plan 504: All Other Students	43471	52	32.94	8.83	0.89	2.96
4	All Students	44206	52	35.18	8.64	0.88	3.04
	Male	22760	52	34.13	8.85	0.88	3.05
	Female	21443	52	36.30	8.26	0.87	3.01
	Gender Not Reported	3	52				
	Hispanic or Latino	3119	52	31.24	9.23	0.87	3.28
	American Indian or Alaskan Native	285	52	32.34	9.42	0.89	3.13
	Asian	1048	52	36.47	8.69	0.88	3.02
	Black or African American	1567	52	30.69	9.66	0.88	3.29
	Native Hawaiian or Pacific Islander	36	52	35.53	7.93	0.84	3.19
	White (Non-Hispanic)	37377	52	35.71	8.38	0.87	3.00
	Two or More Races (non-Hispanic)	532	52	33.95	8.91	0.88	3.13
	No Primary Race/Ethnicity Reported	242	52	34.31	8.73	0.88	3.08
	Currently Receiving LEP services	1512	52	27.04	9.35	0.87	3.41
	Former LEP Student - Monitoring Year 1	369	52	35.58	7.37	0.83	3.07
	Former LEP Student - Monitoring Year 2	60	52	38.03	7.22	0.83	2.94
	LEP: All Other Students	42265	52	35.47	8.48	0.87	3.02
	Students with an IEP	6154	52	25.52	9.46	0.87	3.35
	IEP: All Other Students	38052	52	36.75	7.39	0.84	2.96
	Economically Disadvantaged Students	17164	52	32.02	9.07	0.88	3.19
	SES: All Other Students	27042	52	37.19	7.71	0.86	2.92
	Migrant Students	16	52	28.31	9.55	0.89	3.22
	Migrant: All Other Students	44190	52	35.19	8.64	0.88	3.04
	Students Receiving Title 1 Services	9633	52	31.25	8.65	0.86	3.25
	Title 1: All Other Students	34573	52	36.28	8.31	0.87	2.97
	Plan 504	339	52	33.93	7.57	0.83	3.14
	Plan 504: All Other Students	43867	52	35.19	8.64	0.88	3.04
5	All Students	44031	52	31.23	8.56	0.89	2.90
	Male	22889	52	30.01	8.57	0.89	2.85
	Female	21138	52	32.55	8.35	0.88	2.92
	Gender Not Reported	4	52				
	Hispanic or Latino	2897	52	27.77	9.12	0.88	3.10
	American Indian or Alaskan Native	265	52	28.63	9.20	0.89	2.99
	Asian	915	52	33.17	8.54	0.88	2.91
	Black or African American	1514	52	26.61	9.53	0.90	3.07
	Native Hawaiian or Pacific Islander	36	52	30.08	8.72	0.90	2.81
	White (Non-Hispanic)	37686	52	31.68	8.31	0.88	2.87
	Two or More Races (non-Hispanic)	510	52	29.46	9.71	0.90	3.00
	No Primary Race/Ethnicity Reported	208	52	30.60	10.37	0.92	3.02
	Currently Receiving LEP services	1240	52	23.06	9.01	0.88	3.18
	Former LEP Student - Monitoring Year 1	254	52	31.06	7.36	0.84	2.99
	Former LEP Student - Monitoring Year 2	121	52	33.68	7.60	0.86	2.85
	LEP: All Other Students	42416	52	31.46	8.44	0.88	2.89
	Students with an IEP	6672	52	21.93	8.68	0.88	3.03
	IEP: All Other Students	37359	52	32.89	7.40	0.85	2.85
	Economically Disadvantaged Students	16716	52	28.05	8.84	0.89	2.99
	SES: All Other Students	27315	52	33.17	7.77	0.87	2.83
	Migrant Students	16	52	24.88	8.99	0.88	3.07

	Migrant: All Other Students	44015	52	31.23	8.56	0.89	2.90
	Students Receiving Title 1 Services	8514	52	27.76	8.60	0.87	3.05
	Title 1: All Other Students	35517	52	32.06	8.34	0.88	2.86
	Plan 504	403	52	30.28	7.46	0.86	2.83
	Plan 504: All Other Students	43628	52	31.24	8.57	0.89	2.90
6	All Students	44329	52	31.72	9.24	0.91	2.85
	Male	22935	52	30.43	9.19	0.91	2.81
	Female	21392	52	33.11	9.10	0.90	2.85
	Gender Not Reported	2	52				
	Hispanic or Latino	2720	52	26.88	9.54	0.90	3.02
	American Indian or Alaskan Native	294	52	28.42	9.36	0.90	2.93
	Asian	915	52	33.63	9.33	0.91	2.85
	Black or African American	1515	52	26.58	9.87	0.91	2.99
	Native Hawaiian or Pacific Islander	32	52	31.03	9.44	0.91	2.78
	White (Non-Hispanic)	38175	52	32.27	9.01	0.90	2.83
	Two or More Races (non-Hispanic)	482	52	30.28	9.89	0.92	2.85
	No Primary Race/Ethnicity Reported	196	52	30.81	9.92	0.92	2.87
	Currently Receiving LEP services	1015	52	21.86	9.49	0.89	3.08
	Former LEP Student - Monitoring Year 1	206	52	29.59	8.35	0.87	3.03
	Former LEP Student - Monitoring Year 2	101	52	34.31	7.59	0.86	2.87
	LEP: All Other Students	43007	52	31.96	9.12	0.90	2.84
	Students with an IEP	6816	52	21.44	8.93	0.89	2.98
	IEP: All Other Students	37513	52	33.59	7.99	0.88	2.79
	Economically Disadvantaged Students	16457	52	27.73	9.43	0.90	2.94
	SES: All Other Students	27872	52	34.08	8.28	0.89	2.76
	Migrant Students	16	52	21.69	9.65	0.91	2.97
	Migrant: All Other Students	44313	52	31.73	9.24	0.91	2.85
	Students Receiving Title 1 Services	6860	52	26.72	9.01	0.89	2.99
	Title 1: All Other Students	37469	52	32.64	8.99	0.90	2.82
	Plan 504	500	52	31.03	8.31	0.89	2.81
	Plan 504: All Other Students	43829	52	31.73	9.25	0.91	2.85
7	All Students	46409	52	30.47	9.16	0.89	3.10
	Male	23709	52	28.76	8.97	0.89	3.04
	Female	22696	52	32.26	9.00	0.88	3.09
	Gender Not Reported	4	52				
	Hispanic or Latino	3123	52	25.02	9.08	0.88	3.19
	American Indian or Alaskan Native	349	52	27.45	9.51	0.89	3.11
	Asian	934	52	31.91	9.71	0.89	3.17
	Black or African American	1604	52	25.30	9.36	0.88	3.18
	Native Hawaiian or Pacific Islander	29	52	32.62	8.13	0.84	3.29
	White (Non-Hispanic)	39706	52	31.12	8.92	0.88	3.08
	Two or More Races (non-Hispanic)	456	52	28.23	8.84	0.88	3.08
	No Primary Race/Ethnicity Reported	208	52	30.99	9.75	0.90	3.08
	Currently Receiving LEP services	1003	52	19.75	8.32	0.85	3.20
	Former LEP Student - Monitoring Year 1	107	52	29.99	7.94	0.85	3.08
	Former LEP Student - Monitoring Year 2	102	52	31.87	8.05	0.85	3.16
	LEP: All Other Students	45197	52	30.71	9.04	0.88	3.09
	Students with an IEP	7201	52	20.51	7.91	0.85	3.08
	IEP: All Other Students	39208	52	32.30	8.13	0.86	3.04
	Economically Disadvantaged Students	17064	52	26.40	8.88	0.88	3.14
	SES: All Other Students	29345	52	32.84	8.46	0.87	3.03
	Migrant Students	15	52	22.27	10.24	0.91	3.14
	Migrant: All Other Students	46394	52	30.48	9.16	0.89	3.10

	Students Receiving Title 1 Services	5117	52	25.30	8.68	0.87	3.17
	Title 1: All Other Students	41292	52	31.11	9.01	0.88	3.08
	Plan 504	531	52	29.18	7.84	0.85	3.03
	Plan 504: All Other Students	45878	52	30.49	9.17	0.89	3.10
8	All Students	46456	52	32.75	8.39	0.88	2.92
	Male	24026	52	31.50	8.44	0.88	2.87
	Female	22425	52	34.09	8.12	0.87	2.91
	Gender Not Reported	5	52				
	Hispanic or Latino	2941	52	28.29	8.82	0.88	3.06
	American Indian or Alaskan Native	309	52	30.77	8.51	0.87	3.01
	Asian	930	52	35.15	8.32	0.88	2.89
	Black or African American	1596	52	28.56	8.95	0.88	3.09
	Native Hawaiian or Pacific Islander	37	52	31.70	9.22	0.90	2.90
	White (Non-Hispanic)	39971	52	33.23	8.17	0.87	2.90
	Two or More Races (non-Hispanic)	481	52	31.89	8.34	0.87	2.96
	No Primary Race/Ethnicity Reported	191	52	30.65	9.11	0.89	2.97
	Currently Receiving LEP services	897	52	23.37	8.23	0.85	3.15
	Former LEP Student - Monitoring Year 1	89	52	32.09	8.48	0.87	3.00
	Former LEP Student - Monitoring Year 2	103	52	33.91	7.47	0.84	2.99
	LEP: All Other Students	45367	52	32.93	8.29	0.88	2.91
	Students with an IEP	7314	52	23.83	8.19	0.86	3.05
	IEP: All Other Students	39142	52	34.42	7.30	0.85	2.86
	Economically Disadvantaged Students	16236	52	29.24	8.40	0.87	3.01
	SES: All Other Students	30220	52	34.64	7.75	0.86	2.85
	Migrant Students	14	52	25.93	10.45	0.92	2.88
	Migrant: All Other Students	46442	52	32.75	8.39	0.88	2.92
	Students Receiving Title 1 Services	4499	52	28.07	8.31	0.86	3.06
	Title 1: All Other Students	41957	52	33.25	8.24	0.88	2.90
	Plan 504	569	52	32.21	7.29	0.84	2.87
	Plan 504: All Other Students	45887	52	32.76	8.40	0.88	2.92
11	All Students	32527	52	31.48	8.80	0.88	3.02
	Male	16479	52	29.86	8.92	0.89	3.00
	Female	16039	52	33.16	8.35	0.87	2.99
	Gender Not Reported	9	52				
	Hispanic or Latino	2327	52	27.67	8.74	0.87	3.17
	American Indian or Alaskan Native	135	52	28.83	9.19	0.89	3.05
	Asian	750	52	32.63	9.06	0.89	3.05
	Black or African American	1235	52	27.20	8.82	0.87	3.20
	Native Hawaiian or Pacific Islander	40	52	27.03	8.54	0.88	3.02
	White (Non-Hispanic)	27576	52	32.01	8.66	0.88	3.00
	Two or More Races (non-Hispanic)	271	52	31.25	8.18	0.86	3.11
	No Primary Race/Ethnicity Reported	193	52	27.80	9.23	0.89	3.09
	Currently Receiving LEP services	558	52	20.52	7.74	0.82	3.31
	Former LEP Student - Monitoring Year 1	74	52	27.59	6.26	0.75	3.14
	Former LEP Student - Monitoring Year 2	69	52	28.68	7.02	0.79	3.23
	LEP: All Other Students	31826	52	31.69	8.70	0.88	3.02
	Students with an IEP	4633	52	21.90	8.24	0.86	3.08
	IEP: All Other Students	27894	52	33.07	7.83	0.86	2.96
	Economically Disadvantaged Students	8346	52	27.59	8.81	0.87	3.13
	SES: All Other Students	24181	52	32.83	8.39	0.87	2.98
	Migrant Students	4	52				
	Migrant: All Other Students	32523	52	31.48	8.80	0.88	3.02
	Students Receiving Title 1 Services	2759	52	27.92	8.58	0.86	3.16

	Title 1: All Other Students	29768	52	31.81	8.75	0.88	3.01
	Plan 504	250	52	33.12	7.43	0.84	2.96
	Plan 504: All Other Students	32277	52	31.47	8.81	0.88	3.03

Table Q-3. 2010–11 NECAP: Subgroup Reliabilities – Writing

Grade	*Description*	*Number of Students*	*Raw Score*			*Alpha*	*Standard Error*
			Maximum	*Mean*	*Standard Deviation*		
5	All Students	43956	34	19.37	4.72	0.73	2.48
	Male	22835	34	18.23	4.67	0.73	2.44
	Female	21117	34	20.62	4.45	0.70	2.44
	Gender Not Reported	4	34				
	Hispanic or Latino	2896	34	18.33	4.89	0.75	2.47
	American Indian or Alaskan Native	265	34	18.23	5.01	0.72	2.67
	Asian	915	34	20.77	4.74	0.73	2.48
	Black or African American	1512	34	17.46	5.20	0.76	2.56
	Native Hawaiian or Pacific Islander	36	34	19.22	4.82	0.76	2.35
	White (Non-Hispanic)	37617	34	19.52	4.64	0.72	2.47
	Two or More Races (non-Hispanic)	508	34	18.57	5.19	0.76	2.55
	No Primary Race/Ethnicity Reported	207	34	18.90	5.61	0.77	2.66
	Currently Receiving LEP services	1238	34	15.80	5.13	0.74	2.60
	Former LEP Student - Monitoring Year 1	254	34	20.09	3.72	0.58	2.41
	Former LEP Student - Monitoring Year 2	120	34	21.20	3.93	0.65	2.33
	LEP: All Other Students	42344	34	19.47	4.68	0.72	2.47
	Students with an IEP	6651	34	14.49	4.63	0.72	2.45
	IEP: All Other Students	37305	34	20.24	4.18	0.67	2.39
	Economically Disadvantaged Students	16680	34	17.74	4.74	0.73	2.48
	SES: All Other Students	27276	34	20.37	4.42	0.70	2.43
	Migrant Students	16	34	13.81	4.83	0.69	2.70
	Migrant: All Other Students	43940	34	19.38	4.72	0.72	2.48
	Students Receiving Title 1 Services	8500	34	17.93	4.64	0.72	2.47
	Title 1: All Other Students	35456	34	19.72	4.68	0.72	2.47
	Plan 504	402	34	18.31	4.06	0.65	2.41
	Plan 504: All Other Students	43554	34	19.38	4.73	0.73	2.48
8	All Students	46274	34	21.28	5.32	0.78	2.52
	Male	23908	34	19.77	5.37	0.77	2.57
	Female	22361	34	22.90	4.76	0.76	2.35
	Gender Not Reported	5	34				
	Hispanic or Latino	2919	34	19.29	5.52	0.76	2.71
	American Indian or Alaskan Native	309	34	19.79	5.50	0.77	2.65
	Asian	929	34	23.31	5.06	0.78	2.35
	Black or African American	1589	34	19.16	5.46	0.76	2.67
	Native Hawaiian or Pacific Islander	37	34	20.62	4.98	0.72	2.65
	White (Non-Hispanic)	39825	34	21.49	5.24	0.77	2.50
	Two or More Races (non-Hispanic)	477	34	20.77	4.94	0.74	2.53
	No Primary Race/Ethnicity Reported	189	34	20.16	6.01	0.81	2.65
	Currently Receiving LEP services	893	34	16.55	5.68	0.74	2.88
	Former LEP Student - Monitoring Year 1	90	34	21.86	5.05	0.77	2.42
	Former LEP Student - Monitoring Year 2	103	34	22.46	4.16	0.70	2.29
	LEP: All Other Students	45188	34	21.37	5.27	0.77	2.51
	Students with an IEP	7249	34	15.57	5.23	0.71	2.80
	IEP: All Other Students	39025	34	22.34	4.61	0.74	2.37

	Economically Disadvantaged Students	16124	34	19.17	5.33	0.75	2.64
	SES: All Other Students	30150	34	22.41	4.96	0.76	2.41
	Migrant Students	14	34	16.57	5.24	0.75	2.64
	Migrant: All Other Students	46260	34	21.28	5.32	0.78	2.52
	Students Receiving Title 1 Services	4474	34	19.00	5.28	0.74	2.69
	Title 1: All Other Students	41800	34	21.53	5.26	0.77	2.50
	Plan 504	566	34	20.14	4.91	0.74	2.51
	Plan 504: All Other Students	45708	34	21.30	5.32	0.78	2.52
11	All Students	32409	12	6.30	1.79		
	Male	16391	12	5.98	1.89		
	Female	16009	12	6.62	1.61		
	Gender Not Reported	9	12				
	Hispanic or Latino	2306	12	5.76	1.99		
	American Indian or Alaskan Native	136	12	5.99	1.94		
	Asian	748	12	6.60	1.87		
	Black or African American	1224	12	5.55	2.04		
	Native Hawaiian or Pacific Islander	40	12	5.28	1.87		
	White (Non-Hispanic)	27496	12	6.37	1.73		
	Two or More Races (non-Hispanic)	271	12	6.18	1.74		
	No Primary Race/Ethnicity Reported	188	12	5.52	2.07		
	Currently Receiving LEP services	546	12	4.10	2.33		
	Former LEP Student - Monitoring Year 1	74	12	6.08	1.59		
	Former LEP Student - Monitoring Year 2	70	12	5.97	1.63		
	LEP: All Other Students	31719	12	6.33	1.75		
	Students with an IEP	4581	12	4.60	1.95		
	IEP: All Other Students	27828	12	6.57	1.60		
	Economically Disadvantaged Students	8300	12	5.67	1.91		
	SES: All Other Students	24109	12	6.51	1.69		
	Migrant Students	4	12				
	Migrant: All Other Students	32405	12	6.30	1.79		
	Students Receiving Title 1 Services	2729	12	5.78	1.99		
	Title 1: All Other Students	29680	12	6.34	1.76		
	Plan 504	250	12	6.70	1.40		
	Plan 504: All Other Students	32159	12	6.29	1.79		

Table Q-4. 2010–11 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	6.14	2.59	0.75	1.29
	FA	8	10	6.65	2.27	0.63	1.39
	GM	8	10	5.41	2.47	0.63	1.50
	NO	30	35	23.94	7.15	0.89	2.42
4	DP	8	10	6.26	2.18	0.59	1.40
	FA	9	10	5.62	2.32	0.66	1.35
	GM	11	13	9.23	2.34	0.61	1.46
	NO	27	32	20.56	6.71	0.88	2.34
5	DP	6	10	5.00	2.59	0.59	1.66
	FA	9	13	7.49	3.40	0.69	1.90
	GM	9	13	7.91	2.91	0.61	1.81
	NO	24	30	17.14	6.65	0.86	2.49

Grade	Reporting Category	Number of Items	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
6	DP	6	10	5.64	2.68	0.63	1.62
	FA	9	13	7.94	3.27	0.66	1.89
	GM	12	17	7.64	3.70	0.74	1.88
	NO	21	26	13.94	6.33	0.85	2.41
7	DP	6	10	4.50	2.31	0.63	1.41
	FA	15	20	9.69	4.80	0.82	2.05
	GM	11	15	7.76	3.28	0.63	1.99
	NO	14	19	8.61	4.80	0.78	2.26
8	DP	6	10	3.67	2.37	0.60	1.50
	FA	21	26	14.44	5.85	0.85	2.24
	GM	12	17	7.15	3.86	0.76	1.90
	NO	8	12	6.03	3.22	0.71	1.72
11	DP	6	10	3.67	2.54	0.57	1.67
	FA	20	26	12.37	6.24	0.86	2.33
	GM	15	19	6.54	3.80	0.76	1.87
	NO	5	9	3.14	2.69	0.65	1.60

Table Q-5. 2010–11 NECAP: Reliabilities by Reporting Category – Reading

Grade	Reporting Category	Number of Items	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
3	LA	8	14	7.06	2.43	0.60	1.54
	LI	12	18	10.88	3.79	0.82	1.63
	TI	10	16	8.41	3.04	0.71	1.64
	TL	10	16	9.53	3.14	0.76	1.53
	WV	14	20	14.99	3.64	0.72	1.91
4	LA	8	14	7.98	2.54	0.65	1.49
	LI	14	20	13.04	3.99	0.75	2.00
	TI	10	16	8.95	3.09	0.70	1.70
	TL	12	18	12.07	3.47	0.72	1.82
	WV	12	18	14.17	3.23	0.72	1.71
5	LA	12	21	12.32	3.52	0.74	1.79
	LI	13	22	12.60	3.85	0.75	1.91
	TI	12	21	11.84	3.86	0.76	1.87
	TL	13	22	13.07	3.59	0.74	1.82
	WV	9	9	6.31	2.12	0.67	1.22
6	LA	10	22	11.51	3.95	0.77	1.88
	LI	14	20	12.20	4.06	0.81	1.78
	TI	12	21	11.30	4.26	0.80	1.91
	TL	12	21	12.42	3.82	0.79	1.74
	WV	10	10	8.01	2.07	0.71	1.11
7	LA	11	23	11.78	4.37	0.78	2.07
	LI	13	19	11.50	3.73	0.75	1.87
	TI	12	21	12.21	4.16	0.79	1.91
	TL	12	21	11.07	3.97	0.74	2.04
	WV	10	10	7.19	2.04	0.62	1.25
8	LA	12	24	14.67	4.06	0.76	1.97
	LI	12	18	10.85	3.12	0.70	1.71
	TI	12	21	13.27	3.39	0.70	1.85

Grade	Reporting Category	Number of Items	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
	TL	12	21	12.24	3.86	0.77	1.83
	WV	10	10	7.23	2.26	0.69	1.26
11	LA	11	26	14.06	4.82	0.81	2.10
	LI	13	16	10.04	3.11	0.71	1.67
	TI	12	21	11.43	4.10	0.76	2.01
	TL	12	21	12.67	3.78	0.76	1.85
	WV	10	10	7.39	1.95	0.61	1.22

Table Q-6. 2010–11 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	5.49	1.88		
	MC	10	10	8.15	1.93	0.67	1.11
	SR	3	12	5.74	2.06	0.70	1.12
8	ER	1	12	6.57	1.81		
	MC	10	10	7.15	2.25	0.69	1.25
	SR	3	12	7.57	2.34	0.75	1.16
11	RE	1	12	6.30	1.79		

Appendix R—INTERRATER AGREEMENT

Table R-1. 2010–11 NECAP: Item-Level Interrater Agreement Statistics by Grade – Mathematics

Grade	*Item*	*Number of Score Categories*	*Number of Responses Scored Twice*	*Percent Exact*	*Percent Adjacent*	*Correlation*	*Percent of Third Scores*
3	119688	3	857	95.22	4.67	0.96	0.12
	119731	2	865	98.73	1.27	0.96	0.00
	119838	3	874	95.88	4.12	0.97	0.00
	119889	3	875	96.34	3.31	0.97	0.34
	119918	3	867	96.89	3.00	0.98	0.12
	122872	2	862	99.42	0.58	0.93	0.00
	124459	2	882	98.53	1.47	0.96	0.00
	139583	2	864	98.73	1.27	0.97	0.00
	139590	2	854	98.48	1.52	0.97	0.00
	139603	3	871	94.14	5.51	0.93	0.34
	139645	2	864	99.31	0.69	0.98	0.00
	139653	3	865	93.06	6.82	0.95	0.12
	139655	3	873	96.56	3.44	0.97	0.00
	139669	2	877	97.49	2.51	0.94	0.00
	139673	3	883	97.06	2.83	0.96	0.11
	139689	2	864	92.71	7.29	0.85	0.00
	144612	3	855	97.31	1.99	0.97	0.70
	144617	2	874	99.20	0.80	0.97	0.00
	144623	3	860	92.67	7.09	0.94	0.23
	145249	2	871	98.62	1.38	0.97	0.00
4	120075	3	899	89.32	10.57	0.89	0.11
	120119	2	884	98.87	1.13	0.94	0.00
	120127	3	909	90.32	9.57	0.84	0.11
	120221	2	885	99.66	0.34	0.99	0.00
	120222	3	886	97.07	2.93	0.94	0.00
	120240	2	874	99.20	0.80	0.98	0.00
	120262	2	885	97.74	2.26	0.95	0.00
	120286	3	907	95.48	4.41	0.95	0.11
	124661	3	909	85.48	14.19	0.89	0.33
	124723	3	911	93.08	6.81	0.94	0.11
	139447	2	882	98.75	1.25	0.97	0.00
	139452	3	887	93.35	6.20	0.94	0.45
	139479	2	878	99.09	0.91	0.98	0.00
	139517	2	871	97.93	2.07	0.96	0.00
	139523	2	880	98.30	1.70	0.96	0.00
	139527	3	898	98.11	1.89	0.98	0.00
	139529	3	894	84.23	14.21	0.81	1.57
	145085	2	888	99.10	0.90	0.97	0.00
	145545	2	880	98.07	1.93	0.93	0.00
	145559	3	897	97.88	2.01	0.98	0.11
5	120725	3	886	89.95	9.82	0.90	0.23
	120726	3	862	88.63	9.98	0.91	1.39
	124804	2	890	96.07	3.93	0.91	0.00
	124858	5	903	90.81	8.08	0.97	1.11
	139129	2	894	96.42	3.58	0.93	0.00
	139136	2	899	97.22	2.78	0.94	0.00
	139141	3	903	91.25	8.64	0.92	0.11
	139203	3	882	85.49	14.17	0.83	0.34

Grade	Item	Number of Score Categories	Number of Responses Scored Twice	Percent Exact	Percent Adjacent	Correlation	Percent of Third Scores
	139283	5	907	88.09	10.25	0.96	1.65
	139384	5	923	78.11	19.83	0.93	2.17
	145571	3	900	84.11	14.56	0.88	1.33
	145576	5	957	79.94	16.93	0.92	3.24
	145879	2	885	95.14	4.86	0.88	0.00
	145887	2	887	97.63	2.37	0.95	0.00
	145888	3	897	86.06	12.82	0.89	1.11
	145922	2	889	98.09	1.91	0.96	0.00
6	119195	3	881	82.18	16.91	0.80	0.91
	119230	3	885	91.86	7.34	0.93	0.79
	119306	2	899	97.00	3.00	0.92	0.00
	123489	5	1026	78.75	19.40	0.93	1.95
	139310	2	888	97.07	2.93	0.94	0.00
	139312	2	893	98.43	1.57	0.97	0.00
	139320	5	913	87.40	11.83	0.94	0.77
	139356	5	999	87.79	11.31	0.97	0.90
	139377	3	887	91.09	8.34	0.93	0.56
	139387	5	914	88.62	10.18	0.96	1.20
	144707	3	887	93.46	5.64	0.92	0.90
	145133	2	879	98.29	1.71	0.97	0.00
	145347	2	895	98.55	1.45	0.97	0.00
	145995	2	888	97.30	2.70	0.95	0.00
	146011	3	892	93.27	6.28	0.94	0.45
	197550	3	883	94.11	5.44	0.94	0.45
7	120387	2	939	98.94	1.06	0.98	0.00
	120406	2	919	98.37	1.63	0.97	0.00
	120427	3	897	90.97	8.70	0.93	0.33
	120468	2	929	98.92	1.08	0.98	0.00
	120470	3	912	95.61	4.39	0.96	0.00
	125382	3	952	85.50	14.29	0.90	0.21
	125397	2	948	98.42	1.58	0.96	0.00
	139980	3	923	91.77	7.91	0.91	0.33
	139987	5	1087	80.22	15.55	0.93	4.32
	140223	5	926	86.72	11.77	0.95	1.51
	140241	3	936	94.98	5.02	0.94	0.00
	140251	5	954	74.21	20.86	0.88	4.72
	140280	2	936	98.50	1.50	0.97	0.00
	140292	5	955	84.92	14.14	0.93	0.94
	144756	3	918	95.42	4.58	0.93	0.00
	146230	2	921	97.72	2.28	0.95	0.00
8	120889	2	925	95.03	4.97	0.88	0.00
	120933	2	932	96.67	3.33	0.92	0.00
	120942	5	1001	76.42	21.88	0.90	1.70
	120958	3	910	95.16	4.29	0.95	0.55
	121039	2	905	97.68	2.32	0.94	0.00
	125563	3	929	89.02	10.87	0.91	0.11
	139795	5	977	78.20	20.16	0.92	1.64
	139813	3	912	86.40	13.16	0.87	0.44
	139827	5	933	90.57	8.25	0.95	1.18
	139863	2	922	98.48	1.52	0.96	0.00

Grade	Item	Number of Score Categories	Number of Responses Scored Twice	Percent Exact	Percent Adjacent	Correlation	Percent of Third Scores
	144778	5	1035	76.81	22.22	0.90	0.87
	145660	3	915	88.74	11.04	0.91	0.22
	145661	3	1115	75.52	23.41	0.74	1.17
	146422	2	911	99.01	0.99	0.98	0.00
	146483	3	1736	90.09	9.74	0.90	0.52
11	119457	2	628	99.84	0.16	1.00	0.00
	119481	3	596	95.64	4.19	0.93	0.17
	119492	2	644	99.69	0.31	0.99	0.00
	119543	3	576	96.88	2.78	0.98	0.35
	119606	2	606	98.18	1.82	0.96	0.00
	119616	2	625	76.64	23.36	0.53	0.00
	125772	2	584	99.32	0.68	0.98	0.00
	130121	2	590	98.64	1.36	0.96	0.00
	140031	2	619	94.51	5.49	0.89	0.00
	140035	3	626	93.61	3.99	0.92	2.40
	140135	2	586	97.61	2.39	0.92	0.00
	140203	2	617	99.03	0.97	0.98	0.00
	141291	5	583	90.05	8.58	0.96	1.37
	141307	5	609	83.74	14.45	0.91	1.81
	141322	5	611	88.22	9.33	0.94	2.45
	141365	5	613	87.60	11.75	0.97	0.65
	145227	3	604	91.23	8.61	0.87	0.17
	145465	2	621	99.03	0.97	0.98	0.00
	145484	2	544	98.90	1.10	0.95	0.00
	145492	2	653	99.23	0.77	0.98	0.00
	145498	3	582	97.08	2.75	0.98	0.17
	146944	3	532	90.41	9.21	0.92	0.38

Table R-2. 2010–11 NECAP: Item-Level Interrater Agreement Statistics by Grade – Reading

Grade	Item	Number of Score Categories	Number of Responses Scored Twice	Percent Exact	Percent Adjacent	Correlation	Percent of Third Scores
3	117650	5	883	91.39	6.57	0.95	1.93
	117669	5	872	69.15	30.05	0.81	0.80
	117766	5	892	64.01	33.52	0.70	2.35
	147742	5	885	67.80	30.28	0.79	2.03
	147744	5	884	98.19	1.70	0.99	0.11
	148057	5	881	65.49	30.42	0.70	3.97
4	117946	5	922	65.84	32.00	0.81	2.17
	117970	5	904	61.73	36.06	0.73	2.10
	118005	5	910	98.24	1.65	0.99	0.11
	147914	5	931	55.85	42.00	0.71	2.15
	147915	5	911	97.91	2.09	0.99	0.00
	147988	5	919	70.84	26.22	0.85	2.94
5	118048	5	890	61.80	35.62	0.78	2.58
	118049	5	899	62.51	35.60	0.71	1.78
	118076	5	928	61.53	36.10	0.73	2.26
	118080	5	925	64.54	32.86	0.75	2.16
	148369	5	922	67.03	31.45	0.76	1.52

Grade	Item	Number of Score Categories	Number of Responses Scored Twice	Percent Exact	Percent Adjacent	Correlation	Percent of Third Scores
	148767	5	919	68.44	30.36	0.77	1.20
6	118274	5	902	65.74	31.71	0.74	2.66
	118275	5	888	69.48	29.17	0.81	1.24
	118353	5	923	61.76	35.75	0.77	2.28
	126904	5	900	65.00	33.00	0.77	2.00
	126905	5	901	59.71	37.51	0.78	2.77
	148891	5	923	63.27	34.78	0.69	1.73
7	118542	5	951	60.15	37.54	0.78	2.21
	118557	5	939	61.45	36.10	0.70	2.13
	118558	5	946	63.74	35.10	0.82	0.95
	129228	5	923	61.32	36.73	0.79	1.73
	129229	5	939	60.49	36.74	0.77	2.45
	147224	5	933	63.34	34.62	0.80	2.04
8	118607	5	941	61.96	36.24	0.76	1.81
	118608	5	926	67.82	30.35	0.77	1.62
	118670	5	932	63.73	34.55	0.75	1.72
	118671	5	949	64.38	34.04	0.73	1.48
	147242	5	947	64.41	34.64	0.74	0.95
	147481	5	936	60.15	35.58	0.75	3.85
11	118766	5	661	64.60	33.13	0.74	2.12
	118767	5	650	72.62	25.85	0.77	1.54
	147564	5	658	61.25	35.87	0.72	2.58
	147726	5	655	63.82	35.11	0.78	1.07
	147732	5	624	62.82	34.78	0.75	2.40
	147885	5	674	67.51	29.97	0.76	2.37

Table R-3. 2010–11 NECAP: Item-Level Interrater Agreement Statistics by Grade – Writing

Grade	Item	Number of Score Categories	Number of Responses Scored Twice	Percent Exact	Percent Adjacent	Correlation	Percent of Third Scores
5	150134	5	942	67.62	30.79	0.73	1.59
	150165	7	43423	56.70	38.97	0.70	4.10
	150198	5	941	62.70	35.07	0.74	2.23
	150237	5	947	69.38	29.57	0.77	1.06
8	150433	5	995	68.24	30.65	0.79	1.11
	150789	5	951	62.25	35.33	0.72	2.21
	150803	5	924	68.18	30.41	0.77	1.41
	150854	7	45173	64.84	33.67	0.77	1.36
11	121154	7	31548	64.12	34.32	0.76	1.50

Appendix S—Decision Accuracy and Consistency Results

Table S-1. 2010–11 NECAP: Summary of Decision Accuracy (and Consistency) Results by Subject and Grade—Overall and Conditional on Performance Level

Subject	*Grade*	*Overall*	*Kappa*	*Conditional on Level*			
				Novice	*Nearing Proficiency*	*Proficient*	*Advanced*
Mathematics	3	0.82 (0.75)	0.64	0.83 (0.76)	0.73 (0.64)	0.82 (0.77)	0.89 (0.80)
	4	0.80 (0.73)	0.61	0.83 (0.75)	0.69 (0.59)	0.81 (0.76)	0.88 (0.78)
	5	0.80 (0.73)	0.61	0.83 (0.76)	0.62 (0.51)	0.83 (0.78)	0.88 (0.77)
	6	0.81 (0.74)	0.63	0.83 (0.77)	0.64 (0.53)	0.83 (0.77)	0.90 (0.81)
	7	0.79 (0.72)	0.60	0.81 (0.75)	0.60 (0.49)	0.83 (0.77)	0.89 (0.79)
	8	0.81 (0.73)	0.63	0.83 (0.77)	0.66 (0.55)	0.84 (0.78)	0.89 (0.80)
	11	0.84 (0.77)	0.67	0.88 (0.85)	0.73 (0.63)	0.88 (0.82)	0.82 (0.62)
Reading	3	0.80 (0.72)	0.57	0.77 (0.65)	0.70 (0.61)	0.82 (0.78)	0.86 (0.73)
	4	0.77 (0.69)	0.54	0.77 (0.65)	0.69 (0.59)	0.78 (0.72)	0.87 (0.74)
	5	0.80 (0.72)	0.57	0.77 (0.63)	0.72 (0.62)	0.81 (0.76)	0.87 (0.75)
	6	0.82 (0.75)	0.61	0.78 (0.66)	0.75 (0.67)	0.84 (0.80)	0.87 (0.76)
	7	0.80 (0.72)	0.56	0.78 (0.66)	0.70 (0.61)	0.83 (0.79)	0.85 (0.71)
	8	0.80 (0.72)	0.56	0.75 (0.58)	0.73 (0.64)	0.80 (0.75)	0.87 (0.75)
	11	0.77 (0.69)	0.54	0.77 (0.64)	0.66 (0.56)	0.77 (0.70)	0.87 (0.77)
Writing	5	0.67 (0.56)	0.35	0.67 (0.45)	0.69 (0.62)	0.61 (0.52)	0.79 (0.55)
	8	0.71 (0.60)	0.41	0.70 (0.53)	0.67 (0.59)	0.72 (0.64)	0.81 (0.57)

Table S-2. 2010–11 NECAP: Summary of Decision Accuracy (and Consistency) Results by Subject and Grade—Conditional on Cutpoint

Subject	*Grade*	*Novice/Nearing Proficiency*			*Nearing Proficiency/Proficient*			*Proficient/Advanced*		
		Accuracy (consistency)	*False positive*	*False negative*	*Accuracy (consistency)*	*False positive*	*False negative*	*Accuracy (consistency)*	*False positive*	*False negative*
Mathematics	3	0.96 (0.94)	0.02	0.02	0.93 (0.90)	0.04	0.03	0.93 (0.91)	0.05	0.02
	4	0.95 (0.93)	0.02	0.02	0.92 (0.89)	0.04	0.04	0.93 (0.90)	0.05	0.02
	5	0.94 (0.92)	0.03	0.03	0.92 (0.89)	0.04	0.03	0.93 (0.91)	0.05	0.02
	6	0.95 (0.93)	0.03	0.03	0.93 (0.90)	0.04	0.03	0.93 (0.91)	0.04	0.02
	7	0.93 (0.90)	0.04	0.03	0.92 (0.88)	0.05	0.04	0.94 (0.92)	0.04	0.02
	8	0.94 (0.92)	0.03	0.03	0.92 (0.89)	0.04	0.03	0.94 (0.92)	0.04	0.02
	11	0.92 (0.89)	0.05	0.03	0.93 (0.91)	0.04	0.03	0.98 (0.98)	0.01	0
Reading	3	0.96 (0.95)	0.01	0.02	0.91 (0.88)	0.05	0.04	0.92 (0.89)	0.05	0.02
	4	0.96 (0.94)	0.02	0.02	0.90 (0.87)	0.05	0.04	0.91 (0.87)	0.06	0.03
	5	0.97 (0.96)	0.01	0.02	0.91 (0.88)	0.05	0.04	0.92 (0.88)	0.06	0.03

	6	0.97 (0.96)	0.01	0.02	0.92 (0.89)	0.04	0.04	0.93 (0.90)	0.05	0.02
	7	0.96 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04	0.93 (0.91)	0.05	0.02
	8	0.98 (0.97)	0.01	0.01	0.91 (0.88)	0.04	0.04	0.91 (0.87)	0.06	0.03
	11	0.96 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04	0.90 (0.87)	0.06	0.03
Writing	5	0.95 (0.92)	0.02	0.04	0.82 (0.76)	0.1	0.07	0.89 (0.85)	0.08	0.02
	8	0.95 (0.92)	0.02	0.03	0.85 (0.80)	0.09	0.06	0.91 (0.87)	0.07	0.02

Appendix T—SAMPLE REPORTS

School Name	Code	Grade
Demonstration School 1	DEM-DEA-DEMO1	5
District Name	**Report Type**	
Demonstration District A	Student Report - School Copy	

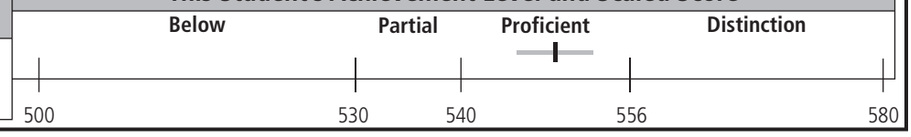
01R063Z

Student	Grade	School	District	State
Kristoff Allick	5	Demonstration School 1	Demonstration District A	NH

Fall 2010 - Beginning of Grade 5 NECAP Test Results

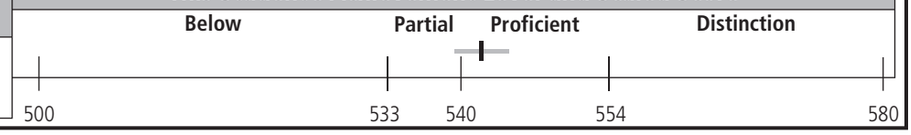
Content Area	Achievement Level	Scaled Score
Reading	Proficient	549

This Student's Achievement Level and Scaled Score



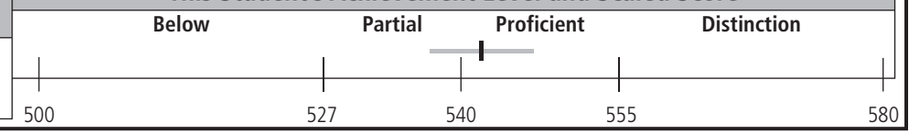
Content Area	Achievement Level	Scaled Score
Mathematics	Proficient	542

This Student's Achievement Level and Scaled Score



Content Area	Achievement Level	Scaled Score
Writing	Proficient	542

This Student's Achievement Level and Scaled Score



Interpretation of Graphic Display

The line (▌) represents the student's score. The bar (▬) surrounding the score represents the probable range of scores for the student if he or she were to be tested many times. This statistic is called the standard error of measurement. See the reverse side for the achievement level descriptions.

This Student's Achievement Level Compared to Other Beginning of Grade 5 Students by School, District, and State

	Reading				Mathematics				Writing			
	Student	School	District	State	Student	School	District	State	Student	School	District	State
Proficient with Distinction		14%	12%	21%		13%	14%	23%		9%	10%	14%
Proficient	✓	61%	59%	57%	✓	56%	52%	50%	✓	44%	36%	42%
Partially Proficient		14%	14%	16%		16%	14%	15%		42%	43%	38%
Substantially Below Proficient		11%	15%	6%		16%	21%	12%		4%	10%	6%

This Student's Performance in Content Area Subcategories

Reading		Possible Points	Student	Average Points Earned			
				School	District	State	Students at Proficient Level
Word ID/ Vocabulary		9	**7**	6.0	6.0	6.6	4.2-6.7
Type of Text*	Literary	22	**16**	12.7	12.4	13.3	9.9-13.5
	Informational	21	**11**	11.5	11.2	12.2	8.0-11.7
Level of Comprehension*	Initial Understanding	22	**15**	12.3	11.9	12.9	8.8-12.6
	Analysis and Interpretation	21	**12**	11.9	11.6	12.6	9.1-12.7

Mathematics	Possible Points	Student	Average Points Earned			
			School	District	State	Students at Proficient Level
Numbers and Operations	30	**19**	16.4	16.6	18.3	12.0-17.0
Geometry and Measurement	13	**6**	8.1	8.2	8.5	5.5-9.2
Functions and Algebra	13	**6**	7.5	7.2	8.0	4.7-8.5
Data, Statistics, and Probability	10	**3**	4.9	4.8	5.4	2.4-5.8

Writing	Possible Points	Student	Average Points Earned			
			School	District	State	Students at Proficient Level
Multiple Choice	10	**8**	8.2	7.9	8.3	7.6-9.8
Short Responses	12	**5**	5.7	5.5	5.9	4.6-6.9
Extended Response	12	**7**	5.6	5.3	5.6	5.5-5.6

Comments about this student's writing performance on the extended response:

Writing has a general purpose with attempted focus.

Writing has some organization.

Writing includes supporting details with sufficient elaboration.

Writing has appropriate word choice and some control of sentence structure.

*With the exception of Word ID/Vocabulary items, reading items are reported in two ways - Type of Text and Level of Comprehension.

CONFIDENTIAL

Fall 2010 - Beginning of Grade 04 NECAP Tests

Grade 04 Students in 2010-2011

Item Analysis Report — Reading

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

		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	WV	IA	II	II	WV	II	WV	II	WV	IA	Word ID/Vocabulary	Literary	Informational	Initial Understanding	Analysis & Interpretation	Total Points Earned	Scaled Score	Achievement Level
	GE Code	3-2	3-3	3-2	3-8	3-7	3-7	3-1	3-7	3-3	3-7	3-3	3-8								
	Depth of Knowledge Code	1	1	2	2	1	1	1	1	1	1	2	2								
	Item Type	MC	MC	MC	MC	MC	MC	CR	MC	MC	MC	MC	CR								
	Correct MC Response	C	C	A	B	B	C		D	C	B	D									
Name/Student ID	Total Possible Points	1	1	1	1	1	1	4	1	1	1	1	4	18	18	16	20	14	52		
Austin, Samantha K	D044005	+	+	+	+	+	+	4	+	+	+	+	2	18	18	14	20	12	50	480	4
Baez, Matthew	D044038	A	D	B	D	A	+	4	C	A	A	B	1	8	7	4	7	4	19	426	1
Caldwell, Tedrique M	D044047													0	0	0	0	0	0		E
Cardoza, Devyn K	D044037	+	+	+	+	+	+	4	+	+	+	+	2	18	13	12	15	10	43	458	4
Loptien, Elizabeth G	D044007	+	+	+	+	D	+	4	+	+	+	+	1	16	12	10	15	7	38	448	3
Luna, Angel	D044018	+	+	+	+	+	+	1	+	+	+	+	2	13	13	11	14	10	37	447	3
Mankey, Sara	D044048	+	+	B	+	C	D	3	A	+	A	B	0	9	6	2	5	3	17	423	1
Martel, Jacob V	D044045	+	+	+	+	+	+	4	+	+	+	+	3	17	16	12	17	11	45	463	4
Mason, Colton	D044050													0	0	0	0	0	0		N
Mastrion, Giana	D044025	+	+	+	+	+	+	4	+	+	+	+	3	17	15	12	16	11	44	460	4
Mcanistan, Joshua	D044023													0	0	0	0	0	0		S
Paris, Jacqulyn	D044031	+	+	+	+	A	+	3	A	+	+	+	3	15	9	10	11	8	34	443	3
Patrick, Andrew M	D044014	+	+	+	+	+	+	4	+	+	+	+	1	16	15	11	17	9	42	455	3
Pelosi, Jenna R	D044029	+	+	+	+	A	+	4	+	+	+	C	1	15	15	11	18	8	41	453	3
Perez, Anthony	D044009													0	0	0	0	0	0		W
Rigdon, Celsi	D044051													0	0	0	0	0	0		A
Rivera, Ruben	D044004													0	0	0	0	0	0		L
Sawyer, Brianna M	D044002	+	+	+	+	D	B	2	+	+	D	+	1	11	5	4	6	3	20	428	1
Schafer, Scott	D044012	+	+	+	+	+	+	4	+	+	+	+	4	18	15	15	18	12	48	476	4
Smithson, Alexys C	D044006	+	+	+	+	+	+	4	+	+	+	+	2	18	13	12	17	8	43	458	4
Veilleux, Vanessa V	D044022	+	+	+	+	A	+	4	+	+	+	+	4	16	16	13	16	13	45	463	4
Villalpando, Jessejames R	D044021	+	B	+	A	+	D	2	A	D	D	C	0	6	4	1	3	2	11	410	1

Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12									
Percent Correct/Average Score: Group	94	88	88	88	56	81	3.4	75	88	75	75	1.9	14.4	12.0	9.6	13.4	8.2				
Percent Correct/Average Score: School	94	88	88	88	56	81	3.4	75	88	75	75	1.9	14.4	12.0	9.6	13.4	8.2				
Percent Correct/Average Score: District	88	79	88	88	60	79	3.2	67	86	67	79	1.8	13.7	12.3	9.3	13.5	8.0				
Percent Correct/Average Score: State	88	79	89	83	64	71	3.3	63	76	77	73	1.9	13.9	12.0	8.9	13.0	7.9				

CONFIDENTIAL

Fall 2010 - Beginning of Grade 03 NECAP Tests

Grade 03 Students in 2010-2011

Item Analysis Report — Mathematics

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

		Released Items																Total Test Results							
	Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Subcategory Points Earned					Total Points Earned	Scaled Score	Achievement Level
	Content Strand	NO	NO	NO	NO	NO	NO	FA	FA	FA	DP	NO	NO	GM	NO	GM	DP	Numbers & Operations	Geometry & Measurement	Functions & Algebra	Data, Statistics, & Probability				
	GE Code	2-1	2-2	2-2	2-2	2-3	2-5	2-1	2-1	2-4	2-2	2-2	2-5	2-1	2-3	2-7	2-4								
	Depth of Knowledge Code	2	2	2	2	1	1	2	2	2	2	2	2	2	2	2	2								
	Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	SA	SA	SA	SA	SA	SA								
	Correct MC Response	C	D	A	B	A	B	B	B	C	D														
Name/Student ID	Total Possible Points	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	35	10	10	10		65		
Bicalho, Melissa	D034015	+	+	+	+	+	+	+	+	A	+	1	1	1	1	1	1	28	8	7	9		52	351	3
Blair, Joseph J	D034053	+	+	+		+	+		+	B	B	1	0	1	1	1	1	27	4	6	6		43	344	3
Blaskovich, Kiley N	D034004																	0	0	0	0		0		A
Brooks, Dylan	D034045	+	+	+	+	+	+	+	+	+	+	1	1	1	0	2	1	33	8	9	9		59	360	4
Brown, Shannon F	D034035	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0		0	300	1
Buus, Charles	D034006	+	C	+	C	C	+	+	+	+	+	1	0	1	2	0	0	19	3	7	3		32	336	2
Caldwell, Diamond	D034046	B		B	+	B	+	+	A	B	B	0	0	0	0	0	0	9	0	3	0		12	318	1
Colvin, Timothy J	D034008	+	+	+	C	+	+	+	+	A	C	1	0	0	1	0	0	22	3	6	3		34	338	2
Connors, Mckayla	D034040	+	C	+	C	+	+	+	+	+	+	1	1	1	2	1	2	28	6	10	9		53	352	3
Lovec, Makayla	D034039	+	C	C	C	C	+	+	+	A	C	0	0	0	0	0	0	16	3	3	6		28	334	2
Lyon, Steven W	D034037	+	C	+	+	+	+	+	+	A	+	1	0	0	0	2	1	29	7	8	9		53	352	3
Madore, Brittany I	D034026	A	+	+	A	+	+	+	+	B	+	1	1	1	0	1	0	22	8	6	4		40	342	3
Mccrobie, Casey L	D034036	+	+	C	+	C	+	+	+	A	+	1	0	1	2	2	2	28	7	7	10		52	351	3
Mckinney, Timothy	D034020	+	C	+	+	C	+	+	+	B	+	1	0	0	0	1	0	25	4	6	8		43	344	3
Mehigan, Jessica	D034043																	0	0	0	0		0		W
Michaud, Cody M	D034052	+	B	+	C	+	+	+	A	A	B	0	0	1	1	1	0	19	4	4	3		30	335	2
Mills, Tayla	D034001	+	+	+	C	+	+	+	+	A	+	1	0	1	1	2	2	22	8	8	10		48	348	3
Mitchell, Kaelin M	D034002	+	+	C	+	D	A	+	+	D	+	0	0	0	1	0	0	20	2	5	7		34	338	2
Moore, Tevin	D034028	D	+	+	+	D	+	+	+	B	+	1	0		0	0	0	25	5	6	6		42	343	3
Pantojadiaz, Ravennebeat	D034049	+	C	D	+	D	A	A	A	B	C	1	0	0	0	1	0	14	4	1	4		23	330	1
Powers, Alexandra C	D034057	+	+	+	+	+	+	+	+	+	+	1	1	1	1	2	2	31	9	8	10		58	358	4
Shea, Timothy	D034059																	0	0	0	0		0		N
Siegel, Malesha	D034016	+	C	B	+	D	+	+	+	D	C	0	0	0	0	0	1	19	3	8	5		35	338	2
Simpson, Jordan	D034044	+	+	+	C	+	+	+	+	A	+	1	1	0	2	1	1	32	5	8	9		54	353	4
Srey, Sophany C	D034033	+	+	+	+	+	+	+	+	+	+	0	0	1	0	1	2	28	6	9	10		53	352	3
Statham, Trevor	D034025	A	C	D	C	D	+	D	C	A	B	1	0	0	0	1	0	12	3	2	1		18	325	1

C O N F I D E N T I A L

Fall 2010 - Beginning of Grade 03 NECAP Tests

Grade 03 Students in 2010-2011

Item Analysis Report — Mathematics

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

	Released Items																Total Test Results							
Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Subcategory Points Earned					Total Points Earned	Scaled Score	Achievement Level
Content Strand	NO	NO	NO	NO	NO	NO	FA	FA	FA	DP	NO	NO	GM	NO	GM	DP	Numbers & Operations	Geometry & Measurement	Functions & Algebra	Data, Statistics, & Probability				
GE Code	2-1	2-2	2-2	2-2	2-3	2-5	2-1	2-1	2-4	2-2	2-2	2-5	2-1	2-3	2-7	2-4								
Depth of Knowledge Code	2	2	2	2	1	1	2	2	2	2	2	2	2	2	2	2								
Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	SA	SA	SA	SA	SA	SA								
Correct MC Response	C	D	A	B	A	B	B	B	C	D														
Name/Student ID Total Possible Points	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	35	10	10	10		65		

Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16					
Percent Correct/Average Score: Group	78	52	65	52	52	87	83	78	22	61	0.7	0.3	0.5	0.7	0.9	0.7	22.1	4.8	6.0	6.1	
Percent Correct/Average Score: School	78	52	65	52	52	87	83	78	22	61	0.7	0.3	0.5	0.7	0.9	0.7	22.1	4.8	6.0	6.1	
Percent Correct/Average Score: District	78	59	73	49	65	84	80	80	33	57	0.7	0.4	0.6	0.8	1.0	0.8	23.6	5.4	6.4	6.3	
Percent Correct/Average Score: State	69	75	73	59	57	89	87	79	39	41	0.7	0.3	0.7	0.8	0.9	0.8	23.9	5.4	6.6	6.2	

CONFIDENTIAL

Fall 2010 - Beginning of Grade 08 NECAP Tests

Grade 08 Students in 2010-2011

Item Analysis Report — Writing

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

		Released Items														Total Test Results					
	Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Subcategory Points Earned			Total Points Earned	Scaled Score	Achievement Level
	Content Strand	SC	SC	SC	SC	SC	SC	SC	SC	SC	SC	NW	LR	PW	RW	Multiple Choice	Short Responses	Extended Response			
	GE Code	7-9	7-9	7-9	7-9	7-9	7-1	7-9	7-9	7-9	7-9	7-5	7-2	7-8	M						
	Depth of Knowledge Code	1	1	1	1	1	2	1	1	1	1	2	2	2	3						
	Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	CR	CR	CR	ER						
	Correct MC Response	D	B	A	C	B	D	A	C	B	C										
Name/Student ID	Total Possible Points	1	1	1	1	1	1	1	1	1	1	4	4	4	12	10	12	12	34		
Barnes, Sawyer	D084018	+	+	C	+	+	B	B	A	+	A	2	1	1	2	5	4	2	11	822	1
Bergeron, Micheal	D084022	+	D	C	D	+	C	B	A	C	D	3	2	2	6	2	7	6	15	829	2
Bernal, Perla	D084019	+	+	C	+	+	B	D	+	+	B	4	3	3	7	6	10	7	23	844	3
Blushi, Thomas	D084047	+	+	+	+	+	+	+	+	+	+	3	2	2	8	10	7	8	25	849	3
Boekankamp, Danielle C	D084051	+	+	+	+	+	+	+	+	+	+	4	3	3	8	10	10	8	28	856	4
Brayton, Christop J	D084060	+	+	+	D	+	+	D	+	C	+	2	2	3	7	7	7	7	21	840	3
Brewer, Katie R	D084038	+	+	+	+	+	+	+	+	C	D	4	3	3	8	8	10	8	26	851	3
Byrd, Jacob M	D084058	+	+	+	+	A	A	+	B	+	+	4	3	3	6	7	10	6	23	844	3
Calderon, Bryan	D084061	+	+	+	+	A	A	+	D	D	D	3	3	3	7	5	9	7	21	840	3
Cambria, Reagan M	D084011	+	+	B	A	+	+	+	+	D	+	4	2	3	7	7	9	7	23	844	3
Carbonell, Jasmine L	D084066															0	0	0	0		N
Chamberland, Matthew	D084063	+	+	+	+	+	+	+	+	+	+	3	2	2	6	10	7	6	23	844	3
Edwards, Trey	D084015	+	+	+	+	+	+	+	+	+	+	2	3	4	6	10	9	6	25	849	3
Eggleston, Rachel N	D084046	+	+	+	+	+	B	+	+	C	B	4	2	3	7	7	9	7	23	844	3
Fennell, Jeffrey	D084002	+	+	+	+	+	B	+	D	A	+	4	3	3	6	7	10	6	23	844	3
Fisher, Arianna J	D084043	+	+	+	+	+	A	+	+	+	+	3	3	3	9	9	9	9	27	853	3
Fong, Eddy N	D084024	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	800	1
Frances, Brittany	D084020	+	+	+	A	+	+	B	A	A	B	2	1	2	6	5	5	6	16	831	2
Godair, Bryan	D084026	+	+	+	+	+	+	B	+	+	B	4	3	2	6	8	9	6	23	844	3
Godes, Amanda M	D084021	+	+	+	A	C	C	+	B	D	A	2	2	1	6	4	5	6	15	829	2
Hanson, Kadeen	D084032	+	+	+	+	A	B	D	+	A	D	3	3	2	7	5	8	7	20	838	2
Harper, Jordan J	D084010	+	+	C	+	+	C	D	D	C	+	2	2	2	2	5	6	2	13	826	1
Hoskins, Tamara L	D084025	A	C	+	A	+	C	+	A	C	A	3	2	2	0	3	7	0	10	821	1
Huish, Alfred P	D084029	+	+	B	+	+	C	D	+	A	A	3	3	3	6	5	9	6	20	838	2
Mekkelson, Olivia R	D084039	+	+	+	+	+	+	+	+	+	A	4	4	4	8	9	12	8	29	859	4
Milekic, Luka	D084014															0	0	0	0		S
Porter, Travis	D084034	+	+	+	+	+	+	+	+	+	+	4	3	3	10	10	10	10	30	863	4
Smith, Christopher	D084017	+	+	C	A	A	+	C	D	+	B	2	1	1	5	4	4	5	13	826	1
Spencer, Joshua	D084012	+	+	D	B	+	C	+	+	C	B	3	4	3	7	5	10	7	22	842	3
Thomas, James A	D084006															0	0	0	0		S

CONFIDENTIAL

Fall 2010 - Beginning of Grade 08 NECAP Tests

Grade 08 Students in 2010-2011

Item Analysis Report — Writing

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

		Released Items														Total Test Results					
	Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Subcategory Points Earned			Total Points Earned	Scaled Score	Achievement Level
	Content Strand	SC	SC	SC	SC	SC	SC	SC	SC	SC	SC	NW	LR	PW	RW	Multiple Choice	Short Responses	Extended Response			
	GE Code	7-9	7-9	7-9	7-9	7-9	7-1	7-9	7-9	7-9	7-9	7-5	7-2	7-8	M						
	Depth of Knowledge Code	1	1	1	1	1	2	1	1	1	1	2	2	2	3						
	Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	CR	CR	CR	ER						
	Correct MC Response	D	B	A	C	B	D	A	C	B	C										
Name/Student ID	Total Possible Points	1	1	1	1	1	1	1	1	1	1	4	4	4	12	**10**	**12**	**12**	**34**		
Tuckerdillon, Kaedon	D084059	+	+	+	+	+	+	+	+	C	A	2	2	2	6	**8**	**6**	**6**	**20**	**838**	**2**
Turavani, Giana	D084030	A	D	B	A	+	C	+	D	D	A	4	2	2	6	**2**	**8**	**6**	**16**	**831**	**2**

Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14			
Percent Correct/Average Score: Group	90	86	66	66	79	45	62	59	41	38	3.0	2.4	2.4	6.0	6.3	7.8	6.0
Percent Correct/Average Score: School	90	86	66	66	79	45	62	59	41	38	3.0	2.4	2.4	6.0	6.3	7.8	6.0
Percent Correct/Average Score: District	86	88	70	65	79	51	72	65	47	46	3.0	2.4	2.4	6.4	6.7	7.9	6.4
Percent Correct/Average Score: State	89	84	80	79	69	62	74	64	49	55	3.0	2.4	2.4	6.7	7.0	7.8	6.7

About The New England Common Assessment Program

This report highlights results from the Fall 2010 New England Common Assessment Program (NECAP) tests. The NECAP tests are administered to students in New Hampshire, Rhode Island, and Vermont as part of each state's statewide assessment program. NECAP test results are used primarily for school improvement and accountability. Achievement level results are used in the state accountability system required under No Child Left Behind (NCLB). More detailed school and district results are used by schools to help improve curriculum and instruction. Individual student results are used to support information gathered through classroom instruction and assessments.

NECAP tests in reading and mathematics are administered to students in grades 3 through 8 and 11 and writing tests are administered to students in grades 5, 8, and 11. The NECAP grade 11 tests are designed to measure student performance on grade span expectations (GSE) developed and adopted by the three states. Specifically, the tests are designed to measure the content and skills that students are expected to have as they begin the school year in their current grade – in other words, the content and skills which students have learned through the end of the previous grade.

Each test contains a mix of multiple-choice and constructed-response questions. Constructed-response questions require students to develop their own answers to questions. On the mathematics test, students may be required to provide the correct answer to a computation or word problem, draw or interpret a chart or graph, or explain how they solved a problem. On the reading test, students may be required to make a list or write a few paragraphs to answer a question related to a literary or informational passage. On the writing test, students are required to provide two extended responses of 1-3 pages.

This report contains a variety of school- and/or district-, and state-level assessment results for the NECAP tests administered at a grade level. Achievement level distributions and mean scaled scores are provided for all students tested as well as for subgroups of students classified by demographics or program participation. The report also contains comparative information on school and district performance on subtopics within each content area tested.

In addition to this report of grade 11 results, schools and districts will also receive Item Analysis Reports, Released Item support materials, and student-level data files containing NECAP results. Together, these reports and data constitute a rich source of information to support local decisions in curriculum, instruction, assessment, and professional development. Over time, this information can also strengthen school's and district's evaluation of their ongoing improvement efforts.

Fall 2010
Beginning of Grade 11
NECAP Tests

Grade 11 Students in 2010-2011

School Results

School: Demonstration School 1

District: Demonstration District A

Code: DA-DEMO1

Fall 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Grade Level Summary Report

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-DEMO1

PARTICIPATION in NECAP	Number									Percentage								
	School			District			State			School			District			State		
Students enrolled on or after October 1	43			85			11,164			100			100			100		
	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing
Students tested	**41**	**42**	**41**	**78**	**79**	**77**	**10,628**	**10,672**	**10,594**	**95**	**98**	**95**	**92**	**93**	**91**	**95**	**96**	**95**
With an approved accommodation	10	10	8	12	18	9	1,745	2,408	1,172	24	24	20	15	23	12	16	23	11
Current LEP Students	2	3	2	5	5	4	271	311	261	5	7	5	6	6	5	3	3	2
With an approved accommodation	0	0	0	0	0	0	36	50	28	0	0	0	0	0	0	13	16	11
IEP Students	4	4	4	5	5	5	1,627	1,625	1,612	10	10	10	6	6	6	15	15	15
With an approved accommodation	3	3	3	3	3	3	820	841	747	75	75	75	60	60	60	50	52	46
Students not tested in NECAP	**2**	**1**	**2**	**7**	**6**	**8**	**536**	**492**	**570**	**5**	**2**	**5**	**8**	**7**	**9**	**5**	**4**	**5**
State Approved	2	1	2	5	3	5	195	156	189	100	100	100	71	50	63	36	32	33
Alternate Assessment	0	0	0	1	1	1	108	108	108	0	0	0	20	33	20	55	69	57
First Year LEP	1	0	1	1	0	1	40	0	41	50	0	50	20	0	20	21	0	22
Withdrew After October 1	1	1	1	1	1	1	34	35	30	50	100	50	20	33	20	17	22	16
Enrolled After October 1	0	0	0	1	0	1	1	0	1	0	0	0	20	0	20	1	0	1
Special Consideration	0	0	0	1	1	1	12	13	9	0	0	0	20	33	20	6	8	5
Other	0	0	0	2	3	3	341	336	381	0	0	0	29	50	38	64	68	67

NECAP RESULTS

	School													District						State					
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
READING	43	2	0	41	9	22	23	56	6	15	3	7	1146	78	23	59	12	6	1147	10,628	28	48	16	8	1147
MATH	43	1	0	42	0	0	9	21	16	38	17	40	1134	79	1	30	32	37	1136	10,672	3	30	29	38	1135
WRITING	43	2	0	41	0	0	24	59	14	34	3	7	6.3	77	0	57	38	5	6.4	10,594	1	50	44	5	6.4

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 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Reading Results

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-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.

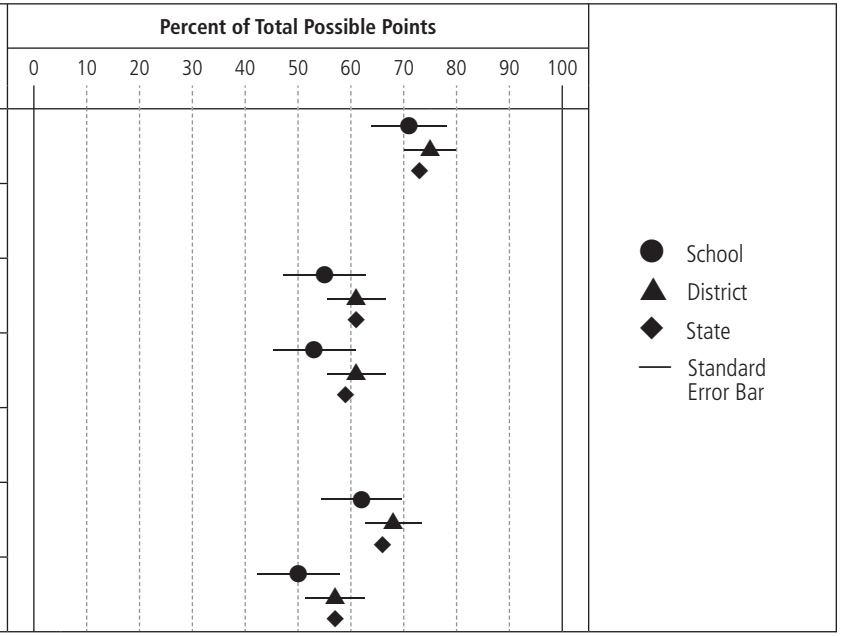
Proficient (Level 3)
Student's performance demonstrates an ability to read and comprehend grade-appropriate text. Student is able to analyze and interpret literary and informational text. Student makes and supports relevant assertions by referencing text. Student uses vocabulary strategies and breadth of vocabulary knowledge to read and comprehend text.

Partially Proficient (Level 2)
Student's performance demonstrates an inconsistent ability to read and comprehend grade-appropriate text. Student attempts to analyze and interpret literary and informational text. Student may make and/or support assertions by referencing text. Student's vocabulary knowledge and use of strategies may be limited and may impact the ability to read and comprehend text.

Substantially Below Proficient (Level 1)
Student's performance demonstrates minimal ability to derive/construct meaning from grade-appropriate text. Student may be able to recognize story elements and text features. Student's limited vocabulary knowledge and use of strategies impacts the ability to read and comprehend text.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2008-09	71	1	18	52	0	0	20	38	17	33	15	29	1136
2009-10	69	3	2	64	14	22	35	55	10	16	5	8	1146
2010-11	**43**	**2**	**0**	**41**	**9**	**22**	**23**	**56**	**6**	**15**	**3**	**7**	**1146**
Cumulative Total	183	6	20	157	23	15	78	50	33	21	23	15	1143
DISTRICT													
2008-09	130	6	0	124	16	13	66	53	31	25	11	9	1143
2009-10	128	6	3	119	33	28	60	50	17	14	9	8	1148
2010-11	**85**	**5**	**2**	**78**	**18**	**23**	**46**	**59**	**9**	**12**	**5**	**6**	**1147**
Cumulative Total	343	17	5	321	67	21	172	54	57	18	25	8	1146
STATE													
2008-09	11,203	179	340	10,684	1,909	18	5,498	51	2,125	20	1,152	11	1145
2009-10	11,189	223	224	10,742	2,466	23	5,416	50	1,875	17	985	9	1146
2010-11	**11,164**	**195**	**341**	**10,628**	**2,964**	**28**	**5,126**	**48**	**1,688**	**16**	**850**	**8**	**1147**
Cumulative Total	33,556	597	905	32,054	7,339	23	16,040	50	5,688	18	2,987	9	1146

Subtopic	Total Possible Points
Word ID/Vocabulary	20
Type of Text	
Literary	42
Informational	42
Level of Comprehension	
Initial Understanding	31
Analysis & Interpretation	53



Fall 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Disaggregated Reading Results

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-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	**43**	**2**	**0**	**41**	**9**	**22**	**23**	**56**	**6**	**15**	**3**	**7**	**1146**	**78**	**23**	**59**	**12**	**6**	**1147**	**10,628**	**28**	**48**	**16**	**8**	**1147**
Gender																									
Male	22	2	0	20	2	10	12	60	5	25	1	5	1143	33	12	67	15	6	1144	5,364	22	50	18	10	1145
Female	19	0	0	19	7	37	9	47	1	5	2	11	1150	42	33	50	10	7	1150	5,255	34	46	14	6	1149
Not Reported	2	0	0	2										3						9					
Race/Ethnicity																									
Hispanic or Latino	7	1	0	6										14	0	64	14	21	1139	1,803	11	48	24	17	1141
Not Hispanic or Latino																									
American Indian or Alaskan Native	1	0	0	1										1						63	16	41	30	13	1142
Asian	1	0	0	1										2						288	26	52	18	4	1148
Black or African American	2	0	0	2										5						862	11	46	27	15	1141
Native Hawaiian or Pacific Islander	1	0	0	1										1						30	3	53	23	20	1140
White	25	0	0	25	7	28	14	56	3	12	1	4	1149	47	32	57	9	2	1151	7,343	34	49	12	5	1150
Two or more races	2	0	0	2										4						166	28	47	20	5	1147
No Race/Ethnicity Reported	4	1	0	3										4						73	5	19	38	37	1134
LEP Status																									
Current LEP student	3	1	0	2										5						271	1	18	32	48	1131
Former LEP student - monitoring year 1	1	0	0	1										1						26	0	54	38	8	1140
Former LEP student - monitoring year 2	0	0	0	0										1						32	3	53	38	6	1141
All Other Students	39	1	0	38	9	24	22	58	6	16	1	3	1147	71	24	62	13	1	1149	10,299	29	49	15	7	1148
IEP																									
Students with an IEP	5	1	0	4										5						1,627	4	32	33	31	1135
All Other Students	38	1	0	37	9	24	21	57	5	14	2	5	1147	73	25	60	11	4	1148	9,001	32	51	13	4	1149
SES																									
Economically Disadvantaged Students	15	1	0	14	2	14	5	36	4	29	3	21	1139	28	11	50	21	18	1141	3,762	13	49	24	14	1142
All Other Students	28	1	0	27	7	26	18	67	2	7	0	0	1150	50	30	64	6	0	1151	6,866	36	48	11	5	1150
Migrant																									
Migrant Students	0	0	0	0										0						0					
All Other Students	43	2	0	41	9	22	23	56	6	15	3	7	1146	78	23	59	12	6	1147	10,628	28	48	16	8	1147
Title I																									
Students Receiving Title I Services	12	1	0	11	1	9	5	45	2	18	3	27	1137	19	5	53	16	26	1138	2,652	13	48	25	15	1142
All Other Students	31	1	0	30	8	27	18	60	4	13	0	0	1150	59	29	61	10	0	1150	7,976	33	48	13	6	1149
504 Plan																									
Students with a 504 Plan	2	0	0	2										2						250	30	55	12	3	1149
All Other Students	41	2	0	39	9	23	22	56	5	13	3	8	1146	76	24	59	11	7	1148	10,378	28	48	16	8	1147

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 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Mathematics Results

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-DEMO1

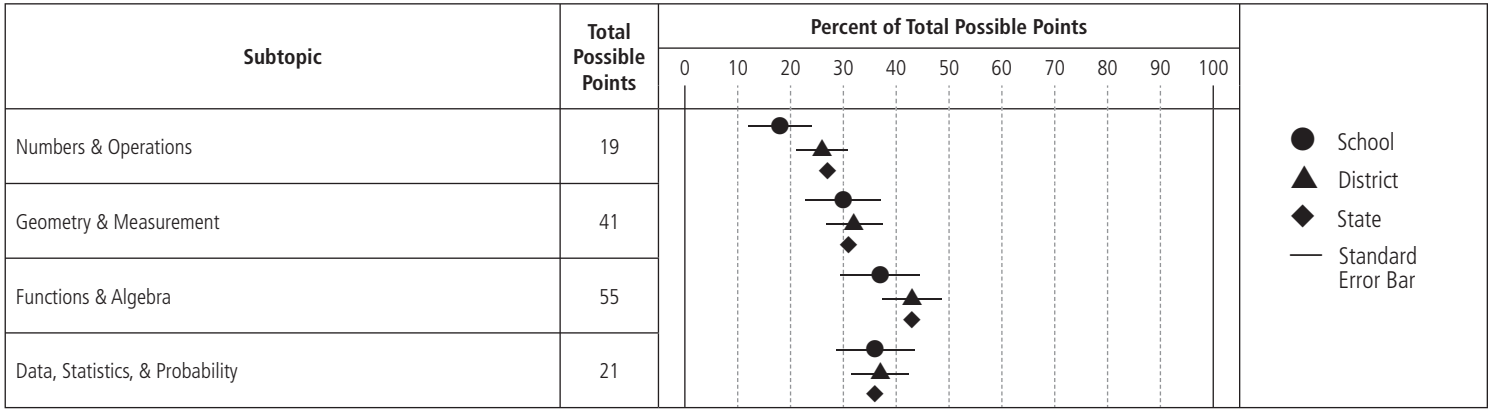
Proficient with Distinction (Level 4)
Student's problem solving demonstrates logical reasoning with strong explanations that include both words and proper mathematical notation. Student's work exhibits a high level of accuracy, effective use of a variety of strategies, and an understanding of mathematical concepts within and across grade level expectations. Student demonstrates the ability to move from concrete to abstract representations.

Proficient (Level 3)
Student's problem solving demonstrates logical reasoning with appropriate explanations that include both words and proper mathematical notation. Student uses a variety of strategies that are often systematic. Computational errors do not interfere with communicating understanding. Student demonstrates conceptual understanding of most aspects of the grade level expectations.

Partially Proficient (Level 2)
Student's problem solving demonstrates logical reasoning and conceptual understanding in some, but not all, aspects of the grade level expectations. Many problems are started correctly, but computational errors may get in the way of completing some aspects of the problem. Student uses some effective strategies. Student's work demonstrates that he or she is generally stronger with concrete than abstract situations.

Substantially Below Proficient (Level 1)
Student's problem solving is often incomplete, lacks logical reasoning and accuracy, and shows little conceptual understanding in most aspects of the grade level expectations. Student is able to start some problems but computational errors and lack of conceptual understanding interfere with solving problems successfully.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2008-09	71	2	18	51	0	0	2	4	10	20	39	76	1128
2009-10	69	3	2	64	3	5	14	22	19	30	28	44	1135
2010-11	**43**	**1**	**0**	**42**	**0**	**0**	**9**	**21**	**16**	**38**	**17**	**40**	**1134**
Cumulative Total	183	6	20	157	3	2	25	16	45	29	84	54	1132
DISTRICT													
2008-09	130	7	0	123	1	1	27	22	34	28	61	50	1132
2009-10	128	5	3	120	4	3	29	24	36	30	51	43	1135
2010-11	**85**	**3**	**3**	**79**	**1**	**1**	**24**	**30**	**25**	**32**	**29**	**37**	**1136**
Cumulative Total	343	15	6	322	6	2	80	25	95	30	141	44	1134
STATE													
2008-09	11,203	161	343	10,699	155	1	2,739	26	3,007	28	4,798	45	1134
2009-10	11,189	195	278	10,716	181	2	2,748	26	2,958	28	4,829	45	1134
2010-11	**11,164**	**156**	**336**	**10,672**	**272**	**3**	**3,224**	**30**	**3,094**	**29**	**4,082**	**38**	**1135**
Cumulative Total	33,556	512	957	32,087	608	2	8,711	27	9,059	28	13,709	43	1134



Fall 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Disaggregated Mathematics Results

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-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	**43**	**1**	**0**	**42**	**0**	**0**	**9**	**21**	**16**	**38**	**17**	**40**	**1134**	**79**	**1**	**30**	**32**	**37**	**1136**	**10,672**	**3**	**30**	**29**	**38**	**1135**
Gender																									
Male	22	1	0	21	0	0	3	14	8	38	10	48	1131	33	0	24	39	36	1134	5,386	3	31	28	38	1135
Female	19	0	0	19	0	0	6	32	7	37	6	32	1137	43	2	37	26	35	1137	5,277	2	29	30	39	1135
Not Reported	2	0	0	2										3						9					
Race/Ethnicity																									
Hispanic or Latino	7	0	0	7										14	0	21	14	64	1130	1,837	<1	13	25	62	1129
Not Hispanic or Latino																									
American Indian or Alaskan Native	1	0	0	1										1						64	2	14	28	56	1132
Asian	1	0	0	1										2						293	4	33	31	32	1137
Black or African American	2	0	0	2										5						865	<1	12	24	63	1129
Native Hawaiian or Pacific Islander	1	0	0	1										1						29	0	21	24	55	1129
White	25	0	0	25	0	0	7	28	11	44	7	28	1136	47	2	38	38	21	1138	7,342	3	37	31	29	1137
Two or more races	2	0	0	2										4						166	1	23	31	45	1135
No Race/Ethnicity Reported	4	1	0	3										5						76	0	8	13	79	1124
LEP Status																									
Current LEP student	3	0	0	3										5						311	0	3	9	89	1122
Former LEP student - monitoring year 1	1	0	0	1										1						26	0	8	15	77	1130
Former LEP student - monitoring year 2	0	0	0	0										1						33	0	12	24	64	1130
All Other Students	39	1	0	38	0	0	9	24	16	42	13	34	1134	72	1	32	35	32	1136	10,302	3	31	30	37	1135
IEP																									
Students with an IEP	5	1	0	4										5						1,625	0	6	13	81	1124
All Other Students	38	0	0	38	0	0	9	24	15	39	14	37	1136	74	1	32	32	34	1137	9,047	3	35	32	31	1137
SES																									
Economically Disadvantaged Students	15	0	0	15	0	0	2	13	4	27	9	60	1130	28	0	18	29	54	1132	3,794	1	16	27	56	1131
All Other Students	28	1	0	27	0	0	7	26	12	44	8	30	1136	51	2	37	33	27	1137	6,878	4	38	30	28	1137
Migrant																									
Migrant Students	0	0	0	0										0						0					
All Other Students	43	1	0	42	0	0	9	21	16	38	17	40	1134	79	1	30	32	37	1136	10,672	3	30	29	38	1135
Title I																									
Students Receiving Title I Services	12	0	0	12	0	0	1	8	2	17	9	75	1126	19	0	16	16	68	1129	2,680	1	15	25	60	1130
All Other Students	31	1	0	30	0	0	8	27	14	47	8	27	1137	60	2	35	37	27	1138	7,992	3	35	30	31	1137
504 Plan																									
Students with a 504 Plan	2	0	0	2										2						250	4	32	37	27	1137
All Other Students	41	1	0	40	0	0	9	23	16	40	15	38	1134	77	1	31	32	35	1136	10,422	3	30	29	39	1135

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 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Writing Results

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-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.

Proficient (Level 3)
Student's writing demonstrates an ability to respond to prompt/task. Focus is clear and maintained throughout the response. Response is organized with a beginning, middle, and end with appropriate transitions. Details are sufficiently elaborated to support focus. Sentence structures and language use are varied. Response demonstrates control of conventions; errors may occur but do not interfere with meaning.

Partially Proficient (Level 2)
Student's writing demonstrates an attempt to respond to prompt/task. Focus may be present but not maintained. Organizational structure is inconsistent with limited use of transitions. Details may be listed and lack elaboration. Sentence structures and language use are unsophisticated and may be repetitive. Response demonstrates inconsistent control of conventions.

Substantially Below Proficient (Level 1)
Student's writing demonstrates a minimal response to prompt/task. Focus is unclear or lacking. Little or no organizational structure is evident. Details are minimal and/or random. Sentence structures and language use are minimal or absent. Frequent errors in conventions may interfere with meaning.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2008-09	71	1	17	53	0	0	6	11	41	77	6	11	5.2
2009-10	69	3	2	64	2	3	32	50	25	39	5	8	6.4
2010-11	**43**	**2**	**0**	**41**	**0**	**0**	**24**	**59**	**14**	**34**	**3**	**7**	**6.3**
Cumulative Total	183	6	19	158	2	1	62	39	80	51	14	9	6.0
DISTRICT													
2008-09	130	6	0	124	3	2	51	41	55	44	15	12	6.0
2009-10	128	6	3	119	8	7	60	50	45	38	6	5	6.8
2010-11	**85**	**5**	**3**	**77**	**0**	**0**	**44**	**57**	**29**	**38**	**4**	**5**	**6.4**
Cumulative Total	343	17	6	320	11	3	155	48	129	40	25	8	6.4
STATE													
2008-09	11,203	156	378	10,669	403	4	4,023	38	5,322	50	921	9	6.1
2009-10	11,189	221	289	10,679	742	7	5,124	48	4,253	40	560	5	6.7
2010-11	**11,164**	**189**	**381**	**10,594**	**119**	**1**	**5,269**	**50**	**4,680**	**44**	**526**	**5**	**6.4**
Cumulative Total	33,556	566	1,048	31,942	1,264	4	14,416	45	14,255	45	2,007	6	6.4

	Types of Writing Reported in the Results Above
2008-09	**Report** Writing that results from gathering, investigating, and organizing facts and thoughts on a focused topic.
2009-10	**Procedure** Writing a procedure is writing to explain a process or to inform an audience of how to do something. A procedure piece presents the steps of the process in a clear, logical, easy-to-follow manner; includes all necessary steps; and defines any terms the audience may not know.
2010-11	**Reflective Essay** A form of writing in which the writer explores and shares the meaning of a personal experience, belief, or idea.

Fall 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Writing Results

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-DEMO1

Average Score Comparison by Type of Writing§

Type of Writing	Tested	School Number Tested	School Mean Score	District Number Tested	District Mean Score	State Number Tested	State Mean Score
Reflective Essay A form of writing in which the writer explores and shares the meaning of a personal experience, belief, or idea.	2010-11 (C)	41	6.3	77	6.4	10,594	6.4
	2009-10	9		17	6.5	1,339	6.5
Response to Literary Text Writing in which the writer analyzes plot/ideas/concepts, making inferences about content, characters, philosophy, theme, author's craft, or other elements within a piece of literature or informational text.	2010-11	7		11	6.2	1,327	6.2
	2009-10	11	5.8	22	6.1	1,323	6.2
Response to Informational Text Writing in which the writer analyzes plot/ideas/concepts, making inferences about content, characters, philosophy, theme, author's craft, or other elements within a piece of literature or informational text.	2010-11	6		7		1,313	6.3
	2009-10	4		9		1,321	6.2
Reflective Essay A form of writing in which the writer explores and shares the meaning of a personal experience, belief, or idea.	2010-11	7		16	6.4	1,321	6.6
	2009-10	9		17	6.5	1,339	6.5
Report Writing that results from gathering, investigating, and organizing facts and thoughts on a focused topic.	2010-11	4		12	6.2	1,328	6.6
	2009-10	7		13	6.8	1,337	6.8
Persuasive Essay Persuasive writing is writing that aims at convincing people to accept a point of view, to change their minds about something, or to act in a certain way. A persuasive essay is a form of writing in which a writer supports an opinion and tries to persuade an audience.	2010-11	6		11	6.5	1,340	6.6
	2009-10	6		11	8	1,326	6.6

(C) This type of writing was administered to all students.
The ● shows this year's score and the black bar (▬) shows the range where most students in this sample scored.
The ▲ shows last year's score and the gray bar (▬) shows the range where most students in this sample scored.
§ The range of 0 to 12 on the graphic display represents the possible score range for the writing prompt. The range of 0 to 12 is a result of adding the two scores assigned to the student's response from the 6-point scoring rubric. The score of 7 represents the score required to be proficient.
Note: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Writing Results

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-DEMO1

Score Distribution

Total Score	Score 1	Score 2	School		District		State
			N	%	N	%	%
12	6	6	0	0	0	0	<1
11	6	5	0	0	0	0	<1
10	5	5	0	0	0	0	1
9	5	4	1	2	3	4	3
8	4	4	10	24	20	26	27
7	4	3	13	32	21	27	20
6	3	3	6	15	18	23	29
5	3	2	2	5	3	4	8
4	2	2	6	15	8	10	7
3	2	1	1	2	1	1	1
2	1	1	2	5	2	3	2
0	0	0	0	0	1	1	2

Scoring Rubric

6	• purpose is clear throughout; strong focus/controlling idea OR strongly stated purpose focuses the writing • intentionally organized for effect • fully developed details; rich and/or insightful elaboration supports purpose • distinctive voice, tone, and style enhance meaning • consistent application of the rules of grade-level grammar, usage, and mechanics
5	• purpose is clear; focus/controlling idea is maintained throughout • well-organized and coherent throughout • details are relevant and support purpose; details are sufficiently elaborated • strong command of sentence structure; uses language to enhance meaning • consistent application of the rules of grade-level grammar, usage, and mechanics
4	• purpose is evident; focus/controlling idea may not be maintained • generally organized and coherent • details are relevant and mostly support purpose • well-constructed sentences; uses language well • may show inconsistent control of grade-level grammar, usage, and mechanics
3	• writing has a general purpose • some sense of organization; may have lapses in coherence • some relevant details support purpose • uses language adequately; may show little variety of sentence structures • may contain some serious errors in grammar, usage, and mechanics
2	• attempted or vague purpose; stays on topic • little evidence of organization; lapses in coherence • generalizes or lists details • lacks sentence control; uses language poorly • errors in grammar, usage, and mechanics are distracting
1	• lack of evident purpose; topic may not be clear • incoherent or underdeveloped organization • random information • rudimentary or deficient use of language • serious and persistent errors in grammar, usage, and mechanics throughout
0	Response is totally incorrect or irrelevant.

Score 1 and Score 2 represent two independent scores assigned to a student's response to the common writing prompt. The two scores added together equal the student's total score on the common writing prompt. If the two scores differ by more than one point, the student's response is scored a third time to resolve the difference.

Note: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2010 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2010-2011
Disaggregated Writing Results

School: Demonstration School 1
District: Demonstration District A
State: Rhode Island
Code: DA-DEMO1

REPORTING CATEGORIES	School													District						State					
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students	**43**	**2**	**0**	**41**	**0**	**0**	**24**	**59**	**14**	**34**	**3**	**7**	**6.3**	**77**	**0**	**57**	**38**	**5**	**6.4**	**10,594**	**1**	**50**	**44**	**5**	**6.4**
Gender																									
Male	22	2	0	20	0	0	9	45	8	40	3	15	5.7	32	0	50	41	9	6.1	5,339	1	43	49	6	6.2
Female	19	0	0	19	0	0	14	74	5	26	0	0	6.9	42	0	62	36	2	6.7	5,246	1	56	39	3	6.7
Not Reported	2	0	0	2										3						9					
Race/Ethnicity																									
Hispanic or Latino	7	1	0	6										13	0	38	46	15	5.5	1,784	<1	37	52	11	5.8
Not Hispanic or Latino																									
American Indian or Alaskan Native	1	0	0	1										1						64	2	42	45	11	6.0
Asian	1	0	0	1										2						288	2	54	41	3	6.6
Black or African American	2	0	0	2										5						852	1	34	56	10	5.7
Native Hawaiian or Pacific Islander	1	0	0	1										1						30	0	17	70	13	5.2
White	25	0	0	25	0	0	15	60	8	32	2	8	6.4	47	0	66	30	4	6.7	7,342	1	55	41	3	6.7
Two or more races	2	0	0	2										4						166	1	41	55	3	6.3
No Race/Ethnicity Reported	4	1	0	3										4						68	0	18	47	35	4.4
LEP Status																									
Current LEP student	3	1	0	2										4						261	0	11	50	39	3.7
Former LEP student - monitoring year 1	1	0	0	1										1						26	0	27	65	8	5.7
Former LEP student - monitoring year 2	0	0	0	0										1						33	0	24	70	6	5.7
All Other Students	39	1	0	38	0	0	24	63	12	32	2	5	6.5	71	0	59	37	4	6.6	10,274	1	51	44	4	6.5
IEP																									
Students with an IEP	5	1	0	4										5						1,612	<1	18	64	18	4.9
All Other Students	38	1	0	37	0	0	23	62	13	35	1	3	6.5	72	0	60	38	3	6.6	8,982	1	55	41	3	6.7
SES																									
Economically Disadvantaged Students	15	1	0	14	0	0	6	43	5	36	3	21	5.2	27	0	41	48	11	5.8	3,740	<1	37	54	8	5.9
All Other Students	28	1	0	27	0	0	18	67	9	33	0	0	6.8	50	0	66	32	2	6.8	6,854	1	57	39	3	6.7
Migrant																									
Migrant Students	0	0	0	0										0						0					
All Other Students	43	2	0	41	0	0	24	59	14	34	3	7	6.3	77	0	57	38	5	6.4	10,594	1	50	44	5	6.4
Title I																									
Students Receiving Title I Services	12	1	0	11	0	0	5	45	4	36	2	18	5.2	18	0	39	44	17	5.2	2,623	<1	37	52	10	5.8
All Other Students	31	1	0	30	0	0	19	63	10	33	1	3	6.7	59	0	63	36	2	6.8	7,971	1	54	42	3	6.6
504 Plan																									
Students with a 504 Plan	2	0	0	2										2						250	2	51	46	2	6.7
All Other Students	41	2	0	39	0	0	23	59	13	33	3	8	6.3	75	0	57	37	5	6.4	10,344	1	50	44	5	6.4

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 2010 NECAP Tests
School Summary
2010-2011 Students

School: Demonstration School 1
District: Demonstration District A
State: Maine
Code: DEMA-DEM1

Reading	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
Demonstration School 1	**213**	**8**	**6**	**199**	**26**	**13**	**123**	**62**	**33**	**17**	**17**	**9**	
Beginning of Grade 3	28	3	1	24	3	13	15	63	2	8	4	17	343
Beginning of Grade 4	31	0	1	30	4	13	17	57	5	17	4	13	446
Beginning of Grade 5	36	1	1	34	4	12	22	65	6	18	2	6	546
Beginning of Grade 6	41	2	1	38	4	11	24	63	9	24	1	3	647
Beginning of Grade 7	36	1	1	34	6	18	20	59	7	21	1	3	748
Beginning of Grade 8	41	1	1	39	5	13	25	64	4	10	5	13	846

Mathematics	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
Demonstration School 1	**213**	**5**	**6**	**202**	**38**	**19**	**90**	**45**	**41**	**20**	**33**	**16**	
Beginning of Grade 3	28	2	1	25	6	24	9	36	1	4	9	36	341
Beginning of Grade 4	31	0	1	30	3	10	13	43	10	33	4	13	441
Beginning of Grade 5	36	1	1	34	5	15	17	50	10	29	2	6	545
Beginning of Grade 6	41	1	1	39	6	15	21	54	5	13	7	18	643
Beginning of Grade 7	36	1	1	34	9	26	13	38	8	24	4	12	745
Beginning of Grade 8	41	0	1	40	9	23	17	43	7	18	7	18	842

Writing	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
Demonstration School 1	**77**	**1**	**2**	**74**	**4**	**5**	**34**	**46**	**31**	**42**	**5**	**7**	
Beginning of Grade 5	36	1	1	34	0	0	19	56	15	44	0	0	541
Beginning of Grade 8	41	0	1	40	4	10	15	38	16	40	5	13	838

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

About The New England Common Assessment Program

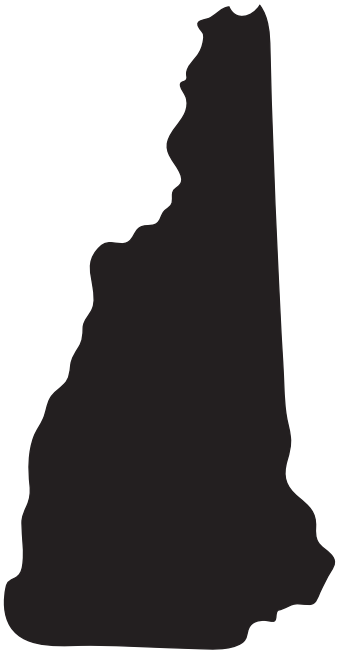
This report highlights results from the Fall 2010 Beginning of Grade New England Common Assessment Program (NECAP) tests. The NECAP tests are administered to students in Maine, New Hampshire, Rhode Island, and Vermont as part of each state's statewide assessment program. NECAP test results are used primarily 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 2010
Beginning of Grade 8
NECAP Tests

Grade 8 Students in 2010-2011

District Results

District: Demonstration District A

Code: DEM-DEA

Fall 2010 - Beginning of Grade 8 NECAP Tests
Grade 8 Students in 2010-2011
Grade Level Summary Report

District: Demonstration District A
State: New Hampshire
Code: DEM-DEA

PARTICIPATION in NECAP	Number									Percentage								
	School			District			State			School			District			State		
Students enrolled on or after October 1				102			15,492						100			100		
	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing
Students tested				**95**	**95**	**94**	**15,173**	**15,208**	**15,125**				**93**	**93**	**92**	**98**	**98**	**98**
With an approved accommodation				16	15	16	2,578	2,559	2,443				17	16	17	17	17	16
Current LEP Students				0	1	0	227	271	225				0	1	0	1	2	1
With an approved accommodation				0	0	0	108	148	120					0		48	55	53
IEP Students				17	16	16	2,424	2,421	2,409				18	17	17	16	16	16
With an approved accommodation				13	13	13	1,830	1,826	1,759				76	81	81	75	75	73
Students not tested in NECAP				**7**	**7**	**8**	**319**	**284**	**367**				**7**	**7**	**8**	**2**	**2**	**2**
State Approved				5	5	6	205	186	209				71	71	75	64	65	57
Alternate Assessment				1	1	1	157	157	157				20	20	17	77	84	75
First Year LEP				1	0	1	23	0	23				20	0	17	11	0	11
Withdrew After October 1				1	1	1	11	13	13				20	20	17	5	7	6
Enrolled After October 1				1	1	1	2	2	1				20	20	17	1	1	<1
Special Consideration				1	2	2	12	14	15				20	40	33	6	8	7
Other				2	2	2	114	98	158				29	29	25	36	35	43

NECAP RESULTS

	District													State												
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
READING	102	5	2	95	21	22	54	57	14	15	6	6	849	15,173	23	55	17	4	850						
MATH	102	5	2	95	23	24	35	37	21	22	16	17	843	15,208	21	45	19	16	843						
WRITING	102	6	2	94	10	11	50	53	29	31	5	5	842	15,125	12	52	29	7	842						

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 2010 - Beginning of Grade 8 NECAP Tests
Grade 8 Students in 2010-2011
Reading Results

District: Demonstration District A
State: New Hampshire
Code: DEM-DEA

Proficient with Distinction (Level 4)
Student's performance demonstrates an ability to read and comprehend grade-appropriate text. Student is able to analyze and interpret literary and informational text. Student offers insightful observations/assertions that are well supported by references to the text. Student uses range of vocabulary strategies and breadth of vocabulary knowledge to read and comprehend a wide variety of texts.

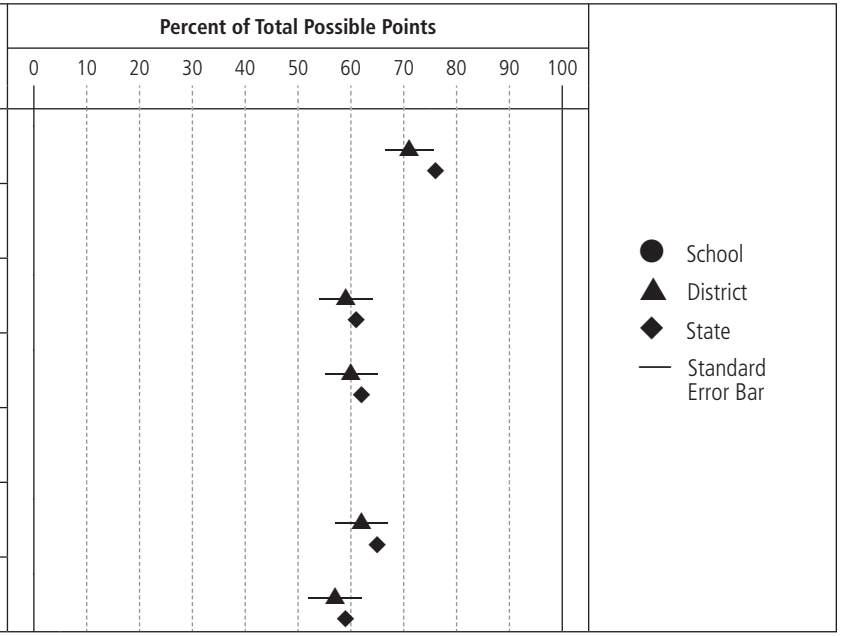
Proficient (Level 3)
Student's performance demonstrates an ability to read and comprehend grade-appropriate text. Student is able to analyze and interpret literary and informational text. Student makes and supports relevant assertions by referencing text. Student uses vocabulary strategies and breadth of vocabulary knowledge to read and comprehend text.

Partially Proficient (Level 2)
Student's performance demonstrates an inconsistent ability to read and comprehend grade-appropriate text. Student attempts to analyze and interpret literary and informational text. Student may make and/or support assertions by referencing text. Student's vocabulary knowledge and use of strategies may be limited and may impact the ability to read and comprehend text.

Substantially Below Proficient (Level 1)
Student's performance demonstrates minimal ability to derive/construct meaning from grade-appropriate text. Student may be able to recognize story elements and text features. Student's limited vocabulary knowledge and use of strategies impacts the ability to read and comprehend text.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2008-09													
2009-10													
2010-11													
Cumulative Total													
DISTRICT													
2008-09	93	0	0	93	9	10	54	58	26	28	4	4	846
2009-10	94	5	3	86	15	17	33	38	28	33	10	12	844
2010-11	**102**	**5**	**2**	**95**	**21**	**22**	**54**	**57**	**14**	**15**	**6**	**6**	**849**
Cumulative Total	289	10	5	274	45	16	141	51	68	25	20	7	846
STATE													
2008-09	15,950	220	90	15,640	2,176	14	8,839	57	3,451	22	1,174	8	846
2009-10	15,671	218	84	15,369	3,326	22	8,305	54	2,874	19	864	6	849
2010-11	**15,492**	**205**	**114**	**15,173**	**3,540**	**23**	**8,365**	**55**	**2,620**	**17**	**648**	**4**	**850**
Cumulative Total	47,113	643	288	46,182	9,042	20	25,509	55	8,945	19	2,686	6	848

Subtopic	Total Possible Points
Word ID/Vocabulary	25
Type of Text	
Literary	56
Informational	49
Level of Comprehension	
Initial Understanding	42
Analysis & Interpretation	63



Fall 2010 - Beginning of Grade 8 NECAP Tests
Grade 8 Students in 2010-2011
Disaggregated Reading Results

District: Demonstration District A
State: New Hampshire
Code: DEM-DEA

REPORTING CATEGORIES	District													State											
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students	**102**	**5**	**2**	**95**	**21**	**22**	**54**	**57**	**14**	**15**	**6**	**6**	**849**	**15,173**	**23**	**55**	**17**	**4**	**850**						
Gender																									
Male	54	1	0	53	11	21	30	57	8	15	4	8	848	7,795	17	57	20	5	848						
Female	48	4	2	42	10	24	24	57	6	14	2	5	850	7,378	30	53	15	3	852						
Not Reported	0	0	0	0										0											
Race/Ethnicity																									
Hispanic or Latino	3	1	0	2										513	11	51	29	10	843						
Not Hispanic or Latino																									
American Indian or Alaskan Native	1	0	1	0										62	18	55	23	5	847						
Asian	4	0	0	4										347	37	50	10	3	854						
Black or African American	2	0	0	2										323	10	51	27	13	843						
Native Hawaiian or Pacific Islander	1	0	0	1										14	7	57	29	7	847						
White	90	4	1	85	18	21	48	56	14	16	5	6	848	13,781	24	56	17	4	850						
Two or more races	0	0	0	0										0											
No Race/Ethnicity Reported	1	0	0	1										133	18	55	21	6	848						
LEP Status																									
Current LEP student	1	1	0	0										227	1	30	43	26	835						
Former LEP student - monitoring year 1	1	0	0	1										29	14	72	14	0	849						
Former LEP student - monitoring year 2	1	0	0	1										29	24	59	17	0	850						
All Other Students	99	4	2	93	21	23	53	57	13	14	6	6	849	14,888	24	55	17	4	850						
IEP																									
Students with an IEP	20	2	1	17	0	0	9	53	5	29	3	18	837	2,424	3	37	41	19	838						
All Other Students	82	3	1	78	21	27	45	58	9	12	3	4	851	12,749	27	59	13	2	852						
SES																									
Economically Disadvantaged Students	23	0	0	23	3	13	12	52	4	17	4	17	843	3,571	12	51	28	9	844						
All Other Students	79	5	2	72	18	25	42	58	10	14	2	3	851	11,602	27	56	14	3	851						
Migrant																									
Migrant Students	0	0	0	0										0											
All Other Students	102	5	2	95	21	22	54	57	14	15	6	6	849	15,173	23	55	17	4	850						
Title I																									
Students Receiving Title I Services	2	0	0	2										437	8	55	31	6	843						
All Other Students	100	5	2	93	21	23	52	56	14	15	6	6	849	14,736	24	55	17	4	850						

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 2010 - Beginning of Grade 8 NECAP Tests
Grade 8 Students in 2010-2011
Mathematics Results

District: Demonstration District A
State: New Hampshire
Code: DEM-DEA

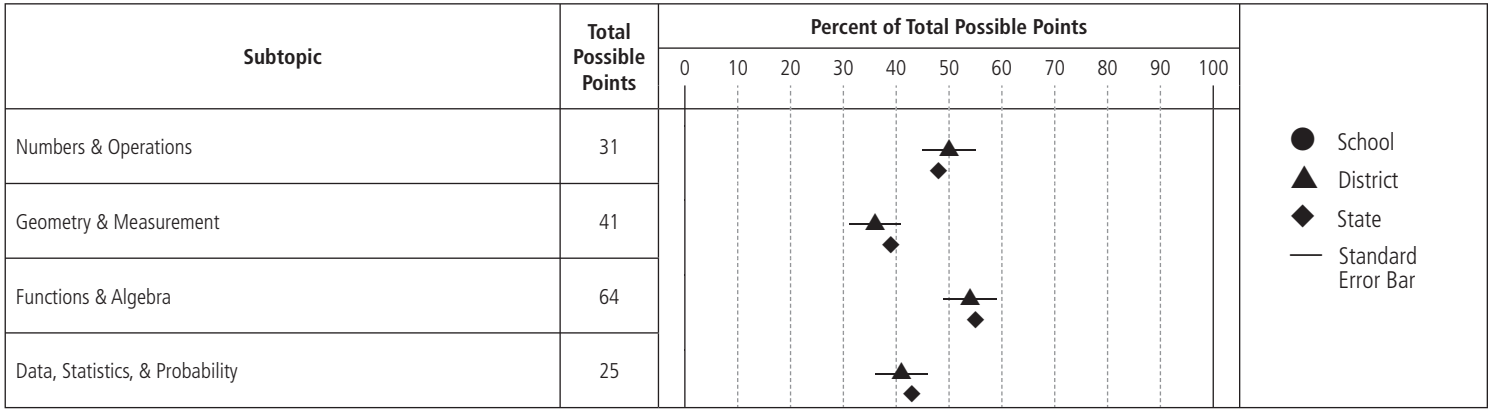
Proficient with Distinction (Level 4)
Student's problem solving demonstrates logical reasoning with strong explanations that include both words and proper mathematical notation. Student's work exhibits a high level of accuracy, effective use of a variety of strategies, and an understanding of mathematical concepts within and across grade level expectations. Student demonstrates the ability to move from concrete to abstract representations.

Proficient (Level 3)
Student's problem solving demonstrates logical reasoning with appropriate explanations that include both words and proper mathematical notation. Student uses a variety of strategies that are often systematic. Computational errors do not interfere with communicating understanding. Student demonstrates conceptual understanding of most aspects of the grade level expectations.

Partially Proficient (Level 2)
Student's problem solving demonstrates logical reasoning and conceptual understanding in some, but not all, aspects of the grade level expectations. Many problems are started correctly, but computational errors may get in the way of completing some aspects of the problem. Student uses some effective strategies. Student's work demonstrates that he or she is generally stronger with concrete than abstract situations.

Substantially Below Proficient (Level 1)
Student's problem solving is often incomplete, lacks logical reasoning and accuracy, and shows little conceptual understanding in most aspects of the grade level expectations. Student is able to start some problems but computational errors and lack of conceptual understanding interfere with solving problems successfully.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2008-09													
2009-10													
2010-11													
Cumulative Total													
DISTRICT													
2008-09	93	0	0	93	15	16	51	55	10	11	17	18	843
2009-10	94	5	2	87	19	22	23	26	23	26	22	25	840
2010-11	**102**	**5**	**2**	**95**	**23**	**24**	**35**	**37**	**21**	**22**	**16**	**17**	**843**
Cumulative Total	289	10	4	275	57	21	109	40	54	20	55	20	842
STATE													
2008-09	15,950	201	94	15,655	2,803	18	7,299	47	3,050	19	2,503	16	843
2009-10	15,671	195	87	15,389	3,224	21	6,885	45	2,760	18	2,520	16	843
2010-11	**15,492**	**186**	**98**	**15,208**	**3,119**	**21**	**6,783**	**45**	**2,901**	**19**	**2,405**	**16**	**843**
Cumulative Total	47,113	582	279	46,252	9,146	20	20,967	45	8,711	19	7,428	16	843



Fall 2010 - Beginning of Grade 8 NECAP Tests
Grade 8 Students in 2010-2011
Disaggregated Mathematics Results

District: Demonstration District A
State: New Hampshire
Code: DEM-DEA

REPORTING CATEGORIES	District													State											
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students	**102**	**5**	**2**	**95**	**23**	**24**	**35**	**37**	**21**	**22**	**16**	**17**	**843**	**15,208**	**21**	**45**	**19**	**16**	**843**						
Gender																									
Male	54	1	0	53	16	30	23	43	7	13	7	13	845	7,811	21	45	18	16	843						
Female	48	4	2	42	7	17	12	29	14	33	9	21	841	7,397	20	45	20	16	843						
Not Reported	0	0	0	0										0											
Race/Ethnicity																									
Hispanic or Latino	3	1	0	2										524	6	31	30	33	837						
Not Hispanic or Latino																									
American Indian or Alaskan Native	1	0	1	0										62	15	39	19	27	840						
Asian	4	0	0	4										366	37	37	13	13	847						
Black or African American	2	0	0	2										333	7	27	25	41	836						
Native Hawaiian or Pacific Islander	1	0	0	1										14	0	50	29	21	840						
White	90	4	1	85	18	21	33	39	19	22	15	18	842	13,777	21	46	19	15	844						
Two or more races	0	0	0	0										0											
No Race/Ethnicity Reported	1	0	0	1										132	14	45	20	22	841						
LEP Status																									
Current LEP student	1	0	0	1										271	1	11	22	66	830						
Former LEP student - monitoring year 1	1	0	0	1										29	14	38	24	24	841						
Former LEP student - monitoring year 2	1	0	0	1										29	7	48	34	10	842						
All Other Students	99	5	2	92	23	25	35	38	20	22	14	15	843	14,879	21	45	19	15	843						
IEP																									
Students with an IEP	20	3	1	16	1	6	4	25	4	25	7	44	835	2,421	3	21	25	51	833						
All Other Students	82	2	1	79	22	28	31	39	17	22	9	11	845	12,787	24	49	18	9	845						
SES																									
Economically Disadvantaged Students	23	1	0	22	0	0	9	41	8	36	5	23	839	3,566	8	36	25	30	838						
All Other Students	79	4	2	73	23	32	26	36	13	18	11	15	845	11,642	24	47	17	11	845						
Migrant																									
Migrant Students	0	0	0	0										0											
All Other Students	102	5	2	95	23	24	35	37	21	22	16	17	843	15,208	21	45	19	16	843						
Title I																									
Students Receiving Title I Services	2	0	0	2										300	4	29	34	33	837						
All Other Students	100	5	2	93	23	25	34	37	21	23	15	16	843	14,908	21	45	19	15	843						

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 2010 - Beginning of Grade 8 NECAP Tests
Grade 8 Students in 2010-2011
Writing Results

District: Demonstration District A
State: New Hampshire
Code: DEM-DEA

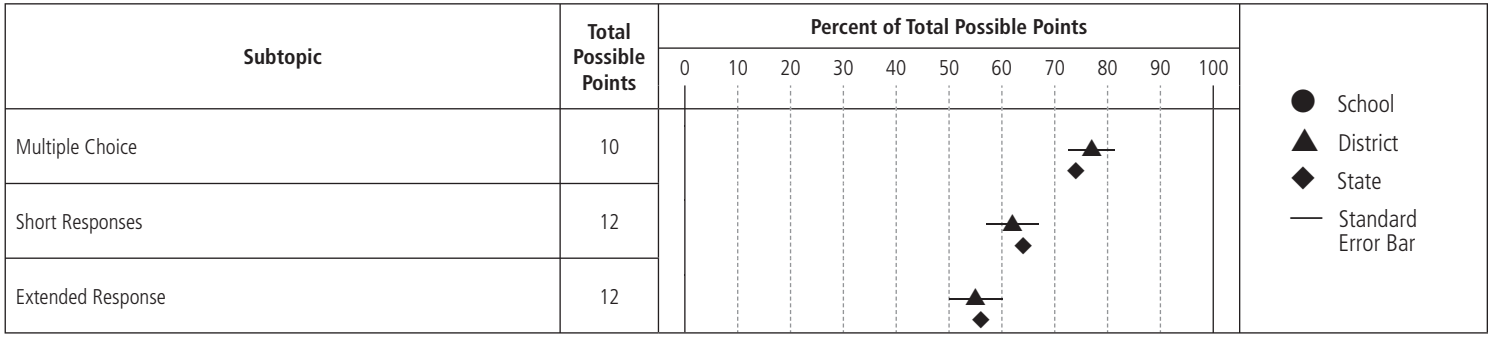
Proficient with Distinction (Level 4)
Student's writing demonstrates an ability to respond to prompt/task with clarity and insight. Focus is well developed and maintained throughout response. Response demonstrates use of strong organizational structures. A variety of elaboration strategies is evident. Sentence structures and language choices are varied and used effectively. Response demonstrates control of conventions; minor errors may occur.

Proficient (Level 3)
Student's writing demonstrates an ability to respond to prompt/task. Focus is clear and maintained throughout the response. Response is organized with a beginning, middle, and end with appropriate transitions. Details are sufficiently elaborated to support focus. Sentence structures and language use are varied. Response demonstrates control of conventions; errors may occur but do not interfere with meaning.

Partially Proficient (Level 2)
Student's writing demonstrates an attempt to respond to prompt/task. Focus may be present but not maintained. Organizational structure is inconsistent with limited use of transitions. Details may be listed and lack elaboration. Sentence structures and language use are unsophisticated and may be repetitive. Response demonstrates inconsistent control of conventions.

Substantially Below Proficient (Level 1)
Student's writing demonstrates a minimal response to prompt/task. Focus is unclear or lacking. Little or no organizational structure is evident. Details are minimal and/or random. Sentence structures and language use are minimal or absent. Frequent errors in conventions may interfere with meaning.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2008-09													
2009-10													
2010-11													
Cumulative Total													
DISTRICT													
2008-09	93	0	0	93	9	10	37	40	37	40	10	11	842
2009-10													
2010-11	**102**	**6**	**2**	**94**	**10**	**11**	**50**	**53**	**29**	**31**	**5**	**5**	**842**
Cumulative Total	195	6	2	187	19	10	87	47	66	35	15	8	842
STATE													
2008-09	15,950	225	145	15,580	1,367	9	6,566	42	5,393	35	2,254	14	841
2009-10													
2010-11	**15,492**	**209**	**158**	**15,125**	**1,829**	**12**	**7,875**	**52**	**4,394**	**29**	**1,027**	**7**	**842**
Cumulative Total	31,442	434	303	30,705	3,196	10	14,441	47	9,787	32	3,281	11	841



No historical data is available for 2009-10 because a pilot test was administered to field-test new writing items for future writing tests.

Fall 2010 - Beginning of Grade 8 NECAP Tests
Grade 8 Students in 2010-2011
Disaggregated Writing Results

District: Demonstration District A
State: New Hampshire
Code: DEM-DEA

REPORTING CATEGORIES	District													State											
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students	**102**	**6**	**2**	**94**	**10**	**11**	**50**	**53**	**29**	**31**	**5**	**5**	**842**	**15,125**	**12**	**52**	**29**	**7**	**842**						
Gender																									
Male	54	1	0	53	3	6	28	53	19	36	3	6	840	7,765	6	46	37	10	839						
Female	48	5	2	41	7	17	22	54	10	24	2	5	844	7,360	18	58	20	3	846						
Not Reported	0	0	0	0										0											
Race/Ethnicity																									
Hispanic or Latino	3	2	0	1										508	6	45	38	12	838						
Not Hispanic or Latino																									
American Indian or Alaskan Native	1	0	1	0										62	13	44	31	13	841						
Asian	4	0	0	4										347	24	56	17	3	847						
Black or African American	2	0	0	2										324	5	41	39	15	838						
Native Hawaiian or Pacific Islander	1	0	0	1										14	0	64	21	14	838						
White	90	4	1	85	10	12	43	51	27	32	5	6	842	13,738	12	53	29	6	843						
Two or more races	0	0	0	0										0											
No Race/Ethnicity Reported	1	0	0	1										132	9	55	27	9	842						
LEP Status																									
Current LEP student	1	1	0	0										225	<1	26	44	30	832						
Former LEP student - monitoring year 1	1	0	0	1										29	3	72	24	0	843						
Former LEP student - monitoring year 2	1	0	0	1										29	14	62	24	0	845						
All Other Students	99	5	2	92	10	11	50	54	27	29	5	5	842	14,842	12	52	29	6	843						
IEP																									
Students with an IEP	20	3	1	16	0	0	5	31	9	56	2	13	834	2,409	1	22	50	27	832						
All Other Students	82	3	1	78	10	13	45	58	20	26	3	4	844	12,716	14	58	25	3	844						
SES																									
Economically Disadvantaged Students	23	1	0	22	1	5	9	41	9	41	3	14	838	3,547	5	41	39	14	838						
All Other Students	79	5	2	72	9	13	41	57	20	28	2	3	843	11,578	14	55	26	5	844						
Migrant																									
Migrant Students	0	0	0	0										0											
All Other Students	102	6	2	94	10	11	50	53	29	31	5	5	842	15,125	12	52	29	7	842						
Title I																									
Students Receiving Title I Services	2	0	0	2										436	3	42	44	10	838						
All Other Students	100	6	2	92	10	11	49	53	28	30	5	5	842	14,689	12	52	29	7	843						

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 2010 NECAP Tests
District Summary
2009-2010 Students

District: Demonstration District A
State: New Hampshire
Code: DEM-DEA

Reading	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
Demonstration District A				**596**	**117**	**20**	**326**	**55**	**95**	**16**	**58**	**10**	
Beginning of Grade 3				91	14	15	57	63	12	13	8	9	346
Beginning of Grade 4				83	24	29	46	55	7	8	6	7	450
Beginning of Grade 5				86	10	12	51	59	12	14	13	15	544
Beginning of Grade 6				68	16	24	38	56	9	13	5	7	648
Beginning of Grade 7				92	15	16	48	52	21	23	8	9	747
Beginning of Grade 8				95	21	22	54	57	14	15	6	6	849
Beginning of Grade 11				81	17	21	32	40	20	25	12	15	1143

Mathematics	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
Demonstration District A				**603**	**132**	**22**	**246**	**41**	**99**	**16**	**126**	**21**	
Beginning of Grade 3				93	22	24	48	52	17	18	6	6	346
Beginning of Grade 4				84	31	37	35	42	6	7	12	14	447
Beginning of Grade 5				87	12	14	45	52	12	14	18	21	544
Beginning of Grade 6				70	19	27	28	40	6	9	17	24	644
Beginning of Grade 7				93	21	23	35	38	20	22	17	18	743
Beginning of Grade 8				95	23	24	35	37	21	22	16	17	843
Beginning of Grade 11				81	4	5	20	25	17	21	40	49	1134

Writing	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
Demonstration District A				**260**	**21**	**8**	**113**	**43**	**104**	**40**	**22**	**8**	
Beginning of Grade 5				86	9	10	31	36	37	43	9	10	540
Beginning of Grade 8				94	10	11	50	53	29	31	5	5	842
Beginning of Grade 11				80	2	3	32	40	38	48	8	10	6.1

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

Achievement Level Summary

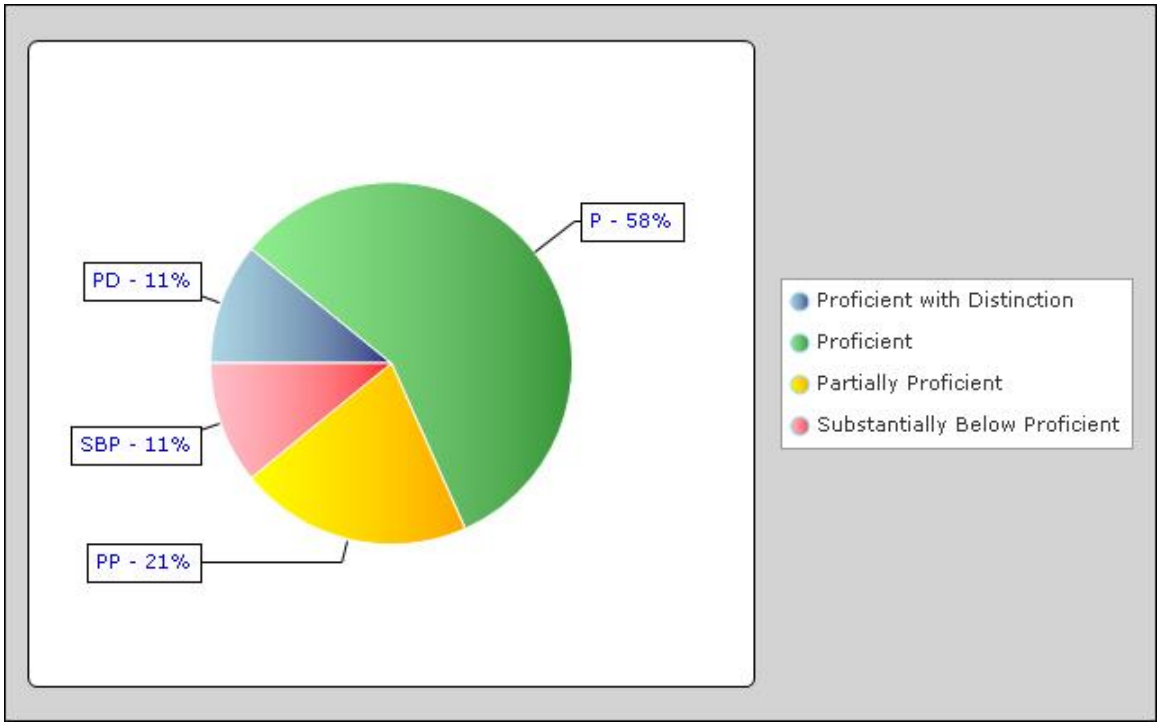
District: Demonstration District A

School: Demonstration School 1

Grade: 06

Date: 6/14/2011 10:49:03 AM

Reading



Achievement Level	Count	Percentage %*
Proficient with Distinction	2	11
Proficient	11	58
Partially Proficient	4	21
Substantially Below Proficient	2	11

*Percentages may not total exactly 100% due to applied rounding.

Achievement Level Summary

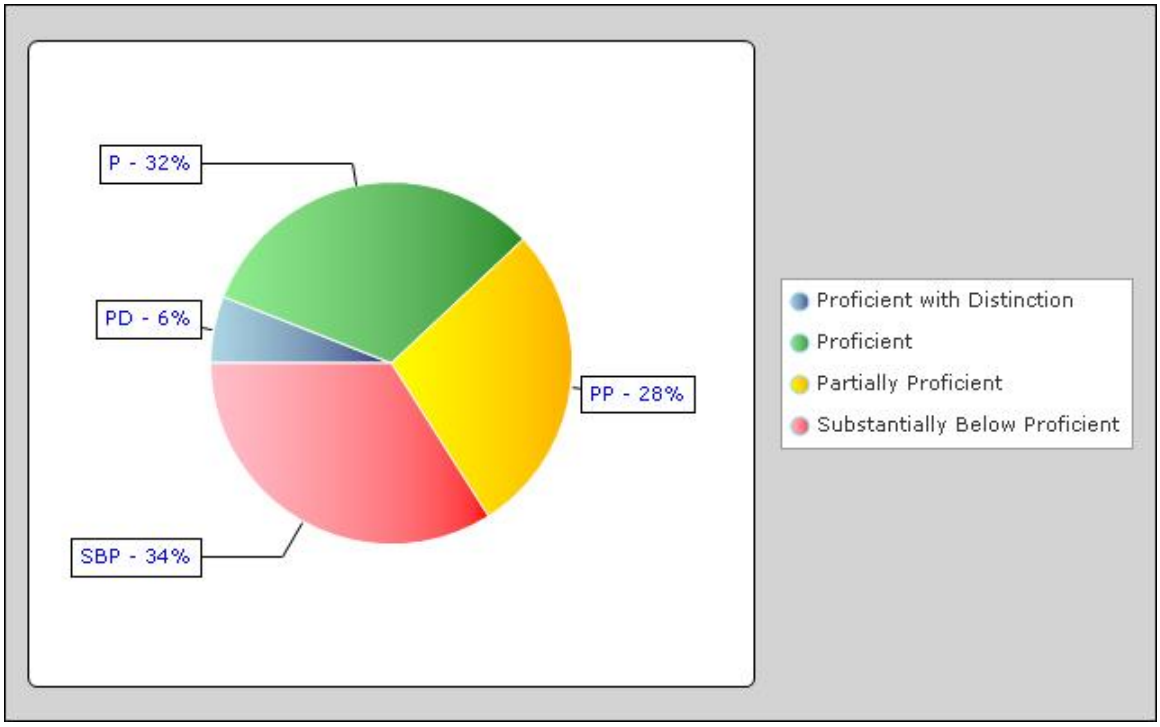
District: Demonstration District A

School: Demonstration School 1

Grade: 11

Date: 6/14/2011 10:49:54 AM

Mathematics



Achievement Level	Count	Percentage %*
Proficient with Distinction	3	6
Proficient	16	32
Partially Proficient	14	28
Substantially Below Proficient	17	34

*Percentages may not total exactly 100% due to applied rounding.

Achievement Level Summary

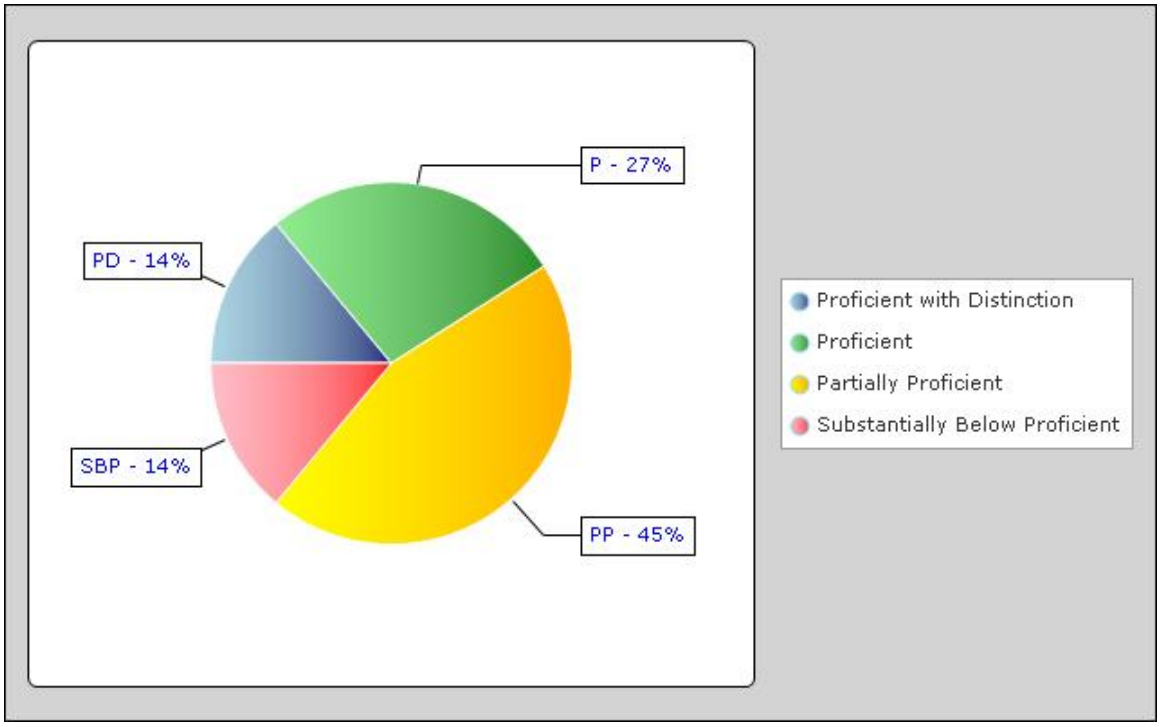
District: Demonstration District A

School: Demonstration School 1

Grade: 05

Date: 6/14/2011 10:49:35 AM

Writing



Achievement Level	Count	Percentage %*
Proficient with Distinction	3	14
Proficient	6	27
Partially Proficient	10	45
Substantially Below Proficient	3	14

*Percentages may not total exactly 100% due to applied rounding.

Reading Released Items Summary Data

District: Demonstration District A
School: Demonstration School 1
Grade: 11
Date: 6/14/2011 10:50:40 AM

Multiple Choice

Released Item	Content Strand	GE Code	Correct (#)	A (#)	B (#)	C (#)	D (#)	IR (#)	Correct Response
1	WV	10-2	35	6	1	8	35	0	D
2	WV	10-3	36	3	0	36	11	0	C
3	WV	10-2	46	1	46	3	0	0	B
4	LA	10-5	37	5	6	37	1	1	C
5	LA	10-5	36	8	2	36	3	1	C
6	LA	10-5	37	0	37	3	10	0	B
8	WV	10-3	39	7	39	2	2	0	B
9	II	10-7	31	12	31	0	7	0	B
10	II	10-7	26	13	26	7	4	0	B
11	II	10-7	27	12	5	6	27	0	D
13	II	10-7	34	1	6	34	8	1	C
14	II	10-7	28	28	3	5	12	2	A
15	WV	10-2	27	15	2	27	5	1	C
16	IA	10-8	22	1	22	17	9	1	B

Open Response

Released Item	Content Strand	GE Code	Point Value	Average Score
7	LI	10-4	4	2.0
12	IA	10-8	4	1.9
17	IA	10-8	4	1.7

Reading Released Items Summary Data

District: Demonstration District A
School: Demonstration School 1
Grade: 11
Date: 6/14/2011 10:50:57 AM

Multiple Choice

Released Item	Content Strand	GE Code	Correct (%)	A (%)	B (%)	C (%)	D (%)	IR (%)	Correct Response
1	WV	10-2	70	12	2	16	70	0	D
2	WV	10-3	72	6	0	72	22	0	C
3	WV	10-2	92	2	92	6	0	0	B
4	LA	10-5	74	10	12	74	2	2	C
5	LA	10-5	72	16	4	72	6	2	C
6	LA	10-5	74	0	74	6	20	0	B
8	WV	10-3	78	14	78	4	4	0	B
9	II	10-7	62	24	62	0	14	0	B
10	II	10-7	52	26	52	14	8	0	B
11	II	10-7	54	24	10	12	54	0	D
13	II	10-7	68	2	12	68	16	2	C
14	II	10-7	56	56	6	10	24	4	A
15	WV	10-2	54	30	4	54	10	2	C
16	IA	10-8	44	2	44	34	18	2	B

Open Response

Released Item	Content Strand	GE Code	Point Value	Average Score
7	LI	10-4	4	2.0
12	IA	10-8	4	1.9
17	IA	10-8	4	1.7

Mathematics Released Items Summary Data

District: Demonstration District A
School: Demonstration School 1
Grade: 08
Date: 6/14/2011 10:52:17 AM

Multiple Choice

Released Item	Content Strand	GE Code	Correct (#)	A (#)	B (#)	C (#)	D (#)	IR (#)	Correct Response
1	GM	7-1	17	4	7	17	2	0	C
2	GM	7-2	14	14	6	3	6	1	A
3	GM	7-5	15	15	8	3	4	0	A
4	GM	7-6	12	9	4	5	12	0	D
5	FA	7-1	21	3	21	6	0	0	B
6	FA	7-2	22	5	2	22	1	0	C
7	FA	7-3	12	6	8	12	3	1	C
8	FA	7-4	15	15	8	4	3	0	A
9	FA	7-4	15	6	5	4	15	0	D
10	DP	7-1	9	10	8	3	9	0	D

Open Response

Released Item	Content Strand	GE Code	Point Value	Average Score
11	GM	7-4	1	0.2
12	FA	7-2	1	0.7
13	FA	7-1	2	0.6
14	DP	7-3	2	0.7
15	NO	7-4	4	1.7

Mathematics Released Items Summary Data

District: Demonstration District A
School: Demonstration School 1
Grade: 08
Date: 6/14/2011 10:51:24 AM

Multiple Choice

Released Item	Content Strand	GE Code	Correct (%)	A (%)	B (%)	C (%)	D (%)	IR (%)	Correct Response
1	GM	7-1	57	13	23	57	7	0	C
2	GM	7-2	47	47	20	10	20	3	A
3	GM	7-5	50	50	27	10	13	0	A
4	GM	7-6	40	30	13	17	40	0	D
5	FA	7-1	70	10	70	20	0	0	B
6	FA	7-2	73	17	7	73	3	0	C
7	FA	7-3	40	20	27	40	10	3	C
8	FA	7-4	50	50	27	13	10	0	A
9	FA	7-4	50	20	17	13	50	0	D
10	DP	7-1	30	33	27	10	30	0	D

Open Response

Released Item	Content Strand	GE Code	Point Value	Average Score
11	GM	7-4	1	0.2
12	FA	7-2	1	0.7
13	FA	7-1	2	0.6
14	DP	7-3	2	0.7
15	NO	7-4	4	1.7

Writing Released Items Summary Data

District: Demonstration District A
School: Demonstration School 1
Grade: 08
Date: 6/14/2011 10:56:24 AM

Multiple Choice

Released Item	Content Strand	GE Code	Correct (#)	A (#)	B (#)	C (#)	D (#)	IR (#)	Correct Response
1	SC	7-9	26	2	0	0	26	1	D
2	SC	7-9	25	0	25	1	2	1	B
3	SC	7-9	19	19	3	5	1	1	A
4	SC	7-9	19	6	1	19	2	1	C
5	SC	7-9	23	4	23	1	0	1	B
6	SC	7-1	13	3	5	7	13	1	D
7	SC	7-9	18	18	4	1	5	1	A
8	SC	7-9	17	4	2	17	5	1	C
9	SC	7-9	12	4	12	8	4	1	B
10	SC	7-9	11	7	6	11	4	1	C

Open Response

Released Item	Content Strand	GE Code	Point Value	Average Score
11	NW	7-5	4	3.0
12	LR	7-2	4	2.4
13	PW	7-8	4	2.4
14	RW	M	12	6.0

Writing Released Items Summary Data

District: Demonstration District A
School: Demonstration School 1
Grade: 08
Date: 6/14/2011 10:56:41 AM

Multiple Choice

Released Item	Content Strand	GE Code	Correct (%)	A (%)	B (%)	C (%)	D (%)	IR (%)	Correct Response
1	SC	7-9	90	7	0	0	90	3	D
2	SC	7-9	86	0	86	3	7	3	B
3	SC	7-9	66	66	10	17	3	3	A
4	SC	7-9	66	21	3	66	7	3	C
5	SC	7-9	79	14	79	3	0	3	B
6	SC	7-1	45	10	17	24	45	3	D
7	SC	7-9	62	62	14	3	17	3	A
8	SC	7-9	59	14	7	59	17	3	C
9	SC	7-9	41	14	41	28	14	3	B
10	SC	7-9	38	24	21	38	14	3	C

Open Response

Released Item	Content Strand	GE Code	Point Value	Average Score
11	NW	7-5	4	3.0
12	LR	7-2	4	2.4
13	PW	7-8	4	2.4
14	RW	M	12	6.0

CONFIDENTIAL

Student Name
Sawyer Barnes

Longitudinal Data Report

Year	Enrolled Grade	School Name	Administration	Test Name	Content Area	Score	Achievement Level
1011	08	Demonstration School 1	NECAP Fall 2010	Grade 08 Mathematics	mat	832	Substantially Below Proficient
1011	08	Demonstration School 1	NECAP Fall 2010	Grade 08 Reading	rea	800	Substantially Below Proficient
1011	08	Demonstration School 1	NECAP Fall 2010	Grade 08 Writing	wri	822	Substantially Below Proficient

Note: This report returns as many years of NECAP data as are available for this student beginning with 08-09.

Appendix U—Decision Rules

Data Analysis and Static Reporting Decision Rules
NECAP
Fall 10-11 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. 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
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	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

		state data.		data.
District	For OUT, OOD, 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	Generate a report for each district with at least one student enrolled using the tested district aggregate denominator. Always report tested year state data.	For OUT, OOD, 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	Generate a report for each district with at least one student enrolled using the teaching district aggregate denominator. Always report tested year state data.
State	Do not include students tested at PRI schools for NH and RI. Include all students for VT 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	n/a.

value of 1, 2, 3 or 4.

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. Fall: The reported years are 0809 and1011 for grades 05 and08 writing. The reported years are 0809, 0910, and 1011 for all other grades and subjects. Spring: The reported years are 0809 , 0910, and 1011.

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 NECAP1011StudentDemographicFileDescription.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	1,2,3, or 4

					Proficient, Proficient, or Proficient with Distinction	
B	Tested Incomplete(%)	✓	✓	✓	Substantially Below Proficient, Partially Proficient, Proficient, or Proficient with Distinction	1,2,3, or 4
C	Tested with Non-Standard Accommodations (%%)	✓	✓	✓	Substantially Below Proficient, Partially Proficient, Proficient, or Proficient with Distinction	1,2,3, or 4
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 grade 11 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 a 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 NECAP0910SInquiryIREF.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 0910 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 1011 Fall grades 05 and 08 writing cumulative total will be calculated using 0809 and 1011 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

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 0 will also receive annotations of FF and 00 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. For students with writing prompt score of 2-12 print at least one, but up to five annotation statements. Grade 11 students with the common writing prompt score of F or 0 will also receive annotations of FF and 00 respectively.

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.

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. 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’.

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

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.
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	NECAP1011FallStudentReleasedItemLayout.xls(one worksheet for grade 11 and one worksheet for 03-08)	NECAP1011FallStateStudentReleasedItem[GS].csv GS=0308 or 11
State Student Raw Data	NECAP1011FallStateStudentRawLayout.xls (one worksheet for each of the 4 unique test designs)	NECAP1011FallStateStudentRaw[GS].csv Gs=0304, 0508, or11
State Student Scored Data	NECAP1011FallStateStudentScoredLayout.xls	NECAP1011FallStateStudentScored.csv
District Student Data	NECAP1011FallStudentReleasedItemLayout.xls(one worksheet for grade 11 and one worksheet for 03-08)	NECAP1011FallTestingDistrictSlice[GR]_[District Code].csv NECAP1011FallTeachingDistrictSlice[GR]_[District Code].csv
School Student Data	NECAP1011FallStudentReleasedItemLayout.xls(one worksheet for grade 11 and one worksheet for 03-08)	NECAP1011FallTestingSchoolSlice[GR]_[District Code][School Code].csv NECAP1011FallTeachingSchoolSlice[GR]_[District Code][School Code]..csv
Common Item Information	NECAP1011FallCommonItemInformationLayout.xls	NECAP1011FallCommonItemInformation.xls
Grade Level Results Report Disaggregated and Historical Data	NECAP1011FallResultsReport DisaggregatedandHistoricaLayout.xls	NECAP1011FallResultsReportTestingDisaggregatedandHistoricalcsv NECAP1011FallResultsReportTeachingDisaggregatedandHistorical.csv
State Standard Deviations and Average Scaled Scores	NECAP1011FallStateStandardDeviationsLayout.xls	NECAP1011FallStateStandardDeviations.csv
Grade Level	NECAP1011FallResultsReportParticipationLayout.xls	NECAP1011FallResultsReportTestingParticipation.csv

Results Report Participation Category Data		
Grade Level Results Report Subtopic Data	NECAP1011FallResultsReport SubtopicLayout.xls	NECAP1011FallResultsReportTestingSubtopic.csv NECAP1011FallResultsReportTeachingSubtopic.csv
Summary Results Data	NECAP1011FallSummaryResultsLayout.xls	NECAP1011FallSummaryResultsTesting.csv NECAP1011FallSummaryResultsTeaching.csv
Released Item Percent Responses Data	NECAP1011FallReleasedItemPercentResponsesLayout.xls	NECAP1011FallReleasedItemPercentResponses.csv
Invalidated Students Original Score	NECAP1011FallStateInvalidatedStudent OriginalScoredLayout.xls	NECAP1011FallStateInvalidatedStudent OriginalScored.csv
Student Questionnaire Summary	NECAP1011FallStudentQuestionnaireSummaryLayout.xls	NECAP1011FallStudentQuestionnaireSummary.csv
TCTA Questionnaire Raw Data	NECAP1011FallTCQuestionnaireRawLayout.xls NECAP1011FallTAQuestionnaireRawLayout.xls	NECAP1011FallTCQuestionnaireRaw.csv NECAP1011FallTAQuestionnaireRaw.csv
TCTA Questionnaire Frequency Distribution	NECAP1011FallTCTAQuestionnaireFreqLayout.xls	NECAP1011FallTCTAQuestionnaireFreq.csv
Scaled Score Lookup	NECAP1011FallScaleScoreLookupLayout.xls	NECAP1011FallScaleScoreLookup.xls NECAP1011FallScaleScoreLookup.csv
Grade 11Writing Score Distribution Results	NECAP1011FallResultsReportGrade11WriScoreDistributionLayout.xls	NECAP1011FallResultsReportGrade11WriScoreDistributionTesting.xls NECAP1011FallResultsReportGrade11WriScoreDistributionTeaching.xls
Grade 11Writing Historical Subtopic Results	NECAP1011FallResultsReportGrade11WriHisoricalSubtopicLayout.xls	NECAP1011FallResultsReportGrade11WriHisoricalSubtopicTesting.xls NECAP1011FallResultsReportGrade11WriHisoricalSubtopicTeaching.xls
Subtopic Average Points Earned (For Project Management)	N/A	NECAP1011FallSubtopicAvgPointsEarned.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	NECAP1011SpringStudentReleasedItemLayout.xls	NECAP1011SpringStateStudentReleasedItem[GR].csv
State Student Raw Data	NECAP1011SpringStateStudentRawLayout.xls	NECAP1011SpringStateStudentRaw[GR].csv
State Student Scored Data	NECAP1011SpringStateStudentScoredLayout.xls	NECAP1011SpringStateStudentScored[GR].csv
District Student Data	NECAP1011SpringStudentReleasedItemLayout.xls	NECAP1011SpringDistrictSlice[GR]_[District Code].csv
School Student Data	NECAP1011SpingStudentReleasedItemLayout.xls	NECAP1011SpringSchoolSlice[GR]_[District Code][School Code].csv
Common Item Information	NECAP1011SpringCommonItemInformationLayout.xls	NECAP1011SpringCommonItemInformation.csv
State Standard Deviations and Average Scaled Scores	NECAP1011SpingStateStandardDeviationsLayout.xls	NECAP1011SpringStateStandardDeviations[GR].csv
Grade Level Results Report Disaggregated and Historical Data	NECAP1011SpringResultsReport DisaggregatedandHistoricaLayout.xls	NECAP1011SpringResultsReport DisaggregatedandHistorical[GR].csv
Grade Level Results Report Participation Category Data	NECAP1011SpringResultsReport ParticipationLayout.xls	NECAP1011SpringResultsReport Participation[GR].csv
Grade Level Results Report Subtopic Data	NECAP1011SpringResultsReport SubtopicLayout.xls	NECAP1011SpringResultsReport Subtopic[GR].csv
Summary Results Data	NECAP1011SpringSummaryResultsLayout.xls	NECAP1011SpringSummaryResults.csv
Released Item Percent Responses Data	NECAP1011SpringReleasedItemPercentResponsesLayout.xls	NECAP1011SpringReleasedItemPercentResponses.csv
Invalidated Students Original Score	NECAP1011SpringStateInvalidatedStudent OriginalScoredLayout.xls	NECAP1011SpringStateInvalidatedStudent OriginalScored.csv
Student Questionnaire Summary	NECAP1011SpringStudentQuestionnaireSummaryLayout.xls	NECAP1011SpringStudentQuestionnaireSummary.csv
TCTA Questionnaire Raw Data	NECAP1011SpringTCQuestionnaireRawLayout.xls NECAP1011SpringTAQuestionnaireRawLayout.xls	NECAP1011SpringTCQuestionnaireRaw.csv NECAP1011SpringTAQuestionnaireRaw.csv
TCTA Questionnaire Frequency	NECAP1011SpringTCTAQuestionnaireFreqLayout.xls	NECAP1011SpringTCTAQuestionnaireFreq.csv

Distribution		
Scaled Score Lookup	NECAP1011SpringScaleScoreLookupLayout.xls	NECAP1011SpringScaleScoreLookup.xls NECAP1011SpringScaleScoreLookup.csv
Subtopic Average Points Earned (For Project Management)	N/A	NECAP1011SpringSubtopicAvgPointsEarned.xls
Item Stats for Inquiry Task Items (For Program Management)	N/A	NECAP1011SpringInquiryItemStats.csv
Memo Shipping Files (For Program Management)	N/A	TBD

VII. Addenda

A. The following rule should replace the annotations rule in section III.K on page 13 and in section IV.A.5 on page 14.

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.

B. The following rule should replace the annotations rule in section III.D.5 on page 10.

1. 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.