OKLAHOMA STATE DEPARTMENT OF EDUCATION

Janet Barresi, State Superintendent of Public Instruction

Oklahoma School Testing Program
Oklahoma Core Curriculum Tests

End-of-Instruction Assessments
2012–2013 Technical Report

FINAL

Mc Graw Hill Education | CTB

Copyright

Revision History

Version 1.0

Table of Contents

List of Tables

List of Figures

List of Appendices

Acronyms and Abbreviations

2PPC Two Parameter Partial Credit model
3PL Three Parameter Logistic model
ACE Achieving Classroom Excellence
AERA American Educational Research Association
APA American Psychological Association
AYP Adequate Yearly Progress
BR Braille
BTC Building Test Coordinator
C^3 Oklahoma's Core Curriculum, the College, Career and Citizen Ready
CCSSO Council of Chief State School Officers
CE Critical Element
CFA Confirmatory Factory Analysis
CR Constructed-Response
CSEM Conditional Standard Error of Measurement
DIF Differential Item Functioning
DOK Depth of Knowledge
DTC District Test Coordinator
EFA Exploratory Factor Analysis
EHS Electronic Handscoring System
ELL English Language Learners
EOI End-of-Instruction
EQ Equivalent
FP False Positive
FN False Negative
GRT General Research Tape
HOSS Highest Obtainable Scale Score
ICC Item Characteristic Curve
IEP Individualized Education Program
IRT Item Response Theory
LIU Language in Use
LOSS Lowest Obtainable Scale Score

MC Multiple-Choice
MH Mantel-Haenszel
NCES National Center for Education Statistics
NCLB No Child Left Behind
NCME National Council on Measurement in Education
NGA National Governors Association Center
NSLP National School Lunch Program
OAAP Oklahoma Alternate Assessment Program
OAC Oklahoma Administrative Code
OCCT Oklahoma Core Curriculum Tests
OE Open-Ended
OMAAP Oklahoma Modified Alternate Assessment Program
OP Operational
OSTP Oklahoma School Testing Program
PASS Priority Academic Student Skills
RIBs Rater Item Blocks
RT Retest
SAS Statistical Analysis System
SD Standard Deviation
SDE Oklahoma State Department of Education
SEM Standard Error of Measurement
SS Scale Score
TA Test Administrator
TAC Technical Advisory Committee
TCC Test Characteristic Curve
TP Test Proctor
TPM Test Preparation Manual
US DOE United States Department of Education
WP Writing Prompt

Introduction

This report summarizes the research data analyses conducted on the Oklahoma Core Curriculum Tests End-of-Instruction (OCCT EOI) 2013 test administrations and provides data evidences in supporting the test validity and reliability of the tests.

The Oklahoma School Testing Program (OSTP) was established to improve academic achievement for all Oklahoma students, and it also meets the requirements of the No Child Left Behind (NCLB) Act (US DOE, 2002), which was introduced by the Federal Government in 2001. The OSTP is a statewide assessment program that, in an attempt to meet the needs of the students of the state of Oklahoma, encompasses three different assessment types—The Oklahoma Core Curriculum Test (OCCT), intended for regular education students; the Oklahoma Modified Alternate Assessment Program (OMAAP), referred to as the modified test and intended for most students enrolled in an Individualized Education Program (IEP) or a 504 Plan, and English Language Learners (ELL); and the Oklahoma Alternate Assessment Program (OAAP), the portfolio assessment for students with the most severe cognitive disabilities in the IEP programs.

The Oklahoma state tests are used to assess student achievement; target student, classroom, and program improvement; and inform parents of student progress. The administration of the OCCT, OMAAP, and OAAP tests fulfills the NCLB Act and state mandates for testing, mathematics and reading, and the test results are used for federal accountability. The scope and general administration of the OSTP is outlined in state law, 70 O.S. § 1210.505. Rules that govern the specifics of test administration and other details are available under Oklahoma Administrative Code (OAC) 210:10-13.

For the OCCT, Reading and Mathematics tests are administered in Grades 3–8; Science, Social Studies, and Writing tests are given in Grade 5; Geography is given in Grade 7; and Science, U.S. History, and Writing are given in Grade 8. English II, English III, Algebra I, Algebra II, Geometry, Biology I, and U.S. History are given as End-of-Instruction (EOI) tests in high school. For the OMAAP, Reading and Math tests are available in Grades 3-8 and Science is available in Grades 5 and 8. English II, Algebra I, Biology I, and U.S. History are available as EOI tests in high school. All students must take the OCCT for content areas in which a modified assessment is not available. The Department of Special Education oversees the implementation of the OAAP, or Portfolio assessment, which includes all of the Grades 3–8 content areas and EOI for Algebra I, Algebra II, Biology I, Geometry, U.S. History, English II, and English III based on the grade level of the student in question.

This document serves to provide detailed descriptions and evidence of reliability and validity of the OCCT EOI, a component of the Oklahoma assessment system. The validity evidence is reflected in the work done by the Oklahoma State Department of Education (SDE) and CTB/McGraw-Hill in the process of the OCCT development. The validity evidence of OCCT can be found in the development of the *Priority Academic Student Skills (PASS)*, most recently Oklahoma's Core curriculum, the College, Career and Citizen Ready (C^3) content standards, the development of the Oklahoma School Testing Program (OSTP) items and operational test forms, the review of the alignment of the content to the test, the administrations of the test, the machine

scoring and handscoring of student responses, the setting of cut scores, and the psychometric analyses (Barton, 2007).

Because the OCCT results are used as part of the state and federal accountability system, CTB/McGraw-Hill follows the *Standards for Educational and Psychological Testing* (1999) by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME). This technical report presents validity and reliability evidence according to the *Standards*. Attention is also given to requirements from the Standards and Assessments Peer Review Guidance (US DOE, 2004) and the Critical Elements (CE) for Peer Review of State summative tests. The detailed documentation is provided in the following sections of this report.

Section 1 – Overview

The Oklahoma End-of-Instruction (EOI) assessments require that students who complete an area of instruction must also take the corresponding standardized test. Each test has the purpose of measuring each student's knowledge relative to the Oklahoma Core curriculum, the College, Career and Citizen Ready (C^3) Oklahoma's content standards. These tests are part of the Achieving Classroom Excellence (ACE) legislation passed in 2005 and amended in 2006, which outlines the curriculum, the competencies, and the testing requirements for students to receive a high school diploma from the state of Oklahoma. Algebra I, English II, Biology I, and U.S. History were existing tests in the program with Algebra II, Geometry, and English III added as operational tests for the 2007–2008 testing cycle. The Spring 2009 administration was the first administration with graduation requirements attached to them for the incoming freshmen students. In order to graduate with a high school diploma from the State of Oklahoma, these students, as well as future incoming freshmen students, are required to score proficient or above on the standardized test assessments for Algebra I and English II, as well as score proficient or above in two of the following five standardized test assessments: Algebra II, Biology I, English III, Geometry, and U.S. History. Students who fail to earn a proficient score are permitted to retake these tests.

All Oklahoma secondary-level students, enrolled in a regular educational program and completing instruction in Algebra I, Algebra II, Biology I, Geometry, English II, English III, and U.S. History, must take the corresponding OCCT EOI tests. The OCCT EOI tests are administered mainly online, with the exception of the Writing Prompt (WP) in English II and English III, which are administered only in a paper/pencil format. All EOI testing administrations have one Writing Prompt (WP) for English II and English III for 2012–2013. These End-of-Instruction standardized assessment tests are administered in Winter/Trimester, Spring, and Summer including other form variations other than Operational (OP), as Braille (BR), Retest (RT), and Equivalent (EQ) forms.

In the Fall of 2012, CTB/McGraw-Hill was contracted by the Oklahoma SDE to develop, administer, and maintain the OSTP OCCT and OMAAP for ACE EOI and Grades 3–8. This technical report provides objective information regarding technical aspects of the Oklahoma OCCT EOI assessments by specifying the technical details of the work accomplished from Summer 2012 (developed, administered, and processed by Pearson) through the end of Spring 2013 on these tests. This volume is intended to be one source of information for Oklahoma K-12 educational stakeholders (including testing coordinators, educators, parents, and other interested citizens) about the development, implementation, scoring, and technical attributes of the Oklahoma OCCT EOI assessments.

Other sources of information regarding the OSTP-ACE EOI tests include the administration manual *OSTP 2012–2013 Test Preparation Manual* found at http://ok.gov/sde/sites/ok.gov.sde/files/TPM_EOI_w12OK_Final.pdf; interpretation manuals, implementation materials, and training materials for administrators, schools, and teachers found at http://www.ok.gov/sde/test-support-teachers-and-administrators; and guides for teachers, students, and parents found at http://ok.gov/sde/assessment-administrator-resources-administrators.

The Summer 2012 OCCT EOI assessments for Algebra I, Algebra II, Biology I, Geometry, English II, English III, and U.S. History were developed by Pearson in collaboration with the Oklahoma SDE and were administered by the SDE. The Winter/Trimester 2012-13 and the Spring 2013 OCCT EOI for Algebra I, Algebra II, Biology I, Geometry, English II, English III, and U.S. History assessments were developed by CTB/McGraw-Hill in collaboration with the SDE and were administered by the SDE.

Section 1.1 – Purpose

This report includes data and analysis results on the operational forms in the Summer 2012, Winter/Trimester 2012–13, and Spring 2013 administrations. A description of the Oklahoma content standards is provided in **Section 1.2. – *Oklahoma C^3* Content Standards**. All operational and field test items for the OCCT EOI Winter/Trimester 2012-13 and Spring 2013 were subjected to cycles of reviews by the SDE and CTB/McGraw-Hill (Summer 2012 by Pearson and the SDE). The item development and alignment process and test development, is detailed in **Section 2 – Item and Test Development**. The test administrations processes can be found in **Section 3 – Administration**. Discussion of the operational population and the research samples utilized in the analysis is found in **Section 5 – Sampling Plan and Field Test Design**. Note that relevant information from the Summer 2012 administration is occasionally shown in these sections but is not the subject of CTB/McGraw-Hill's analysis in this report.

The Summer 2012 OCCT EOI scores were based on a pre-equating design for all content areas. The Winter/Trimester 2012–13 OCCT EOI scores were mostly based on a pre-equating design, where full post-equating analyses were only conducted for English II and English III to assure comparability and stability of the pre- and post-equating results. The Spring 2013 OCCT EOI scores were based on a pre-equating design for Algebra I, Algebra II, and Geometry, where full post-equating analyses were conducted for Biology I, English II, English III, and U.S. History. The Winter/Trimester 2012–13 and Spring OCCT EOI operational and field test items were analyzed and processed separately. A complete description of the operational and field test item analyses and the calibration/scaling and equating analyses is found in **Section 6 – Methods** and **Section 7 – Results**. A summary of reliability and validity for different levels of analyses is found in **Section 8 – Summary of Reliability and Validity**.

Section 1.2 – *Oklahoma C^3* Content Standards

CTB/McGraw-Hill developed the Winter/Trimester 2012–13 and the Spring 2013 Oklahoma OCCT EOI assessments to measure the *Oklahoma C^3* content standards, which are shown in Table 1.1. The objectives associated with the content and/or process standards tested are provided in Appendix A.

Table 1.1 Oklahoma Content Standards by Subject

Algebra I	
Standard 1.	Number Sense and Algebraic Operations
Standard 2.	Relations and Functions
Standard 3.	Data Analysis, Probability & Statistics
Algebra II	
Standard 1.	Number Sense and Algebraic Operations
Standard 2.	Relations and Functions
Standard 3.	Data Analysis, Probability & Statistics
Geometry	
Standard 1.	Logical Reasoning
Standard 2.	Properties of 2-Dimensional Figures
Standard 3.	Triangles and Trigonometric Ratios
Standard 4.	Properties of 3-Dimensional Figures
Standard 5.	Coordinate Geometry
Biology I	
Process/Inquiry Standards and Objectives:	
Process 1.	Observe and Measure
Process 2.	Classify
Process 3.	Experiment
Process 4.	Interpret and Communicate
Process 5.	Model
Content Standards and Objectives:	
Standard 1.	The Cell
Standard 2.	The Molecular Basis of Heredity
Standard 3.	Biological Diversity
Standard 4.	The Interdependence of Organisms
Standard 5.	Matter/Energy/Organization in Living Systems
English II	
Reading/Literature:	
Standard 1.	Vocabulary
Standard 2.	Comprehension
Standard 3.	Literature
Standard 4.	Research and Information
Writing/Grammar/Usage and Mechanics:	
Standard 1/2.	Writing (Writing Prompt)
Standard 3.	Grammar/Usage and Mechanics

Table 1.1 Oklahoma Content Standards by Subject (*continued*)

English III	
Reading/Literature:	
Standard 1.	Vocabulary
Standard 2.	Comprehension
Standard 3.	Literature
Standard 4.	Research and Information
Writing/Grammar/Usage and Mechanics:	
Standard 1/2.	Writing (Writing Prompt)
Standard 3.	Grammar/Usage and Mechanics
U.S. History	
Standard 1.	Post-Reconstruction to the Progressive Era, 1878-1900
Standard 2.	Expanding Role of the United States in International Affairs
Standard 3.	Cycles of Economic Boom and Bust in the 1920s and 1930s
Standard 4.	Role of the U.S. in International Affairs and World War II, 1933-1946
Standard 5.	U.S. Foreign and Domestic Policies during the Cold War, 1945-1975

Section 2 – Item and Test Development

In the Summer 2012 and Winter/Trimester 2012–2013 administrations, there was one operational form with embedded sets of field test items for the tests administered for Algebra I, Algebra II, Geometry, English II, English III, Biology I, and U.S. History. In the Spring 2013 administration, there were two Core Operational forms (A, B) and each form was embedded with sets of field test items to add to the item pool. This resulted in 8 field test forms for Algebra I, Algebra II, English II, English III, and Geometry; 7 field test forms for Biology I; and 10 field test forms for U.S. History. For each administration, a Braille form, an Equivalent form, and a Retest form are produced.

The Braille form is usually a mirror of the operational form. The Equivalent, designated as a breach form, and the Retest forms usually are a reproduction of past administration forms, except for the open-ended items or Writing Prompts. A student could receive an Equivalent form for various reasons, such as becoming ill during test administration or experiencing any kind of security breach. The Oklahoma State Department of Education Office of Accountability and Assessments determines eligibility for an Equivalent form on a case-by-case basis.

Test Design

For Summer 2012 and Winter/Trimester 2012–13, CTB/McGraw-Hill Content Development selected repurposed forms approved by the SDE for operational use. CTB Research analyzed the selected forms and provided feedback to CTB Content Development regarding item position. Adjustments were made by Content Development based on Research feedback. For Spring 2013, Content Development selected items from the available item pools that had been field tested previously and approved by the SDE staff for usage on operational assessments. Field test items were selected from items approved by the SDE and Oklahoma teachers. CTB Research analyzed the selected items and provided feedback to Content Development regarding the best set of items for the Spring 2013 operational form.

Specifically, Research reviewed the forms for comparability of blueprints; total test information; cut score test information; standard errors of measurement; raw score to scale score stability (particularly at the cut scores); item locations (difficulty parameters) for all items within a form and item information levels; test characteristic curves (TCCs) for each form selected compared to a reference form, the Spring 2012 operational form.

Table 2.1 to Table 2.3 provide overviews of the number of operational and field test items that composed the Summer 2012, Winter/Trimester 2012–13, and Spring 2013 OCCT EOI assessments. The Summer 2012 and Winter/Trimester 2012–13 tests were comprised of one core operationally-scored form for each subject. Field test items were embedded in the operational test for all content areas. The Spring 2013 test was comprised of two Core Operational forms (A, B), for each subject. Field test items were embedded in the operational test forms for all content areas to build the item bank for future use. The forms in the Spring 2013 assessments were randomly assigned within classrooms to obtain randomly-equivalent samples of examinees for the field test items. Although most items were unique to each form, approximately 17 items were

common across the core forms. The number of common linking items per subject is presented in Table 2.4.

Table 2.1. Configuration of the OCCT EOI Tests for Summer 2012

| | | Item Counts (Per Form) | | | Maximum Possible Points on Test Items (Per Form) | | | |
| | | | | | OP | | FT | |
Subject	Forms	OP	FT	Test	MC	OE	MC	OE
Algebra I	1	55	10	65	55		10	
Algebra II	1	55	10	65	55		10	
Biology I	1	60	20	80	60		20	
English II	1	61	20	81	60	6	20	
English III	1	62	20	82	61	10	20	
Geometry	1	55	10	65	55		10	
U.S. History	1	60	20	80	60		20	

Note: OP = Operational; FT = Field Test; MC = Multiple-Choice; OE = Open-Ended

Table 2.2. Configuration of the OCCT EOI Tests for Winter/Trimester 2012-13

| | | Item Counts (Per Form) | | | Maximum Possible Points on Test Items (Per Form) | | | |
| | | | | | OP | | FT | |
Subject	Forms	OP	FT	Test	MC	OE	MC	OE
Algebra I	1	55	10	65	55		10	
Algebra II	1	55	10	65	55		10	
Biology I	1	60	15	75	60		15	
English II	1	61	15	76	61	6	15	
English III	1	62	15	77	62	10	15	
Geometry	1	55	10	65	55		10	
U.S. History	1	60	10	70	60		10	

Note: OP = Operational; FT = Field Test; MC = Multiple-Choice; OE = Open-Ended

Table 2.3. Configuration of the OCCT EOI Tests for Spring 2013

| Subject | Forms | Item Counts (Per Form) | | | Maximum Possible Points on Test Items (Per Form) | | | |
| | | | | | OP | | FT | |
		OP	FT	Test	MC	OE	MC	OE
Algebra I	8	55	10	65	55		10	
Algebra II	8	55	10	65	55		10	
Biology I	7	60	15	75	60		15	
English II	8	61	15	76	60	6	15	
English III	8	62**	15	78	61	10	15	
Geometry	8	55	10	65	55		10	
U.S. History	10	60	15	75	60		15	

Note: OP = Operational; FT = Field Test; MC = Multiple-Choice; OE = Open-Ended;
** = English III has one suppressed item, reducing the total possible points to 71.

Table 2.4. Number of Common Linking Items per Subject for Spring 2013

Subject	No. of CL Items	Total No. of Items*
Algebra I	17	38
Algebra II	17	38
Biology I	16	44
English II	19	41****
English III	18**	43
Geometry	17	38
U.S. History	17	43***

Note: No. = Number; CL = common linking;
* = Number of unique operational items per form.
** = English III has one suppressed common item.
*** = U.S. History Form B has one suppressed unique item.
**** = English II has one suppressed unique item.

Section 2.1 – Aligning Test to *Oklahoma C³* Content Standards

In general, alignment is a process that provides experts the opportunity to make item-level judgments about the grade level, standards, and indicators to which items should be aligned. There are multiple points in the alignment process at which assessment items are either created or evaluated for alignment to content. Most tests, particularly high-stakes, large-scale assessments, are built via rigorous and well-researched methodologies. They are guided by well-defined content and by the boundaries within the content that can be reasonably assessed in a testing environment. Such guidance is typically in the form of item specifications and test blueprints. The item specifications help define which content standards can be assessed by a test (and which content standards are better assessed in the classroom), the breadth and depth of the content that may be limited for the test, and the format and types of items appropriate for the content being assessed (e.g., multiple-choice or open-ended item). The test blueprint defines the proportion of the content to be covered on the test that best reflects the proportional importance and coverage of the standards in the classroom.

A list of the assessable standards for each subject is provided in Table 2.5 for Algebra I, Algebra II, Geometry, English II, English III, Biology I, and U.S. History. In addition to the test blueprints provided by the SDE, Table 2.6 describes four criteria for test alignment with the Oklahoma C^3 Content Standards and objectives.

Table 2.5. Testable Standards for OCCT EOI

Algebra I	
Standard 1.	Number Sense and Algebraic Operations
Standard 2.	Relations and Functions
Standard 3.	Data Analysis, Probability & Statistics
Algebra II	
Standard 1.	Number Sense and Algebraic Operations
Standard 2.	Relations and Functions
Standard 3.	Data Analysis, Probability & Statistics
Geometry	
Standard 1.	Logical Reasoning
Standard 2.	Properties of 2-Dimensional Figures
Standard 3.	Triangles and Trigonometric Ratios
Standard 4.	Properties of 3-Dimensional Figures
Standard 5.	Coordinate Geometry
English II	
Reading/Literature	
Standard 1.	Vocabulary
Standard 2.	Comprehension
Standard 3.	Literature
Standard 4.	Research and Information
Writing/Grammar/Usage/Mechanics	
Standard 1. and 2.	Writing
Standard 3.	Grammar/Usage and Mechanics
English III	
Reading/Literature	
Standard 1.	Vocabulary
Standard 2.	Comprehension
Standard 3.	Literature
Standard 4.	Research and Information
Writing/Grammar/Usage/Mechanics	
Standard 1. and 2.	Writing
Standard 3.	Grammar/Usage and Mechanics

Table 2.5. Testable Standards for OCCT EOI *(continued)*

Biology I	
Standard 1.	The Cell
Standard 2	The Molecular Basis of Heredity
Standard 3.	Biological Diversity
Standard 4.	The Interdependence of Organisms
Standard 5.	Matter/Energy/Organization in Living Systems
Process 1.	Observe and Measure
Process 2.	Classify
Process 3.	Experimental Design
Process 4.	Interpret and Communicate
Process 5.	Model

U.S. History	
Standard 1.	Transformation of the United States from Post-Reconstruction to the Progressive Era, 1878-1900
Standard 2	Expanding Role of the United States in International Affairs
Standard 3.	Cycles of Economic Boom and Bust in the 1920s and 1930s
Standard 4.	Role of the U.S. in International Affairs and World War II, 1933-1946
Standard 5.	U.S. Foreign and Domestic Policies during the Cold War, 1945-1975
Standard 6.	U.S. Foreign and Domestic Policies, 1976 to the Present

Table 2.6. Criteria for Aligning the Test with *Oklahoma C^3* Standards and Objectives

1. Categorical Concurrence	The test is constructed so that there are at least six items measuring each *Oklahoma C^3* standard with the content category consistent with the related standard. The number of items, six, is based on estimating the number of items that could produce a reasonably reliable estimate of a student's mastery of the content measured.
2. Range-of-Knowledge	The test is constructed so that at least 50% of the objectives for an *Oklahoma C^3* standard have at least one corresponding assessment item.
3. Balance-of-Representation	The test is constructed according to the alignment blueprint, which reflects the degree of representation given on the test to each *Oklahoma C^3* standard and objective in terms of the percent of total test items measuring each standard and the number of test items measuring each objective.
4. Source-of-Challenge	Each test item is constructed in such a way that the major cognitive demand comes directly from the targeted *Oklahoma C^3* skill or concept being assessed, not from specialized knowledge or cultural background that the test-taker may bring to the testing situation.

Data review represents a critical step in the test development cycle. At the data review meeting, the SDE and CTB/McGraw-Hill staff had the opportunity to review actual student performance on the newly-developed and field tested Multiple-Choice items across the seven subjects based on the Winter/Trimester 2012–13 and Spring 2013 field test administrations. The data review focused on the content validity, curricular alignment, and statistical functioning of field tested items prior to selection for operational test forms. The field test results used in the data review provided evidence that the items were designed to yield valid results and were accessible for use by the widest possible range of students. The review of student performance should provide evidence regarding the fulfillment of requirement 200.2(b)(2)of NCLB. The purpose of the review meeting was to ensure that psychometrically-sound, fair, and aligned items are used in the construction of the ACE EOI assessments and entered into the respective item banks. CTB/McGraw-Hill provided technical and psychometric expertise and a clear explanation about the items' content, the field test process, the scoring process, and the resulting field test data to ensure the success of these meetings and the defensibility of the program.

Data review meetings were a collaborative effort between the SDE and CTB/McGraw-Hill. The SDE administrators and content specialists attended the training facilitated by CTB/McGraw-Hill content specialists and research scientists on best practices involved in interpreting and reviewing the field test data. Meeting materials included a document explaining the flagging criteria, a document containing flagged items, and the item images. CTB/McGraw-Hill discussed with the SDE the analyses performed and the criteria for flagging the items. Each of the flagged items was then reviewed, and the decision was made to accept the item, accept the item for future re-field testing with revisions, or reject the item. Review of the data included presentation of p-value, point-biserial correlation, point-biserial correlation by response option, response distributions, mean overall score by response option, and indications of item differential item functioning (DIF) and item response theory (IRT) misfit. Items failing to meet the requirements of sound technical data were carefully considered for rejection by the review panel, thereby enhancing the reliability and improving the validity of the items remaining in the bank for future use. Although the panel used the data as a tool to inform their judgments, the panel (and not the data alone) made the final determination as to the appropriateness or fairness of the assessment items. The flagging criteria for the ACE EOI assessments are as follows:

- p-value $< .25$ or $> .90$
- point-biserial correlation $< .15$
- distractor point-biserial correlation $> .05$
- Differential Item Functioning (DIF): test item biases for subgroups
- IRT misfit as flagged by the Q_1 index (see **Section – 6.3 Calibration & Item Fit**)

Bias and Sensitivity

One aspect of the data review meetings was to assess potential bias based on DIF results and item content. Although bias in the items had been deflected by writer training and review processes, there is always the potential for bias to be detected through statistical analysis. This step in the development cycle is essential because the SDE and CTB/McGraw-Hill seek to avoid inclusion of items biased in any manner against a group, because these items may lead to inequitable test results. As described earlier, all field test items were analyzed statistically for DIF using the field test data. A CTB/McGraw-Hill research scientist explained the significance,

in terms of level, and the direction of the DIF flags. The data review panel reviewed the item content, the percentage of students selecting each response option, and the point-biserial correlation for each response option by gender and ethnicity for all items flagged for DIF. The data review panel was then asked if there was context (e.g., cultural barriers) or language in an item that might result in bias and provide an explanation for the existence of the statistical DIF flag.

Once items were written, they were reviewed to assure the items were appropriate for and aligned to the grade level, the *Oklahoma C³* Standard and objective, and the DOK intended. The items were also reviewed to assure they were accurate, written at an appropriate reading level for the grade, written at an appropriate level of difficulty, and did not contain sensitive or potentially biased issues.

Statistical bias analyses were performed as part of the development, review, and fairness efforts. Field test items were analyzed for statistical bias utilizing the Mantel-Haenszel method (MH; Mantel & Haenszel, 1959; Holland & Thayer, 1988; Michaelides, 2008). The results for Winter/Trimester 2012–13 and Spring 2013 are found in **Section 7** – **Results** of this report.

Section 2.2 – Item Pool Development and Selection

The source of the operational items included a pool of previously field tested or operationally-administered items ranging from the Spring 2005 through the Spring 2012 administrations for Algebra I, Biology I, English II, and U.S. History and from the census Spring 2007 field test through the Spring 2011 embedded field test for Algebra II, Geometry, and English III. The items were calibrated live using data from the operational administrations to estimate the items' parameters.

The ACE EOI tests for the Winter/Trimester 2012-13 and Spring 2013 cycle were built by including previously field tested and operational items. Content experts targeted the percentage of items measuring various Depth of Knowledge (DOK) levels for assembling the tests. Table 2.7 provides the DOK level percentages for the Summer 2012, Winter/Trimester 2012–13 and Spring 2013 operational assessments. During test construction, every effort was made to construct test forms that met the target percentages as closely as possible.

Blueprints

Text and Item Development Process

To ensure content validity of the Oklahoma OCCT EOI tests, CTB/McGraw-Hill content experts carefully studied the *Oklahoma C³* Standards and/or the *Priority Academic Student Skills (PASS)* content standards. They worked with Oklahoma content area specialists, teachers, and assessment experts to gather a pool of existing items that measure Oklahoma's Assessment Frameworks (i.e., *Oklahoma C³*) for each subject. Once the need for field test items was determined, based on the items' availability for future test construction, a pool of items was developed to measure *Oklahoma C³* in each subject. These items were developed under universal design guidelines set by the SDE and carefully reviewed and discussed by Content and

Bias/Sensitivity Review Committees. These committees, comprised of Oklahoma teachers and SDE staff, evaluated items' content for validity, plain language, and the quality and appropriateness. The committees' recommendations were used to select and/or revise items from the item pool used to construct the field test portions of the Winter/Trimester 2012–13 and the Spring 2013 assessments.

Item selection and form development for the Spring 2013 cycle was completed as a collaborative effort between the SDE staff and CTB/McGraw-Hill Content Development and psychometricians (Research). The primary criterion for the selection of items was to meet the content specifications represented by test blueprints and statistical guidelines. Within the limits set by these requirements, such as classical and IRT statistics, described in **Section 6 – Methods**, editors selected items with the best content-relevant and statistical characteristics.

The OCCT EOI Operational tests for the Winter/Trimester 2012–13 and the Spring 2013 cycle were built by including previously field tested and operational items. Content experts also targeted the percentage of items measuring various DOK levels when assembling the tests.

Table 2.7 provides the DOK level percentages for the Summer 2012, Winter/Trimester 2012–13, and Spring 2013 operational assessments.

Table 2.7. Percentage of Items by Depth of Knowledge Levels

Test Session	DOK Level	Target DOK %	Actual %			
			Algebra I	Algebra II	Biology I[1]	English II
Summer	1	10-15	15	13	17	15
2012	2	60-70	65	65	63	60
	3/4	15-25	20	22	20	25
Winter/	1	10-15	13	15	12	7
Trimester	2	60-70	67	69	45	70
2012-13	3/4	15-25	20	16	43	23
Spring 2013	1	10-15	16	16	13	10
Core A	2	60-70	69	62	50	74
	3/4	15-25	15	24	37	16
Spring 2013	1	10-15	15	16	13	16
Core B	2	60-70	71	60	53	64*
	3/4	15-25	16	24	33	18

Note: For Biology I, the target DOK percentages are 10–15 for DOK level 1, 55–65 for DOK level 2, and 25–35 for DOK level 3 for the school year of 2012–13.
* = English II one DOK level 2 item was suppressed in Form B.

Table 2.7. Percentage of Items by Depth of Knowledge Levels (cont.)

Test Session	DOK Level	Target DOK %	Actual %		
			English III	Geometry	U.S. History
Summer	1	10-15	8	18	10
2012	2	60-70	76	64	67
	3/4	15-25	16	18	23
Winter/	1	10-15	13	15	7
Trimester	2	60-70	68	69	68
2012-13	3/4	15-25	19	16	25
Spring 2013	1	10-15	6	16	7
Core A	2	60-70	76*	64	73
	3/4	15-25	16	20	20
Spring 2013	1	10-15	5	16	5
Core B	2	60-70	73*	62	75**
	3/4	15-25	21	22	18

Note: For Biology I, the target DOK percentages are 10–15 for DOK level 1, 55–65 for DOK level 2, and 25–35 for DOK level 3 for the school year of 2012–13.
*= English III: one DOK level 2 item was suppressed in Form A and in Form B.
**= U.S. History: one DOK level 2 item was suppressed in Form B.

Section 3 – Administration

To ensure a valid and reliable assessment, the OCCT EOI tests are first constructed in alignment with the *Oklahoma C³* Standards (now called the Oklahoma Academic Standards) by the Oklahoma SDE in collaboration with CTB/McGraw-Hill. The tests are then administered and scored according to sound measurement principles for the purpose of evaluating validity. Additionally, best practices require that the test administering and scoring entities perform their tasks in a consistent manner throughout the state so that all students have a fair and equitable opportunity for a score that reflects their achievement in each subject.

Schools play a key role in administering the OCCT EOI assessments in a manner that is consistent with established procedures, monitoring the fair administration of the assessment, and working with the SDE office to address deviations from established assessment administration best practice procedures. School faculty members play a vital role in the success of OCCT EOI assessments by ensuring fairness in administration of the test.

Section 3.1 – Packaging and Shipping

In order to provide secure and dependable services for the shipping of the OCCT EOI assessment materials, CTB/McGraw-Hill's Transportation Department maintains the quality and security of material distribution and return by hiring reputable carriers that possess the ability to trace shipments. CTB/McGraw-Hill uses all available tracking capabilities to provide status information and early opportunities for corrective action.

Materials are packaged by school and delivered to the district test coordinators. Each shipment to a district contains a shipping document set that includes a packing list for each school's materials.

Materials are packaged using information provided by the test coordinators through the CTB/McGraw-Hill Precode Utility (EOI) or the Oklahoma WAVE system (Grades 3-8). Oklahoma educators also use these systems to provide CTB/McGraw-Hill with the precode information needed to print student barcode labels, which are affixed on answer documents or consumable test books. The bar-coding of all secure materials at the time of production allows for accurate tracking of these materials through the entire packing, delivery, and return process. This allows CTB/McGraw-Hill to inventory all materials throughout the packaging and delivery process.

Section 3.2 – Materials Return

The Test Preparation Manual and Materials Return poster provide clear instructions on how to assemble, box, label, and return testing materials after test administration. CTB/McGraw-Hill utilizes double-column boxes to distribute and collect test materials, and makes additional cartons available for order in order to meet the various return needs of the districts.

Stack cards and paper bands are provided to group and secure used student response booklets for scoring. Color-coded return labels with pre-printed return information are also provided. These

labels facilitate the sorting of each carton and its contents upon receipt at CTB/McGraw-Hill's Data Processing Facility.

Section 3.3 – Materials Discrepancies Process

The scanning process allows CTB/McGraw-Hill to capture multiple choice responses and student writing images. Test security form information is also captured electronically via a secure database. All scorable material discrepancies are captured, investigated by the CTB/McGraw-Hill Oklahoma Help Desk, reported, and the results are subsequently reported to the Oklahoma SDE.

A pre-determined date is set by the SDE and CTB/McGraw-Hill in order to account for any materials that arrive after the scheduled deadline. Late-arriving material is processed up to the agreed-upon date, at which point the Oklahoma SDE must be notified of any late-arriving documents and render a processing decision. Following an initial call campaign to all districts with outstanding secure material, the CTB/McGraw-Hill Oklahoma Program Management team notifies the SDE regarding unresolved material discrepancies presented in a preliminary file. A subsequent call or email campaign may be conducted based on the results of the initial effort. Final missing inventory reports are then provided to the SDE. CTB/McGraw-Hill takes test security seriously and makes every effort to recover missing material.

Section 4 – Scoring

The OCCT EOI Spring 2013 test books included MC items that were machine scored and extended writing prompt items that were scored by trained human or "hand" scorers (raters). The MC items were scanned and scored as correct or incorrect according to predefined answer keys. Items that had multiple marked answers or were blank were treated as incorrect.

The Writing test is one portion of the OCCT EOI English II and English III tests. Each writing response receives two types of scores. First, a series of analytic scores focus on specific writing traits. These traits receive scores of 1 to 4. Next, a composite score is derived by providing a differential weight or percentage to the score in each of the analytic traits and applying a formula to obtain the final Writing score, which ranges from 1 to 6 for English II and from 1 to 10 for English III. Condition codes are used if the student's writing response is unscorable. Students do not receive separate reports for English II and English III Writing; the results are reported with the MC results.

Scoring Rubrics

Scoring rubrics were provided by the Oklahoma SDE. The rubrics focus on five specific writing skills: Ideas and Development; Organization, Unity, and Coherence; Word Choice; Sentences and Paragraphs; and Grammar, Usage, and Mechanics. Each trait is rated from 4 (the highest score) to 1 (the lowest score).

Anchor Papers

The OCCT EOI English II and English III writing prompts underwent field testing by the previous contractor. The SDE provided CTB/McGraw-Hill with approved anchor sets for these field tested items during the contract transition period. The English II OCCT and OCCT Equivalent writing prompts were newly developed versions of the English II Form A and English II Equivalent prompts. Anchor paper candidates were selected by Handscoring staff and submitted to the SDE for review and approval, and used in rater training and scoring of the OCCT writing responses.

Anchor sets for English II and English III writing prompts were presented to raters by trait, with three examples for each of the four score points. The OCCT prompts utilize a four-point analytic rubric for each of five traits.

Section 4.1 – Hand Scoring

Handscoring involves training and qualifying team leaders and raters, monitoring scoring accuracy and production, and ensuring the security of both the test materials and the scoring facilities. An explanation of the training and qualification procedures follows.

Training, Qualification, and Checkset Materials

All raters were trained and qualified in specific rater item blocks (RIBs), each of which consisted of a single writing prompt. Raters and team leaders were trained using the following steps:

- Provide a general introduction to OCCT EOI
- Introduce and review the writing prompts and scoring rubrics
- Review anchor papers and training papers, and answer questions arising from established scores
- Explain scoring strategies, followed by a question-and-answer period
- Administer Qualifying Round 1
- Review Qualifying Round 1 established scores, and answer questions arising from the scores.
- Administer Qualifying Round 2 (if necessary)
- Explain condition codes and sensitive paper procedures
- Explain unscannable image procedures

All raters were trained and qualified using the same procedures and criteria used for the team leaders, who had been trained prior to the training of the raters. The qualification process was conducted through the Online Training System and proctored by Handscoring Supervisors and team leaders. The Online Training System enabled supervisors to determine whether a rater had qualified upon completion of the set. The CTB/McGraw-Hill Handscoring Supervisors proctored the training of the team leaders.

Throughout the course of handscoring, calibration sets of pre-scored papers (checksets/validity sets) were administered daily to each rater to monitor scoring accuracy and to maintain a consistent focus on the established rubrics and guidelines. Checksets were executed via imaging software that provided images in a manner so that the rater did not know when a checkset was being administered.

The CTB/McGraw-Hill Data Monitoring staff ran inter-rater reliability reports throughout live scoring to look for any raters who were struggling and in need of retraining. Retraining involved a one-on-one discussion between the team leader (or Handscoring Supervisor) and the rater, who discussed the scoring concerns as well as the scoring guides and, if necessary, training papers. If the rater's accuracy on checkset scores did not meet the quality standards after this retraining, they were dismissed from the project immediately.

In addition to the checkset process, CTB/McGraw-Hill's handscoring protocol included the use of read-behinds (spot-checks during live scoring). The read-behind was another valuable rater-reliability monitoring technique that allowed a team leader to review a rater's scored documents and provide feedback and counseling as appropriate.

Selection of Handscorers

CTB/McGraw-Hill and Kelly Services, Inc., strive to develop a highly qualified, experienced core of raters so that the integrity of all projects is appropriately maintained.

CTB/McGraw-Hill requires that all content experts, team leaders, and raters possess a bachelor's degree or higher. Kelly Services, Inc., carefully screened all new applicants and required them to produce either a transcript or a copy of the degree. Kelly Services, Inc., also required a one- to

two-hour interview/screening process. Individuals who did not present proper documentation or had less than desirable work records were eliminated during this process. Kelly Services, Inc., verified that 100% of all potential raters met the degree requirement. All experienced raters and team leaders had already successfully completed the screening process.

All potential raters completed a pre-interview activity. For some parts of the pre-interview activity, applicants were shown examples of test responses and were supplied with a scoring guide. In a brief introduction, they became acquainted with the application of a rubric. After the introduction, applicants applied the scoring guide to score the sample responses. Each applicant's scores were used for discussion during the interview process to determine the applicant's trainability as well as an ability to understand and implement the standards set forth in the sample scoring guide.

Kelly Services, Inc., interviewed each applicant and determined the applicant's suitability for a specific content area and grade level. Applicants with strong leadership skills were interviewed further to determine whether they were qualified to be team leaders.

When Kelly Services, Inc., determined that applicants were qualified, they were recommended for employment. All assignments were made according to availability and suitability. Before being hired, all employees were required to read, agree to, and sign a nondisclosure agreement outlining CTB/McGraw-Hill business ethics and security procedures.

Security guards were on-site whenever employees were present in the building. All employees were issued identification badges and required to wear them in plain view at all times. Visitors and employees who presented at the building entrance without their issued ID badges were issued temporary visitors' badges good for that one day only and were required to wear them in plain view. In addition, employees were advised to arrive the following day with their previously-issued ID badges worn in plain view. All employees and visitors were subject to inspection of their personal effects.

Handscoring Process

Writing prompts were evaluated on each of the five analytic traits and in accordance with Oklahoma's rubric. Using CTB/McGraw-Hill's Electronic Handscoring System (EHS), all writing responses were scored independently by two raters. The EHS employed an automated, random distribution of papers for first reads, second reads, and resolution reads across all readers designated to score that item. No student biographical or identifiable information was available to raters; all imaged items were scored as blind reads.

Rater Reliability

Section 8 – Reliability and Validity describes the outcomes of inter-rater percentage of perfect and adjacent agreements. The inter-rater results for the operational writing prompts are presented in Tables 4.1 and 4.2 for English II and English III, respectively.

Table 4.1. Inter-rater Percentage of Perfect and Adjacent Agreement for English II

| Item | Form | Data Point | Score Points | % of Agreement | | | Checkset Average Agreement Percentages |
				Perfect	Adjacent	Perfect + Adjacent	
English II II Writing	A	A	1-4	61.5	36.6	98.1	84.8
		B	1-4	61.0	37.0	98.0	84.9
		C	1-4	60.2	37.8	98.0	84.9
		D	1-4	59.5	38.2	97.7	83.7
		E	1-4	58.5	39.0	97.5	83.7
English II Writing	B	A	1-4	55.1	41.0	96.1	73.5
		B	1-4	55.4	40.8	96.2	74.2
		C	1-4	55.4	41.0	96.4	74.1
		D	1-4	55.6	41.0	96.6	72.6
		E	1-4	55.2	41.2	96.4	73.7
English II EQ Writing		A	1-4	67.4	31.2	98.6	*na*
		B	1-4	64.4	34.8	99.2	*na*
		C	1-4	61.5	37.0	98.5	*na*
		D	1-4	60.0	38.6	98.6	*na*
		E	1-4	60.0	38.6	98.6	*na*

Note: "*na*": Too few documents distributed among multiple rating Supervisors to trigger EHS-generated checkset/validity papers.

Table 4.2. Inter-rater Percentage of Perfect and Adjacent Agreement for English III

Item	Form	Data Point	Score Points	% of Agreement			Checkset Average Agreement Percentages
				Perfect	Adjacent	Perfect + Adjacent	
English III Writing	A	A	1-4	63.5	35.2	98.7	83.1
		B	1-4	63.0	35.6	95.6	83.2
		C	1-4	62.0	36.4	98.4	82.6
		D	1-4	61.4	36.8	98.2	81.8
		E	1-4	60.8	37.4	98.2	82.0
English III Writing	B	A	1-4	61.9	36.2	98.1	79.9
		B	1-4	62.0	36.0	98.0	80.1
		C	1-4	61.6	36.6	98.2	80.7
		D	1-4	61.5	36.4	97.9	80.8
		E	1-4	60.3	37.4	97.7	80.2
English III EQ Writing		A	1-4	48.6	46.0	94.6	*na*
		B	1-4	53.2	42.4	96.6	*na*
		C	1-4	57.7	35.2	92.9	*na*
		D	1-4	57.7	36.0	93.7	*na*
		E	1-4	59.5	34.2	93.7	*na*

Note: *na*= Too few documents distributed among multiple rating Supervisors to trigger EHS-generated checkset/validity papers.

Section 5 – Sampling Plan and Field Test Design

Section 5.1 – Sampling Plan

A sample representative of the population of Oklahoma students was used for the Spring 2013 English II and English III post-equating because final scale scores and performance levels should be reported within two weeks of the closed testing window. Due to the reporting schedule, some students' data were prioritized in the scanning and scoring process and used throughout item level analyses, calibration, and equating. Once the data was available, CTB/McGraw-Hill Research conducted a data integrity check and compared the sample selection to the 2012 population to ensure that the sample was representative. Table 5.1 shows the Spring 2013 calibration samples by form used for post-equating of the English II and English III tests.

Table 5.1. English Calibration Sample for Spring 2013 and Respective Percentage of the Population

English II Forms	Population	Sample N Counts	Percent	English III Forms	Population	Sample N Counts	Percent
AA	9776	5844	60	AA	10096	5809	58
AB	9470	4849	51	AB	9579	5915	62
BA	8911	5528	62	BA	8556	5250	61
BB	8718	4416	51	BB	8554	5440	64

Table 8, in the **Tables** section, provides the proportion of students in the English II and English III samples and within the Spring 2013 population that came from each of the subgroups: gender, ethnicity, special population (ELL, IEP, Section 504, and accommodated), and socio-economic status (SES Low). SES Low is for students who have participated in the National School Lunch Program (NSLP). It is clear from these tables that the sample is also representative of the state's Spring 2013 population, even across most of the subgroups. The differences between the sample and the state tend to be less than +/-5% with a median difference of 0.09 (absolute value).

No sampling decreases or increases were required since the sample received was well representative of the target or expected sample and therefore, representative of the population of students in Oklahoma.

Section 5.2 – Field Test Design

New items are field tested to build up the item bank for future form selections. An embedded field test design was used in which newly developed field test items were embedded throughout the test. The advantage of an embedded field test design is that test-takers do not know where the field test items are located and therefore students' motivation for operational and field test items are the same. Table 5.2 shows the number of field test items for each content area. Ten to fifteen MC field test items per form were placed in common positions across all contents. Algebra I, Algebra II, and Geometry had 10 items each in the four field test forms per form A and B. Biology I had 15 items each in the three field test forms per form A and B. English II and

English III had 15 items each in the 4 field test forms per form combination (AA, AB, BA, BB). English II and III had some common field test items between field test forms. U.S. History had 15 items each in the five field test forms per form A and B. In total, across contents, there were 714 unique field test items.

Table 5.2 Number of Field Test Items for Each Content Area

Content	N of FT Forms	FT Items per Form	Total
Algebra I	8	10	80
Algebra II	8	10	80
Biology	6	15	90
English II	8	15	120
English III	8	15	120
Geometry	8	10	80
U.S. History	10	15	150
Total			720

Section 5.3 – Data Receipt Activities

5.3.1 Suppressed/Omitted/Invalidated cases

Eliminate suppressed, omitted, and invalidated cases that were flagged in the WinScore files. Cases that had five or less valid attempts were eliminated as well.

5.3.2. Duplicate cases

Any duplicate cases were eliminated by checking student ID (if available), first and last name, middle initial, GIS_CD (GIS code normally containing the district and school ID), teacher name, school, birthday, gender, and response vectors.

5.3.3. Non-public schools

The non-public schools were excluded. Those schools are:

• Oklahoma School for the Deaf
• Oklahoma School for the Blind
• Riverside Indian School
• Sequoia Indian School
• Jones Academy

5.3.4 Second-timers

Students who took the test for the second time were excluded as well.

Section 6 – OCCT Scaling and Equating Methods

The Winter 2012 OCCT EOI and Spring 2013 OCCT EOI programs were based on the application of pre-equating for Algebra I, Algebra II, Geometry and U.S. History; while post-equating was applied to English II and English III. For Biology I, a new scale was set up and a standard setting was performed during Summer 2013.

Verification of the equating samples was described in **Section 5 – Sampling Plan and Field Test Design**. A series of item-level analyses were conducted. These analyses were highly scrutinized to confirm that score keys were accurately and systematically applied and that the summary statistics, such as the item difficulties (*p*-values) and reliabilities (point biserial correlations), were comparable across administrations. CTB/McGraw-Hill Content Development completed a review of all items flagged for possible mis-keys and approved the score keys that were applied. The items were then scaled using the IRT models chosen for OCCT EOI: the three-parameter logistic (3PL) model for MC items and the two-parameter partial credit (2PPC) model for the writing prompts. The following section describes the IRT methods used in the analyses of the operational test items.

Section 6.1 – Classical Item Analyses

Item Level Analyses

Each Winter 2012 and Spring 2013 operational test item was first reviewed in terms of classical raw score statistics. Each item was reviewed for frequency distribution (number of students responding for each answer choice or score level), overall *p*-value (proportion of students choosing the correct answer), and point biserial or item-test correlation (how correlated each individual item is with the test as a whole based on the correct response). Typically, *p*-values should range between 0.25 and 0.90. Items with a *p*-value less than 0.25 are considered more difficult because fewer than 25% of the students are achieving the correct answer. Values greater than 0.90 indicate a fairly easy item because more than 90% of students are achieving the correct answer. A small number of easy items are included to motivate low-performing students, and a small number of difficult items are included to motivate high-performing students. With newly-tested content, the *p*-values may dip lower than 0.25, at which point the item should be evaluated in light of the newness of content or students' opportunity to learn the content. Point biserials or item-test correlations are usually in the range of 0.30 and above, although some items can be acceptable when as low as 0.15. The point biserials of each item's distractors, or incorrect responses, were also analyzed, as well as any distractor with a positive point biserial, either of which was reviewed for the possibility of an additional correct response or no correct response.

Item omit rate (percentage of students that didn't respond to an item) was also examined. Omitted items are scored as zero. The rate of omission often provides information about test speededness, particularly if there is a high omit rate on an item at the end of a test session. High omit rate on an item might also indicate other problems associated with the item such as an unclear question or a confusing presentation. When more than 5% of students omitted an item, the item was reviewed by both CTB/McGraw-Hill Research and Publishing.

A summary comparison of the classical statistics between the Spring 2011, Spring 2012, and Spring 2013 OCCT EOI results is presented in Table 9. Typically, differences less than about │0.05│ are desirable and, as can be seen, *p*-values and mean item-test correlation differences were within expectation.

A detailed summary of the item level classical raw score statistics and omission rates for Spring 2013 and a comparison to Spring 2012 are provided in **Section 7—Results**.

Section 6.2 – Differential Item Functioning (DIF) Analysis

One of the goals of the OCCT EOI assessments is to assemble a set of items that provides a measure of a student's achievement that is as fair and accurate as possible for all subgroups within the population. Differential item functioning (DIF) analysis refers to statistical procedures that assess whether items are differentially difficult for matched-achievement students across reference and focal subgroups (the latter being the group of interest). DIF procedures typically control for overall between-group differences on a criterion, usually total test scores. Between-group performance on each item is then compared within sets of examinees having the same total test scores. If the item is differentially more difficult for an identifiable subgroup when conditioned on achievement, the item may be measuring something different from the intended construct. However, it is important to recognize that the flagging of items for DIF might be related to actual differences in relevant knowledge, skills or statistical Type I error. As a result, DIF statistics are used only to identify potential sources of item bias. Subsequent review by content experts and bias committees are required to determine the source and meaning of performance differences. OCCT EOI conducts DIF analyses across gender (males/females) and ethnicity— focal subgroups African American (not Hispanic), American Indian/Alaskan Native, and Hispanic versus the reference group White (not Hispanic).

The Mantel-Haenszel (MH) DIF statistic was used for the OCCT EOI operational tests. It matches students across the reference and focal groups based on their overall test performance and provides a chi-square to test whether the odds of answering an item correctly are similar for both the reference and focal groups. The items were classified into three categories on the basis of the MH DIF chi-square statistics and the MH delta (Δ) value of A, B, or C for either dichotomous or polytomous items (see Dorans & Holland, 1993; Zieky, 1993; and Michaelides, 2008), where items classified as A are interpreted as having no DIF and items classified as C are interpreted as having potentially severe DIF. The item flag classifications are made as follows:

- The item is classified into the C category if MH DIF is significantly different from zero ($p < 0.05$), and the absolute value of MH delta is greater than or equal to 1.5.
- The item is classified into the B category if MH DIF is significantly different from zero ($p < 0.05$), and the absolute value of MH delta is between 1.0 and 1.5.
- The item is classified into the A category if MH DIF is not significantly different from zero ($p \geq 0.05$), or if the absolute value of MH delta is less than 1.0.

Section 6.3 – Calibration & Item Fit

Item Response Theory (IRT) Models

IRT Models and Rationale for OCCT EOI Applications

Item response theory (IRT) allows comparisons between items and examinees, even those from different test forms, by using a common scale for all items and examinees (i.e., as if there were a hypothetical test that contained items from all forms). The three-parameter logistic (3PL) model (Lord & Novick, 1968; Lord, 1980) was used to analyze item responses for the MC items. For analysis of the Constructed-Response (CR) items, the two-parameter partial credit model (2PPC) (Muraki, 1992; Yen, 1993) was used.

IRT is a statistical methodology that takes into account the fact that not all test items are alike and that all items do not provide the same amount of information in determining how much a student knows or can do. Computer programs that implement IRT models use actual student data to estimate the characteristics of the items on a test, called "parameters." The parameter estimation process is called "item calibration."

IRT models typically vary according to the number of parameters estimated. For the OCCT EOI tests, three parameters are estimated: the discrimination parameter, the difficulty parameter(s), and, for MC items, the guessing parameter. The discrimination parameter is an index of how well an item differentiates between high-performing and low-performing students. An item that cannot be answered correctly by low-performing students, but can be answered correctly by high-performing students, will have a high discrimination value. The difficulty parameter is an index of how easy or difficult an item is. The higher the difficulty parameter, the more difficult the item is. The guessing parameter is the probability that a student with very low ability will answer the item correctly.

Because the characteristics of MC and CR items are different, two IRT models were used in item calibration. The three-parameter logistic (3PL) model (Lord & Novick, 1968; Lord, 1980) was used in the analysis of MC items. In this model, the probability that a student with ability θ responds correctly to item i is

$$P_i(\theta) = c_i + \frac{1 - c_i}{1 + \exp[-1.7a_i(\theta - b_i)]},$$

(1)

where a_i is the item discrimination, b_i is the item difficulty, and c_i is the probability of a correct response by a very low-scoring student.

For analysis of the CR items, the 2PPC model was used. The 2PPC model is a special case of Bock's (1972) nominal model. Bock's model states that the probability of an examinee with ability θ having a score $(k - 1)$ at the kth level of the jth item is

$$P_{jk}(\theta) = P(x_j = k-1 \mid \theta) = \frac{\exp Z_{jk}}{\sum_{i=1}^{m_j} \exp Z_{ji}} \ , \ k = 1 \ldots m_j \tag{2}$$

where

$$Z_{jk} = A_{jk}\theta + C_{jk} \tag{3}$$

and k is the item response category ($k = 1, 2, \ldots m_j$). The m_j denotes the number of score levels for the jth item, and typically the highest score level is assigned ($m_j - 1$) score points. For the special case of the 2PPC model used here, the following constraints were used:

$$A_{jk} = \alpha_j(k-1) \ , \tag{4}$$

and

$$C_{jk} = -\sum_{i=0}^{k-1} \gamma_{ji} \ , \tag{5}$$

where,

$$\gamma_{j0} = 0 \ , \tag{6}$$

and α_j and γ_{ji} are the free parameters to be estimated from the data.

Each item has ($m_j - 1$) independent γ_{ji} parameters and one α_j parameter; a total of m_j parameters are estimated for each item.

The IRT model parameters were estimated using CTB/McGraw-Hill's PARDUX software (Burket, 2002). PARDUX estimates parameters simultaneously for MC and CR items using marginal maximum likelihood procedures implemented via the expectation-maximization algorithm (Bock & Aitkin, 1981; Thissen, 1982; CTB, 2011). Simulation studies have compared PARDUX with MULTILOG (Thissen, 1991), PARSCALE (Muraki & Bock, 1991), and BIGSTEPS (Wright & Linacre, 1992). PARSCALE, MULTILOG, and BIGSTEPS are among the most widely known and used IRT programs. PARDUX was found to perform as well as these other programs (Fitzpatrick, 1990; Fitzpatrick, 1994; Fitzpatrick & Julian, 1996).

Assessment of Item Fit to the IRT Model

Item-Model Fit

Item fit statistics discern the appropriateness of using an item in the 3PL or 2PPC model. A procedure described by Yen (1981) was used to measure fit to the 3PL model. Students are rank-ordered on the basis of $\hat{\theta}$ values and sorted into ten cells with 10% of the sample in each cell. For each item, the number of students in cell k who answered item i, N_{ik}, and the number of students in that cell who answered item i correctly, R_{ik}, were determined. The observed proportion in cell k passing item i, O_{ik}, is R_{ik}/N_{ik}. The fit index for item i is

$$Q_{1i} = \sum_{k=1}^{10} \frac{N_{ik}(O_{ik} - E_{ik})^2}{E_{ik}(1 - E_{ik})} \quad , \tag{7}$$

with

$$E_{ik} = \frac{1}{N_{ik}} \sum_{j \varepsilon \, \text{cell} \, k}^{N_{ik}} P_i(\hat{\theta}_j) \quad . \tag{8}$$

A modification of this procedure was used to measure fit to the 2PPC model. For the 2PPC model, Q_{1j} was assumed to have approximately a chi-square distribution with the following degree of freedom:

$$df = I(m_j - 1) - m_j \quad , \tag{9}$$

where I is the total number of cells (usually 10) and m_j is the possible number of score levels for item j.

To adjust for differences in degrees of freedom among items, Q_1 was transformed to Z_{Q1} where

$$Z_{Q_1} = (Q_1 - df)/(2df)^{1/2} \quad . \tag{10}$$

The value of Z will increase with sample size, all else being equal. To use this standardized statistic to flag items for potential misfit, it has been CTB/McGraw-Hill's practice to vary the critical value for Z as a function of sample size. For the OP tests, which have large calibration sample sizes, the criterion $Z_{Q_1}Crit$ used to flag items was calculated using the expression

$$Z_{Q_1}Crit = \left(\frac{N}{1500}\right) * 4 \quad , \tag{11}$$

where N is the calibration sample size.

Items were considered to have poor model fit if the value of the obtained Z_{Q1} was greater than the value of Z_{Q1} critical. If the obtained Z_{Q1} was less than Z_{Q1} critical, the items were rated as having acceptable fit.

Section 6.4 – Equating

Test Scaling and Equating

Once all item-level analyses were conducted, each Spring 2013 OCCT EOI English II and English III form was calibrated and equated using the Stocking and Lord procedure (Stocking & Lord, 1983), a standard method of equating a new test form onto an existing scale. The Stocking and Lord procedure is based on the test characteristic curve (TCC) from the anchor items, which were selected to be representative of reference forms and Spring 2013 operational forms by statistics and content. CTB PARDUX software was applied to equating (2011). The same process was applied to both Winter 2012 and Spring 2013 English II and English III. TCC and IRT standard error of measurement (SEM) plots showing the quality of the test equating for Spring 2013 OCCT EOI are found in Figures 19–36.

Stability of Anchor Items

The stability of the anchor items is important for the equating procedure. The following method was applied to drop anchor items prior to equating:

1.) Items flagged using the TCC method are considered for exclusion when the correlation between the input and estimated item parameters is below 0.80 for the *a*-parameter and below 0.90 for the *b*-parameter. If the exclusion of an anchor outlier item increases the *a*-parameter correlation to above 0.80 or increases the *b*-parameter correlation to above 0.90, then the anchor is a candidate for removal.
2.) An anchor is a candidate for removal when the item is flagged on four of the seven statistics considered when examining the severe differences between the IRT regression curves: Item characteristic curves (ICCs) for anchor items prior to and after equating.
3.) An outlier for *a*-parameter or *b*-parameter can be a candidate based on an anchor item plot, which shows the relationships of anchor item parameters before and after equating (Kolen & Brennan, 2005).
4.) Removal of the item may not significantly alter the content distribution of the anchor set. The distribution of items across the content standards must remain within 10% of the test blueprint for the content area.
5.) The mean difference and standard deviation ratio are also referenced.
6.) It is important to recognize that differential item performance in two test administrations does not necessarily indicate item flaws and may be affected by population differences, differences in teaching strategies, curriculum changes, etc. Therefore, CTB/McGraw-Hill recommended that the Oklahoma SDE consider item content-related factors in addition to statistical evidence of differential item performance in two test administrations.

Items removed from the anchor set based on the flags from the evaluation procedure are still scored as part of the whole test. After an anchor item is removed from the anchor set based on the previous criteria, the anchor file needs to be adjusted and a second version of the calibration and equating must be produced. All outputs in the second version need to be evaluated following the same guidelines as the original calibration runs.

Section 6.5 – Writing Test Scoring

Writing prompts were administered as a part of the English II and English III in the Winter 2012 and Spring 2013 administrations. The writing score is a weighted composite of five analytic scores that focus on specific domains of writing skills. The steps for calculating the English II Writing scores follow and are illustrated for an example in Table 6.1.

Steps to Calculate OCCT EOI English II Writing Scores

STEP 1:   Average the trait scores from the two raters to obtain each of the five analytic trait scores. Average the scores in Column C and Column D, and write the results in Column E.

STEP 2:   Apply the weights to the trait scores. Multiply the numbers in Column B and Column E. Write the results in Column F.

STEP 3:   Sum all the weighted trait scores in Column F (lower right corner).

STEP 4:   Transform the sum of the weighted trait scores. That is, multiply the weighted sum of the trait scores by 1.7 and subtract 1.025 as shown following the table.

STEP 5:   Round the transformed weighted composite score to the nearest whole number to obtain the final Writing score. After calculation, the final writing score value will range from 1 to 6.

Table 6.1. Calculating Writing Composite Scores for English II

A	B	C	D	E	F
Analytic Traits	Weights	Trait Scores from Rater 1	Trait Scores from Rater 2	Average (C+D)/2	Weighted Trait Scores (B X E)
Ideas and Development	0.30	3	2	(3+2)/2=2.5	.30 X 2.5 = 0.75
Organization, Unity, and Coherence	0.25	3	3	(3+3)/2=3.0	.25 X 3.0 = 0.75
Word Choice	0.15	3	2	(3+2)/2=2.5	.15 X 2.5 = 0.375
Sentences and Paragraphs	0.15	2	3	(2+3)/2=2.5	.15 X 2.5 = 0.375
Grammar/Usage and Mechanics	0.15	3	2	(3+2)/2=2.5	.15 X 2.5 = 0.375
					Sum Above
					= 2.625

Transformed Writing Score = 2.625 X 1.7 – 1.025 = 3.4375
Final Writing Score = 3

The steps for calculating the English III Writing scores follow and are illustrated for an example in Table 6.2.

Steps to Calculate OCCT English III Writing Scores

The steps that follow show the calculation of the ACE English III Writing scores based on the trait scores for a writing prompt. The table shows an example of the calculation of the ACE English III Writing scores.

STEP 1:   Average the trait scores from the two raters to obtain each of the five analytic trait scores. Average the scores in Column C and Column D, and write the results in Column E.

STEP 2:   Multiply the weights by 5 to give new weights. Multiply the numbers in Column B by 5, and write the results in Column F.

STEP 3:   Multiply each trait score by the new weight to give the weighted score. Multiply Column E by Column F, and write the results in Column G.

STEP 4:   Sum all the weighted scores in Column G (lower right corner).

STEP 5:   Transform the sum of the weighted trait scores. Multiply the weighted sum of the trait scores by .58 and subtract 1.67843 as shown following the table.

STEP 6:   Round the transformed score to the nearest whole number to obtain the final English III Writing score. After calculation, the final ACE English III Writing score value will range from 1 to 10.

Table 6.2. Calculating Writing Composite Scores for English III

A	B	C	D	E	F	G
Analytic Traits	Weights	Trait Scores from Rater 1	Trait Scores from Rater 2	Average Trait (C+D)/2	New Weight (B X 5)	Weighted Trait Scores (E X F)
Ideas and Development	0.30	2	2	2	(.30 X 5) = 1.5	(2 X 1.5) = 3
Organization, Unity, and Coherence	0.25	1	2	1.5	(.25 X 5) = 1.25	(1.5 X 1.25) = 1.875
Word Choice	0.15	2	3	2.5	(.15 X 5) = 0.75	(2.5 X .75) = 1.875
Sentences and Paragraphs	0.15	3	3	3	(.15 X 5) = 0.75	(3 X .75) = 2.25
Grammar/Usage and Mechanics	0.15	4	3	3.5	(.15 X 5) = 0.75	(3.5 X .75) = 2.625
						Sum Above
						11.625

Transformed ACE English III Writing Score = 11.625 x .58 - 1.67843 = 5.06407
Final Writing Score = 5

Section 7 – Results

This section provides the data analysis results for the Winter 2012 and Spring 2013 OCCT EOI. Item level analyses for operational and field test items are presented below. Standard, test, and proficiency level student performances are summarized and presented as well. Three MC items were suppressed in the OCCT EOI Spring 2013 operational tests, due to their extremely poor item statistics. One item in English II form B, one item in English III forms A and B, and one item in U.S. History form B were suppressed in Spring 2013. These suppressions were approved by the SDE. The suppressed items did not contribute to students' scores and are not included in the analyses results in this section.

Section 7.1 – Item Level Performance

A summary comparison of the classical item statistics between the Spring 2011, Spring 2012, and Spring 2013 OCCT EOI is presented in Table 9. Typically, differences of less than about │0.05│are expected. As can be seen in Table 9, item p-values had a slight increase across grades and content areas, with the largest difference seen in Biology I and English III (0.05) between Spring 2012 and Spring 2013. The mean item-test correlation differences range from -0.01 to 0.01, except for U.S. History (0.02).

A summary of the range of p-values and item-test correlations of all operational and field test items for Spring 2013 is presented in Table 10. (Item-test correlations were calculated by correlating the correct response of the focal item to the remainder of the items in the test, focal item excluded.) As shown in Table 10, the average p-values for the operational test items are in the mid 0.60s for Algebra I, low 0.60s for Algebra II, high 0.60s for Biology I, mid 0.70s for English II, high 0.60s to low 0.70s for English III, around 0.70 for Geometry, and in the mid 0.60s for U.S. History. The range of the p-values dips below 0.25 in Algebra I and English III for all forms. Item-test correlations across content areas for operational items are within typical and acceptable ranges. For Biology I form B and most forms in English II and English III, one or more items show item-test correlation lower than 0.15. For the field test items, the average p-values for the items are in the low and mid 0.50s for Algebra I, low 0.50s for Algebra II, low to mid 0.50s for Biology I, low to mid 0.60s for English II, around 0.60 for English III, low to mid 0.50s for Geometry, and mid to high 0.50s for U.S. History. Average item-test correlations for field test items are in the low to mid 0.30s for most content areas.

The item omission rates for operational and field test items are presented in Table 11. The operational items show less than 0.85% (well below the 5% criteria) omission rate across contents, indicating acceptable administration times for the number of items in each test session. The MC field test items for all contents show omission rates well below the 5% criteria.

The Spring 2013 OCCT EOI DIF results are reported for all contents in Table 12 for gender and Tables 13 and 14 for ethnicity. There were no items flagged for moderate or severe DIF for the American Indian/ Alaskan Native subgroup in either operational or field test items. The results indicate that the majority of operational test items did not exhibit potential bias. For operational items on gender DIF, there were a total of 35 items (5.1%) flagged for moderate "B" DIF and 13 items (1.9%) flagged for severe "C" DIF. For operational test items in the African American (not

Hispanic) and Hispanic ethnicity groups included in the DIF analyses, there were respectively 3.5% and 2.6% of the items flagged for moderate "B" DIF, and 0.7% and 0.6% of the items flagged for severe "C" DIF.

DIF results for the field test in all contents for gender DIF show that there were a total of 35 items (5%) flagged for moderate "B" DIF and 2 items (0.3%) flagged for severe "C" DIF. For field test items in the African American (not Hispanic) and Hispanic ethnicity groups included in the DIF analyses, there were respectively 4.9% and 2.6% of the items flagged for moderate "B" DIF, and 0.07% and 0.4% of items flagged for severe "C" DIF.

All of the items flagged were reviewed by CTB/McGraw-Hill content experts who cross-referenced all teacher judgments and comments from across the content reviews, bias and sensitivity reviews, as well as alignment workshops to make decisions with the SDE about suppressions from operational scoring and use of the flagged operational and/or field test items in future test forms. No Spring 2013 operational items were suppressed due to DIF.

Items with Poor Statistics

Piloting or field testing items is the best way to find potentially problematic items in the item pool. However, even during an operational administration, there are times that items become unstable or do not exhibit the highest expected qualities. Therefore, the evaluation of items across administrations from the content reviews, bias and sensitivity reviews, alignment workshops, and the various statistical analyses can be exhaustive and sensitive to the test blueprints, which can sometimes result in the suppression of some operational items from student scores and of some field test items from the item pool. Sometimes, OE items that do not show enough/adequate case counts at a given score level, resulting in score collapses, and items that do not converge during scaling or that exhibit extreme misfit are also suppressed.

During the Spring 2013 OCCT EOI operational and field test administration, items were reviewed for their classical statistics, and when those statistics were outside the range of difficulty (p-values less than 0.25 or greater than 0.90) or showed low item-test-correlations (less than 0.15) for a specific item, the item was used or kept as a "good item" in the pool only when the content of the item justified its use (e.g., it was a new standard or new approach that was expected to be difficult). For the Spring 2013 OCCT EOI operational test, three items were suppressed before scoring began. Items with less than desirable p-values and item-test correlations were reviewed by CTB/McGraw-Hill content experts and Research, and field test items considered to have less than desirable statistics were suppressed from the item pool.

Section 7.2 – Performance at Standards Level

A review of the item difficulty across standards within each content area is provided to illustrate for which standards items were more or less difficult for students. The summaries are presented in Tables 15–20. The tables provide the number of operational items, the reliability (coefficient alpha) and the standard error of measurement (SEM) (See **Section 8—Summary of Reliability and Validity**), and the average difficulty or IRT location (b parameter) value. The tables also provide the average p-values for the state and for each proficiency level for each standard.

As shown in Tables 15–20, the reliability at each standard, which is influenced by the number of items contributing to each standard, ranges from 0.54 to 0.86 in Algebra I, from 0.57 to 0.85 in Algebra II, from 0.45 to 0.79 in Biology I, from 0.31 to 0.68 in English II, from 0.45 to 0.77 in English III, from 0.47 to 0.83 in Geometry, and from 0.54 to 0.77 in U.S. History. Across the content areas, the standard errors are no greater than 2.43 and the maximum amount of IRT information is about 0.66.

IRT locations and *p*-values should be reviewed within each content area by standard shown in Tables 15–20. The IRT scale locations provide an indication of the average *b* parameters or location values of the set of items contributing to each of the standards. Different from the average *p*-values, the IRT locations provide information on the location of the items along the scale score continuum, such that higher values indicate a higher probability for a student with estimated higher ability to answer those items correctly. The *p*-values provide only the proportion of students in each group answering the items correctly, averaged across items within each of the standards.

Section 7.3 – Test Level Performance

Total Group Scale Scores

The Spring 2013 OCCT EOI applies a number-correct to scale score scoring method based on the 3PL IRT model. In this method, all students who have the same raw score get the same scale score regardless of which items are correct.

Tables 21 to 23 provide the state-level distribution of the scale scores across grades and content areas for Spring 2012, Winter 2012, and Spring 2013, respectively. Tables 22 and 23 provide the state-level distribution of the scale scores across grades and content areas as well as the distribution across the 25th, 50th, and 75th percentiles for Winter 2012 and Spring 2013. (Spring 2013 scale score reliability as coefficient alpha and standard error of measurement are also provided and further explained in **Section 8—Summary of Reliability and Validity**.) Provided as a reference only, Table 21 shows those results for Spring 2012. Histograms and associated skewness and kurtosis of the data for Spring 2013 are provided in Figures 1–18. The data are close to normally distributed with a very minimal positive skew in most content areas and grades.

Subgroup Scale Scores and Mean Differences

Subgroup-level scale score performance data (scale score means and standard deviations, minimum and maximum scale scores, reliability and standard error of measurement) are provided along with state-level data in Table 24 for Winter 2012 and in Table 25 for Spring 2013. For Spring 2013, mean differences were subjected to independent sample *t*-tests for gender, IEP, Low SES, ELL, Section 504, and accommodated students subgroups, in each content area. One way Analysis of Variance (ANOVA) tests were conducted across ethnicities, for which equal variances were not assumed and the level of significance was set at 0.05. Spring 2013 results of the *t*-tests and ANOVA are found in Tables 26–31 and Table 32, respectively.

As shown in Table 26, females outperform males in all forms in English II and English III as well as in Algebra I in both forms, and Algebra II and Geometry in form A; males outperform females in Algebra II form B, Geometry form B, Biology I in both forms, and U.S. History both forms. Mean differences are not statistically significant for Geometry form B only.

Results of the *t*-tests within each category indicate that IEP, Low SES, ELL, and accommodated students all score significantly lower than the rest of the population in all content areas (ranging from 15 to 103 fewer scale score points), as expected. IEP and ELL students tend to have the lowest performance of the "special population" subgroups. For the Section 504 group, the same is true in all content areas, but the mean differences compared with the rest of the state are less than 22 scale score points and not significant at most contents and forms, except Algebra I form B and English II from AB.

Statistically significant differences exist between the ethnicity groups in all content areas and grades as presented in the ANOVA results in Tables 32.

In comparing ethnicities across all content areas, students identified as Asian and White tended to outperform the other ethnicities in Spring 2013. A *post-hoc Dunnett's C pair-wise* comparison analysis was conducted to identify potential pairs of significant differences ($p = 0.05$), the results of which are found in Table 33. As shown in Table 33, Asian outperformed all other ethnicities in Algebra I and II, Biology I, and Geometry as well as most other ethnicities in English II and English III and U.S. History, with the exception of Pacific Islander, White, and the Other subgroup. White outperformed most other ethnicities, except Asian, in most content areas. Most pairs were significantly different with some exceptions, which were **not** significant, between African American, Native American, Hispanic, Pacific Islander, White and the Other subgroups. For example, White mean differences were not significantly different from Other subgroup in Algebra I forms A and B; and Pacific Islander in Biology I form B, Geometry form B, and U.S. History forms A and B, English II all forms, and English III forms AA, BA, and BB.

Section 7.4 – Proficiency Level Performance

Table 7.1 shows the Spring 2013 scale score cut points for each proficiency performance level and the scale bounds. The lowest obtainable scale score (LOSS) and highest obtainable scale score (HOSS) values are shown for all content areas.

Table 7.1 Spring 2013 Scale Score Cuts and Scale Bounds

Content Area	LOSS	Cut1	Cut2	Cut3	HOSS
Algebra I	490	662	700	762	999
Algebra II	440	654	700	783	999
Biology I	440	651	700	773	999
English II	440	609	700	817	999
English III	440	670	700	802	999
Geometry	440	635	700	777	999
U.S. History	440	627	700	773	999

Table 34 shows the scale score means and standard deviations for the state and for students in each proficiency level. Table 35 provides the statewide distribution (or "impact data") of students within each proficiency level (*Unsatisfactory, Limited Knowledge, Proficient, and Advanced*) and the overall pass rates defined as the total percentage of students in both the *Proficient and Advanced* proficiency levels for Spring 2013. Tables 34 and Table 35 do not include the number of students considered *Undetermined* (invalid) in the denominator of the calculation.

Impact data across proficiency levels are also provided for each gender, ethnicity, and special population subgroups in Table 36, where comparative performance across subgroups mimics what was provided for the scale score descriptions.

Section 8 – Summary of Reliability and Validity

This section summarizes some of the evidence in the earlier sections and provides additional evidence to support the degree to which the OCCT EOI tests are reliable and valid. For the OCCT EOI, several measures of reliability are available. First, the tests are administered in standard fashion to all students. When students needed accommodations, such accommodations were provided with specific guidance from the *OSTP 2012-2013 Test Preparation Manual* (http://ok.gov/sde/sites/ok.gov.sde/files/documents/files/2705543-W_tpm_w13OK.pdf). The General Guidance section describes details about the tests as well as specific administration policies, procedures, and accommodation guidelines.

Section 8.1 – Item Level Reliability

Item-specific reliability statistics include inter-rater reliability, item-test correlations, and differential item functioning (DIF) or item bias. The inter-rater reliabilities of OE items rely heavily on the solid and consistent training of the Handscorers, as described in **Section 4 – Scoring**. Tables 4.1 and 4.2, for English II and English III, respectively, provide the relevant inter-rater statistics, which are presented in terms of the percentage of perfect and adjacent agreement and checkset average agreement.

The point biserial, or item-test correlation, is one measure of reliability, computed using the correlation between each item and the overall test. We discussed the item-test correlation in **Section 6 and in Section 7.** The item-test correlations for each content area and item type are shown in Table 37. The operational item correlations ranged from 0.17 to 0.58 (Algebra I), from 0.19 to 0.55 (Algebra II), from 0.12 to 0.54 (Biology I), from 0.11 to 0.58 (English II), from 0.06 to 0.62 (English III), from 0.22 to 0.64 (Geometry), and from 0.15 to 0.57 (U.S. History). Several items in the Spring 2013 operational OCCT EOI presented item-test correlation less than 0.15. Those items were investigated by Content Development for scoring key errors and found to be correctly scored. Any operational items with extremely low point biserial that may remain in the OCCT EOI item pool will be avoided on future operational forms.

DIF statistics (described in **Section 6 – Methods** and **Section 7 – Results**) provide a measure of the systematic errors by subgroups that are specifically attributed to potential bias or systematic over- or under-representation of subgroup performance when compared to total group performance. As shown in Tables 12–14 (last rows), the percentage of operational items that exhibited DIF at the moderate and severe levels was about 7.0% for gender and between 0.6% and 3.5% for the two ethnicity groups.

Section 8.2 – Test Level Reliability

Total test reliability statistics (alpha and conditional standard errors of measurement, CSEMs) measure the level of consistency (reliability) of performance over all test questions in a given form, the results of which imply how well the questions measure the content domain and could continue to do so over repeated administrations. Total test reliability coefficients, measured by Cronbach's alpha (Cronbach, 1951) range from 0.00 to 1.00, where 1.00 refers to a perfectly reliable test. The OCCT EOI reliability data are based on the Oklahoma student population and

the results for 2013 are typical of the results obtained for all previous OCCT EOI operational tests. The total test reliabilities of the operational forms were evaluated first by Cronbach's alpha (1951) index of internal consistency. The specific calculation for Cronbach's alpha, α, is

$$\hat{\alpha} = \frac{k}{k-1}\left(1 - \frac{\sum \hat{\sigma}_i^2}{\hat{\sigma}_X^2}\right), \tag{12}$$

where k is the number of items on the test form, $\hat{\sigma}_i^2$ is the variance of item i, $\hat{\sigma}_X^2$ is the total test variance, and the summation is over all the items ($i = 1, \ldots, k$) on the test. Achievement tests are typically considered of sound reliability when their reliability coefficients are in the range of 0.80 and above.

Table 38 shows the reliability coefficients for each scored operational test form for each content area and grade for both Spring 2012 and Spring 2013. The alpha reliability coefficients for Spring 2012 and Spring 2013 are quite similar. The reliability coefficients for Spring 2013 ranged between 0.86 (English form BA) and 0.93 (Geometry form A). Such a range is indicative of the high reliability of the Spring 2013 OCCT EOI operational tests. As is evident in Table 25, for Spring 2013 state and subgroup data, the coefficients are quite high and similar to the state values, even at the subgroup levels. The mean of the state-level reliability coefficients for each content area in Table 38 are as follows: 0.91 (Algebra I), 0.90 (Algebra II), 0.89 (Biology I), 0.87 (English II), 0.88 (English III), 0.92 (Geometry), and 0.90 (U.S. History). At the subgroup level in Table 25, the lowest reliability (0.76) was found for the Section 504 students in English II form BA.

The SEM is another measure of reliability and is a direct estimate of the degree of measurement error in students' total scores (per the alpha reliability coefficient). The SEM represents the number of score points about which a given score can vary, similar to the standard deviation of a score; the smaller the SEM, the smaller the variability of the estimate, and the higher the reliability. The total SEMs are computed with the following formula:

$$SEM = SD\_TT(\sqrt{1-\hat{\alpha}}), \tag{13}$$

where SD_TT is the standard deviation for the total test and $\hat{\alpha}$ is the result of the calculation of Cronbach's α in Equation 12.

The CSEMs conditional on each scale score are computed with the following formula:

$$CSEM = SD\_SS(\sqrt{1-\hat{\alpha}}), \tag{14}$$

SD_SS is the standard deviation of the scale score. The total test SEMs for each test form are provided for each content area and grade at the state and subgroup levels in Table 25. Scale score specific SEMs are given in Tables 39–43, which also provide the raw scores associated with each scale score.

Section 8.3 – Test Level Validity

Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were conducted to investigate potential evidence to further support the validity of the OCCT EOI test scores for the total population and then for the accommodated, ELL, and IEP subgroups. The subgroups were chosen such that the students within each group may have characteristics that could contribute to issues of access and/or for whom the test measures construct irrelevant variances. A variety of criteria are used conjunctively to evaluate the assumption that each test for each content area measures a single (unidimensional) construct (e.g., Algebra I, English II, U.S. History). In factor analyses, the "construct" is referred to as a factor. The analyses help to organize the data such that relationships defined as factors are illuminated. If the data are essentially unidimensional, a single factor should account for most of the variation in the data.

Accordingly, a unidimensional factor model was tested using polychoric correlation coefficients against the obtained covariance matrix using maximum likelihood estimation (Bentler & Bonett, 1980, Jöreskog, & Sorbom, 1989) for each content area and grade for the total population and each subgroup using SAS version 9.1. The polychoric correlation is most appropriate when variables are dichotomous or ordinal and together are assumed to reflect a single, underlying construct (Byrne, 1998).

First, the factorability of the correlation matrix was examined before conducting the CFA (Is the data adequately correlated and thus analyzable or "factorable" to move forward?). The Kaiser-Meyer Olkin (KMO; Kaiser, 1970, 1974) measure of sampling adequacy was used through an EFA procedure to evaluate the strength of the linear relationship among the items within each correlation matrix. KMO values in the 0.90 and greater range are considered "marvelous" according to Kaiser's (1974) criteria. As shown in Table 44 and Table 45, KMO values for the total group ranged from 0.95 to 0.98, and for each subgroup: from 0.86 to 0.95 (Accommodated), from 0.61 to 0.92 (ELL), and from 0.84 to 0.95 (IEP). That most of the KMO values are in the "marvelous" range suggests that the matrix is appropriate for CFA for each analysis.

As a rough estimate of the number of factors (dimensions or constructs) that might be present in the data, the Kaiser criterion of computing the eigenvalues for the correlation matrix was examined next. Eigenvalues represent how much variability is accounted for by each factor not in sum but out of the total amount of variance. Table 44 and Table 45 also show the total amount of variance that exists in each form, as well as the percentage of variance accounted for by the initial eigenvalue. For the total group analyses, the first eigenvalue's measure of the amount of variance in relation to the total variance is 91 and 89% (Algebra I), 91 and 95% (Algebra II), 97 and 99% (Biology I), 87–91% (English I), 72–91% (English III), 88 and 91% (Geometry), and 98 and 99% (U.S. History) respectively. The range of variance by the first eigenvalue in each content area and subgroup is as follows:

- Accommodated: 79% and 67% (Algebra I), 75% and 59% (Algebra II), 79% and 62% (Biology I), 52–60% (English II), 52–71% (English III), 81% and 73% (Geometry), and 80% and 65% (U.S. History).

- ELL: 70% and 63% (Algebra I), 52% and 41% (Algebra II), 63% and 44% (Biology I), 32–35% (English II), 27–30% (English III), 66% and 59% (Geometry), and 58% and 41% (U.S. History).
- IEP: 77% and 70% (Algebra I), 75% and 67% (Algebra II), 79% and 65% (Biology I), 49–60% (English II), 55–73% (English III), 81% and 76% (Geometry), and 80% and 67% (U.S. History).

Such values indicate one major factor is present in each of the content assessments. It is interesting to note that the range of variance for the total population for U.S. History is mostly higher than the other content areas. At the subgroup level, the variance is higher for accommodated and IEP in Geometry and for ELL in Algebra I.

As a rule, "essential unidimensionality" is assumed when the ratio of the first eigenvalue to the second eigenvalue is at least three. The final column of Tables 44 and 45 provides the ratio of the first and second eigenvalues. All grades and content areas for the total population and each subgroup have no ratios less than three; therefore, the OCCT EOI tests demonstrate essential unidimensionality per the eigenvalue ratio criterion.

An additional available criterion used in EFA to judge the number of factors present is the scree test (Cattell, 1966) of eigenvalues plotted against factors. Examination of the scree plots (Figures 37–108) for the content areas for the total population and each subgroup indicates a single factor model is present and similar patterns between the total population and subgroups.

Summary inspection across all the criteria—variance, ratio of eigenvalues, and scree plots—seems to indicate that the tests for each content area and grade, and for each subgroup, are essentially unidimensional. It is important to review the relationships of factors in conjunction with all other data, particularly where items may be dependent (for example, where several items share a passage or stimulus).

Section 8.4 – Performance Level Reliability

Proficiency Level Reliability

One of the cornerstones of the federal Elementary and Secondary Education Act (ESEA) emphasizes the need for all students to score in the "Proficient" category on English Language Arts, Mathematics and Science. Because of a heavy emphasis on moving all students to or above the "Proficient" category, the consistency and accuracy of the classification of students into these proficiency categories is of particular interest. The statistical quality of cut scores that define the proficiency levels in which students are classified based on their performance serves as additional validity evidence. Details about the OCCT standard setting workshop and the Bookmark Standard Setting Procedure used to set the cut scores are given in the Oklahoma School Testing Program Bookmark Standard Setting Technical Report for End-of-Instruction Biology I (CTB/McGraw-Hill, 2013). The Bookmark Procedure (Mitzel, Lewis, Patz, & Green, 2001) is a well-documented and highly regarded procedure that has been demonstrated by independent research to produce reasonable cut scores on tests across the country.

It is also important to review the specific scale score SEM for each cut score. Table 46 shows the Spring 2013 SEMs estimated for each of the cut scores for each content area and grade. Comparison of the SEMs for the cut scores to the SEMs associated with other OCCT EOI scale scores for each test (shown in Tables 39–43) reveal that the SEMs for the cut scores are almost always among the lowest, which means that the OCCT EOI tests tend to measure most accurately near the cut scores. This is a desirable quality when cut scores are used to classify examinees. (Not every scale score possible, sometimes including the cut score, is shown in Tables 39–43; there are more scale scores possible at each raw score than can be shown in these tables.)

Not only it is important that the amount of measurement error around the cut scores be minimal; but it is also important to have the expected consistency with which students would be classified into performance levels if a test were given on repeated occasions.

Classification consistency is defined as the extent to which two classifications of a single student agree from two independent administrations of the same test (or two parallel forms of the test). Classification consistency and accuracy are additional measures of test reliability as well as validity. Reliability coefficients, such as Cronbach's alpha, are used to check for the internal consistency within a single test. Test-retest reliability requires two administrations of the same test, which requires another test as an external reference. Consistency in the classification sense represents how well two forms of an assessment with equal difficulty agree (Livingston & Lewis, 1995). It is estimated using actual response data and total test reliability from an administered form of an assessment, from which two parallel forms of the assessment are statistically modeled and classifications compared.

Classification accuracy is defined as the agreement between the actual classifications using observed cut scores and true classifications based on known true cut scores (Livingston & Lewis, 1995). It is common to estimate classification accuracy by utilizing a psychometric model to find true scores corresponding to observed scores.

In other words, classification consistency refers to the agreement between two observed classification results, while classification accuracy refers to the agreement between the observed classification outcome and the true classification result. A straightforward approach to classification consistency estimation can be expressed in terms of a contingency table representing the probability of a particular classification outcome under specific scenarios. For example, the following is a contingency table of $(H+1) \times (H+1)$, where H is the number of cut scores such that two cut scores yield a 3x3 contingency table below.

	Level 1	Level 2	Level 3	Sum
Level 1	P_{11}	P_{21}	P_{31}	$P_{.1}$
Level 2	P_{12}	P_{22}	P_{32}	$P_{.2}$
Level 3	P_{13}	P_{23}	P_{33}	$P_{.3}$
Sum	$P_{1.}$	$P_{2.}$	$P_{3.}$	1.0

To report classification consistency, Swaminathan, Hambleton, and Algina (1974) suggest using Cohen's kappa (1960):

$$\text{kappa} = \frac{P - P_c}{1 - P_c}, \tag{15}$$

where P is defined as the sum of the diagonal values of the contingency table (the values shaded in the above table), and P_c is the chance probability of a consistent classification under two completely random assignments. This probability, P_c, is the sum of the probabilities obtained by multiplying the marginal probability of the first administration and the corresponding marginal probability of the second administration:

$$P_c = (P_{1.} \times P_{.1}) + (P_{2.} \times P_{.2}) + (P_{3.} \times P_{.3}). \tag{16}$$

The Livingston and Lewis (1995) method, based on the binomial error model and the four-parameter beta true score distribution, was applied to OCCT EOI. Tables 47 and 48 show the classification consistency and classification accuracy indices. The values of all indices depend on several factors, such as the reliability of the actual test form, the distribution of scores, the number of cut scores, and the location of each cut score. The probability of a correct classification (Consistency) is the probability that the classification the student received is consistent with the classification that the student would have received on a parallel form; in other words, that the classification is correct. This is akin to the exact agreement rate in inter-rater reliability, and the expectation is that this probability would be high.

Table 47 shows that the average consistency is 0.73 across content areas and ranges from 0.64 (Biology I form B) to 0.78 (English III form BB). The average accuracy is 0.80 across all content areas and ranges from 0.73 (Biology I form B) to 0.84 (English III form BB). Cohen's kappa (Kappa) provides the same type of reliability or agreement statistic as in the inter-rater reliabilities. In this context, it represents the agreement of the classifications between the two parallel forms with consideration of the probability of a correct classification by chance (Consistency–Chance[1])/(1–Chance). In general, the value of Kappa is lower than the value of Consistency because the probability of a correct classification by chance is greater than zero. This is true of the OCCT EOI data in Table 47. The average Kappa is 0.58 over all content areas and ranges from 0.51 (Biology I form B) to 0.63 (Geometry form A).

Consistency and accuracy are important to consider together. The probability of accuracy (Accuracy) represents the agreement between the observed classification, based on the actual test form, and true classification given the modeled form. Table 48 shows consistency and accuracy at the cut score level. The average consistency across content areas and cut score levels is 0.91, ranging from 0.85 (Biology I form B, at the Proficient and Advanced proficiency levels) to 0.99 (English II both forms, at the Unsatisfactory and Limited Knowledge proficiency levels). The average accuracy across content areas and cut score levels is 0.93, ranging from 0.89 (Biology I form B, at the Proficient and Advanced proficiency levels) to 0.99 (English II both forms, at the

---

[1] The probability of a correct classification by chance (Chance) is the probability that the classification is correct and is due to chance alone. The probability of Chance is estimated under a complete random assignment procedure using the marginal distribution of each form. The Chance probabilities are expected to be low.

Unsatisfactory and Limited Knowledge proficiency levels). Finally, Table 49 provides the probability of false positives (FP) and false negatives (FN) as measures of error in the data table, and these are low (no greater than 0.06), as expected.

Section 9 – Online Disruption in Spring 2013 Administration

Background

Online testing was administered to all Oklahoma OCCT EOI tests in the Spring 2013 administration. While the online testing window started in mid-April and was progressing well, on Monday, April 29th and Tuesday, April 30th, students taking online assessments in the state experienced system interruptions, which led some local districts to temporarily suspend testing on those days. CTB's diagnostic efforts revealed that two separate issues with CTB/McGraw-Hill's Online Assessment System (OAS) caused the interruptions.

CTB/McGraw-Hill utilizes industry standard tools and current state software engineering practices for building and evaluating our systems. In preparation for the 2013 testing season, additional capacity was added and extensive load testing was performed. However, the factors that triggered the system interruptions on April 29th and 30th had not been detected in testing or prior year production use.

The key difference between what was seen in actual use versus CTB's performance evaluation was high utilization sustained over a longer period of time, which consumed more physical memory in the first outage and subsequently more virtual memory in the second. This memory utilization was more than the load tests predicted.

CTB engineers worked quickly to address both issues. In the first event, CTB restarted systems and added physical memory. In the second event, CTB identified necessary memory configuration changes and implemented them as rapidly as possible to allow the testing to continue. Changes included increased physical memory on nodes hosting applications, increased virtual memory Heap size on each node, and finally after analyzing utilization, adding more nodes on the following weekend. Once all these upgrades were completed, the OAS application was able to manage the volume of requests without degraded performance.

On April 30th, while CTB engineers were trying to restore the system health back to normal, one of the activity was to "cut over" to the disaster recovery site. While the system remained accessible, this "cut over" caused interruptions for almost all students who were active in the system. Also, as the system was moved from the working servers to the disaster recovery servers, some student responses were not immediately accessible to students when they logged back into that test session. All of the students' responses had been saved, but they were not immediately available due to the time it takes for the disaster recovery systems to be synchronized with live testing data from the primary servers. Based on the severity of the interruption and a recommendation, the State Superintendent requested that students complete their current test and then schools suspend online testing. For OCCT EOI testing, some students who were taking operational forms switched to an alternate operational form or equivalent form. On May 1st, online testing returned to normal, and students did not experience further interruptions.

Section 9.1 – Interruption Study

The SDE requested an independent third-party study to investigate the impact of the disruptions on the test scores, and HumRRO was selected by the SDE to perform the study. Technical Advisory Committee (TAC) consists of researchers and assessment experts selected by the SDE. TAC members participated in a conference call with HumRRO researchers to recommend a design for the disruption study.

HumRRO applied several different methods using propensity matching between disrupted students and non-disrupted students and linear regression. Their study indicated that in general, students appeared to do as well on the test regardless of disruptions in testing (For detailed description, see "A Statistical Investigation of Oklahoma Computer Disruptions" in Appendix B).

Section 9.2 – Actions for Disrupted Students

There were two types of disruption issues: 1) There were students whose performance was considered to be impacted and 2) There were students whose immediate scores (displayed on the online test delivery client when the test is finished) and two-week scores were different.

For the first issue, the SDE decided on retaining all scores of impacted students who scored Proficient or Advanced. Though students as a whole did not experience depressions in scores, the SDE wanted no lasting effect on student performance. Also, the SDE decided to not report scores for impacted students who scored Limited Knowledge or Unsatisfactory. Though the HumRRO study suggested no systematic impact of the disruption on test scores, in the event a particular student was impacted, not reporting the scores would ensure that the individual student was not wrongly scored. The SDE selected this approach for the benefit of the students and schools.

For the second issue, about 1,400 students received an immediate raw score that differed from the score they received on the two-week preliminary reports. Several factors contributed to this issue, one of which was caused by local sites downloading the wrong test system, and the other where it appears that during a short time-span on the second day of the interruptions (4/30), some of the student responses were not saved for inclusion in the two-week reports. Students received the higher of the two scores, the immediate and the two-week report scores. Students with higher immediate raw scores received a letter from CTB indicating their proficiency level. They were included in the report card, but not in the school or district summary report. Students with a higher two-week score received a traditional score report that included performance level and performance by content standard.

References

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

Barton, K. (2007). Validity and accommodations: The journey toward accessible assessments. In C. Cahalan Laitusis & L. Cook (Eds). *Large-scale assessment and accommodations: What works?* Council for Exceptional Children.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588–606.

Bock, R.D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37:29–51.

Bock, R.D. and M. Aitkin. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika* 46:443–459.

Burket, G.R. (2002). *PARDUX* [Computer program].

Byrne, B. M. (1998). *Structural equation modeling with LISRELA, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245–276.

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

CTB/McGraw-Hill. (2011). *PARDUX Software version 1.66.* Monterey, CA: Author.

CTB/McGraw-Hill. (2013). *Oklahoma School Testing Program: 2013-2014 Test preparation manual.* Monterey, CA: Author.

CTB/McGraw-Hill. (2013). *Oklahoma School Testing Program: Bookmark standard setting technical report for End-of-Instruction Biology I.* Monterey, CA: Author.

Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), *Differential Item functioning* (pp. 35–66). Hillsdale, NJ: Lawrence Erlbaum.

Fitzpatrick, A.R. (1990). *Status report on the results of preliminary analysis of dichotomous and multi-level items using the PARMATE program.*

Fitzpatrick, A.R. (1994). *Two studies comparing parameter estimates produced by PARDUX and BIGSTEPS*.

Fitzpatrick, A.R. & M.W. Julian (1996). *Two studies comparing the parameter estimates produced by PARDUX and PARSCLE*. Monterey, CA: CTB/McGraw-Hill**.**

Holland, P. W., & Thayer, D.T. (1988). Differential Item Performance and the Mantel-Haenszel Procedure. (ETS RR-86-31). Princeton, NJ: Educational Testing Service.

Jöreskog, K., & Sorbom, D. (1989). *LISREL 7 User's Reference Guide*, Chicago: Scientific Software International.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika, 35*(4), 401–415.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: methods and practices (2nd ed.)*. New York: Springer.

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. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

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

Lord, F. M., & Wingersky, M. S. (1984). Comparison of IRT true-score and equipercentile observed-score "equatings." *Applied Psychological Measurement, 8,* 453–461.

Mantel, N., & Haenszel, W. (1959) Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22, 719–748.

Michaelides, M. P. (2008). An Illustration of a Mantel-Haenszel Procedure to Flag Misbehaving Common Items in Test Equating. *Practical Assessment Research & Evaluation, 13*(7). Available online: http://pareonline.net/pdf/v13n7.pdf.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 249–281). Mahwah, NJ: Lawrence Erlbaum.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement* 16:159–176.

Muraki, E. (1997). The generalized partial credit model. In W.J. van der Linden & R.K. Hambleton (Eds.), *Handbook of Modern Item Response Theory* (pp. 153–164). New York: Springer Verlag.

Muraki, E. & R.D. Bock (1991). *PARSCALE*: *Parameter Scaling of Rating Data* [Computer program]. Chicago, IL: Scientific Software, Inc.

Oklahoma State Department of Education. (2013). *Test Support (Teachers and Administrators).* Available online: http://www.ok.gov/sde/test-support-teachers-and-administrators.

Pearson. (2011). *Oklahoma School Testing Program: Oklahoma Core Curriculum Tests Grades 3 to 8 Assessments 2011 Technical Report.* Pearson.

Pearson. (2012). *Oklahoma School Testing Program: Oklahoma Core Curriculum Tests Grades 3 to 8 Assessments 2012 Technical Report.* Pearson.

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

Swaminathan, H., Hambleton, R. K., & Algina, J. (1974). Reliability of criterion-referenced tests: A decision-theoretic formulation. *Journal of Educational Measurement*, *11*(4), 263–267.

Thissen, D. (1982). Marginal maximum likelihood estimation for the one-parameter logistic model. *Psychometrika* 47:175–186.

Thissen, D. (1991). MULTILOG: Multiple category item analysis and test scoring using item response theory [Computer software]. Chicago, IL: Scientific Software International, Inc.

Thissen, D., Chen, W-H., & Bock, R. D. (2003). *MUTILOG; Multiple-category item analysis and test scoring using item response theory (Version 7)* [Computer software]. Lincolnwood, IL: Scientific Software International.

Thissen, D., Pommerich, M., Billeaud, K., & Williams, V. S. L. (1995). Item response theory for scores on tests including polytomous items with ordered responses. *Applied Psychological Measurement, 19,* 39–49.

United States Department of Education. (2002). *No Child Left Behind Act of 2001*. (U.S. Department of Education, Office of Elementary and Secondary Education). Washington, DC.

United States Department of Education. (2004). *Standards and assessment peer review guidance: Information and examples for meeting requirements for the No Child Left Behind Act of 2001*. (U.S. Department of Education, Office of Elementary and Secondary Education). Washington, DC.

Wright, B.D. & J. M. Linacre. (1992). *BIGSTEPS Rasch Analysis* [Computer program]. Chicago, IL: MESA Press.

Yen, W. M. (1981). Using simulation results to choose a latent trait model. *Applied Psychological Measurement, 5,* 245–262.

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement, 8,* 125–145.

Yen, W.M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement* 30:187–213.

Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Hillsdale, NJ: Lawrence Erlbaum.

Tables

Table 8. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2013

Content	Form	Female			Male		
		Sample	State	Diff.	Sample	State	Diff.
English II	AA	50.91	51.19	-0.28	49.09	48.81	0.28
	AB	51.37	51.02	0.35	48.61	48.98	-0.37
	BA	50.72	51.78	-1.06	49.28	48.22	1.06
	BB	51.4	51.81	-0.41	48.6	48.19	0.41
English III	AA	48.6	48.46	0.13	51.39	51.5	-0.11
	AB	48.94	49.19	-0.25	50.99	50.8	0.19
	BA	49.87	50.04	-0.17	50.13	49.96	0.17
	BB	50.31	50.61	-0.3	49.69	49.39	0.3

Table 8. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2013 *(continued)*

Content	Form	American Indian/Alaskan			African American (Not Hispanic)		
		Sample	State	Diff.	Sample	State	Diff.
English II	AA	16.39	15.83	0.56	9.19	9.02	0.17
	AB	16.97	15.92	1.05	9.84	9.10	0.73
	BA	16.14	15.63	0.50	9.19	8.91	0.28
	BB	16.69	15.74	0.95	10.46	9.07	1.39
English III	AA	16.99	16.32	0.67	9.33	9.26	0.07
	AB	16.16	16.64	-0.48	10.09	9.52	0.57
	BA	16.40	16.35	0.05	9.81	9.47	0.34
	BB	15.24	15.2	0.04	10.24	9.18	1.06

Table 8. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2013 *(continued)*

Content	Form	Asian			Hispanic		
		Sample	State	Diff.	Sample	State	Diff.
English II	AA	2.70	2.39	0.31	12.13	11.71	0.42
	AB	2.33	2.32	0.01	14.54	12.12	2.42
	BA	2.77	2.47	0.30	12.25	11.73	0.52
	BB	2.63	2.60	0.02	12.98	11.70	1.28
English III	AA	2.48	2.10	0.38	11.46	11.37	0.09
	AB	2.40	2.14	0.26	11.31	11.19	0.12
	BA	3.10	2.76	0.35	11.26	10.71	0.55
	BB	2.70	2.51	0.19	11.62	11.06	0.56

Table 8. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2013 *(continued)*

Content	Form	White (Not Hispanic)			Multiracial (Other)			Pacific Islander		
		Sample	State	Diff.	Sample	State	Diff.	Sample	State	Diff.
English II	AA	52.36	54.45	-2.09	6.78	6.18	0.60	0.44	0.41	0.04
	AB	49.06	53.64	-4.58	6.29	6.22	0.07	0.97	0.67	0.30
	BA	53.36	54.57	-1.21	5.77	6.17	-0.40	0.52	0.52	0.01
	BB	49.21	54.12	-4.91	7.27	6.27	0.99	0.77	0.49	0.28
English III	AA	52.37	54.96	-2.60	6.80	5.51	1.29	0.57	0.48	0.09
	AB	53.24	54.90	-1.66	6.20	5.09	1.11	0.59	0.51	0.08
	BA	52.13	54.79	-2.66	6.76	5.46	1.30	0.53	0.47	0.07
	BB	53.64	55.85	-2.21	6.03	5.70	0.32	0.53	0.50	0.03

Table 8. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2013 *(continued)*

Content	Form	ELL			IEP			Section 504		
		Sample	State	Diff.	Sample	State	Diff.	Sample	State	Diff.
English II	AA	2.40	2.34	0.05	5.27	6.36	-1.09	0.84	0.89	-0.05
	AB	2.47	2.28	0.19	5.11	5.97	-0.85	0.64	0.77	-0.13
	BA	2.24	2.10	0.14	4.38	4.29	0.09	0.76	0.77	-0.01
	BB	2.31	2.18	0.13	4.33	4.61	-0.29	0.86	0.92	-0.06
English III	AA	2.05	2.37	-0.32	12.29	15.71	-3.42	0.81	0.84	-0.03
	AB	2.16	2.28	-0.11	11.87	14.39	-2.52	0.93	0.86	0.07
	BA	2.06	1.86	0.20	8.65	8.59	0.06	0.76	0.69	0.07
	BB	2.06	1.93	0.13	8.40	7.98	0.42	0.94	0.96	-0.02

Note: ELL = English Language Learner; IEP = Individualized Education Program

Table 8. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2013 *(continued)*

Content	Form	SES-Low			SES-High			Accommodated		
		Sample	State	Diff.	Sample	State	Diff.	Sample	State	Diff.
English II	AA	54.38	54.71	-0.32	45.62	45.29	0.32	5.77	6.45	-0.69
	AB	50.36	53.26	-2.90	49.64	46.74	2.90	5.63	6.00	-0.37
	BA	53.11	53.00	0.11	46.89	47.00	-0.11	4.36	4.10	0.26
	BB	50.75	53.89	-3.14	49.25	46.11	3.14	4.62	4.38	0.24
English III	AA	54.55	54.33	0.22	45.45	45.67	-0.22	10.81	14.17	-3.36
	AB	56.45	55.19	1.26	43.55	44.81	-1.26	10.33	12.72	-2.39
	BA	57.52	57.75	-0.23	42.48	42.25	0.23	7.31	6.97	0.35

Table 9. Summary of P-values and Item-Test Correlations Statistics for Operational Test Forms, Spring 2011 to Spring 2013

Subject	Operational Mean P-values*				Operational Mean Item-Test Correlations*			
	2011	2012	2013	Diff. S13-S12	2011	2012	2013	Diff. S13-S12
Algebra I	0.53	0.65	0.66	0.01	0.39	0.42	0.41	-0.01
Algebra II	0.57	0.61	0.63	0.02	0.46	0.40	0.40	0.00
Biology I	0.64	0.64	0.69	0.05	0.40	0.37	0.37	0.00
English II	0.72	0.74	0.76	0.02	0.39	0.33	0.33	0.00
English III	0.64	0.65	0.70	0.05	0.42	0.35	0.35	0.00
Geometry	0.64	0.70	0.71	0.01	0.42	0.43	0.44	0.01
U.S. History	0.65	0.64	0.65	0.01	0.39	0.37	0.39	0.02

*Census Data.
Note: Suppressed items are not included in data; Spring 2013 Biology I was based on new content standards.

Table 10. Summary of Range of P-values and Item-Test Correlations Statistics for Operational and Field Test, Spring 2013

Content	Form	Mean P-values*						Mean Item-Test Correlations*					
		Operational Items			Field Test Items			Operational Items			Field Test Items		
		Low	Mean	High	Low	Mean	High	Low	Mean	High	Low	Mean	High
Algebra I	A	0.19	0.66	0.92	0.20	0.54	0.87	0.17	0.41	0.58	0.00	0.34	0.54
	B	0.20	0.66	0.94	0.18	0.51	0.83	0.23	0.40	0.55	0.12	0.34	0.54
Algebra II	A	0.25	0.62	0.91	0.13	0.53	0.89	0.22	0.40	0.52	0.11	0.34	0.57
	B	0.31	0.63	0.95	0.15	0.51	0.94	0.19	0.39	0.55	0.03	0.31	0.51
Biology I	A	0.35	0.68	0.87	0.15	0.51	0.92	0.16	0.38	0.54	0.06	0.29	0.56
	B	0.38	0.69	0.98	0.15	0.57	0.97	0.12	0.35	0.54	0.12	0.32	0.52
English II	AA	0.32	0.75	0.99	0.18	0.65	0.93	0.16	0.32	0.48	0.12	0.32	0.52
	AB	0.33	0.75	0.99	0.18	0.65	0.93	0.14	0.32	0.58	0.10	0.31	0.51
	BA	0.36	0.76	0.96	0.22	0.64	0.96	0.13	0.33	0.48	0.05	0.30	0.49
	BB	0.37	0.76	0.96	0.22	0.63	0.95	0.11	0.34	0.58	0.04	0.30	0.49
English III	AA	0.23	0.68	0.91	0.07	0.59	0.97	0.07	0.34	0.57	0.01	0.32	0.50
	AB	0.23	0.68	0.91	0.09	0.59	0.97	0.06	0.34	0.62	0.05	0.31	0.49
	BA	0.24	0.72	0.96	0.20	0.60	0.91	0.14	0.36	0.59	0.09	0.31	0.49
	BB	0.24	0.72	0.96	0.23	0.60	0.91	0.14	0.37	0.62	0.11	0.32	0.47
Geometry	A	0.33	0.69	0.94	0.18	0.52	0.91	0.28	0.45	0.64	0.10	0.36	0.62
	B	0.40	0.72	0.94	0.12	0.54	0.94	0.22	0.42	0.57	0.08	0.36	0.57
U.S. History	A	0.27	0.64	0.94	0.20	0.54	0.89	0.15	0.40	0.57	0.06	0.33	0.53
	B	0.27	0.65	0.96	0.15	0.58	0.91	0.20	0.37	0.55	0.02	0.33	0.54

*Census Data.
Note: Suppressed items are not included in data.

Table 11. Summary of Range of Omission Rates for Operational and Field Test by Item Type, Spring 2013

Content	Form	Item Type	Omission Rates*					
			Operational Items			Field Test Items		
			Low	Mean	High	Low	Mean	High
Algebra I	A	MC	0.01%	0.21%	0.85%	0.00%	0.22%	0.80%
	B	MC	0.01%	0.17%	0.72%	0.00%	0.17%	0.67%
Algebra II	A	MC	0.01%	0.11%	0.27%	0.00%	0.12%	0.39%
	B	MC	0.01%	0.12%	0.34%	0.00%	0.13%	0.37%
Biology I	A	MC	0.01%	0.09%	0.16%	0.00%	0.08%	0.24%
	B	MC	0.04%	0.09%	0.17%	0.00%	0.08%	0.23%
English II	AA	CR	0.11%	0.30%	0.48%	.	.	.
		MC	0.01%	0.08%	0.22%	0.00%	0.06%	0.18%
	AB	CR	0.13%	0.31%	0.49%	.	.	.
		MC	0.03%	0.08%	0.23%	0.00%	0.07%	0.19%
	BA	CR	0.13%	0.32%	0.52%	.	.	.
		MC	0.01%	0.09%	0.29%	0.00%	0.09%	0.18%
	BB	CR	0.06%	0.20%	0.34%	.	.	.
		MC	0.02%	0.06%	0.14%	0.00%	0.05%	0.23%
English III	AA	CR	0.13%	0.40%	0.66%	.	.	.
		MC	0.04%	0.14%	0.28%	0.00%	0.19%	0.44%
	AB	CR	0.12%	0.41%	0.71%	.	.	.
		MC	0.02%	0.10%	0.24%	0.00%	0.12%	0.27%
	BA	CR	0.11%	0.38%	0.66%	.	.	.
		MC	0.01%	0.12%	0.26%	0.00%	0.16%	0.33%
	BB	CR	0.16%	0.43%	0.71%	.	.	.
		MC	0.01%	0.08%	0.18%	0.00%	0.11%	0.27%
Geometry	A	CR	0.01%	0.14%	0.31%	0.00%	0.14%	0.28%
	B	MC	0.01%	0.09%	0.21%	0.00%	0.09%	0.24%
U.S. History	A	CR	0.03%	0.06%	0.13%	0.00%	0.06%	0.20%
	B	MC	0.01%	0.04%	0.14%	0.00%	0.04%	0.14%

*Census Data.
Note: Suppressed items are not included in data.

Table 12. Spring 2013 Summary of Operational and Field Test Items Flagged for Mantel-Haenszel Differential Item Functioning, by Item Type: Gender Male/Female

Content	Item Type	Operational Items		Field Test Items		Total DIF Flags B+C
		B	C	B	C	
Algebra I	MC	4	1	5	1	11
Algebra II	MC	2	.	4	.	6
Biology I	MC	5	.	5	1	11
English II	MC	9	2	3	.	14
	CR	.	2	.	.	2
English III	MC	7	2	7	.	16
	CR	.	2	.	.	2
Geometry	MC	6	1	5	.	12
U.S. History	MC	2	3	6	.	11
Total Items Flagged		35	13	35	2	85
Total Items Tested		693		696		1389
Percentage of Items Flagged		5.05%	1.88%	5.03%	0.29%	6.12%

Table 13. Spring 2013 Summary of Operational and Field Test Items Flagged for Mantel-Haenszel Differential Item Functioning, by Item Type: Ethnicity African American/White

Content	Item Type	Operational Items		Field Test Items		Total DIF Flags B+C
		B	C	B	C	
Algebra I	MC	3	.	5	1	9
Algebra II	MC	3	2	4	.	9
Biology I	MC	2	.	2	.	4
English II	MC	3	2	6	4	15
	CR	.	.	.	.	0
English III	MC	7	1	4	.	12
	CR	.	.	.	.	0
Geometry	MC	2	.	1	.	3
U.S. History	MC	4	.	12	.	16
Total Items Flagged		24	5	34	5	68
Total Items Tested		693		696		1389
Percentage of Items Flagged		3.46%	0.72%	4.89%	0.72%	4.90%

Table 14. Spring 2013 Summary of Operational and Field Test Items Flagged for Mantel-Haenszel Differential Item Functioning, by Item Type: Ethnicity Hispanic/White

Content	Item Type	Operational Items		Field Test Items		Total DIF Flags B+C
		B	C	B	C	
Algebra I	MC	.	.	4	.	4
Algebra II	MC	4	.	2	.	6
Biology I	MC	.	.	.	1	1
English II	MC	3	1	4	2	10
	CR	.	.	.	.	0
English III	MC	7	1	2	.	10
	CR	1	.	.	.	1
Geometry	MC	1	.	2	.	3
U.S. History	MC	2	2	4	.	8
Total Items Flagged		18	4	18	3	43
Total Items Tested		693		696		1389
Percentage of Items Flagged		2.60%	0.58%	2.59%	0.43%	3.10%

Note: No DIF flags found for American Indian/White

Table 15. Algebra I & Algebra II Standards Level Summary Data, Spring 2013

Content	Form		Standard Reference	No. of Items	Average Difficulty (IRT Loc)	Average IRT Information	Objective % Correct State Mean	Average P-value					
								State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass
Algebra I	A	1	Number Sense and Algebraic Operations	15	739.47	0.20	63.97	0.64	0.26	0.38	0.60	0.86	0.70
		2	Relations and Functions	31	728.68	0.19	67.05	0.67	0.28	0.41	0.65	0.87	0.73
		3	Data Analysis, Probability, and Statistics	9	741.78	0.17	65.37	0.65	0.31	0.44	0.62	0.84	0.71
	B	1	Number Sense and Algebraic Operations	15	733.20	0.27	61.39	0.62	0.22	0.34	0.57	0.85	0.68
		2	Relations and Functions	31	730.68	0.22	68.48	0.69	0.30	0.43	0.66	0.88	0.74
		3	Data Analysis, Probability, and Statistics	9	740.78	0.12	64.60	0.65	0.32	0.44	0.61	0.83	0.69
Algebra II	A	1	Number Sense and Algebraic Operations	15	758.93	0.07	62.29	0.62	0.31	0.43	0.62	0.85	0.70
		2	Relations and Functions	31	749.61	0.09	61.80	0.62	0.31	0.43	0.61	0.84	0.70
		3	Data Analysis, Probability, and Statistics	9	747.56	0.07	64.74	0.65	0.27	0.43	0.66	0.88	0.74
	B	1	Number Sense and Algebraic Operations	15	755.20	0.08	61.62	0.62	0.31	0.42	0.60	0.83	0.69
		2	Relations and Functions	31	762.74	0.08	61.97	0.62	0.29	0.40	0.61	0.85	0.70
		3	Data Analysis, Probability, and Statistics	9	720.33	0.06	70.58	0.71	0.40	0.55	0.70	0.88	0.77

Note: P.L. = Performance Level

Table 15. Algebra I & Algebra II Standards Level Summary Data, Spring 2013 *(continued)*

Content	Form		Standard Reference	No. of Items	Alpha	SEM
Algebra I	A	1	Number Sense and Algebraic Operations	15	0.75	1.64
		2	Relations and Functions	31	0.86	2.25
		3	Data Analysis, Probability, and Statistics	9	0.56	1.28
	B	1	Number Sense and Algebraic Operations	15	0.78	1.60
		2	Relations and Functions	31	0.85	2.25
		3	Data Analysis, Probability, and Statistics	9	0.54	1.30
Algebra II	A	1	Number Sense and Algebraic Operations	15	0.73	1.69
		2	Relations and Functions	31	0.84	2.37
		3	Data Analysis, Probability, and Statistics	9	0.69	1.27
	B	1	Number Sense and Algebraic Operations	15	0.72	1.66
		2	Relations and Functions	31	0.85	2.43
		3	Data Analysis, Probability, and Statistics	9	0.57	1.20

Table 16. Biology I Standards Level Summary Data, Spring 2013

Form		Standard Reference	No. of Items	Average Difficulty (IRT Loc)	Average IRT Information	Objective % Correct State Mean	Average P-value					
							State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass
A	1	The Cell	12	676.00	0.04	68.66	0.69	0.47	0.65	0.77	0.90	0.81
	2	The Molecular Basis of Heredity	12	676.33	0.08	66.77	0.67	0.43	0.61	0.77	0.92	0.81
	3	Biological Diversity	12	684.67	0.06	67.39	0.67	0.43	0.64	0.77	0.89	0.81
	4	The Interdependence of Organisms	8	688.38	0.08	67.22	0.67	0.41	0.62	0.78	0.91	0.82
	5	Matter/Energy/Organization in Living Systems	13	699.39	0.07	64.93	0.65	0.38	0.59	0.77	0.91	0.81
	P1	Observe and Measure	6	680.83	0.07	70.49	0.71	0.45	0.67	0.81	0.93	0.85
	P2	Classify	8	658.50	0.05	72.07	0.72	0.49	0.68	0.81	0.93	0.85
	P3	Experiment	17	661.12	0.04	68.89	0.69	0.49	0.65	0.77	0.89	0.81
	P4	Interpret and Communicate	21	682.43	0.07	66.68	0.67	0.41	0.62	0.78	0.92	0.82
	P5	Model	8	712.50	0.08	60.80	0.61	0.35	0.53	0.72	0.88	0.76
B	1	The Cell	12	667.33	0.07	69.80	0.70	0.48	0.64	0.78	0.93	0.82
	2	The Molecular Basis of Heredity	12	700.25	0.06	66.93	0.67	0.43	0.61	0.77	0.92	0.81
	3	Biological Diversity	13	680.54	0.06	66.24	0.66	0.42	0.62	0.76	0.89	0.79
	4	The Interdependence of Organisms	8	693.38	0.06	66.26	0.66	0.46	0.61	0.74	0.88	0.78
	5	Matter/Energy/Organization in Living Systems	12	672.33	0.04	70.81	0.71	0.51	0.67	0.79	0.90	0.82
	P1	Observe and Measure	6	649.17	0.06	71.85	0.72	0.50	0.67	0.81	0.93	0.84
	P2	Classify	7	682.14	0.07	69.48	0.70	0.48	0.63	0.79	0.93	0.83
	P3	Experiment	17	667.82	0.05	68.16	0.68	0.50	0.63	0.75	0.89	0.79
	P4	Interpret and Communicate	22	676.23	0.06	71.00	0.71	0.47	0.67	0.81	0.92	0.84
	P5	Model	8	721.00	0.06	60.09	0.60	0.38	0.53	0.68	0.86	0.73

Note: P.L. = Performance Level

Table 16. Biology I Standards Level Summary Data, Spring 2013 *(continued)*

Form		Standard Reference	No. of Items	Alpha	SEM
A	1	The Cell	12	0.58	1.49
	2	The Molecular Basis of Heredity	12	0.68	1.44
	3	Biological Diversity	12	0.64	1.44
	4	The Interdependence of Organisms	8	0.60	1.18
	5	Matter/Energy/Organization in Living Systems	13	0.72	1.54
	P1	Observe and Measure	6	0.52	1.01
	P2	Classify	8	0.54	1.17
	P3	Experiment	17	0.64	1.74
	P4	Interpret and Communicate	21	0.79	1.92
	P5	Model	8	0.57	1.25
B	1	The Cell	12	0.63	1.41
	2	The Molecular Basis of Heredity	12	0.66	1.48
	3	Biological Diversity	13	0.67	1.48
	4	The Interdependence of Organisms	8	0.46	1.23
	5	Matter/Energy/Organization in Living Systems	12	0.56	1.42
	P1	Observe and Measure	6	0.45	0.97
	P2	Classify	7	0.48	1.11
	P3	Experiment	17	0.61	1.72
	P4	Interpret and Communicate	22	0.77	1.92
	P5	Model	8	0.50	1.27

Table 17. English II Standards Level Summary Data, Spring 2013

Form		Standard Reference	No. of Items	Average Difficulty (IRT Loc)	Average IRT Information	Objective % Correct State Mean	Average P-value					
							State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass
AA	1	Vocabulary	6	731.91	0.07	77.92	0.78	0.42	0.59	0.77	0.92	0.81
	2	Comprehension	18	721.36	0.07	73.75	0.74	0.32	0.52	0.73	0.90	0.78
	3	Literature	18	716.09	0.06	77.77	0.78	0.34	0.55	0.78	0.92	0.82
	3	Writing/Grammar/Usage and Mechanics	12	761.04	0.07	68.60	0.69	0.30	0.45	0.67	0.87	0.73
	4	Research and Information	6	710.58	0.06	76.23	0.76	0.32	0.55	0.77	0.89	0.80
AB	1	Vocabulary	6	731.91	0.07	77.89	0.78	0.40	0.59	0.78	0.93	0.81
	2	Comprehension	18	721.36	0.07	73.64	0.74	0.32	0.52	0.74	0.90	0.78
	3	Literature	18	716.09	0.06	78.03	0.78	0.36	0.55	0.79	0.93	0.82
	3	Writing/Grammar/Usage and Mechanics	12	761.04	0.07	68.38	0.68	0.31	0.45	0.68	0.88	0.72
	4	Research and Information	6	710.58	0.06	76.57	0.77	0.32	0.56	0.77	0.90	0.80
BA	1	Vocabulary	6	698.09	0.06	81.91	0.82	0.37	0.60	0.82	0.95	0.85
	2	Comprehension	17	741.61	0.09	73.75	0.74	0.29	0.50	0.73	0.89	0.77
	3	Literature	19	710.54	0.07	78.03	0.78	0.32	0.54	0.78	0.92	0.81
	3	Writing/Grammar/Usage and Mechanics	12	745.08	0.10	72.31	0.72	0.31	0.45	0.71	0.91	0.76
	4	Research and Information	5	688.30	0.06	-99.00	0.83	0.37	0.61	0.83	0.94	0.86
BB	1	Vocabulary	6	698.09	0.06	81.75	0.82	0.36	0.60	0.81	0.94	0.85
	2	Comprehension	17	741.61	0.09	73.69	0.74	0.29	0.50	0.73	0.89	0.77
	3	Literature	19	710.54	0.07	78.27	0.78	0.34	0.54	0.78	0.92	0.82
	3	Writing/Grammar/Usage and Mechanics	12	745.08	0.10	72.43	0.72	0.29	0.46	0.71	0.91	0.76
	4	Research and Information	5	688.30	0.06	-99.00	0.83	0.35	0.61	0.83	0.93	0.86

Note: P.L. = Performance Level

Table 17. English II Standards Level Summary Data, Spring 2013 *(continued)*

Form		Standard Reference	No. of Items	Alpha	SEM
AA	1	Vocabulary	6	0.39	0.93
	2	Comprehension	18	0.67	1.67
	3	Literature	18	0.68	1.62
	3	Writing/Grammar/Usage and Mechanics	12	0.60	1.42
	4	Research and Information	6	0.36	0.95
AB	1	Vocabulary	6	0.38	0.93
	2	Comprehension	18	0.67	1.67
	3	Literature	18	0.68	1.61
	3	Writing/Grammar/Usage and Mechanics	12	0.59	1.43
	4	Research and Information	6	0.34	0.94
BA	1	Vocabulary	6	0.38	0.88
	2	Comprehension	17	0.63	1.61
	3	Literature	19	0.66	1.68
	3	Writing/Grammar/Usage and Mechanics	12	0.65	1.33
	4	Research and Information	5	0.31	0.79
BB	1	Vocabulary	6	0.40	0.88
	2	Comprehension	17	0.63	1.61
	3	Literature	19	0.67	1.66
	3	Writing/Grammar/Usage and Mechanics	12	0.66	1.33
	4	Research and Information	5	0.32	0.79

Table 18. English III Standards Level Summary Data, Spring 2013

Form		Standard Reference	No. of Items	Average Difficulty (IRT Loc)	Average IRT Information	Objective % Correct State Mean	Average P-value					
							State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass
AA	1	Vocabulary	7	724.19	0.13	73.27	0.73	0.34	0.46	0.72	0.93	0.78
	2	Comprehension	18	757.63	0.13	65.04	0.65	0.32	0.44	0.64	0.82	0.69
	3	Literature	17	750.08	0.10	66.49	0.67	0.31	0.44	0.65	0.84	0.71
	3	Writing/Grammar/ Usage and Mechanics	13	785.41	0.08	62.19	0.62	0.30	0.39	0.60	0.82	0.66
	4	Research and Information	6	774.39	0.13	65.51	0.66	0.30	0.39	0.62	0.89	0.70
AB	1	Vocabulary	7	724.19	0.13	73.91	0.74	0.34	0.45	0.72	0.92	0.79
	2	Comprehension	18	757.63	0.13	65.21	0.65	0.33	0.43	0.63	0.81	0.69
	3	Literature	17	750.08	0.10	67.06	0.67	0.33	0.43	0.65	0.83	0.71
	3	Writing/Grammar/ Usage and Mechanics	13	785.41	0.08	62.65	0.63	0.29	0.39	0.59	0.80	0.67
	4	Research and Information	6	774.39	0.13	65.87	0.66	0.28	0.36	0.61	0.89	0.70
BA	1	Vocabulary	7	727.00	0.15	79.10	0.79	0.33	0.49	0.76	0.95	0.82
	2	Comprehension	16	723.10	0.09	76.23	0.76	0.32	0.49	0.74	0.90	0.79
	3	Literature	19	741.00	0.08	74.33	0.75	0.33	0.48	0.72	0.88	0.77
	3	Writing/Grammar/ Usage and Mechanics	13	780.41	0.66	60.62	0.61	0.26	0.33	0.54	0.83	0.63
	4	Research and Information	6	757.44	0.14	67.39	0.68	0.29	0.39	0.63	0.84	0.70
BB	1	Vocabulary	7	727.00	0.15	79.22	0.79	0.32	0.49	0.76	0.95	0.82
	2	Comprehension	16	723.10	0.09	76.39	0.76	0.33	0.48	0.73	0.90	0.79
	3	Literature	19	741.00	0.08	74.54	0.75	0.33	0.47	0.72	0.88	0.77
	3	Writing/Grammar/ Usage and Mechanics	13	780.41	0.66	61.04	0.61	0.25	0.33	0.54	0.83	0.64
	4	Research and Information	6	757.44	0.14	67.87	0.68	0.31	0.39	0.63	0.85	0.70

Table 18. English III Standards Level Summary Data, Spring 2013 *(continued)*

Form		Standard Reference	No. of Items	Alpha	SEM
AA	1	Vocabulary	7	0.61	1.02
	2	Comprehension	18	0.71	1.71
	3	Literature	17	0.69	1.71
	3	Writing/Grammar/Usage and Mechanics	13	0.63	1.58
	4	Research and Information	6	0.55	1.04
AB	1	Vocabulary	7	0.61	1.01
	2	Comprehension	18	0.70	1.70
	3	Literature	17	0.67	1.69
	3	Writing/Grammar/Usage and Mechanics	13	0.63	1.57
	4	Research and Information	6	0.56	1.04
BA	1	Vocabulary	7	0.59	0.95
	2	Comprehension	16	0.68	1.56
	3	Literature	19	0.68	1.73
	3	Writing/Grammar/Usage and Mechanics	13	0.77	1.46
	4	Research and Information	6	0.45	0.99
BB	1	Vocabulary	7	0.59	0.94
	2	Comprehension	16	0.68	1.55
	3	Literature	19	0.69	1.71
	3	Writing/Grammar/Usage and Mechanics	13	0.77	1.45
	4	Research and Information	6	0.45	0.99

Table 19. Geometry Standards Level Summary Data, Spring 2013

Form		Standard Reference	No. of Items	Average Difficulty (IRT Loc)	Average IRT Information	Objective % Correct State Mean	State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass
A	1	Logical Reasoning	6	725.83	0.08	68.28	0.68	0.31	0.47	0.66	0.86	0.75
	2	Properties of 2-Dimensional Figures	20	707.85	0.14	75.61	0.76	0.30	0.53	0.76	0.93	0.84
	3	Triangles and Trigonometric Ratios	12	775.58	0.17	58.95	0.59	0.20	0.28	0.52	0.86	0.68
	4	Properties of 3-Dimensional Figures	10	719.20	0.13	67.93	0.68	0.29	0.42	0.64	0.89	0.76
	5	Coordinate Geometry	7	730.00	0.07	67.12	0.67	0.26	0.42	0.66	0.85	0.75
B	1	Logical Reasoning	6	718.33	0.07	70.66	0.71	0.31	0.48	0.66	0.86	0.76
	2	Properties of 2-Dimensional Figures	20	712.45	0.13	75.29	0.75	0.29	0.47	0.72	0.92	0.82
	3	Triangles and Trigonometric Ratios	12	775.67	0.14	62.35	0.62	0.24	0.33	0.54	0.83	0.69
	4	Properties of 3-Dimensional Figures	10	711.50	0.12	76.95	0.77	0.32	0.48	0.74	0.93	0.83
	5	Coordinate Geometry	7	742.43	0.09	68.77	0.69	0.28	0.43	0.64	0.85	0.75

Table 19. Geometry Standards Level Summary Data, Spring 2013 *(continued)*

Form		Standard Reference	No. of Items	Alpha	SEM
A	1	Logical Reasoning	6	0.52	0.98
	2	Properties of 2-Dimensional Figures	20	0.83	1.65
	3	Triangles and Trigonometric Ratios	12	0.80	1.44
	4	Properties of 3-Dimensional Figures	10	0.71	1.26
	5	Coordinate Geometry	7	0.61	1.06
B	1	Logical Reasoning	6	0.47	1.00
	2	Properties of 2-Dimensional Figures	20	0.82	1.68
	3	Triangles and Trigonometric Ratios	12	0.73	1.47
	4	Properties of 3-Dimensional Figures	10	0.70	1.16
	5	Coordinate Geometry	7	0.56	1.10

Table 20. U.S. History Standards Level Summary Data, Spring 2013

Form		Standard Reference	No. of Items	Average Difficulty (IRT Loc)	Average IRT Information	Objective % Correct State Mean	Average P-value					
							State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass
A	1	Post-Reconstruction to the Progressive Era, 1878-1900	8	724.75	0.08	66.25	0.66	0.31	0.46	0.65	0.84	0.74
	2	Expanding Role of the United States in International Affairs	9	764.22	0.09	59.12	0.59	0.24	0.36	0.58	0.80	0.67
	3	Cycles of Economic Boom and Bust in the 1920s and 1930s	11	698.73	0.11	72.72	0.73	0.33	0.52	0.73	0.91	0.81
	4	Role of U.S. in International Affairs and WW II, 1933-1946	11	765.18	0.09	61.68	0.62	0.27	0.39	0.59	0.84	0.70
	5	U.S. Foreign & Domestic Policies during the Cold War, 1945-1975	21	751.10	0.09	62.55	0.63	0.28	0.43	0.62	0.81	0.70
B	1	Post-Reconstruction to the Progressive Era, 1878-1900	9	737.00	0.08	68.87	0.69	0.32	0.48	0.66	0.85	0.75
	2	Expanding Role of the United States in International Affairs	8	742.13	0.09	66.45	0.66	0.26	0.42	0.63	0.85	0.73
	3	Cycles of Economic Boom and Bust in the 1920s and 1930s	11	736.82	0.07	67.32	0.67	0.34	0.49	0.64	0.83	0.73
	4	Role of U.S. in International Affairs and WW II, 1933-1946	11	782.18	0.11	56.90	0.57	0.23	0.33	0.52	0.77	0.63
	5	U.S. Foreign & Domestic Policies during the Cold War, 1945-1975	20	749.75	0.10	66.41	0.66	0.31	0.45	0.64	0.83	0.72

Table 20. U.S. History Standards Level Summary Data, Spring 2013 *(continued)*

Form		Standard Reference	No. of Items	Alpha	SEM
A	1	Post-Reconstruction to the Progressive Era, 1878-1900	8	0.54	1.22
	2	Expanding Role of the United States in International Affairs	9	0.62	1.29
	3	Cycles of Economic Boom and Bust in the 1920s and 1930s	11	0.70	1.28
	4	Role of U.S. in International Affairs and WW II, 1933-1946	11	0.68	1.42
	5	U.S. Foreign & Domestic Policies during the Cold War, 1945-1975	21	0.77	1.90
B	1	Post-Reconstruction to the Progressive Era, 1878-1900	9	0.54	1.26
	2	Expanding Role of the United States in International Affairs	8	0.57	1.22
	3	Cycles of Economic Boom and Bust in the 1920s and 1930s	11	0.55	1.39
	4	Role of U.S. in International Affairs and WW II, 1933-1946	11	0.64	1.46
	5	U.S. Foreign & Domestic Policies during the Cold War, 1945-1975	20	0.74	1.83

Table 21. Spring 2012 Scale Score Statistics

Content	Form	N Count	Mean	SD	LOSS	N Min.	Scale Score Percentile 25th	50th	75th	N Max.	HOSS	Alpha	SEM
Algebra I	A	19956	739.40	54.75	490	110	707	741	768	132	999	0.91	16.42
	B	18910	742.36	54.34	490	83	711	744	771	153	999	0.91	16.30
Algebra II	A	16742	733.92	85.93	440	218	689	742	787	52	999	0.91	25.78
	B	15636	738.08	80.57	440	185	693	743	785	62	999	0.91	24.17
Biology	A	19867	742.86	78.21	440	99	700	747	794	46	999	0.89	25.94
	B	18646	747.56	74.20	440	52	704	748	797	53	999	0.89	24.61
English II	AA	10458	765.89	71.82	440	15	724	763	817	38	999	0.86	26.87
	AB	10434	766.19	71.11	440	12	723	768	808	42	999	0.86	26.61
	BA	7695	766.52	71.19	440	9	724	767	817	15	999	0.84	28.48
	BB	7770	769.18	71.28	440	11	723	773	817	21	999	0.84	28.51
English III	AA	9879	754.09	63.75	440	20	717	757	794	2	999	0.88	22.09
	AB	9747	753.22	65.15	440	21	716	757	796	3	999	0.88	22.57
	BA	8511	758.87	59.65	440	9	725	762	796	3	999	0.88	20.66
	BB	8639	760.16	58.76	440	9	724	762	802	1	999	0.88	20.36
Geometry	A	20131	749.46	76.43	440	81	706	757	798	138	999	0.92	21.62
	B	18011	753.86	69.03	440	28	712	758	794	82	999	0.91	20.71
U.S. History	A	17721	735.26	73.97	440	91	693	739	782	34	999	0.89	24.53
	B	16841	739.52	73.99	440	100	700	744	787	43	999	0.90	23.40

Note: Statistics re-calculated by CTB/McGraw-Hill based on 2012 GRT

Table 22. Winter 2012 Scale Score Statistics

Content	N Count	Mean	SD	LOSS	N Min.	Scale Score Percentile			N Max.	HOSS	Alpha	SEM
						25th	50th	75th				
Algebra I	1028	718.53	57.34	490	14	691	719	752	4	999	0.91	17.38
Algebra II	1006	735.62	92.06	440	20	689	747	787	3	999	0.91	26.95
Biology I	1214	738.26	77.25	440	6	693	747	787	2	999	0.88	26.65
English II	1184	753.52	74.01	440	3	714	754.5	803	2	999	0.87	26.69
English III	1362	749.22	62.96	440	2	713	757	788	2	999	0.85	24.21
Geometry	1348	741.62	75.48	440	11	700.5	746	791	3	999	0.92	21.06
U.S. History	1302	733.99	76.06	440	10	687	739	782	3	999	0.89	24.88

Table 23. Spring 2013 Scale Score Statistics

Content	Form	N Count	Mean	SD	LOSS	N Min.	Scale Score Percentile 25th	50th	75th	N Max.	HOSS	Alpha	SEM
Algebra I	A	20891	739.89	53.98	490	113	711	741	769	90	999	0.91	16.24
	B	18677	740.20	52.90	490	87	712	740	769	90	999	0.91	16.00
Algebra II	A	16023	741.33	83.56	440	189	702	747	787	59	999	0.90	25.77
	B	14230	746.07	77.13	440	79	702	747	794	38	999	0.90	24.61
Biology I	A	22060	694.48	81.67	440	323	654	701	747	17	999	0.90	25.91
	B	15051	700.46	77.95	440	87	656	703	746	21	999	0.88	26.80
English II	AA	9776	765.46	70.03	440	46	726	770	808	11	999	0.87	25.44
	AB	9470	765.50	68.93	440	35	727	770	806	31	999	0.87	25.02
	BA	8911	771.97	64.87	440	16	733	771	812	20	999	0.86	24.43
	BB	8718	771.83	65.37	440	8	733	771	809	44	999	0.87	23.92
English III	AA	10096	763.72	64.95	440	49	730	771	807	3	999	0.88	22.34
	AB	9579	766.34	61.26	440	25	735	772	807	2	999	0.88	20.87
	BA	8556	775.45	55.94	440	9	745	780	809	17	999	0.88	19.68
	BB	8554	776.67	55.35	440	7	746	781	810	5	999	0.88	19.01
Geometry	A	20232	752.85	78.61	440	101	709	754	801	231	999	0.93	21.14
	B	17329	763.20	72.64	440	32	724	764	809	208	999	0.92	20.97
U.S. History	A	17691	741.07	78.25	440	125	700	748	789	27	999	0.91	23.32
	B	15721	749.72	70.53	440	53	706	751	794	11	999	0.89	23.36

Note: Census Data; Suppressed items are not included in data.

Table 24. Winter 2012, State and Subgroup Scale Score Descriptive Data

Content	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
Algebra I	Whole State	1028	718.53	57.34	490	999	0.91	17.38
	Female	489	719.24	54.60	490	999	0.91	16.58
	Male	539	717.88	59.76	490	999	0.91	18.02
	Native American	138	712.14	61.08	490	999	0.90	19.37
	African American	98	690.95	64.86	490	821	0.89	21.60
	Asian	15	812.93	85.39	711	999	0.91	25.06
	Hispanic	109	707.58	56.31	490	837	0.90	18.13
	White	582	727.44	51.03	490	999	0.91	15.48
	Other	84	696.23	46.99	490	794	0.84	18.56
	Pacific Islander	2	743.50	28.99	723	764	0.79	13.25
	IEP	36	686.97	46.83	617	821	0.88	16.25
	Low SES	420	710.76	59.03	490	865	0.91	17.71
	ELL	28	701.11	59.81	565	837	0.91	17.62
	Section 504	15	683.80	67.41	490	745	0.89	21.99
	Accommodated	18	687.61	53.02	565	782	0.88	18.56
Algebra II	Whole State	1006	735.62	92.06	440	999	0.91	26.95
	Female	500	730.31	85.18	440	999	0.90	26.78
	Male	506	740.88	98.19	440	999	0.92	27.01
	Native American	155	728.73	97.87	440	917	0.91	28.81
	African American	70	639.74	109.32	440	967	0.88	37.19
	Asian	20	815.35	98.74	623	999	0.93	26.78
	Hispanic	66	693.39	102.49	440	889	0.92	29.04
	White	640	751.27	77.93	440	999	0.90	24.46
	Other	52	716.10	87.19	440	889	0.91	26.81
	Pacific Islander	3	727.00	63.46	654	769	0.90	19.74
	IEP	62	664.79	109.21	440	917	0.91	32.66
	Low SES	343	699.69	98.67	440	999	0.91	30.23
	ELL	14	639.36	105.94	440	775	0.87	37.71
	Section 504	13	716.31	110.68	440	855	0.93	28.89
	Accommodated	13	643.39	112.93	440	787	0.90	36.54

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program;

Subgroups with sample sizes lower than 10 cases were not reported

Table 24. Winter 2012, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
	Whole State	1214	738.26	77.25	440	999	0.88	26.65
	Female	587	735.18	74.32	440	950	0.87	27.12
	Male	627	741.15	79.86	440	950	0.89	26.21
	Native American	197	737.12	71.70	440	891	0.86	26.40
	African American	88	671.50	82.12	440	874	0.85	31.61
	Asian	14	802.43	78.27	653	950	0.90	24.26
	Hispanic	86	703.76	95.20	440	914	0.90	30.17
Biology	White	763	751.85	69.03	440	950	0.86	25.46
	Other	63	704.03	78.46	481	861	0.88	27.53
	Pacific Islander	3	725.00	86.85	634	807	0.92	24.10
	IEP	65	686.79	90.94	440	874	0.89	30.46
	Low SES	438	712.42	82.88	440	914	0.88	28.40
	ELL	21	631.00	89.37	440	760	0.83	36.41
	Section 504	16	751.50	95.66	440	849	0.90	29.74
	Accommodated	17	652.59	110.78	440	831	0.91	33.28

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Subgroups with sample sizes lower than 10 cases were not reported

Table 24. Winter 2012, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
English II	Whole State	1184	753.52	74.01	440	999	0.87	26.69
	Female	592	760.01	71.50	440	999	0.86	27.11
	Male	592	747.04	75.94	440	935	0.88	26.29
	Native American	168	748.21	69.78	461	935	0.86	26.44
	African American	103	729.43	74.08	440	975	0.85	28.38
	Asian	17	777.35	76.10	627	886	0.90	24.35
	Hispanic	84	723.37	80.12	530	869	0.89	26.18
	White	722	765.28	70.17	440	999	0.86	26.34
	Other	88	720.81	81.37	440	869	0.88	28.71
	Pacific Islander	2	700.50	62.93	656	745	0.87	22.57
	IEP	53	708.91	79.80	461	886	0.87	28.40
	Low SES	466	732.66	73.43	440	975	0.86	27.26
	ELL	19	641.16	61.65	530	739	0.77	29.87
	Section 504	10	655.60	46.86	583	726	0.64	28.01
	Accommodated	13	654.00	52.74	546	739	0.71	28.41
English III	Whole State	1362	749.22	62.96	440	999	0.85	24.21
	Female	671	754.99	60.87	440	923	0.85	23.89
	Male	687	744.11	64.21	440	923	0.86	24.37
	Native American	200	744.60	64.95	440	877	0.86	24.58
	African American	115	720.38	54.50	484	812	0.80	24.59
	Asian	17	757.82	54.99	657	848	0.85	21.14
	Hispanic	81	728.52	62.50	582	866	0.86	23.03
	White	847	758.06	61.24	440	923	0.85	23.82
	Other	99	732.59	66.88	567	923	0.84	26.67
	Pacific Islander	3	723.67	82.25	631	788	0.91	24.26
	IEP	160	682.56	69.82	440	826	0.85	27.10
	Low SES	516	736.42	58.12	484	889	0.84	23.27
	ELL	17	680.71	30.76	623	732	0.43	23.23
	Section 504	21	726.52	60.24	605	840	0.86	22.66
	Accommodated	16	706.06	41.76	645	819	0.72	22.18

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Subgroups with sample sizes lower than 10 cases were not reported

Table 24. Winter 2012, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
Geometry	Whole State	1348	741.62	75.48	440	999	0.92	21.06
	Female	653	742.63	77.62	440	918	0.92	21.30
	Male	695	740.68	73.45	440	999	0.92	20.78
	Native American	184	741.60	76.18	440	999	0.93	20.58
	African American	124	693.22	81.43	440	918	0.90	25.31
	Asian	18	798.00	74.49	701	999	0.89	24.91
	Hispanic	110	724.01	76.04	440	875	0.92	22.13
	White	839	753.70	69.50	440	999	0.92	20.15
	Other	71	695.93	72.69	440	836	0.90	22.52
	Pacific Islander	2	762.50	30.41	741	784	0.67	17.51
	IEP	121	674.86	85.99	440	852	0.90	27.36
	Low SES	500	722.67	70.88	440	999	0.91	21.83
	ELL	18	698.00	74.24	555	836	0.91	21.65
	Section 504	10	736.50	55.75	650	824	0.89	18.43
	Accommodated	12	717.00	88.99	555	836	0.94	21.06
U.S. History	Whole State	1302	733.99	76.06	440	999	0.89	24.88
	Female	661	726.44	72.50	440	999	0.88	24.71
	Male	639	741.96	78.91	440	999	0.90	24.99
	Native American	221	739.39	72.50	440	938	0.88	24.60
	African American	96	680.71	68.69	440	820	0.85	27.00
	Asian	19	745.05	53.54	630	820	0.84	21.46
	Hispanic	83	707.25	98.45	440	999	0.92	28.09
	White	797	743.40	72.31	440	999	0.89	24.37
	Other	84	714.85	75.08	440	883	0.89	25.22
	Pacific Islander	2	752.50	51.62	716	789	0.85	20.16
	IEP	69	690.78	97.98	440	883	0.91	29.69
	Low SES	521	715.54	71.43	440	999	0.87	25.44
	ELL	13	645.54	96.44	440	782	0.90	31.10
	Section 504	23	754.09	76.06	598	938	0.90	24.04
	Accommodated	17	717.24	83.41	584	852	0.91	24.44

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Subgroups with sample sizes lower than 10 cases were not reported

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data

Content	Form	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
Algebra I	A	Whole State	20891	739.89	53.98	490	999	0.91	16.24
		Female	10351	742.68	51.22	490	999	0.90	15.83
		Male	10534	737.20	56.36	490	999	0.91	16.56
		Native American	3288	733.51	50.32	490	999	0.90	15.93
		African American	1952	715.30	59.38	490	999	0.91	18.22
		Asian	441	778.05	64.60	545	999	0.92	18.56
		Hispanic	2552	730.28	52.11	490	999	0.90	16.13
		White	11211	746.24	51.16	490	999	0.91	15.70
		Other	1351	744.75	57.01	490	999	0.91	16.83
		Pacific Islander	96	728.64	70.96	490	999	0.92	20.18
		IEP	1381	701.01	60.17	490	871	0.90	19.13
		Low SES	10239	727.48	52.44	490	999	0.90	16.49
		ELL	851	707.65	56.66	490	999	0.89	18.46
		Section 504	177	734.65	56.30	490	999	0.90	18.20
		Accommodated	1566	699.58	58.87	490	999	0.89	19.55
Algebra I	B	Whole State	18677	740.20	52.90	490	999	0.91	16.00
		Female	9507	741.21	51.37	490	999	0.91	15.82
		Male	9170	739.15	54.44	490	999	0.91	16.14
		Native American	2976	733.85	49.27	490	999	0.90	15.50
		African American	1751	716.18	57.20	490	999	0.90	17.75
		Asian	442	781.51	63.17	599	999	0.91	18.98
		Hispanic	2263	727.64	51.00	490	999	0.90	16.06
		White	10009	747.05	50.39	490	999	0.90	15.60
		Other	1138	743.81	53.77	490	999	0.91	16.05
		Pacific Islander	98	724.05	48.33	599	870	0.89	15.70
		IEP	756	712.41	61.21	490	999	0.92	17.83
		Low SES	9147	727.45	51.46	490	999	0.90	16.12
		ELL	564	714.75	55.64	490	999	0.90	17.41
		Section 504	142	727.71	50.09	490	870	0.90	16.04
		Accommodated	748	707.27	56.23	490	999	0.90	17.56

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Form	Subgroup	Sample Size	Scale Score Mean	SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
Algebra II	A	Whole State	16023	741.33	83.56	440	999	0.90	25.77
		Female	8208	743.14	80.13	440	999	0.90	25.39
		Male	7809	739.50	86.93	440	999	0.91	26.06
		Native American	2467	728.54	81.20	440	999	0.90	26.11
		African American	1373	711.47	84.64	440	999	0.89	28.52
		Asian	458	802.35	91.32	440	999	0.92	25.52
		Hispanic	1837	724.57	79.34	440	999	0.89	26.06
		White	9030	750.29	80.93	440	999	0.90	25.16
		Other	773	736.04	88.32	440	999	0.91	26.06
		Pacific Islander	85	724.58	86.21	440	999	0.90	26.58
		IEP	1401	650.57	98.25	440	999	0.87	35.13
		Low SES	6760	718.06	84.07	440	999	0.89	27.34
		ELL	360	683.58	96.44	440	969	0.90	30.81
		Section 504	126	731.17	81.88	440	891	0.91	24.01
		Accommodated	1351	655.47	97.76	440	999	0.88	34.13
Algebra II	B	Whole State	14230	746.07	77.13	440	999	0.90	24.61
		Female	7340	744.76	74.79	440	999	0.89	24.38
		Male	6890	747.47	79.52	440	999	0.90	24.78
		Native American	2148	733.43	73.46	440	999	0.89	24.70
		African American	1287	710.67	80.08	440	970	0.88	27.60
		Asian	439	799.51	81.85	440	999	0.91	23.88
		Hispanic	1596	730.02	76.35	440	999	0.89	25.42
		White	7954	756.09	74.49	440	999	0.90	24.02
		Other	730	741.93	68.75	440	999	0.88	23.50
		Pacific Islander	76	722.38	84.25	440	858	0.90	26.15
		IEP	687	677.68	92.12	440	999	0.88	31.71
		Low SES	5966	726.52	75.86	440	999	0.89	25.57
		ELL	238	707.62	84.90	440	999	0.89	27.93
		Section 504	111	749.44	67.93	440	922	0.88	23.94
		Accommodated	556	684.99	85.86	440	894	0.87	30.49

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Form	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
Biology	A	Whole State	22060	694.48	81.67	440	999	0.90	25.91
		Female	11211	690.50	76.87	440	999	0.89	25.57
		Male	10834	698.87	85.86	440	999	0.91	26.09
		Native American	3559	687.59	76.96	440	999	0.89	25.79
		African American	1999	653.73	84.59	440	999	0.89	27.72
		Asian	505	721.00	94.55	440	999	0.92	26.45
		Hispanic	2792	664.86	86.33	440	999	0.90	27.46
		White	11893	709.92	75.36	440	999	0.89	25.13
		Other	1186	687.81	84.53	440	999	0.90	26.53
		Pacific Islander	126	691.10	100.74	440	893	0.93	26.29
		IEP	1521	624.00	93.50	440	999	0.90	29.60
		Low SES	10612	671.51	80.83	440	999	0.89	26.49
		ELL	803	608.14	91.47	440	893	0.89	31.00
		Section 504	165	693.91	87.83	440	960	0.91	27.00
		Accommodated	1702	615.45	91.26	440	999	0.89	30.45
Biology	B	Whole State	15051	700.46	77.95	440	999	0.88	26.80
		Female	7705	695.29	74.03	440	999	0.87	26.46
		Male	7346	705.87	81.51	440	999	0.89	27.03
		Native American	2361	692.16	72.84	440	999	0.87	26.40
		African American	1413	657.80	78.64	440	999	0.87	28.64
		Asian	394	733.69	82.00	440	999	0.90	26.54
		Hispanic	1753	673.31	79.07	440	999	0.88	27.66
		White	8254	714.15	74.03	440	999	0.87	26.46
		Other	798	702.70	77.63	440	955	0.88	26.33
		Pacific Islander	78	694.62	83.53	440	837	0.90	26.71
		IEP	626	643.67	90.72	440	955	0.90	29.36
		Low SES	7033	678.06	77.39	440	999	0.87	27.47
		ELL	356	614.03	84.82	440	955	0.86	31.57
		Section 504	126	700.52	75.18	476	955	0.87	26.84
		Accommodated	630	632.08	89.18	440	955	0.88	30.32

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Form	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
English II	AA	Whole State	9776	765.46	70.03	440	999	0.87	25.44
		Female	5004	775.26	65.79	440	999	0.85	25.11
		Male	4772	755.18	72.82	440	999	0.88	25.65
		Native American	1548	757.66	68.52	440	999	0.86	25.22
		African American	882	731.43	75.89	440	999	0.87	27.17
		Asian	234	777.56	80.74	440	999	0.89	26.36
		Hispanic	1145	742.21	70.55	440	965	0.87	25.57
		White	5323	777.64	65.50	440	999	0.85	25.14
		Other	604	766.07	68.32	440	965	0.86	25.39
		Pacific Islander	40	781.85	66.96	521	928	0.86	24.89
		IEP	622	694.14	90.15	440	928	0.89	30.02
		Low SES	4428	746.56	70.48	440	999	0.87	25.77
		ELL	229	682.34	80.98	440	928	0.86	30.57
		Section 504	87	754.75	69.79	440	999	0.84	27.59
		Accommodated	631	687.29	90.36	440	999	0.89	30.60
English II	AB	Whole State	9470	765.50	68.93	440	999	0.87	25.02
		Female	4832	775.24	64.16	440	999	0.86	24.32
		Male	4638	755.36	72.19	440	999	0.88	25.51
		Native American	1508	756.71	62.87	440	999	0.85	24.29
		African American	862	738.46	71.10	440	999	0.87	25.61
		Asian	220	782.16	80.95	440	999	0.89	26.66
		Hispanic	1148	740.60	72.49	440	947	0.87	25.75
		White	5080	777.16	65.29	440	999	0.86	24.83
		Other	589	768.60	70.18	440	999	0.87	25.04
		Pacific Islander	63	773.06	85.65	471	947	0.91	25.22
		IEP	565	692.69	85.63	440	891	0.89	28.70
		Low SES	4426	747.37	67.58	440	999	0.86	24.89
		ELL	216	669.39	87.27	440	891	0.87	31.23
		Section 504	73	743.44	66.23	572	873	0.88	23.25
		Accommodated	568	683.99	86.36	440	891	0.88	29.32

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Form	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
English II	BA	Whole State	8911	771.97	64.87	440	999	0.86	24.43
		Female	4614	779.01	64.17	440	999	0.85	24.58
		Male	4297	764.42	64.78	440	999	0.86	24.13
		Native American	1393	766.38	57.87	524	999	0.83	23.52
		African American	794	741.08	69.16	440	959	0.88	24.43
		Asian	220	788.23	67.58	440	999	0.87	24.23
		Hispanic	1045	747.53	69.81	440	959	0.87	24.74
		White	4863	782.77	61.94	440	999	0.84	24.77
		Other	550	775.91	59.87	561	999	0.83	24.34
		Pacific Islander	46	764.41	77.42	544	895	0.91	22.78
		IEP	382	720.61	74.66	440	921	0.89	24.79
		Low SES	4188	753.96	63.94	440	999	0.86	23.86
		ELL	187	685.55	85.92	440	999	0.89	28.18
		Section 504	69	768.86	45.89	623	859	0.76	22.26
		Accommodated	365	708.24	80.88	440	921	0.90	26.00
English II	BB	Whole State	8718	771.83	65.37	440	999	0.87	23.92
		Female	4517	779.42	64.25	440	999	0.86	24.01
		Male	4201	763.67	65.59	440	999	0.87	23.68
		Native American	1372	765.12	60.29	440	999	0.85	23.68
		African American	791	742.07	64.18	440	940	0.87	23.16
		Asian	227	794.73	83.32	527	999	0.91	24.78
		Hispanic	1020	746.96	69.46	440	999	0.88	23.77
		White	4718	782.85	62.14	440	999	0.85	24.15
		Other	547	771.46	59.90	588	999	0.84	23.76
		Pacific Islander	43	797.67	74.57	563	940	0.91	22.12
		IEP	402	723.57	67.88	465	940	0.88	23.51
		Low SES	4020	753.38	63.50	440	999	0.87	23.28
		ELL	190	671.35	80.52	440	867	0.89	27.26
		Section 504	80	759.74	65.11	465	852	0.88	23.01
		Accommodated	382	706.12	77.61	440	940	0.90	24.76

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Form	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
English III	AA	Whole State	10096	763.72	64.95	440	999	0.88	22.34
		Female	4893	771.88	61.30	440	999	0.88	21.45
		Male	5199	756.16	67.13	440	999	0.88	22.83
		Native American	1648	756.50	64.35	440	958	0.88	22.32
		African American	935	731.05	69.51	440	908	0.87	25.42
		Asian	212	787.41	63.78	440	958	0.89	21.52
		Hispanic	1148	748.75	59.57	440	999	0.86	22.20
		White	5549	773.31	63.26	440	999	0.88	21.94
		Other	556	766.53	59.61	440	892	0.87	21.69
		Pacific Islander	48	759.90	60.06	596	879	0.89	20.31
		IEP	1586	685.14	72.98	440	868	0.83	29.83
		Low SES	4611	745.90	65.14	440	958	0.87	23.38
		ELL	239	698.29	60.91	440	824	0.80	27.32
		Section 504	85	762.95	49.29	596	892	0.82	20.91
		Accommodated	1431	684.83	72.37	440	892	0.83	29.67
English III	AB	Whole State	9579	766.34	61.26	440	999	0.88	20.87
		Female	4712	774.67	57.02	440	952	0.88	19.96
		Male	4866	758.28	64.09	440	999	0.89	21.45
		Native American	1594	758.41	61.24	440	952	0.88	21.40
		African American	912	737.06	63.22	440	952	0.88	22.15
		Asian	205	782.04	57.14	567	890	0.88	19.82
		Hispanic	1072	747.98	59.22	440	925	0.87	21.68
		White	5259	776.71	58.88	440	999	0.88	20.37
		Other	488	770.96	55.41	510	905	0.87	19.76
		Pacific Islander	49	745.22	77.76	440	905	0.91	23.82
		IEP	1378	691.87	67.34	440	877	0.84	26.69
		Low SES	4292	747.52	62.34	440	952	0.88	21.58
		ELL	218	695.31	58.68	510	836	0.82	24.57
		Section 504	82	765.46	48.75	610	905	0.83	20.01
		Accommodated	1218	691.26	66.90	440	877	0.84	26.65

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Form	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
English III	BA	Whole State	8556	775.45	55.94	440	999	0.88	19.68
		Female	4281	783.16	54.23	440	999	0.87	19.48
		Male	4275	767.73	56.56	440	999	0.88	19.70
		Native American	1399	769.34	54.15	440	950	0.87	19.50
		African American	810	751.90	58.79	440	950	0.88	20.13
		Asian	236	792.95	59.05	554	999	0.89	19.82
		Hispanic	916	757.59	54.04	440	950	0.87	19.77
		White	4688	784.01	53.53	440	999	0.87	19.42
		Other	467	774.89	57.75	440	999	0.87	20.48
		Pacific Islander	40	774.80	69.61	440	878	0.89	23.43
		IEP	735	707.26	67.39	440	950	0.89	22.82
		Low SES	3615	759.44	56.96	440	999	0.88	20.02
		ELL	159	706.03	66.03	440	842	0.87	23.76
		Section 504	59	767.86	58.10	626	950	0.88	20.00
		Accommodated	596	704.96	67.52	440	950	0.88	23.40
English III	BB	Whole State	8554	776.67	55.35	440	999	0.88	19.01
		Female	4329	782.16	52.59	440	999	0.87	18.61
		Male	4225	771.05	57.51	440	999	0.89	19.27
		Native American	1300	772.08	53.45	440	994	0.88	18.56
		African American	785	749.68	56.35	490	943	0.88	19.21
		Asian	215	795.36	56.86	619	999	0.88	19.31
		Hispanic	946	756.74	53.99	440	943	0.87	19.12
		White	4777	785.61	53.33	440	999	0.87	18.94
		Other	488	775.59	53.43	533	999	0.87	19.42
		Pacific Islander	43	773.12	59.76	490	877	0.88	21.09
		IEP	683	707.33	66.21	440	914	0.89	22.37
		Low SES	3663	759.76	55.07	440	994	0.88	19.04
		ELL	165	708.60	59.86	440	842	0.86	22.54
		Section 504	82	773.01	49.95	600	893	0.85	19.13
		Accommodated	582	706.08	65.58	440	863	0.88	22.75

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Form	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
Geometry	A	Whole State	20232	752.85	78.61	440	999	0.93	21.14
		Female	10129	754.38	75.76	440	999	0.92	20.80
		Male	10093	751.41	81.30	440	999	0.93	21.41
		Native American	3363	742.71	73.49	440	999	0.92	20.57
		African American	1925	712.35	78.32	440	999	0.92	22.68
		Asian	474	808.05	86.77	440	999	0.93	22.91
		Hispanic	2382	733.61	75.37	440	999	0.92	21.38
		White	10901	764.93	76.16	440	999	0.93	20.75
		Other	1077	753.80	77.66	440	999	0.92	21.36
		Pacific Islander	110	744.88	86.53	440	913	0.94	21.54
		IEP	2902	669.95	78.40	440	999	0.89	25.70
		Low SES	9582	730.68	76.42	440	999	0.92	21.52
		ELL	636	688.35	81.69	440	999	0.91	24.39
		Section 504	149	757.87	80.87	440	999	0.93	21.72
		Accommodated	2734	671.24	78.37	440	999	0.89	25.66
Geometry	B	Whole State	17329	763.20	72.64	440	999	0.92	20.97
		Female	8776	762.89	71.91	440	999	0.92	20.86
		Male	8553	763.51	73.39	440	999	0.92	21.04
		Native American	2707	753.98	66.62	440	999	0.91	20.46
		African American	1647	720.75	71.53	440	999	0.91	21.93
		Asian	450	815.50	76.97	595	999	0.90	24.18
		Hispanic	1991	743.57	69.75	440	999	0.91	20.66
		White	9482	774.57	69.97	440	999	0.91	20.84
		Other	960	764.64	75.27	440	999	0.92	21.22
		Pacific Islander	92	775.83	65.00	619	999	0.91	19.55
		IEP	1200	686.02	80.14	440	999	0.91	23.97
		Low SES	7911	742.67	70.08	440	999	0.91	20.97
		ELL	367	707.49	83.42	440	999	0.92	23.81
		Section 504	127	755.24	77.61	440	999	0.92	22.58
		Accommodated	1021	685.30	79.64	440	999	0.91	24.45

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 25. Spring 2013, State and Subgroup Scale Score Descriptive Data *(continued)*

Content	Form	Subgroup	Sample Size	Scale Score Mean	Scale Score SD	Min Scale Score Obtained	Max Scale Score Obtained	Coefficient Alpha	SEM
U.S. History	A	Whole State	17691	741.07	78.25	440	999	0.91	23.32
		Female	8886	727.89	73.36	440	999	0.90	23.10
		Male	8798	754.54	80.52	440	999	0.92	23.39
		Native American	2793	734.86	74.17	440	999	0.90	23.11
		African American	1662	703.03	82.92	440	987	0.90	25.72
		Asian	469	750.84	91.75	440	999	0.93	24.54
		Hispanic	2071	716.94	80.67	440	999	0.91	24.59
		White	9641	753.24	73.56	440	999	0.91	22.59
		Other	953	749.74	79.83	440	999	0.92	22.98
		Pacific Islander	102	744.88	79.55	480	932	0.92	22.70
		IEP	1207	687.66	93.86	440	987	0.92	27.06
		Low SES	7868	720.41	77.84	440	999	0.90	24.10
		ELL	469	655.48	88.57	440	901	0.89	29.70
		Section 504	157	730.38	90.86	440	987	0.92	25.20
		Accommodated	1193	675.54	94.45	440	987	0.91	27.85
U.S. History	B	Whole State	15721	749.72	70.53	440	999	0.89	23.36
		Female	7909	738.31	67.11	440	999	0.88	23.38
		Male	7812	761.28	72.01	440	999	0.90	23.26
		Native American	2463	745.33	67.96	440	999	0.88	23.05
		African American	1497	716.38	70.24	440	976	0.87	25.44
		Asian	411	765.22	75.26	440	999	0.90	23.96
		Hispanic	1818	731.06	70.25	440	928	0.88	24.34
		White	8612	760.09	68.18	440	999	0.89	22.86
		Other	827	748.99	72.05	440	976	0.89	23.35
		Pacific Islander	93	746.25	71.32	555	867	0.91	21.84
		IEP	683	706.36	82.50	440	928	0.89	27.05
		Low SES	6925	730.48	69.41	440	976	0.88	24.18
		ELL	297	673.94	80.89	440	854	0.86	30.43
		Section 504	126	746.46	72.60	578	999	0.89	23.93
		Accommodated	600	691.36	88.66	440	999	0.89	28.81

*SEM=Standard Error of Measurement; ELL=English Language Learner; IEP=Individualized Education Program

Table 26. Spring 2013, Subgroup Scale Score Mean Differences, t-test: Male/Female

Content	Form	t	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Algebra I	A	7.36	20756.86	<.0001	5.48	53.87
	B	2.67	18511.49	0.008	2.07	52.90
Algebra II	A	2.75	15745.17	0.006	3.64	83.51
	B	-2.09	14010.90	0.037	-2.71	77.12
Biology I	A	-7.61	21593.95	<.0001	-8.36	81.41
	B	-8.32	14745.92	<.0001	-10.58	77.77
English II	AA	14.28	9563.15	<.0001	20.07	69.31
	AB	14.14	9237.26	<.0001	19.88	68.21
	BA	10.68	8909.00	<.0001	14.59	64.46
	BB	11.32	8716.00	<.0001	15.74	64.89
English III	AA	12.29	10080.93	<.0001	15.71	64.37
	AB	13.23	9508.33	<.0001	16.38	60.71
	BA	12.88	8537.81	<.0001	15.43	55.41
	BB	9.32	8443.48	<.0001	11.12	55.07
Geometry	A	2.69	20109.97	0.007	2.97	78.57
	B	-0.56	17327.00	0.574	-0.62	72.64
U.S. History	A	-23.00	17497.41	<.0001	-26.65	77.00
	B	-20.69	15612.47	<.0001	-22.98	69.59

Table 27. Spring 2013, Subgroup Scale Score Mean Differences, t-test: IEP/Non IEP

Content	Form	t	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Algebra I	A	-25.05	1532.02	<.0001	-41.63	52.98
	B	-12.81	802.01	<.0001	-28.96	52.60
Algebra II	A	-36.83	1567.11	<.0001	-99.45	78.69
	B	-20.11	732.41	<.0001	-71.86	75.57
Biology I	A	-30.79	1681.41	<.0001	-75.70	79.39
	B	-16.09	664.03	<.0001	-59.25	77.05
English II	AA	-20.70	666.58	<.0001	-76.17	67.52
	AB	-21.11	605.99	<.0001	-77.43	66.45
	BA	-13.83	406.02	<.0001	-53.66	63.96
	BB	-15.36	8716.00	<.0001	-50.60	64.50
English III	AA	-48.67	1889.20	<.0001	-93.23	55.38
	AB	-45.84	1645.22	<.0001	-86.98	53.11
	BA	-29.26	812.31	<.0001	-74.60	51.89
	BB	-29.04	750.88	<.0001	-75.36	51.44
Geometry	A	-62.51	3705.88	<.0001	-96.79	70.91
	B	-34.90	1333.28	<.0001	-82.92	69.53
U.S. History	A	-20.73	1322.83	<.0001	-57.32	76.91
	B	-14.14	726.38	<.0001	-45.33	69.92

Table 28. Spring 2013, Subgroup Scale Score Mean Differences, t-test: Free Lunch SES/Non Free Lunch

Content	Form	t	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Algebra I	A	-33.44	20889.00	<.0001	-24.34	52.59
	B	-33.21	18675.00	<.0001	-24.99	51.41
Algebra II	A	-30.70	14023.59	<.0001	-40.25	81.16
	B	-26.30	14228.00	<.0001	-33.66	75.32
Biology I	A	-41.70	21688.52	<.0001	-44.27	78.62
	B	-34.15	14528.10	<.0001	-42.04	75.07
English II	AA	-24.88	9160.25	<.0001	-34.55	67.89
	AB	-24.74	9468.00	<.0001	-34.03	66.81
	BA	-25.52	8685.40	<.0001	-34.00	62.62
	BB	-25.26	8716.00	<.0001	-34.24	63.10
English III	AA	-25.96	9541.13	<.0001	-32.80	62.86
	AB	-27.87	8709.19	<.0001	-34.09	58.87
	BA	-23.03	7377.61	<.0001	-27.72	54.24
	BB	-25.16	7642.03	<.0001	-29.58	53.38
Geometry	A	-39.51	20230.00	<.0001	-42.13	75.74
	B	-35.29	17327.00	<.0001	-37.76	70.17
U.S. History	A	-32.19	16529.11	<.0001	-37.21	76.04
	B	-31.18	14686.20	<.0001	-34.39	68.43

Table 29. Spring 2013, Subgroup Scale Score Mean Differences, t-test: ELL/Non ELL

Content	Form	t	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Algebra I	A	-16.99	915.35	<.0001	-33.61	53.57
	B	-11.64	18675.00	<.0001	-26.24	52.71
Algebra II	A	-11.53	371.26	<.0001	-59.08	83.10
	B	-7.06	243.65	<.0001	-39.10	76.96
Biology I	A	-27.37	848.36	<.0001	-89.60	79.93
	B	-19.50	369.15	<.0001	-88.52	76.78
English II	AA	-15.77	235.90	<.0001	-85.12	68.84
	AB	-16.45	220.92	<.0001	-98.36	67.35
	BA	-13.97	190.32	<.0001	-88.28	63.63
	BB	-17.46	194.22	<.0001	-102.70	63.63
English III	AA	-15.96	10094.00	<.0001	-67.02	64.14
	AB	-17.59	9577.00	<.0001	-72.68	60.30
	BA	-13.42	162.16	<.0001	-70.74	55.12
	BB	-16.20	8552.00	<.0001	-69.41	54.52
Geometry	A	-21.26	20230.00	<.0001	-66.59	77.74
	B	-12.97	377.87	<.0001	-56.92	72.18
U.S. History	A	-21.28	487.27	<.0001	-87.92	76.97
	B	-16.34	304.48	<.0001	-77.25	69.74

Table 30. Spring 2013, Mathematics Subgroup Scale Score Mean Differences, t-test: Section 504/Non Section 504

Content	Form	t	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Algebra I	A	-1.30	20889.00	0.194	-5.29	53.98
	B	-2.82	18675.00	0.005	-12.58	52.89
Algebra II	A	-1.37	16021.00	0.171	-10.24	83.55
	B	0.46	14228.00	0.644	3.40	77.13
Biology I	A	-0.09	22058.00	0.928	-0.58	81.67
	B	0.01	15049.00	0.993	0.06	77.95
English II	AA	-1.43	9774.00	0.152	-10.81	70.03
	AB	-2.75	9468.00	0.006	-22.23	68.90
	BA	-0.56	70.15	0.574	-3.14	64.88
	BB	-1.66	8716.00	0.097	-12.20	65.36
English III	AA	-0.14	86.50	0.887	-0.77	64.95
	AB	-0.16	83.23	0.871	-0.88	61.26
	BA	-1.05	8554.00	0.296	-7.64	55.94
	BB	-0.60	8552.00	0.547	-3.70	55.35
Geometry	A	0.78	20230.00	0.434	5.06	78.61
	B	-1.24	17327.00	0.215	-8.02	72.64
U.S. History	A	-1.48	158.07	0.140	-10.79	78.25
	B	-0.52	15719.00	0.602	-3.29	70.53

Table 31. Spring 2013, Mathematics Subgroup Scale Score Mean Differences, t-test: Accommodated/Non Accommodated

Content	Form	t	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Algebra I	A	-5.44	20889.00	<.0001	-14.79	53.94
	B	-5.42	18675.00	<.0001	-15.26	52.86
Algebra II	A	-4.80	182.14	<.0001	-34.14	83.48
	B	-2.44	187.90	0.016	-15.49	77.11
Biology I	A	-7.64	329.10	<.0001	-40.44	81.53
	B	-7.38	221.61	<.0001	-45.95	77.76
English II	AA	-7.43	9774.00	<.0001	-44.80	69.84
	AB	-8.96	9468.00	<.0001	-56.97	68.64
	BA	-6.76	131.34	<.0001	-49.07	64.61
	BB	-8.60	137.58	<.0001	-61.73	64.92
English III	AA	-7.76	121.76	<.0001	-35.92	64.83
	AB	-6.56	9577.00	<.0001	-36.12	61.13
	BA	-6.24	102.84	<.0001	-39.85	55.77
	BB	-6.32	8552.00	<.0001	-33.51	55.22
Geometry	A	-6.61	286.18	<.0001	-35.21	78.50
	B	-5.71	244.92	<.0001	-31.45	72.55
U.S. History	A	-6.33	229.86	<.0001	-36.30	78.15
	B	-6.72	15719.00	<.0001	-33.02	70.43

Table 32. Spring 2013, Subgroup Mean Differences, ANOVA: Ethnicity

Content	Form	Dependent	Categories	Sum of Squares	DF	Mean Square	F	Sig.
Algebra I	A	SS	Between Groups	2688627	6	448105	161	<.0001
			Within Groups	58176756	20884	2786	.	.
			Total	60865383	20890	.	.	.
	B	SS	Between Groups	2751765	6	458628	173	<.0001
			Within Groups	49519073	18670	2652	.	.
			Total	52270838	18676	.	.	.
Algebra II	A	SS	Between Groups	4619533	6	769922	115	<.0001
			Within Groups	107238395	16016	6696	.	.
			Total	111857929	16022	.	.	.
	B	SS	Between Groups	4474710	6	745785	132	<.0001
			Within Groups	80164017	14223	5636	.	.
			Total	84638727	14229	.	.	.
Biology	A	SS	Between Groups	9184449	6	1530741	245	<.0001
			Within Groups	137949924	22053	6255	.	.
			Total	147134373	22059	.	.	.
	B	SS	Between Groups	6015740	6	1002623	177	<.0001
			Within Groups	85421736	15044	5678	.	.
			Total	91437476	15050	.	.	.

Table 32. Spring 2013, Subgroup Mean Differences, ANOVA: Ethnicity *(continued)*

Content	Form	Dependent	Categories	Sum of Squares	DF	Mean Square	F	Sig.
English II	AA	SS	Between Groups	2568988	6	428165	92	<.0001
			Within Groups	45369619	9769	4644	.	.
			Total	47938607	9775	.	.	.
	AB	SS	Between Groups	2218888	6	369815	82	<.0001
			Within Groups	42769163	9463	4520	.	.
			Total	44988051	9469	.	.	.
	BA	SS	Between Groups	2061806	6	343634	86	<.0001
			Within Groups	35434942	8904	3980	.	.
			Total	37496748	8910	.	.	.
	BB	SS	Between Groups	2114056	6	352343	87	<.0001
			Within Groups	35130763	8711	4033	.	.
			Total	37244819	8717	.	.	.
English III	AA	SS	Between Groups	1975029	6	329171	82	<.0001
			Within Groups	40605122	10089	4025	.	.
			Total	42580151	10095	.	.	.
	AB	SS	Between Groups	1892063	6	315344	89	<.0001
			Within Groups	34051369	9572	3557	.	.
			Total	35943432	9578	.	.	.
	BA	SS	Between Groups	1209896	6	201649	67	<.0001
			Within Groups	25557881	8549	2990	.	.
			Total	26767777	8555	.	.	.
	BB	SS	Between Groups	1433046	6	238841	82	<.0001
			Within Groups	24767903	8547	2898	.	.
			Total	26200949	8553	.	.	.

Table 32. Spring 2013, Subgroup Mean Differences, ANOVA: Ethnicity *(continued)*

Content	Form	Dependent	Categories	Sum of Squares	DF	Mean Square	F	Sig.
Geometry	A	SS	Between Groups	7427751	6	1237959	213	<.0001
			Within Groups	117576386	20225	5813	.	.
			Total	125004137	20231	.	.	.
	B	SS	Between Groups	6438550	6	1073092	219	<.0001
			Within Groups	85002387	17322	4907	.	.
			Total	91440937	17328	.	.	.
U.S. History	A	SS	Between Groups	5265435	6	877572	151	<.0001
			Within Groups	103061092	17684	5828	.	.
			Total	108326527	17690	.	.	.
	B	SS	Between Groups	3370039	6	561673	118	<.0001
			Within Groups	74824213	15714	4762	.	.
			Total	78194251	15720	.	.	.

Table 33. Ethnic Pair-wise Dunnett's C Post-hoc Comparison

Content	Form	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Algebra I	A	SS	African American	Native American	18.21(*)	2.58	14.32	22.10
				Hispanic	14.98(*)	2.58	10.89	19.08
				Asian	62.75(*)	2.58	55.57	69.92
				Pacific Islander	13.34	2.58	-0.89	27.56
				White	30.94(*)	2.58	27.61	34.28
				Other	29.45(*)	2.58	24.64	34.27
			Native American	African American	-18.21(*)	2.60	-22.14	-14.28
				Hispanic	-3.22	2.60	-6.85	0.40
				Asian	44.54(*)	2.60	37.57	51.51
				Pacific Islander	-4.87	2.60	-19.11	9.36
				White	12.74(*)	2.60	10.01	15.46
				Other	11.24(*)	2.60	6.80	15.69
			Hispanic	African American	-14.98(*)	2.59	-19.10	-10.87
				Native American	3.22	2.59	-0.39	6.83
				Asian	47.76(*)	2.59	40.70	54.82
				Pacific Islander	-1.65	2.59	-15.88	12.58
				White	15.96(*)	2.59	12.96	18.96
				Other	14.47(*)	2.59	9.86	19.07
			Asian	African American	-62.75(*)	2.44	-69.54	-55.96
				Native American	-44.54(*)	2.44	-51.07	-38.01
				Hispanic	-47.76(*)	2.44	-54.41	-41.12
				Pacific Islander	-49.41(*)	2.44	-63.92	-34.91
				White	-31.8(*)	2.44	-38.06	-25.55
				Other	-33.3(*)	2.44	-40.36	-26.23
			Pacific Islander	African American	-13.34(*)	2.23	-25.66	-1.02
				Native American	4.87	2.23	-7.33	17.07
				Hispanic	1.65	2.23	-10.60	13.90
				Asian	49.41(*)	2.23	36.14	62.69
				White	17.61(*)	2.23	5.53	29.69
				Other	16.12(*)	2.23	3.67	28.56
			White	African American	-30.94(*)	2.63	-34.35	-27.54
				Native American	-12.74(*)	2.63	-15.49	-9.99
				Hispanic	-15.96(*)	2.63	-19.00	-12.92
				Asian	31.8(*)	2.63	25.07	38.54
				Pacific Islander	-17.61(*)	2.63	-31.83	-3.39
				Other	-1.49	2.63	-5.49	2.50
			Other	African American	-29.45(*)	2.55	-34.22	-24.68
				Native American	-11.24(*)	2.55	-15.60	-6.89
				Hispanic	-14.47(*)	2.55	-19.00	-9.93
				Asian	33.3(*)	2.55	25.91	40.68
				Pacific Islander	-16.12(*)	2.55	-30.35	-1.89
				White	1.49	2.55	-2.39	5.37

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
Algebra I	B	SS	African American	Native American	17.67(*)	2.58	13.67	21.67
				Hispanic	11.46(*)	2.58	7.23	15.68
				Asian	65.33(*)	2.58	58.26	72.40
				Pacific Islander	7.87	2.58	-5.91	21.66
				White	30.87(*)	2.58	27.43	34.31
				Other	27.63(*)	2.58	22.58	32.69
			Native American	African American	-17.67(*)	2.61	-21.71	-13.63
				Hispanic	-6.21(*)	2.61	-9.96	-2.47
				Asian	47.66(*)	2.61	40.82	54.50
				Pacific Islander	-9.80	2.61	-23.57	3.98
				White	13.2(*)	2.61	10.40	16.00
				Other	9.96(*)	2.61	5.28	14.64
			Hispanic	African American	-11.46(*)	2.59	-15.71	-7.21
				Native American	6.21(*)	2.59	2.49	9.94
				Asian	53.87(*)	2.59	46.92	60.81
				Pacific Islander	-3.59	2.59	-17.36	10.19
				White	19.41(*)	2.59	16.31	22.52
				Other	16.17(*)	2.59	11.32	21.03
			Asian	African American	-65.33(*)	2.45	-72.05	-58.61
				Native American	-47.66(*)	2.45	-54.09	-41.22
				Hispanic	-53.87(*)	2.45	-60.43	-47.30
				Pacific Islander	-57.46(*)	2.45	-71.55	-43.36
				White	-34.46(*)	2.45	-40.59	-28.32
				Other	-37.7(*)	2.45	-44.77	-30.62
			Pacific Islander	African American	-7.87	2.24	-19.87	4.12
				Native American	9.80	2.24	-2.06	21.66
				Hispanic	3.59	2.24	-8.34	15.51
				Asian	57.46(*)	2.24	44.55	70.36
				White	23(*)	2.24	11.27	34.73
				Other	19.76(*)	2.24	7.59	31.93
			White	African American	-30.87(*)	2.63	-34.38	-27.37
				Native American	-13.2(*)	2.63	-16.03	-10.37
				Hispanic	-19.41(*)	2.63	-22.56	-16.26
				Asian	34.46(*)	2.63	27.88	41.04
				Pacific Islander	-23(*)	2.63	-36.74	-9.26
				Other	-3.24	2.63	-7.47	0.99
			Other	African American	-27.63(*)	2.55	-32.63	-22.64
				Native American	-9.96(*)	2.55	-14.53	-5.39
				Hispanic	-16.17(*)	2.55	-20.94	-11.41
				Asian	37.7(*)	2.55	30.34	45.05
				Pacific Islander	-19.76(*)	2.55	-33.57	-5.95
				White	3.24	2.55	-0.86	7.34

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Algebra II	A	SS	African American	Native American	17.07(*)	2.58	9.97	24.16
				Hispanic	13.1(*)	2.58	5.58	20.62
				Asian	90.89(*)	2.58	79.51	102.26
				Pacific Islander	13.11	2.58	-10.45	36.67
				White	38.82(*)	2.58	32.71	44.92
				Other	24.57(*)	2.58	15.09	34.04
			Native American	African American	-17.07(*)	2.61	-24.25	-9.89
				Hispanic	-3.97	2.61	-10.54	2.61
				Asian	73.82(*)	2.61	62.97	84.67
				Pacific Islander	-3.96	2.61	-27.48	19.56
				White	21.75(*)	2.61	16.91	26.59
				Other	7.50	2.61	-1.29	16.29
			Hispanic	African American	-13.1(*)	2.59	-20.67	-5.53
				Native American	3.97	2.59	-2.57	10.50
				Asian	77.78(*)	2.59	66.71	88.86
				Pacific Islander	0.01	2.59	-23.53	23.54
				White	25.72(*)	2.59	20.29	31.15
				Other	11.47(*)	2.59	2.37	20.56
			Asian	African American	-90.89(*)	2.48	-101.82	-79.95
				Native American	-73.82(*)	2.48	-84.13	-63.51
				Hispanic	-77.78(*)	2.48	-88.37	-67.20
				Pacific Islander	-77.78(*)	2.48	-101.72	-53.84
				White	-52.07(*)	2.48	-61.78	-42.36
				Other	-66.32(*)	2.48	-78.27	-54.37
			Pacific Islander	African American	-13.11	2.25	-33.66	7.44
				Native American	3.96	2.25	-16.32	24.24
				Hispanic	-0.01	2.25	-20.40	20.39
				Asian	77.78(*)	2.25	56.06	99.49
				White	25.71(*)	2.25	5.67	45.74
				Other	11.46	2.25	-9.55	32.47
			White	African American	-38.82(*)	2.63	-45.05	-32.59
				Native American	-21.75(*)	2.63	-26.64	-16.86
				Hispanic	-25.72(*)	2.63	-31.22	-20.21
				Asian	52.07(*)	2.63	41.76	62.37
				Pacific Islander	-25.71(*)	2.63	-49.15	-2.27
				Other	-14.25(*)	2.63	-22.31	-6.19
			Other	African American	-24.57(*)	2.53	-33.88	-15.26
				Native American	-7.50	2.53	-16.04	1.03
				Hispanic	-11.47(*)	2.53	-20.35	-2.59
				Asian	66.32(*)	2.53	54.11	78.53
				Pacific Islander	-11.46	2.53	-35.12	12.20
				White	14.25(*)	2.53	6.49	22.01

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
Algebra II	B	SS	African American	Native American	22.76(*)	2.58	15.94	29.58
				Hispanic	19.36(*)	2.58	12.11	26.61
				Asian	88.84(*)	2.58	78.15	99.54
				Pacific Islander	11.72	2.58	-11.13	34.56
				White	45.42(*)	2.58	39.61	51.24
				Other	31.26(*)	2.58	22.29	40.23
			Native American	African American	-22.76(*)	2.60	-29.65	-15.87
				Hispanic	-3.41	2.60	-9.87	3.06
				Asian	66.08(*)	2.60	55.84	76.32
				Pacific Islander	-11.05	2.60	-33.87	11.78
				White	22.66(*)	2.60	17.91	27.41
				Other	8.5(*)	2.60	0.12	16.88
			Hispanic	African American	-19.36(*)	2.59	-26.64	-12.07
				Native American	3.41	2.59	-3.02	9.83
				Asian	69.49(*)	2.59	59.01	79.97
				Pacific Islander	-7.64	2.59	-30.47	15.19
				White	26.07(*)	2.59	20.73	31.40
				Other	11.91(*)	2.59	3.22	20.60
			Asian	African American	-88.84(*)	2.48	-99.14	-78.54
				Native American	-66.08(*)	2.48	-75.84	-56.32
				Hispanic	-69.49(*)	2.48	-79.53	-59.45
				Pacific Islander	-77.13(*)	2.48	-100.28	-53.98
				White	-43.42(*)	2.48	-52.56	-34.29
				Other	-57.58(*)	2.48	-68.84	-46.33
			Pacific Islander	African American	-11.72	2.24	-31.59	8.16
				Native American	11.05	2.24	-8.61	30.70
				Hispanic	7.64	2.24	-12.13	27.41
				Asian	77.13(*)	2.24	56.21	98.05
				White	33.71(*)	2.24	14.30	53.11
				Other	19.55	2.24	-0.75	39.84
			White	African American	-45.42(*)	2.63	-51.35	-39.49
				Native American	-22.66(*)	2.63	-27.46	-17.86
				Hispanic	-26.07(*)	2.63	-31.48	-20.65
				Asian	43.42(*)	2.63	33.75	53.10
				Pacific Islander	-33.71(*)	2.63	-56.45	-10.96
				Other	-14.16(*)	2.63	-21.79	-6.53
			Other	African American	-31.26(*)	2.53	-40.08	-22.45
				Native American	-8.5(*)	2.53	-16.65	-0.35
				Hispanic	-11.91(*)	2.53	-20.41	-3.41
				Asian	57.58(*)	2.53	46.09	69.07
				Pacific Islander	-19.55	2.53	-42.48	3.38
				White	14.16(*)	2.53	6.80	21.52

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
Biology	A	SS	African American	Native American	33.87(*)	2.58	28.17	39.56
				Hispanic	11.13(*)	2.58	5.16	17.10
				Asian	67.28(*)	2.58	57.12	77.43
				Pacific Islander	37.37(*)	2.58	18.65	56.09
				White	56.19(*)	2.58	51.27	61.12
				Other	34.08(*)	2.58	26.61	41.55
			Native American	African American	-33.87(*)	2.61	-39.63	-28.10
				Hispanic	-22.74(*)	2.61	-27.95	-17.53
				Asian	33.41(*)	2.61	23.61	43.21
				Pacific Islander	3.50	2.61	-15.18	22.19
				White	22.33(*)	2.61	18.39	26.27
				Other	0.21	2.61	-6.70	7.12
			Hispanic	African American	-11.13(*)	2.60	-17.14	-5.11
				Native American	22.74(*)	2.60	17.55	27.93
				Asian	56.15(*)	2.60	46.22	66.08
				Pacific Islander	26.24(*)	2.60	7.54	44.94
				White	45.07(*)	2.60	40.75	49.38
				Other	22.95(*)	2.60	15.83	30.07
			Asian	African American	-67.28(*)	2.45	-76.93	-57.63
				Native American	-33.41(*)	2.45	-42.62	-24.20
				Hispanic	-56.15(*)	2.45	-65.52	-46.78
				Pacific Islander	-29.91(*)	2.45	-49.20	-10.62
				White	-11.08(*)	2.45	-19.88	-2.28
				Other	-33.2(*)	2.45	-43.49	-22.90
			Pacific Islander	African American	-37.37(*)	2.26	-53.77	-20.97
				Native American	-3.50	2.26	-19.69	12.68
				Hispanic	-26.24(*)	2.26	-42.50	-9.98
				Asian	29.91(*)	2.26	12.13	47.69
				White	18.83(*)	2.26	2.84	34.82
				Other	-3.29	2.26	-20.02	13.44
			White	African American	-56.19(*)	2.63	-61.22	-51.17
				Native American	-22.33(*)	2.63	-26.30	-18.36
				Hispanic	-45.07(*)	2.63	-49.44	-40.70
				Asian	11.08(*)	2.63	1.64	20.53
				Pacific Islander	-18.83(*)	2.63	-37.44	-0.21
				Other	-22.12(*)	2.63	-28.45	-15.79
			Other	African American	-34.08(*)	2.54	-41.44	-26.72
				Native American	-0.21	2.54	-6.94	6.52
				Hispanic	-22.95(*)	2.54	-29.91	-15.99
				Asian	33.2(*)	2.54	22.53	43.86
				Pacific Islander	3.29	2.54	-15.52	22.10
				White	22.12(*)	2.54	16.00	28.23

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
Biology	B	SS	African American	Native American	34.36(*)	2.58	27.82	40.90
				Hispanic	15.51(*)	2.58	8.56	22.46
				Asian	75.89(*)	2.58	64.81	86.96
				Pacific Islander	36.82(*)	2.58	14.21	59.43
				White	56.35(*)	2.58	50.76	61.95
				Other	44.9(*)	2.58	36.29	53.51
			Native American	African American	-34.36(*)	2.61	-40.96	-27.75
				Hispanic	-18.85(*)	2.61	-25.04	-12.66
				Asian	41.53(*)	2.61	30.84	52.21
				Pacific Islander	2.46	2.61	-20.14	25.05
				White	21.99(*)	2.61	17.41	26.58
				Other	10.54(*)	2.61	2.50	18.58
			Hispanic	African American	-15.51(*)	2.59	-22.49	-8.53
				Native American	18.85(*)	2.59	12.69	25.01
				Asian	60.38(*)	2.59	49.49	71.27
				Pacific Islander	21.31	2.59	-1.29	43.91
				White	40.84(*)	2.59	35.71	45.98
				Other	29.39(*)	2.59	21.05	37.73
			Asian	African American	-75.89(*)	2.46	-86.47	-65.31
				Native American	-41.53(*)	2.46	-51.64	-31.42
				Hispanic	-60.38(*)	2.46	-70.73	-50.02
				Pacific Islander	-39.07(*)	2.46	-62.09	-16.05
				White	-19.53(*)	2.46	-29.11	-9.96
				Other	-30.99(*)	2.46	-42.42	-19.55
			Pacific Islander	African American	-36.82(*)	2.24	-56.47	-17.16
				Native American	-2.46	2.24	-21.91	16.99
				Hispanic	-21.31(*)	2.24	-40.87	-1.75
				Asian	39.07(*)	2.24	18.12	60.02
				White	19.54(*)	2.24	0.31	38.76
				Other	8.08	2.24	-11.97	28.13
			White	African American	-56.35(*)	2.63	-62.06	-50.65
				Native American	-21.99(*)	2.63	-26.62	-17.37
				Hispanic	-40.84(*)	2.63	-46.05	-35.63
				Asian	19.53(*)	2.63	9.32	29.75
				Pacific Islander	-19.54	2.63	-42.07	3.00
				Other	-11.45(*)	2.63	-18.80	-4.11
			Other	African American	-44.9(*)	2.54	-53.37	-36.44
				Native American	-10.54(*)	2.54	-18.37	-2.72
				Hispanic	-29.39(*)	2.54	-37.56	-21.23
				Asian	30.99(*)	2.54	19.22	42.76
				Pacific Islander	-8.08	2.54	-30.76	14.59
				White	11.45(*)	2.54	4.37	18.54

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
English II	AA	SS	African American	Native American	26.23(*)	2.58	18.82	33.64
				Hispanic	10.78(*)	2.58	2.91	18.65
				Asian	46.13(*)	2.58	33.21	59.04
				Pacific Islander	50.42(*)	2.58	22.03	78.81
				White	46.21(*)	2.58	39.82	52.59
				Other	34.64(*)	2.58	25.37	43.92
			Native American	African American	-26.23(*)	2.61	-33.72	-18.74
				Hispanic	-15.45(*)	2.61	-22.38	-8.53
				Asian	19.9(*)	2.61	7.44	32.35
				Pacific Islander	24.19	2.61	-4.25	52.63
				White	19.97(*)	2.61	14.85	25.10
				Other	8.41	2.61	-0.11	16.93
			Hispanic	African American	-10.78(*)	2.59	-18.69	-2.86
				Native American	15.45(*)	2.59	8.57	22.34
				Asian	35.35(*)	2.59	22.68	48.02
				Pacific Islander	39.64(*)	2.59	11.22	68.05
				White	35.43(*)	2.59	29.67	41.18
				Other	23.86(*)	2.59	14.98	32.75
			Asian	African American	-46.13(*)	2.46	-58.43	-33.82
				Native American	-19.9(*)	2.46	-31.64	-8.16
				Hispanic	-35.35(*)	2.46	-47.36	-23.34
				Pacific Islander	4.29	2.46	-24.34	32.93
				White	0.08	2.46	-11.10	11.26
				Other	-11.49	2.46	-24.37	1.40
			Pacific Islander	African American	-50.42(*)	2.22	-74.86	-25.98
				Native American	-24.19	2.22	-48.40	0.02
				Hispanic	-39.64(*)	2.22	-63.96	-15.32
				Asian	-4.29	2.22	-30.15	21.57
				White	-4.21	2.22	-28.20	19.78
				Other	-15.78	2.22	-40.46	8.90
			White	African American	-46.21(*)	2.63	-52.72	-39.69
				Native American	-19.97(*)	2.63	-25.15	-14.80
				Hispanic	-35.43(*)	2.63	-41.26	-29.59
				Asian	-0.08	2.63	-12.04	11.89
				Pacific Islander	4.21	2.63	-24.22	32.64
				Other	-11.57(*)	2.63	-19.26	-3.87
			Other	African American	-34.64(*)	2.55	-43.82	-25.46
				Native American	-8.41(*)	2.55	-16.75	-0.07
				Hispanic	-23.86(*)	2.55	-32.60	-15.12
				Asian	11.49	2.55	-1.89	24.87
				Pacific Islander	15.78	2.55	-12.59	44.15
				White	11.57(*)	2.55	4.10	19.03

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
English II	AB	SS	African American	Native American	18.24(*)	2.58	10.85	25.64
				Hispanic	2.13	2.58	-5.67	9.94
				Asian	43.69(*)	2.58	30.61	56.77
				Pacific Islander	34.6(*)	2.58	11.99	57.20
				White	38.69(*)	2.58	32.31	45.07
				Other	30.13(*)	2.58	20.87	39.39
			Native American	African American	-18.24(*)	2.61	-25.72	-10.76
				Hispanic	-16.11(*)	2.61	-22.97	-9.25
				Asian	25.45(*)	2.61	12.81	38.09
				Pacific Islander	16.36	2.61	-6.17	38.88
				White	20.45(*)	2.61	15.31	25.58
				Other	11.89(*)	2.61	3.38	20.40
			Hispanic	African American	-2.13	2.59	-9.99	5.72
				Native American	16.11(*)	2.59	9.28	22.94
				Asian	41.56(*)	2.59	28.73	54.39
				Pacific Islander	32.47(*)	2.59	9.91	55.02
				White	36.56(*)	2.59	30.86	42.25
				Other	28(*)	2.59	19.16	36.84
			Asian	African American	-43.69(*)	2.45	-56.11	-31.27
				Native American	-25.45(*)	2.45	-37.32	-13.58
				Hispanic	-41.56(*)	2.45	-53.66	-29.45
				Pacific Islander	-9.09	2.45	-32.59	14.41
				White	-5.00	2.45	-16.33	6.33
				Other	-13.56(*)	2.45	-26.55	-0.56
			Pacific Islander	African American	-34.6(*)	2.27	-54.53	-14.67
				Native American	-16.36	2.27	-35.99	3.28
				Hispanic	-32.47(*)	2.27	-52.23	-12.71
				Asian	9.09	2.27	-12.73	30.91
				White	4.09	2.27	-15.27	23.45
				Other	-4.47	2.27	-24.71	15.78
			White	African American	-38.69(*)	2.63	-45.20	-32.18
				Native American	-20.45(*)	2.63	-25.63	-15.27
				Hispanic	-36.56(*)	2.63	-42.33	-30.78
				Asian	5.00	2.63	-7.17	17.17
				Pacific Islander	-4.09	2.63	-26.49	18.31
				Other	-8.56(*)	2.63	-16.25	-0.86
			Other	African American	-30.13(*)	2.55	-39.29	-20.97
				Native American	-11.89(*)	2.55	-20.22	-3.57
				Hispanic	-28(*)	2.55	-36.68	-19.32
				Asian	13.56(*)	2.55	0.02	27.09
				Pacific Islander	4.47	2.55	-18.24	27.18
				White	8.56(*)	2.55	1.10	16.01

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
English II	BA	SS	African American	Native American	25.3(*)	2.58	18.08	32.53
				Hispanic	6.45	2.58	-1.20	14.10
				Asian	47.15(*)	2.58	34.77	59.54
				Pacific Islander	23.34	2.58	-1.31	47.98
				White	41.69(*)	2.58	35.47	47.91
				Other	34.83(*)	2.58	25.81	43.85
			Native American	African American	-25.3(*)	2.61	-32.61	-17.99
				Hispanic	-18.85(*)	2.61	-25.58	-12.13
				Asian	21.85(*)	2.61	9.93	33.78
				Pacific Islander	-1.97	2.61	-26.60	22.66
				White	16.39(*)	2.61	11.39	21.38
				Other	9.53(*)	2.61	1.25	17.80
			Hispanic	African American	-6.45	2.59	-14.15	1.25
				Native American	18.85(*)	2.59	12.16	25.55
				Asian	40.71(*)	2.59	28.58	52.84
				Pacific Islander	16.89	2.59	-7.75	41.52
				White	35.24(*)	2.59	29.66	40.82
				Other	28.38(*)	2.59	19.77	37.00
			Asian	African American	-47.15(*)	2.46	-58.96	-35.35
				Native American	-21.85(*)	2.46	-33.10	-10.61
				Hispanic	-40.71(*)	2.46	-52.20	-29.21
				Pacific Islander	-23.82	2.46	-48.95	1.31
				White	-5.47	2.46	-16.15	5.22
				Other	-12.32	2.46	-24.69	0.04
			Pacific Islander	African American	-23.34(*)	2.24	-44.78	-1.89
				Native American	1.97	2.24	-19.23	23.16
				Hispanic	-16.89	2.24	-38.19	4.42
				Asian	23.82(*)	2.24	0.89	46.75
				White	18.35	2.24	-2.60	39.30
				Other	11.49	2.24	-10.21	33.20
			White	African American	-41.69(*)	2.63	-48.04	-35.34
				Native American	-16.39(*)	2.63	-21.43	-11.35
				Hispanic	-35.24(*)	2.63	-40.89	-29.58
				Asian	5.47	2.63	-5.97	16.90
				Pacific Islander	-18.35	2.63	-42.92	6.21
				Other	-6.86	2.63	-14.32	0.60
			Other	African American	-34.83(*)	2.55	-43.75	-25.91
				Native American	-9.53(*)	2.55	-17.63	-1.43
				Hispanic	-28.38(*)	2.55	-36.85	-19.91
				Asian	12.32	2.55	-0.50	25.15
				Pacific Islander	-11.49	2.55	-36.18	13.19
				White	6.86	2.55	-0.38	14.09

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
							Lower Bound	Upper Bound
English II	BB	SS	African American	Native American	23.04(*)	2.58	15.74	30.35
				Hispanic	4.89	2.58	-2.86	12.64
				Asian	52.66(*)	2.58	40.34	64.98
				Pacific Islander	55.6(*)	2.58	29.98	81.23
				White	40.78(*)	2.58	34.49	47.07
				Other	29.39(*)	2.58	20.29	38.49
			Native American	African American	-23.04(*)	2.61	-30.43	-15.66
				Hispanic	-18.15(*)	2.61	-24.99	-11.31
				Asian	29.61(*)	2.61	17.76	41.47
				Pacific Islander	32.56(*)	2.61	6.94	58.18
				White	17.74(*)	2.61	12.66	22.81
				Other	6.35	2.61	-2.02	14.72
			Hispanic	African American	-4.89	2.59	-12.69	2.91
				Native American	18.15(*)	2.59	11.35	24.96
				Asian	47.76(*)	2.59	35.69	59.84
				Pacific Islander	50.71(*)	2.59	25.09	76.33
				White	35.89(*)	2.59	30.21	41.57
				Other	24.5(*)	2.59	15.78	33.22
			Asian	African American	-52.66(*)	2.46	-64.43	-40.89
				Native American	-29.61(*)	2.46	-40.81	-18.41
				Hispanic	-47.76(*)	2.46	-59.24	-36.29
				Pacific Islander	2.95	2.46	-23.05	28.95
				White	-11.88(*)	2.46	-22.50	-1.25
				Other	-23.26(*)	2.46	-35.61	-10.92
			Pacific Islander	African American	-55.6(*)	2.23	-77.82	-33.39
				Native American	-32.56(*)	2.23	-54.53	-10.59
				Hispanic	-50.71(*)	2.23	-72.80	-28.62
				Asian	-2.95	2.23	-26.55	20.65
				White	-14.82	2.23	-36.56	6.91
				Other	-26.21(*)	2.23	-48.68	-3.74
			White	African American	-40.78(*)	2.63	-47.19	-34.37
				Native American	-17.74(*)	2.63	-22.86	-12.62
				Hispanic	-35.89(*)	2.63	-41.65	-30.12
				Asian	11.88(*)	2.63	0.53	23.22
				Pacific Islander	14.82	2.63	-10.75	40.40
				Other	-11.39(*)	2.63	-18.93	-3.85
			Other	African American	-29.39(*)	2.55	-38.40	-20.39
				Native American	-6.35	2.55	-14.54	1.84
				Hispanic	-24.5(*)	2.55	-33.08	-15.92
				Asian	23.26(*)	2.55	10.48	36.05
				Pacific Islander	26.21(*)	2.55	0.57	51.85
				White	11.39(*)	2.55	4.07	18.70

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
English III	AA	SS	African American	Native American	25.46(*)	2.58	18.75	32.16
				Hispanic	17.7(*)	2.58	10.49	24.91
				Asian	56.36(*)	2.58	43.91	68.81
				Pacific Islander	28.85(*)	2.58	4.62	53.07
				White	42.26(*)	2.58	36.47	48.04
				Other	35.48(*)	2.58	26.72	44.25
			Native American	African American	-25.46(*)	2.61	-32.23	-18.68
				Hispanic	-7.76(*)	2.61	-14.12	-1.40
				Asian	30.91(*)	2.61	18.84	42.98
				Pacific Islander	3.39	2.61	-20.83	27.61
				White	16.8(*)	2.61	12.16	21.44
				Other	10.03(*)	2.61	1.91	18.14
			Hispanic	African American	-17.7(*)	2.59	-24.94	-10.45
				Native American	7.76(*)	2.59	1.44	14.08
				Asian	38.66(*)	2.59	26.37	50.95
				Pacific Islander	11.15	2.59	-13.07	35.37
				White	24.56(*)	2.59	19.23	29.89
				Other	17.79(*)	2.59	9.29	26.28
			Asian	African American	-56.36(*)	2.44	-68.16	-44.57
				Native American	-30.91(*)	2.44	-42.22	-19.59
				Hispanic	-38.66(*)	2.44	-50.26	-27.07
				Pacific Islander	-27.51(*)	2.44	-52.30	-2.73
				White	-14.11(*)	2.44	-24.96	-3.25
				Other	-20.88(*)	2.44	-33.39	-8.36
			Pacific Islander	African American	-28.85(*)	2.24	-49.89	-7.80
				Native American	-3.39	2.24	-24.21	17.43
				Hispanic	-11.15	2.24	-32.10	9.80
				Asian	27.51(*)	2.24	4.79	50.24
				White	13.41	2.24	-7.20	34.02
				Other	6.64	2.24	-14.75	28.03
			White	African American	-42.26(*)	2.63	-48.15	-36.36
				Native American	-16.8(*)	2.63	-21.48	-12.12
				Hispanic	-24.56(*)	2.63	-29.97	-19.15
				Asian	14.11(*)	2.63	2.43	25.78
				Pacific Islander	-13.41	2.63	-37.59	10.77
				Other	-6.77	2.63	-14.19	0.65
			Other	African American	-35.48(*)	2.54	-44.12	-26.84
				Native American	-10.03(*)	2.54	-17.94	-2.12
				Hispanic	-17.79(*)	2.54	-26.12	-9.45
				Asian	20.88(*)	2.54	7.86	33.90
				Pacific Islander	-6.64	2.54	-30.90	17.63
				White	6.77	2.54	-0.40	13.95

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
							Lower Bound	Upper Bound
English III	AB	SS	African American	Native American	21.35(*)	2.58	14.96	27.75
				Hispanic	10.92(*)	2.58	3.98	17.86
				Asian	44.98(*)	2.58	33.08	56.89
				Pacific Islander	8.17	2.58	-14.42	30.75
				White	39.65(*)	2.58	34.13	45.18
				Other	33.9(*)	2.58	25.26	42.53
			Native American	African American	-21.35(*)	2.61	-27.81	-14.89
				Hispanic	-10.43(*)	2.61	-16.58	-4.29
				Asian	23.63(*)	2.61	12.09	35.17
				Pacific Islander	-13.19	2.61	-35.75	9.38
				White	18.3(*)	2.61	13.85	22.75
				Other	12.54(*)	2.61	4.49	20.59
			Hispanic	African American	-10.92(*)	2.59	-17.88	-3.96
				Native American	10.43(*)	2.59	4.33	16.54
				Asian	34.06(*)	2.59	22.28	45.84
				Pacific Islander	-2.76	2.59	-25.33	19.82
				White	28.73(*)	2.59	23.55	33.91
				Other	22.97(*)	2.59	14.53	31.41
			Asian	African American	-44.98(*)	2.45	-56.26	-33.71
				Native American	-23.63(*)	2.45	-34.46	-12.80
				Hispanic	-34.06(*)	2.45	-45.19	-22.94
				Pacific Islander	-36.82(*)	2.45	-60.02	-13.62
				White	-5.33	2.45	-15.72	5.06
				Other	-11.09	2.45	-23.23	1.06
			Pacific Islander	African American	-8.17	2.25	-27.85	11.51
				Native American	13.19	2.25	-6.28	32.65
				Hispanic	2.76	2.25	-16.85	22.36
				Asian	36.82(*)	2.25	15.48	58.16
				White	31.49(*)	2.25	12.23	50.75
				Other	25.73(*)	2.25	5.62	45.84
			White	African American	-39.65(*)	2.63	-45.28	-34.03
				Native American	-18.3(*)	2.63	-22.78	-13.82
				Hispanic	-28.73(*)	2.63	-33.99	-23.48
				Asian	5.33	2.63	-5.83	16.49
				Pacific Islander	-31.49(*)	2.63	-53.99	-8.98
				Other	-5.76	2.63	-13.18	1.66
			Other	African American	-33.9(*)	2.54	-42.38	-25.41
				Native American	-12.54(*)	2.54	-20.37	-4.72
				Hispanic	-22.97(*)	2.54	-31.23	-14.71
				Asian	11.09	2.54	-1.50	23.68
				Pacific Islander	-25.73(*)	2.54	-48.40	-3.06
				White	5.76	2.54	-1.40	12.92

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
English III	BA	SS	African American	Native American	17.44(*)	2.58	11.21	23.67
				Hispanic	5.68	2.58	-1.12	12.49
				Asian	41.05(*)	2.58	30.61	51.48
				Pacific Islander	22.9(*)	2.58	0.05	45.75
				White	32.11(*)	2.58	26.74	37.48
				Other	22.99(*)	2.58	14.79	31.19
			Native American	African American	-17.44(*)	2.61	-23.73	-11.14
				Hispanic	-11.75(*)	2.61	-17.81	-5.69
				Asian	23.61(*)	2.61	13.58	33.64
				Pacific Islander	5.46	2.61	-17.40	28.32
				White	14.67(*)	2.61	10.33	19.02
				Other	5.55	2.61	-2.07	13.17
			Hispanic	African American	-5.68	2.59	-12.51	1.14
				Native American	11.75(*)	2.59	5.74	17.77
				Asian	35.36(*)	2.59	25.04	45.69
				Pacific Islander	17.21	2.59	-5.64	40.07
				White	26.43(*)	2.59	21.32	31.54
				Other	17.31(*)	2.59	9.26	25.35
			Asian	African American	-41.05(*)	2.47	-51.04	-31.05
				Native American	-23.61(*)	2.47	-33.12	-14.10
				Hispanic	-35.36(*)	2.47	-45.23	-25.50
				Pacific Islander	-18.15	2.47	-41.25	4.95
				White	-8.94	2.47	-17.95	0.08
				Other	-18.06(*)	2.47	-28.85	-7.27
			Pacific Islander	African American	-22.9(*)	2.23	-42.65	-3.15
				Native American	-5.46	2.23	-25.01	14.09
				Hispanic	-17.21	2.23	-36.91	2.48
				Asian	18.15	2.23	-2.70	39.00
				White	9.21	2.23	-10.15	28.57
				Other	0.09	2.23	-20.00	20.18
			White	African American	-32.11(*)	2.63	-37.58	-26.64
				Native American	-14.67(*)	2.63	-19.05	-10.29
				Hispanic	-26.43(*)	2.63	-31.62	-21.23
				Asian	8.94	2.63	-0.65	18.53
				Pacific Islander	-9.21	2.63	-32.04	13.61
				Other	-9.12(*)	2.63	-16.10	-2.15
			Other	African American	-22.99(*)	2.54	-31.06	-14.92
				Native American	-5.55	2.54	-12.97	1.87
				Hispanic	-17.31(*)	2.54	-25.20	-9.41
				Asian	18.06(*)	2.54	6.97	29.15
				Pacific Islander	-0.09	2.54	-22.97	22.79
				White	9.12(*)	2.54	2.38	15.86

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
English III	BB	SS	African American	Native American	22.41(*)	2.58	16.13	28.68
				Hispanic	7.06(*)	2.58	0.36	13.77
				Asian	45.68(*)	2.58	35.00	56.37
				Pacific Islander	23.44(*)	2.58	1.69	45.19
				White	35.93(*)	2.58	30.59	41.28
				Other	25.91(*)	2.58	17.91	33.92
			Native American	African American	-22.41(*)	2.61	-28.75	-16.07
				Hispanic	-15.34(*)	2.61	-21.34	-9.35
				Asian	23.28(*)	2.61	12.95	33.60
				Pacific Islander	1.03	2.61	-20.70	22.77
				White	13.53(*)	2.61	9.14	17.91
				Other	3.51	2.61	-3.94	10.95
			Hispanic	African American	-7.06(*)	2.59	-13.80	-0.33
				Native American	15.34(*)	2.59	9.38	21.30
				Asian	38.62(*)	2.59	28.08	49.15
				Pacific Islander	16.38	2.59	-5.37	38.12
				White	28.87(*)	2.59	23.91	33.83
				Other	18.85(*)	2.59	11.08	26.62
			Asian	African American	-45.68(*)	2.46	-55.88	-35.48
				Native American	-23.28(*)	2.46	-33.03	-13.52
				Hispanic	-38.62(*)	2.46	-48.63	-28.61
				Pacific Islander	-22.24(*)	2.46	-44.38	-0.11
				White	-9.75(*)	2.46	-18.99	-0.51
				Other	-19.77(*)	2.46	-30.62	-8.92
			Pacific Islander	African American	-23.44(*)	2.24	-42.33	-4.55
				Native American	-1.03	2.24	-19.73	17.66
				Hispanic	-16.38	2.24	-35.19	2.43
				Asian	22.24(*)	2.24	2.09	42.39
				White	12.49	2.24	-5.99	30.97
				Other	2.47	2.24	-16.72	21.66
			White	African American	-35.93(*)	2.63	-41.38	-30.48
				Native American	-13.53(*)	2.63	-17.95	-9.10
				Hispanic	-28.87(*)	2.63	-33.91	-23.83
				Asian	9.75	2.63	-0.12	19.62
				Pacific Islander	-12.49	2.63	-34.17	9.19
				Other	-10.02(*)	2.63	-16.75	-3.29
			Other	African American	-25.91(*)	2.54	-33.81	-18.02
				Native American	-3.51	2.54	-10.78	3.77
				Hispanic	-18.85(*)	2.54	-26.48	-11.21
				Asian	19.77(*)	2.54	8.56	30.98
				Pacific Islander	-2.47	2.54	-24.26	19.32
				White	10.02(*)	2.54	3.51	16.53

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Geometry	A	SS	African American	Native American	30.36(*)	2.58	24.74	35.99
				Hispanic	21.27(*)	2.58	15.24	27.30
				Asian	95.7(*)	2.58	85.62	105.79
				Pacific Islander	32.54(*)	2.58	13.25	51.82
				White	52.58(*)	2.58	47.72	57.44
				Other	41.45(*)	2.58	33.97	48.94
			Native American	African American	-30.36(*)	2.61	-36.05	-24.68
				Hispanic	-9.1(*)	2.61	-14.42	-3.77
				Asian	65.34(*)	2.61	55.59	75.09
				Pacific Islander	2.17	2.61	-17.09	21.43
				White	22.22(*)	2.61	18.30	26.14
				Other	11.09(*)	2.61	4.13	18.05
			Hispanic	African American	-21.27(*)	2.59	-27.32	-15.21
				Native American	9.1(*)	2.59	3.81	14.39
				Asian	74.44(*)	2.59	64.50	84.38
				Pacific Islander	11.27	2.59	-8.00	30.54
				White	31.32(*)	2.59	26.85	35.79
				Other	20.19(*)	2.59	12.93	27.45
			Asian	African American	-95.7(*)	2.45	-105.29	-86.12
				Native American	-65.34(*)	2.45	-74.51	-56.17
				Hispanic	-74.44(*)	2.45	-83.84	-65.03
				Pacific Islander	-63.17(*)	2.45	-82.96	-43.38
				White	-43.12(*)	2.45	-51.90	-34.35
				Other	-54.25(*)	2.45	-64.56	-43.94
			Pacific Islander	African American	-32.54(*)	2.25	-49.36	-15.71
				Native American	-2.17	2.25	-18.80	14.46
				Hispanic	-11.27	2.25	-28.01	5.47
				Asian	63.17(*)	2.25	45.00	81.34
				White	20.05(*)	2.25	3.60	36.49
				Other	8.92	2.25	-8.26	26.10
			White	African American	-52.58(*)	2.63	-57.54	-47.63
				Native American	-22.22(*)	2.63	-26.17	-18.26
				Hispanic	-31.32(*)	2.63	-35.85	-26.78
				Asian	43.12(*)	2.63	33.72	52.53
				Pacific Islander	-20.05(*)	2.63	-39.25	-0.84
				Other	-11.13(*)	2.63	-17.53	-4.73
			Other	African American	-41.45(*)	2.54	-48.82	-34.09
				Native American	-11.09(*)	2.54	-17.86	-4.32
				Hispanic	-20.19(*)	2.54	-27.29	-13.09
				Asian	54.25(*)	2.54	43.59	64.91
				Pacific Islander	-8.92	2.54	-28.28	10.45
				White	11.13(*)	2.54	4.95	17.31

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
Geometry	B	SS	African American	Native American	33.23(*)	2.58	27.58	38.88
				Hispanic	22.82(*)	2.58	16.80	28.84
				Asian	94.74(*)	2.58	85.13	104.36
				Pacific Islander	55.07(*)	2.58	35.71	74.44
				White	53.82(*)	2.58	49.00	58.65
				Other	43.89(*)	2.58	36.55	51.23
			Native American	African American	-33.23(*)	2.61	-38.93	-27.53
				Hispanic	-10.41(*)	2.61	-15.80	-5.02
				Asian	61.51(*)	2.61	52.22	70.80
				Pacific Islander	21.84(*)	2.61	2.50	41.19
				White	20.59(*)	2.61	16.61	24.57
				Other	10.66(*)	2.61	3.81	17.52
			Hispanic	African American	-22.82(*)	2.59	-28.86	-16.77
				Native American	10.41(*)	2.59	5.05	15.77
				Asian	71.93(*)	2.59	62.45	81.40
				Pacific Islander	32.26(*)	2.59	12.90	51.61
				White	31(*)	2.59	26.53	35.48
				Other	21.07(*)	2.59	13.94	28.20
			Asian	African American	-94.74(*)	2.46	-103.92	-85.57
				Native American	-61.51(*)	2.46	-70.30	-52.73
				Hispanic	-71.93(*)	2.46	-80.93	-62.92
				Pacific Islander	-39.67(*)	2.46	-59.41	-19.93
				White	-40.92(*)	2.46	-49.25	-32.60
				Other	-50.85(*)	2.46	-60.71	-41.00
			Pacific Islander	African American	-55.07(*)	2.24	-71.92	-38.23
				Native American	-21.84(*)	2.24	-38.51	-5.17
				Hispanic	-32.26(*)	2.24	-49.03	-15.49
				Asian	39.67(*)	2.24	21.68	57.66
				White	-1.25	2.24	-17.73	15.22
				Other	-11.18	2.24	-28.35	5.98
			White	African American	-53.82(*)	2.63	-58.74	-48.91
				Native American	-20.59(*)	2.63	-24.60	-16.58
				Hispanic	-31(*)	2.63	-35.54	-26.46
				Asian	40.92(*)	2.63	32.04	49.81
				Pacific Islander	1.25	2.63	-18.04	20.54
				Other	-9.93(*)	2.63	-16.17	-3.69
			Other	African American	-43.89(*)	2.54	-51.12	-36.67
				Native American	-10.66(*)	2.54	-17.35	-3.98
				Hispanic	-21.07(*)	2.54	-28.06	-14.08
				Asian	50.85(*)	2.54	40.69	61.02
				Pacific Islander	11.18	2.54	-8.24	30.60
				White	9.93(*)	2.54	3.90	15.96

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
U.S. History	A	SS	African American	Native American	31.83(*)	2.58	25.73	37.93
				Hispanic	13.91(*)	2.58	7.42	20.39
				Asian	47.81(*)	2.58	37.51	58.10
				Pacific Islander	41.85(*)	2.58	21.77	61.94
				White	50.21(*)	2.58	44.98	55.44
				Other	46.71(*)	2.58	38.71	54.71
			Native American	African American	-31.83(*)	2.61	-37.99	-25.67
				Hispanic	-17.92(*)	2.61	-23.69	-12.15
				Asian	15.98(*)	2.61	6.05	25.90
				Pacific Islander	10.03	2.61	-10.02	30.08
				White	18.38(*)	2.61	14.11	22.66
				Other	14.88(*)	2.61	7.42	22.34
			Hispanic	African American	-13.91(*)	2.59	-20.42	-7.39
				Native American	17.92(*)	2.59	12.18	23.66
				Asian	33.9(*)	2.59	23.78	44.02
				Pacific Islander	27.95(*)	2.59	7.88	48.01
				White	36.31(*)	2.59	31.51	41.10
				Other	32.8(*)	2.59	25.06	40.55
			Asian	African American	-47.81(*)	2.46	-57.64	-37.97
				Native American	-15.98(*)	2.46	-25.37	-6.59
				Hispanic	-33.9(*)	2.46	-43.52	-24.28
				Pacific Islander	-5.95	2.46	-26.51	14.60
				White	2.41	2.46	-6.49	11.30
				Other	-1.10	2.46	-11.71	9.51
			Pacific Islander	African American	-41.85(*)	2.25	-59.40	-24.30
				Native American	-10.03	2.25	-27.37	7.32
				Hispanic	-27.95(*)	2.25	-45.40	-10.50
				Asian	5.95	2.25	-12.84	24.75
				White	8.36	2.25	-8.77	25.48
				Other	4.86	2.25	-13.07	22.78
			White	African American	-50.21(*)	2.63	-55.54	-44.88
				Native American	-18.38(*)	2.63	-22.70	-14.07
				Hispanic	-36.31(*)	2.63	-41.17	-31.45
				Asian	-2.41	2.63	-11.89	7.08
				Pacific Islander	-8.36	2.63	-28.33	11.61
				Other	-3.50	2.63	-10.32	3.31
			Other	African American	-46.71(*)	2.54	-54.58	-38.84
				Native American	-14.88(*)	2.54	-22.15	-7.62
				Hispanic	-32.8(*)	2.54	-40.39	-25.22
				Asian	1.10	2.54	-9.83	12.02
				Pacific Islander	-4.86	2.54	-25.04	15.32
				White	3.50	2.54	-3.07	10.08

(*)Significant differences

Table 33. Ethnic Pair-Wise Dunnett's C Post-hoc Comparison *(continued)*

Content	Form	Dependent Variable	(I)Ethnicity	(J)Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval Lower Bound	Upper Bound
U.S. History	B	SS	African American	Native American	28.95(*)	2.58	23.11	34.79
				Hispanic	14.68(*)	2.58	8.47	20.90
				Asian	48.84(*)	2.58	38.92	58.76
				Pacific Islander	29.87(*)	2.58	10.84	48.90
				White	43.71(*)	2.58	38.72	48.69
				Other	32.61(*)	2.58	24.89	40.32
			Native American	African American	-28.95(*)	2.61	-34.84	-23.06
				Hispanic	-14.27(*)	2.61	-19.83	-8.71
				Asian	19.89(*)	2.61	10.31	29.47
				Pacific Islander	0.92	2.61	-18.07	19.91
				White	14.76(*)	2.61	10.65	18.86
				Other	3.66	2.61	-3.57	10.88
			Hispanic	African American	-14.68(*)	2.59	-20.92	-8.44
				Native American	14.27(*)	2.59	8.74	19.80
				Asian	34.16(*)	2.59	24.39	43.92
				Pacific Islander	15.19	2.59	-3.83	34.20
				White	29.02(*)	2.59	24.41	33.64
				Other	17.93(*)	2.59	10.43	25.43
			Asian	African American	-48.84(*)	2.46	-58.31	-39.37
				Native American	-19.89(*)	2.46	-28.95	-10.83
				Hispanic	-34.16(*)	2.46	-43.44	-24.87
				Pacific Islander	-18.97	2.46	-38.49	0.55
				White	-5.13	2.46	-13.72	3.45
				Other	-16.23(*)	2.46	-26.49	-5.97
			Pacific Islander	African American	-29.87(*)	2.26	-46.52	-13.22
				Native American	-0.92	2.26	-17.38	15.54
				Hispanic	-15.19	2.26	-31.75	1.38
				Asian	18.97(*)	2.26	1.08	36.86
				White	13.84	2.26	-2.40	30.08
				Other	2.74	2.26	-14.30	19.78
			White	African American	-43.71(*)	2.63	-48.79	-38.63
				Native American	-14.76(*)	2.63	-18.90	-10.61
				Hispanic	-29.02(*)	2.63	-33.71	-24.34
				Asian	5.13	2.63	-4.02	14.29
				Pacific Islander	-13.84	2.63	-32.75	5.07
				Other	-11.1(*)	2.63	-17.70	-4.50
			Other	African American	-32.61(*)	2.54	-40.19	-25.03
				Native American	-3.66	2.54	-10.69	3.37
				Hispanic	-17.93(*)	2.54	-25.26	-10.59
				Asian	16.23(*)	2.54	5.67	26.79
				Pacific Islander	-2.74	2.54	-21.88	16.40
				White	11.1(*)	2.54	4.73	17.47

(*)Significant differences

Table 34. Spring 2013, Mean Scale Score and Standard Deviations for State and Each Proficiency Level

Content	Form	N Count	Total Mean	Total SD	Pass Mean	Pass SD	Unsatisfactory Mean	Unsatisfactory SD	Limited Knowledge Mean	Limited Knowledge SD	Proficient Mean	Proficient SD	Advanced Mean	Advanced SD
Algebra I	A	20891	739.89	53.98	755.98	39.70	616.48	49.69	681.80	10.10	732.39	17.77	795.15	34.80
	B	18677	740.20	52.90	755.41	40.06	621.25	46.90	682.50	9.96	731.87	17.62	794.88	35.89
Algebra II	A	16023	741.33	83.56	775.16	55.06	583.80	66.46	678.51	12.92	742.12	23.72	832.68	46.05
	B	14230	746.07	77.13	775.24	53.98	596.43	59.04	678.36	12.83	742.18	23.54	831.55	43.59
Biology I	A	22060	694.48	81.67	754.40	46.50	588.67	58.17	676.24	13.25	730.60	19.75	811.22	42.55
	B	15051	700.46	77.95	757.56	46.66	598.59	49.88	677.79	13.09	734.80	21.55	817.94	41.23
English II	AA	9776	765.46	70.03	783.61	50.40	540.46	59.70	668.55	23.58	760.11	30.65	849.42	33.82
	AB	9470	765.50	68.93	783.31	50.07	542.16	58.27	669.31	23.38	762.70	31.44	855.31	33.82
	BA	8911	771.97	64.87	785.42	51.17	549.66	55.62	672.00	22.60	762.43	30.48	855.56	35.24
	BB	8718	771.83	65.37	786.15	51.92	559.26	49.01	672.50	23.35	761.56	30.29	853.92	36.55
English III	AA	10096	763.72	64.95	781.07	44.22	612.97	60.08	684.84	7.77	758.78	28.23	834.50	26.17
	AB	9579	766.34	61.26	781.48	43.93	621.00	52.49	685.90	8.08	756.60	26.48	830.52	26.97
	BA	8556	775.45	55.94	785.83	43.70	620.68	49.66	685.51	9.69	761.68	25.84	833.40	30.75
	BB	8554	776.67	55.35	785.87	44.23	625.58	44.35	684.01	7.88	761.41	26.56	833.84	30.59
Geometry	A	20232	752.85	78.61	781.75	56.30	579.59	56.56	671.14	18.39	741.08	21.67	828.72	46.60
	B	17329	763.20	72.64	784.65	55.80	588.81	50.16	673.18	18.65	742.84	21.43	828.02	46.62
U.S. History	A	17691	741.07	78.25	773.39	52.06	565.98	57.00	667.91	19.87	736.95	21.19	819.91	41.67
	B	15721	749.72	70.53	774.17	50.52	573.98	56.90	670.73	18.37	736.61	20.98	817.97	38.10

*Pass = Proficiency Levels Proficient + Advanced

Table 35. Spring 2013, State Proficiency Level Impact Data

Content	Form	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
Algebra I	A	20891	83.30	5.70	11.10	52.00	31.30
	B	18677	83.70	5.40	10.80	52.40	31.30
Algebra II	A	16023	76.00	11.20	12.80	48.30	27.70
	B	14230	77.80	9.30	13.00	49.00	28.80
Biology I	A	22060	51.10	24.80	24.20	36.00	15.10
	B	15051	52.20	23.90	23.90	37.90	14.30
English II	AA	9776	87.00	2.50	10.40	64.10	22.90
	AB	9470	87.00	2.30	10.70	67.60	19.40
	BA	8911	89.80	1.50	8.80	67.60	22.20
	BB	8718	88.90	1.40	9.70	65.20	23.70
English III	AA	10096	87.40	7.30	5.40	61.70	25.70
	AB	9579	88.50	6.30	5.30	58.70	29.80
	BA	8556	91.80	3.30	4.90	60.90	30.90
	BB	8554	93.00	3.40	3.60	61.60	31.40
Geometry	A	20232	78.70	5.90	15.40	42.20	36.50
	B	17329	83.50	3.70	12.80	42.50	41.00
U.S. History	A	17691	75.80	6.70	17.50	42.50	33.30
	B	15721	80.30	4.20	15.60	43.20	37.10

Note: Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
Algebra I	A	Whole State	20891	83.30	5.70	11.10	52.00	31.30
		Female	10351	85.30	4.50	10.20	52.90	32.40
		Male	10534	81.40	6.80	11.90	51.10	30.30
		Native American	3288	80.70	5.80	13.50	54.80	25.90
		African American	1952	69.40	12.80	17.80	52.30	17.10
		Asian	441	95.50	2.30	2.30	39.70	55.80
		Hispanic	2552	79.20	7.50	13.30	54.40	24.80
		White	11211	86.80	4.10	9.20	51.40	35.40
		Other	1351	85.10	5.30	9.60	49.10	36.00
		Pacific Islander	96	74.00	12.50	13.50	49.00	25.00
		IEP	1381	59.70	19.60	20.80	48.00	11.70
		Low SES	10239	77.20	7.90	14.90	55.20	22.00
		ELL	851	64.70	16.10	19.20	52.50	12.20
		Section 504	177	83.00	6.20	10.70	61.00	22.00
		Accommodated	1566	58.80	19.70	21.50	49.40	9.40
	B	Whole State	18677	83.70	5.40	10.80	52.40	31.30
		Female	9507	84.80	4.80	10.40	53.50	31.30
		Male	9170	82.60	6.10	11.30	51.30	31.30
		Native American	2976	82.00	6.70	11.30	56.50	25.50
		African American	1751	69.90	12.70	17.30	52.90	17.00
		Asian	442	95.50	0.70	3.80	35.50	60.00
		Hispanic	2263	77.20	7.60	15.20	55.30	21.90
		White	10009	87.50	3.60	8.90	51.50	36.00
		Other	1138	84.90	4.70	10.50	49.50	35.40
		Pacific Islander	98	76.60	11.20	12.20	58.20	18.40
		IEP	756	65.60	15.70	18.70	47.50	18.10
		Low SES	9147	77.60	8.00	14.50	55.10	22.50
		ELL	564	67.40	12.80	19.90	51.60	15.80
		Section 504	142	76.00	6.30	17.60	56.30	19.70
		Accommodated	748	62.30	16.70	21.00	48.40	13.90

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status;
Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data *(continued)*

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
Algebra II	A	Whole State	16023	76.00	11.20	12.80	48.30	27.70
		Female	8208	77.40	10.00	12.60	49.70	27.70
		Male	7809	74.50	12.50	13.00	46.70	27.80
		Native American	2467	71.60	13.90	14.50	49.80	21.80
		African American	1373	64.30	17.80	17.80	49.60	14.70
		Asian	458	91.00	3.50	5.50	33.80	57.20
		Hispanic	1837	70.10	15.20	14.80	50.70	19.40
		White	9030	79.90	8.70	11.40	48.20	31.70
		Other	773	72.10	14.60	13.30	45.80	26.30
		Pacific Islander	85	62.40	16.50	21.20	41.20	21.20
		IEP	1401	33.00	45.50	21.60	27.60	5.40
		Low SES	6760	67.30	16.80	15.90	49.30	18.00
		ELL	360	50.30	33.30	16.40	38.90	11.40
		Section 504	126	68.30	15.90	15.90	38.90	29.40
		Accommodated	1351	34.50	44.00	21.50	28.10	6.40
	B	Whole State	14230	77.80	9.30	13.00	49.00	28.80
		Female	7340	77.60	8.80	13.60	50.10	27.50
		Male	6890	77.90	9.80	12.30	47.80	30.10
		Native American	2148	74.00	11.80	14.20	51.70	22.30
		African American	1287	61.50	17.90	20.60	47.00	14.50
		Asian	439	90.20	3.60	6.20	31.90	58.30
		Hispanic	1596	72.10	13.90	14.00	51.60	20.50
		White	7954	82.10	6.60	11.30	48.70	33.40
		Other	730	75.60	8.60	15.80	52.30	23.30
		Pacific Islander	76	69.80	18.40	11.80	48.70	21.10
		IEP	687	46.80	33.90	19.20	38.40	8.40
		Low SES	5966	70.10	13.60	16.30	50.70	19.40
		ELL	238	58.80	25.60	15.50	43.70	15.10
		Section 504	111	79.30	2.70	18.00	50.50	28.80
		Accommodated	556	49.50	30.00	20.50	41.00	8.50

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status; Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data *(continued)*

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
Biology I	A	Whole State	22060	51.10	24.80	24.20	36.00	15.10
		Female	11211	48.10	25.80	26.10	35.50	12.60
		Male	10834	54.20	23.60	22.20	36.60	17.60
		Native American	3559	46.70	26.20	27.00	35.10	11.60
		African American	1999	31.30	43.80	25.00	25.70	5.60
		Asian	505	67.50	16.40	16.00	39.40	28.10
		Hispanic	2792	35.80	37.90	26.20	27.50	8.30
		White	11893	58.80	18.10	23.10	39.80	19.00
		Other	1186	48.50	27.20	24.40	36.00	12.50
		Pacific Islander	126	59.50	25.40	15.10	38.90	20.60
		IEP	1521	21.60	58.40	19.90	17.70	3.90
		Low SES	10612	39.00	34.40	26.50	30.90	8.10
		ELL	803	15.30	64.30	20.40	12.60	2.70
		Section 504	165	48.50	26.10	25.50	32.10	16.40
		Accommodated	1702	17.50	62.50	20.00	14.60	2.90
	B	Whole State	15051	52.20	23.90	23.90	37.90	14.30
		Female	7705	49.20	25.50	25.40	37.50	11.70
		Male	7346	55.30	22.30	22.40	38.30	17.00
		Native American	2361	46.50	26.80	26.70	35.90	10.60
		African American	1413	29.70	43.50	26.80	25.00	4.70
		Asian	394	71.00	14.20	14.70	42.60	28.40
		Hispanic	1753	37.10	36.10	26.80	30.30	6.80
		White	8254	59.60	17.40	23.00	41.80	17.80
		Other	798	54.60	25.90	19.40	38.70	15.90
		Pacific Islander	78	60.30	24.40	15.40	51.30	9.00
		IEP	626	28.80	51.10	20.10	23.50	5.30
		Low SES	7033	39.90	33.40	26.70	31.90	8.00
		ELL	356	14.90	63.20	21.90	13.50	1.40
		Section 504	126	50.80	27.00	22.20	38.10	12.70
		Accommodated	630	23.20	56.50	20.30	19.70	3.50

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status;
Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data *(continued)*

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
English II	AA	Whole State	9776	87.00	2.50	10.40	64.10	22.90
		Female	5004	90.20	1.30	8.50	63.40	26.80
		Male	4772	83.70	3.80	12.40	64.90	18.80
		Native American	1548	84.90	2.60	12.40	65.80	19.10
		African American	882	73.50	6.00	20.50	65.10	8.40
		Asian	234	84.60	2.60	12.80	51.30	33.30
		Hispanic	1145	79.60	4.80	15.60	66.60	13.00
		White	5323	91.40	1.50	7.10	63.60	27.80
		Other	604	88.20	1.80	9.90	63.20	25.00
		Pacific Islander	40	95.00	2.50	2.50	62.50	32.50
		IEP	622	56.40	15.10	28.50	51.30	5.10
		Low SES	4428	81.10	3.70	15.20	66.90	14.20
		ELL	229	45.40	16.20	38.40	42.80	2.60
		Section 504	87	86.20	1.10	12.60	72.40	13.80
		Accommodated	631	52.50	16.60	30.90	48.20	4.30
	AB	Whole State	9470	87.00	2.30	10.70	67.60	19.40
		Female	4832	90.00	1.20	8.80	67.30	22.70
		Male	4638	83.90	3.50	12.60	68.00	15.90
		Native American	1508	84.80	2.10	13.10	70.30	14.50
		African American	862	78.40	4.40	17.20	67.60	10.80
		Asian	220	87.30	1.80	10.90	58.20	29.10
		Hispanic	1148	77.70	4.40	17.90	67.20	10.50
		White	5080	91.00	1.60	7.40	67.50	23.50
		Other	589	88.00	2.20	9.80	66.60	21.40
		Pacific Islander	63	88.90	7.90	3.20	63.50	25.40
		IEP	565	56.10	15.20	28.70	52.60	3.50
		Low SES	4426	80.90	3.40	15.70	69.40	11.50
		ELL	216	41.20	20.80	38.00	38.90	2.30
		Section 504	73	78.10	6.80	15.10	69.90	8.20
		Accommodated	568	49.80	18.00	32.20	47.20	2.60

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status;
Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data *(continued)*

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
English II	BA	Whole State	8911	89.80	1.50	8.80	67.60	22.20
		Female	4614	91.70	1.10	7.30	66.20	25.50
		Male	4297	87.70	1.90	10.40	69.10	18.60
		Native American	1393	89.10	1.10	9.90	71.40	17.70
		African American	794	79.70	4.00	16.20	69.00	10.70
		Asian	220	91.30	1.40	7.30	59.50	31.80
		Hispanic	1045	81.80	4.00	14.20	69.60	12.20
		White	4863	92.90	0.60	6.50	66.00	26.90
		Other	550	92.80	0.90	6.40	69.30	23.50
		Pacific Islander	46	84.70	6.50	8.70	63.00	21.70
		IEP	382	66.80	7.10	26.20	59.70	7.10
		Low SES	4188	84.40	2.30	13.30	70.90	13.50
		ELL	187	51.30	18.70	29.90	49.70	1.60
		Section 504	69	92.70	.	7.20	73.90	18.80
		Accommodated	365	60.30	10.70	29.00	54.50	5.80
	BB	Whole State	8718	88.90	1.40	9.70	65.20	23.70
		Female	4517	90.80	0.90	8.30	63.30	27.50
		Male	4201	86.70	2.00	11.30	67.20	19.50
		Native American	1372	87.80	1.10	11.10	68.60	19.20
		African American	791	78.40	3.00	18.60	67.40	11.00
		Asian	227	88.10	2.60	9.30	47.10	41.00
		Hispanic	1020	80.10	3.90	16.00	67.00	13.10
		White	4718	92.70	0.70	6.60	63.90	28.80
		Other	547	90.20	0.70	9.10	70.60	19.60
		Pacific Islander	43	88.40	2.30	9.30	41.90	46.50
		IEP	402	70.90	6.00	23.10	63.70	7.20
		Low SES	4020	83.30	2.30	14.30	68.30	15.00
		ELL	190	35.30	18.90	45.80	33.70	1.60
		Section 504	80	83.80	2.50	13.80	62.50	21.30
		Accommodated	382	59.40	10.50	30.10	53.10	6.30

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status;
Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data *(continued)*

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
English III	AA	Whole State	10096	87.40	7.30	5.40	61.70	25.70
		Female	4893	90.30	5.00	4.70	60.60	29.70
		Male	5199	84.70	9.30	6.00	62.70	22.00
		Native American	1648	85.50	8.60	5.90	64.60	20.90
		African American	935	75.90	13.70	10.40	66.80	9.10
		Asian	212	92.90	2.80	4.20	51.90	41.00
		Hispanic	1148	83.70	8.50	7.80	69.20	14.50
		White	5549	90.20	5.80	4.10	58.50	31.70
		Other	556	90.20	6.10	3.60	64.70	25.50
		Pacific Islander	48	83.40	8.30	8.30	56.30	27.10
		IEP	1586	46.70	34.80	18.50	44.10	2.60
		Low SES	4611	82.30	10.30	7.50	67.20	15.10
		ELL	239	54.00	25.10	20.90	52.30	1.70
		Section 504	85	96.50	1.20	2.40	76.50	20.00
		Accommodated	1431	45.70	35.20	19.10	43.10	2.60
	AB	Whole State	9579	88.50	6.30	5.30	58.70	29.80
		Female	4712	91.40	4.10	4.50	57.40	34.00
		Male	4866	85.50	8.40	6.10	59.80	25.70
		Native American	1594	86.40	7.70	5.90	61.40	25.00
		African American	912	77.90	13.00	9.00	64.10	13.80
		Asian	205	93.10	2.90	3.90	55.10	38.00
		Hispanic	1072	83.40	8.70	7.90	66.50	16.90
		White	5259	91.50	4.40	4.10	55.40	36.10
		Other	488	91.00	4.30	4.70	58.00	33.00
		Pacific Islander	49	79.60	10.20	10.20	61.20	18.40
		IEP	1378	50.70	32.50	16.80	47.60	3.10
		Low SES	4292	82.30	9.90	7.70	63.70	18.60
		ELL	218	50.90	28.40	20.60	49.10	1.80
		Section 504	82	95.10	3.70	1.20	74.40	20.70
		Accommodated	1218	49.50	32.50	18.10	46.50	3.00

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status;
Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data *(continued)*

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
English III	BA	Whole State	8556	91.80	3.30	4.90	60.90	30.90
		Female	4281	93.90	2.40	3.70	58.00	35.90
		Male	4275	89.70	4.20	6.10	63.80	25.90
		Native American	1399	91.20	3.50	5.40	65.30	25.90
		African American	810	84.00	7.40	8.50	65.90	18.10
		Asian	236	92.40	3.00	4.70	44.10	48.30
		Hispanic	916	88.10	5.00	6.90	69.80	18.30
		White	4688	94.10	2.20	3.70	58.00	36.10
		Other	467	91.20	3.40	5.40	59.50	31.70
		Pacific Islander	40	92.50	2.50	5.00	57.50	35.00
		IEP	735	58.20	24.50	17.30	53.70	4.50
		Low SES	3615	87.10	5.50	7.50	66.50	20.60
		ELL	159	58.50	20.10	21.40	54.10	4.40
		Section 504	59	86.40	3.40	10.20	62.70	23.70
		Accommodated	596	56.70	24.50	18.80	53.00	3.70
	BB	Whole State	8554	93.00	3.40	3.60	61.60	31.40
		Female	4329	94.90	2.30	2.80	60.30	34.60
		Male	4225	91.00	4.60	4.40	62.90	28.10
		Native American	1300	92.90	4.30	2.80	66.10	26.80
		African American	785	84.60	7.50	7.90	69.60	15.00
		Asian	215	95.80	1.90	2.30	49.80	46.00
		Hispanic	946	87.70	5.50	6.80	69.00	18.70
		White	4777	95.10	2.30	2.60	57.50	37.60
		Other	488	94.50	2.70	2.90	66.80	27.70
		Pacific Islander	43	90.70	4.70	4.70	62.80	27.90
		IEP	683	60.60	25.20	14.20	55.80	4.80
		Low SES	3663	88.70	5.70	5.60	68.60	20.10
		ELL	165	61.90	17.60	20.60	55.80	6.10
		Section 504	82	91.50	2.40	6.10	61.00	30.50
		Accommodated	582	60.30	23.90	15.80	55.00	5.30

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status;
Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data *(continued)*

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
Geometry	A	Whole State	20232	78.70	5.90	15.40	42.20	36.50
		Female	10129	79.90	4.80	15.30	43.30	36.60
		Male	10093	77.60	6.90	15.50	41.10	36.50
		Native American	3363	75.60	6.50	17.80	44.80	30.80
		African American	1925	61.10	12.80	26.10	43.30	17.80
		Asian	474	92.70	1.90	5.50	30.00	62.70
		Hispanic	2382	72.10	7.80	20.10	46.30	25.80
		White	10901	83.50	4.10	12.40	40.70	42.80
		Other	1077	80.70	5.50	13.80	43.70	37.00
		Pacific Islander	110	75.40	11.80	12.70	40.90	34.50
		IEP	2902	35.10	27.30	37.70	28.20	6.90
		Low SES	9582	70.30	8.80	20.80	45.60	24.70
		ELL	636	47.30	20.90	31.80	36.00	11.30
		Section 504	149	77.20	2.70	20.10	37.60	39.60
		Accommodated	2734	35.70	26.40	37.90	28.80	6.90
	B	Whole State	17329	83.50	3.70	12.80	42.50	41.00
		Female	8776	83.50	3.30	13.20	43.10	40.40
		Male	8553	83.60	4.00	12.40	42.00	41.60
		Native American	2707	81.70	3.60	14.60	46.80	34.90
		African American	1647	64.60	9.60	25.80	46.10	18.50
		Asian	450	95.80	0.90	3.30	27.60	68.20
		Hispanic	1991	77.20	5.80	17.00	48.10	29.10
		White	9482	88.10	2.30	9.60	40.40	47.70
		Other	960	83.20	4.00	12.90	41.60	41.60
		Pacific Islander	92	88.10	2.20	9.80	35.90	52.20
		IEP	1200	43.20	22.90	33.80	31.40	11.80
		Low SES	7911	76.20	5.70	18.10	47.70	28.50
		ELL	367	56.60	16.10	27.20	41.10	15.50
		Section 504	127	84.30	4.70	11.00	46.50	37.80
		Accommodated	1021	43.40	22.20	34.40	33.30	10.10

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status;
Pass = Proficiency Levels Proficient + Advanced

Table 36. Spring 2013, State and Subgroup Proficiency Level Impact Data *(continued)*

Content	Form	Subgroup	N Count	Pass	Unsatisfactory	Limited Knowledge	Proficient	Advanced
		Whole State	17691	75.80	6.70	17.50	42.50	33.30
		Female	8886	70.90	7.90	21.20	44.90	26.00
		Male	8798	80.90	5.50	13.60	40.10	40.80
		Native American	2793	73.70	6.90	19.40	44.80	28.90
		African American	1662	59.10	14.70	26.10	41.50	17.60
		Asian	469	76.70	9.60	13.60	36.20	40.50
		Hispanic	2071	65.00	11.60	23.40	42.30	22.70
	A	White	9641	81.50	4.20	14.30	42.90	38.60
		Other	953	77.10	4.80	18.00	38.40	38.70
		Pacific Islander	102	78.40	6.90	14.70	38.20	40.20
		IEP	1207	49.10	22.60	28.30	31.60	17.50
		Low SES	7868	67.00	10.00	23.00	43.50	23.50
		ELL	469	34.30	32.80	32.80	26.20	8.10
		Section 504	157	72.00	8.30	19.70	41.40	30.60
U.S.		Accommodated	1193	43.50	27.00	29.50	29.50	14.00
History		Whole State	15721	80.30	4.20	15.60	43.20	37.10
		Female	7909	76.00	4.80	19.20	46.00	30.00
		Male	7812	84.70	3.50	11.90	40.40	44.30
		Native American	2463	78.60	3.90	17.50	45.20	33.40
		African American	1497	65.40	9.10	25.50	46.30	19.10
		Asian	411	85.40	2.90	11.70	38.20	47.20
		Hispanic	1818	74.00	7.00	19.00	48.20	25.80
	B	White	8612	84.60	2.80	12.70	41.60	43.00
		Other	827	79.40	4.70	15.80	40.60	38.80
		Pacific Islander	93	77.40	7.50	15.10	39.80	37.60
		IEP	683	59.90	15.10	25.00	40.70	19.20
		Low SES	6925	72.90	6.50	20.60	46.60	26.30
		ELL	297	44.80	25.30	30.00	36.70	8.10
		Section 504	126	80.10	6.30	13.50	47.60	32.50
		Accommodated	600	52.30	21.20	26.50	37.50	14.80

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status; Pass = Proficiency Levels Proficient + Advanced

Table 37. Spring 2013, P-values and Item-Test Correlations Statistics for Operational Test Forms

Content	Form	Item Type	P-Values			Item-Test Correlation		
			Low	Mean	High	Low	Mean	High
Algebra I	A	MC	0.19	0.66	0.92	0.17	0.41	0.58
	B	MC	0.20	0.66	0.94	0.23	0.40	0.55
Algebra II	A	MC	0.25	0.62	0.91	0.22	0.40	0.52
	B	MC	0.31	0.63	0.95	0.19	0.39	0.55
Biology I	A	MC	0.35	0.68	0.87	0.16	0.38	0.54
	B	MC	0.38	0.69	0.98	0.12	0.35	0.54
English II	AA	CR	0.63	0.63	0.63	0.48	0.48	0.48
		MC	0.32	0.76	0.99	0.16	0.32	0.47
	AB	CR	0.64	0.64	0.64	0.58	0.58	0.58
		MC	0.33	0.75	0.99	0.14	0.32	0.46
	BA	CR	0.63	0.63	0.63	0.48	0.48	0.48
		MC	0.36	0.77	0.96	0.13	0.33	0.46
	BB	CR	0.63	0.63	0.63	0.58	0.58	0.58
		MC	0.37	0.77	0.96	0.11	0.33	0.46
English III	AA	CR	0.66	0.66	0.66	0.57	0.57	0.57
		MC	0.23	0.68	0.91	0.07	0.33	0.54
	AB	CR	0.64	0.64	0.64	0.62	0.62	0.62
		MC	0.23	0.68	0.91	0.06	0.33	0.53
	BA	CR	0.66	0.66	0.66	0.58	0.58	0.58
		MC	0.24	0.72	0.96	0.14	0.36	0.59
	BB	CR	0.65	0.65	0.65	0.62	0.62	0.62
		MC	0.24	0.72	0.96	0.14	0.36	0.60
Geometry	A	MC	0.33	0.69	0.94	0.28	0.45	0.64
	B	MC	0.40	0.72	0.94	0.22	0.42	0.57
U.S. History	A	MC	0.27	0.64	0.94	0.15	0.40	0.57
	B	MC	0.27	0.65	0.96	0.20	0.37	0.55

Table 38. Spring 2012 and Spring 2013 Test Reliability Data

Content	Form	Coefficient Alpha	
		Spring 2012	Spring 2013
Algebra I	A	0.91	0.91
	B	0.91	0.91
Algebra II	A	0.91	0.90
	B	0.91	0.90
Biology I	A	0.89	0.90
	B	0.89	0.88
English II	AA	0.86	0.87
	AB	0.86	0.87
	BA	0.84	0.86
	BB	0.84	0.87
English III	AA	0.88	0.88
	AB	0.88	0.88
	BA	0.88	0.88
	BB	0.88	0.88
Geometry	A	0.92	0.93
	B	0.91	0.92
U.S. History	A	0.89	0.91
	B	0.90	0.89

Table 39. Algebra I & II, Raw Score to Scale Score Conversions and SEMs

Raw Score	Algebra I Form A		Algebra I Form B		Algebra II Form A		Algebra II Form B	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
0	490	178	490	175	440	215	440	217
1	490	178	490	175	440	215	440	217
2	490	178	490	175	440	215	440	217
3	490	178	490	175	440	215	440	217
4	490	178	490	175	440	215	440	217
5	490	178	490	175	440	215	440	217
6	490	178	490	175	440	215	440	217
7	490	178	490	175	440	215	440	217
8	490	178	490	175	440	215	440	217
9	490	178	490	175	440	215	440	217
10	490	178	490	175	440	215	440	217
11	490	178	490	175	440	215	440	217
12	545	123	568	97	440	215	440	217
13	590	78	599	66	474	181	463	193
14	612	56	617	48	536	119	531	126
15	627	42	631	38	569	86	566	91
16	639	34	641	31	592	65	590	68
17	648	28	650	27	609	53	608	55
18	657	25	658	24	624	45	623	46
19	**664**	**22**	**665**	**22**	636	39	636	40
20	670	20	671	20	647	35	647	36
21	676	19	677	18	**657**	**32**	**657**	**32**
22	681	18	682	17	666	29	666	30
23	686	17	687	16	674	27	674	27
24	691	16	691	15	681	26	682	26
25	695	15	696	15	689	24	689	24

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 39. Algebra I & II, Raw Score to Scale Score Conversions and SEMs *(continued)*

Raw Score	Algebra I Form A Scale Score	SEM	Algebra I Form B Scale Score	SEM	Algebra II Form A Scale Score	SEM	Algebra II Form B Scale Score	SEM
26	**700**	**14**	**700**	**14**	695	23	696	23
27	704	14	704	14	**702**	**22**	**702**	**22**
28	708	14	708	13	708	21	708	21
29	711	13	712	13	714	21	714	20
30	715	13	715	12	720	20	720	20
31	719	13	719	12	725	19	726	19
32	723	12	722	12	731	19	731	19
33	726	12	726	12	736	19	737	19
34	730	12	729	12	742	18	742	18
35	733	12	733	12	747	18	747	18
36	737	12	737	12	753	18	753	18
37	741	12	740	12	758	18	758	18
38	744	12	744	12	764	18	764	18
39	748	12	748	12	769	18	769	18
40	752	12	752	12	775	18	775	18
41	756	12	756	12	781	18	781	18
42	760	12	760	12	**787**	**19**	**787**	**19**
43	**764**	**13**	**764**	**13**	794	19	794	19
44	769	13	769	13	801	20	801	20
45	774	13	774	13	808	20	808	20
46	779	14	779	14	816	21	816	21
47	785	14	785	14	824	22	825	22
48	791	15	791	15	833	23	834	24
49	798	16	798	16	844	25	846	26
50	806	17	805	17	856	28	858	28
51	815	19	815	18	871	31	874	32
52	827	22	826	21	891	37	894	37
53	843	27	842	27	919	47	922	45
54	871	39	870	41	969	73	970	66
55	999	167	999	170	999	94	999	83

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 40. Geometry, Raw Score to Scale Score Conversions and SEMs

Raw Score	Geometry I Form A		Geometry I Form B	
	Scale Score	SEM	Scale Score	SEM
0	440	194	440	192
1	440	194	440	192
2	440	194	440	192
3	440	194	440	192
4	440	194	440	192
5	440	194	440	192
6	440	194	440	192
7	440	194	440	192
8	440	194	440	192
9	440	194	440	192
10	440	194	440	192
11	472	161	468	165
12	533	101	528	104
13	562	72	558	74
14	582	53	579	56
15	597	43	595	46
16	609	36	608	40
17	619	32	619	35
18	629	29	629	32
19	**637**	**27**	**638**	**29**
20	645	26	646	27
21	652	24	654	25
22	659	23	661	24
23	666	23	668	23
24	673	22	674	22
25	679	21	680	21

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 40. Geometry, Raw Score to Scale Score Conversions and SEMs *(continued)*

Raw Score	Geometry I Form A		Geometry I Form B	
	Scale Score	SEM	Scale Score	SEM
26	685	21	686	20
27	691	20	692	20
28	697	19	698	19
29	**703**	**19**	**703**	**18**
30	709	18	708	18
31	714	18	713	18
32	719	17	719	17
33	724	17	724	17
34	729	16	729	17
35	734	16	734	16
36	739	16	739	16
37	744	16	744	16
38	749	15	749	16
39	754	15	754	16
40	759	15	759	16
41	764	15	764	16
42	770	16	770	16
43	775	16	775	16
44	**781**	**16**	**781**	**17**
45	787	17	787	17
46	794	17	794	18
47	801	18	801	19
48	808	19	809	20
49	817	20	818	21
50	828	22	828	23
51	840	25	841	26
52	855	29	857	30
53	877	35	879	37
54	913	51	917	55
55	999	124	999	126

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 41. English II, Raw Score to Scale Score Conversions and SEMs

Raw Score	English II Form AA Scale Score	SEM	English II Form AB Scale Score	SEM	English II Form BA Scale Score	SEM	English II Form BB Scale Score	SEM
0	440	152	440	152	440	162	440	162
1	440	152	440	152	440	162	440	162
2	440	152	440	152	440	162	440	162
3	440	152	440	152	440	162	440	162
4	440	152	440	152	440	162	440	162
5	440	152	440	152	440	162	440	162
6	440	152	440	152	440	162	440	162
7	440	152	440	152	440	162	440	162
8	440	152	440	152	440	162	440	162
9	440	152	440	152	440	162	440	162
10	440	152	440	152	440	162	440	162
11	440	152	440	152	440	162	440	162
12	440	152	440	152	440	162	440	162
13	440	152	440	152	440	162	440	162
14	440	152	440	152	440	162	440	162
15	440	152	440	152	440	162	440	162
16	440	152	440	152	457	145	465	137
17	440	152	440	152	496	106	502	100
18	465	127	471	121	524	79	527	76
19	497	95	502	91	544	63	547	61
20	521	74	524	72	561	53	563	51
21	540	61	543	60	574	45	577	44
22	556	53	558	52	586	40	588	39
23	569	47	572	47	597	36	599	36
24	581	43	584	43	606	33	608	33
25	593	40	595	39	**615**	**31**	**617**	**30**
26	603	37	605	37	623	29	625	29
27	**613**	**35**	**615**	**35**	631	27	632	27
28	622	33	624	33	638	26	639	26
29	631	32	632	31	645	25	646	25
30	639	30	640	30	651	24	652	24
31	647	29	648	29	658	24	659	23
32	654	28	656	28	664	23	665	23
33	662	27	663	27	670	22	671	22
34	669	26	670	26	676	22	677	22

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 41. English II, Raw Score to Scale Score Conversions and SEMs *(continued)*

Raw Score	English II Form AA		English II Form AB		English II Form BA		English II Form BB	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
35	676	25	677	25	682	21	682	21
36	683	24	684	24	687	21	688	21
37	689	24	690	23	693	21	694	21
38	696	23	696	23	699	20	699	20
39	**702**	**23**	**703**	**22**	**704**	**20**	**705**	**20**
40	708	22	709	22	710	20	710	20
41	714	22	715	21	716	20	716	20
42	720	21	721	21	721	20	722	20
43	726	21	727	21	727	20	727	20
44	732	21	733	21	733	20	733	20
45	738	21	739	21	739	20	739	20
46	745	21	745	21	745	20	745	20
47	751	21	751	20	751	21	751	20
48	757	21	757	21	758	21	757	21
49	764	21	763	21	764	21	764	21
50	770	21	770	21	771	22	771	21
51	777	22	776	21	779	22	778	22
52	784	22	783	21	786	23	785	22
53	792	22	790	22	794	23	793	23
54	800	23	798	22	803	24	801	23
55	808	24	806	23	812	25	809	24
56	**817**	**25**	814	24	**822**	**26**	**819**	**25**
57	827	26	**824**	**25**	833	27	829	26
58	838	28	834	26	845	29	840	28
59	850	30	845	28	859	32	852	30
60	864	32	858	30	875	35	867	33
61	881	36	873	33	895	41	885	37
62	901	41	891	38	921	49	907	44
63	928	49	914	44	959	64	940	57
64	965	61	947	56	999	83	999	89
65	999	75	999	79	999	83	999	89
66	999	75	999	79	.	.	.	.

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 42. English III, Raw Score to Scale Score Conversions and SEMs

Raw Score	English III Form AA		English III Form AB		English III Form BA		English III Form BB	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
0	440	195	440	196	440	184	440	185
1	440	195	440	196	440	184	440	185
2	440	195	440	196	440	184	440	185
3	440	195	440	196	440	184	440	185
4	440	195	440	196	440	184	440	185
5	440	195	440	196	440	184	440	185
6	440	195	440	196	440	184	440	185
7	440	195	440	196	440	184	440	185
8	440	195	440	196	440	184	440	185
9	440	195	440	196	440	184	440	185
10	440	195	440	196	440	184	440	185
11	440	195	440	196	440	184	440	185
12	440	195	440	196	440	184	440	185
13	440	195	440	196	485	139	490	135
14	440	195	440	196	529	95	533	92
15	505	130	510	126	554	70	557	67
16	541	94	545	91	572	55	575	53
17	564	71	567	68	587	46	589	44
18	582	58	584	56	598	40	600	39
19	596	50	598	48	609	36	610	35
20	608	44	610	42	618	32	619	32
21	619	39	621	38	626	30	627	29
22	629	36	630	35	634	28	635	27
23	637	33	639	32	640	26	642	25
24	645	31	647	30	647	25	648	24
25	653	29	654	28	653	23	654	23
26	660	27	661	27	659	22	660	22
27	666	26	668	25	664	22	665	21
28	**673**	**25**	**674**	**24**	**670**	**21**	**671**	**21**
29	679	24	680	23	675	20	676	20
30	685	23	686	22	680	20	681	20
31	690	22	691	22	685	19	686	19
32	695	21	697	21	689	19	690	19
33	**701**	**21**	**702**	**20**	694	19	695	19
34	706	20	707	20	699	19	**700**	**18**
35	711	20	712	19	**703**	**18**	704	18

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 42. English III, Raw Score to Scale Score Conversions and SEMs *(continued)*

Raw Score	English III Form AA Scale Score	SEM	English III Form AB Scale Score	SEM	English III Form BA Scale Score	SEM	English III Form BB Scale Score	SEM
36	716	19	716	19	708	18	709	18
37	720	19	721	19	713	18	713	18
38	725	18	726	18	717	18	718	18
39	730	18	730	18	722	18	723	18
40	734	18	735	18	727	18	727	18
41	739	18	740	17	731	18	732	18
42	743	18	744	17	736	18	737	18
43	748	17	749	17	741	18	741	17
44	753	17	753	17	745	18	746	17
45	757	17	758	17	750	17	751	17
46	762	17	762	17	754	17	755	17
47	767	17	767	17	759	16	760	16
48	771	17	772	17	763	16	764	15
49	776	17	777	17	768	15	768	15
50	781	17	782	17	772	15	772	15
51	786	17	786	17	776	15	777	14
52	791	17	791	17	780	15	781	14
53	796	17	797	17	785	15	785	15
54	801	17	**802**	**17**	789	15	790	15
55	**807**	**17**	807	17	794	15	794	15
56	812	17	812	17	799	16	799	16
57	818	18	818	17	**804**	**16**	**804**	**16**
58	824	18	824	18	809	16	810	16
59	830	18	830	18	815	17	815	16
60	836	19	836	19	820	17	821	17
61	843	20	843	20	827	18	827	18
62	850	21	850	21	834	20	834	20
63	858	23	858	22	842	23	842	23
64	868	25	867	24	852	27	852	26
65	879	28	877	27	864	31	863	30
66	892	31	890	30	878	35	877	34
67	908	37	905	35	896	40	893	38
68	929	44	925	41	919	47	914	44
69	958	54	952	51	950	60	943	55
70	999	73	999	72	999	85	994	82
71	999	73	999	72	999	85	999	85

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 43. U.S. History & Biology I, Raw Score to Scale Score Conversions and SEMs

Raw Score	U.S. History Form A		U.S. History Form B		Biology Form A		Biology Form B	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
0	440	196	440	207	440	130	440	107
1	440	196	440	207	440	130	440	107
2	440	196	440	207	440	130	440	107
3	440	196	440	207	440	130	440	107
4	440	196	440	207	440	130	440	107
5	440	196	440	207	440	130	440	107
6	440	196	440	207	440	130	440	107
7	440	196	440	207	440	130	440	107
8	440	196	440	207	440	130	440	107
9	440	196	440	207	440	130	440	107
10	440	196	440	207	440	130	440	107
11	440	196	440	207	440	130	440	107
12	440	196	440	207	440	130	440	107
13	440	196	440	207	440	130	440	107
14	480	156	451	196	440	130	440	107
15	527	109	520	126	440	130	440	107
16	556	80	555	92	440	130	440	107
17	577	62	578	68	459	112	451	96
18	594	51	596	55	487	85	476	77
19	608	44	611	46	509	69	497	65
20	620	39	623	40	527	58	514	57
21	**630**	**35**	**634**	**36**	543	50	529	51
22	640	31	644	32	556	45	543	46
23	648	29	653	30	568	40	555	42
24	656	27	661	28	578	37	567	39
25	664	25	668	26	588	34	577	36
26	670	24	675	24	597	32	587	34
27	677	23	682	23	605	30	596	32
28	683	22	688	22	613	28	604	30
29	689	21	694	21	621	27	613	29
30	695	20	**700**	**21**	628	25	621	28

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 43. U.S. History & Biology I, Raw Score to Scale Score Conversions and SEMs
(continued)

Raw Score	U.S. History Form A Scale Score	SEM	U.S. History Form B Scale Score	SEM	Biology Form A Scale Score	SEM	Biology Form B Scale Score	SEM
31	**700**	**19**	706	20	635	24	628	27
32	706	19	712	20	641	24	635	26
33	711	19	717	20	648	23	643	25
34	716	18	723	19	**654**	**22**	649	25
35	722	18	728	19	660	22	**656**	**24**
36	727	18	734	19	666	21	663	24
37	732	18	740	19	672	21	670	23
38	737	18	745	19	678	21	676	23
39	742	18	751	19	684	20	683	23
40	748	18	757	19	689	20	690	22
41	753	18	763	19	695	20	696	22
42	759	18	769	19	**701**	**20**	**703**	**22**
43	764	18	**775**	**19**	707	20	710	22
44	770	18	781	19	713	20	717	22
45	**776**	**19**	788	20	720	20	724	22
46	782	19	794	20	726	21	731	22
47	789	19	801	20	733	21	738	23
48	796	20	809	21	740	21	746	23
49	803	20	816	21	747	22	754	23
50	811	21	825	22	755	23	763	24
51	820	22	834	23	764	24	772	25
52	829	23	843	24	**774**	**26**	**782**	**26**
53	839	25	854	26	785	28	793	28
54	851	27	867	28	797	31	805	30
55	864	29	882	31	812	34	820	33
56	880	33	901	36	831	40	837	37
57	901	39	928	46	856	48	861	44
58	932	50	976	70	893	63	894	57
59	987	81	999	86	960	98	955	89
60	999	90	.	.	999	125	999	119

*SEM=Standard Error of Measurement; BOLD=Scale Score at or closest to cut scores

Table 44. Spring 2013, Total Group Factor Analysis Results: Eigenvalues

Content	Form	KMO Statistic	Initial Eigenvalue		Ratio 1st / 2nd Eigenvalue
			Total	Variance	
Algebra I	A	0.97	11.37	0.91	9.86
	B	0.97	11.42	0.89	10.68
Algebra I	A	0.97	10.56	0.91	10.61
	B	0.97	9.78	0.95	12.88
Biology I	A	0.97	10.34	0.97	13.83
	B	0.97	8.42	0.99	11.75
English II	AA	0.95	8.07	0.87	7.73
	AB	0.95	8.06	0.87	7.48
	BA	0.95	7.68	0.91	9.96
	BB	0.95	8.23	0.91	10.13
English III	AA	0.97	10.52	0.91	10.53
	AB	0.97	10.24	0.90	10.92
	BA	0.95	13.83	0.73	3.60
	BB	0.95	13.84	0.72	3.57
Geometry	A	0.98	14.66	0.88	7.04
	B	0.98	12.83	0.91	9.61
U.S. History	A	0.98	12.01	0.98	11.45
	B	0.97	9.50	0.99	10.42

Note: KMO = Kaiser's Measure of Sampling Adequacy.

Table 45. Spring 2013, Subgroup Factor Analysis Results: Eigenvalues

Content	Form	Subgroup	KMO Statistic	Initial Eigenvalue		Ratio 1st / 2nd Eigenvalue
				Total	Variance	
Algebra I	A	Accommodated	0.94	9.71	0.79	9.73
		ELL	0.92	11.07	0.70	8.48
		IEP	0.94	10.70	0.77	8.81
	B	Accommodated	0.92	12.12	0.67	8.92
		ELL	0.90	12.60	0.63	8.37
		IEP	0.93	13.60	0.70	10.19
Algebra II	A	Accommodated	0.92	8.38	0.75	8.36
		ELL	0.85	11.95	0.52	6.04
		IEP	0.92	8.12	0.75	8.00
	B	Accommodated	0.86	9.17	0.59	7.87
		ELL	0.78	13.11	0.41	5.85
		IEP	0.90	9.54	0.67	10.04
Biology I	A	Accommodated	0.93	9.49	0.79	9.80
		ELL	0.89	9.60	0.63	7.97
		IEP	0.94	10.72	0.79	10.74
	B	Accommodated	0.88	9.61	0.62	10.53
		ELL	0.78	8.84	0.44	6.13
		IEP	0.90	10.68	0.65	10.56

Note: KMO = Kaiser's Measure of Sampling Adequacy; ELL = English Language Learners; IEP = Individualized Education Program.

Table 45. Spring 2013, Subgroup Factor Analysis Results: Eigenvalues *(continued)*

Content	Form	Subgroup	KMO Statistic	Initial Eigenvalue		Ratio 1st / 2nd Eigenvalue
				Total	Variance	
English II	AA	Accommodated	0.89	10.79	0.60	8.30
		ELL	0.70	11.03	0.32	4.21
		IEP	0.89	11.16	0.60	8.13
	AB	Accommodated	0.87	10.51	0.57	7.42
		ELL	0.70	12.74	0.32	4.85
		IEP	0.87	10.98	0.56	7.66
	BA	Accommodated	0.86	13.17	0.52	7.44
		ELL	0.72	17.13	0.35	5.96
		IEP	0.84	12.32	0.49	7.00
	BB	Accommodated	0.86	13.62	0.52	8.10
		ELL	0.72	15.54	0.35	6.00
		IEP	0.84	11.42	0.50	6.98
English III	AA	Accommodated	0.90	7.72	0.71	7.55
		ELL	0.63	8.65	0.28	3.83
		IEP	0.91	7.74	0.73	7.70
	AB	Accommodated	0.89	7.59	0.68	7.93
		ELL	0.61	10.41	0.27	4.26
		IEP	0.90	7.69	0.70	7.83
	BA	Accommodated	0.88	13.34	0.53	3.14
		ELL	0.67	19.10	0.30	3.11
		IEP	0.90	13.54	0.57	3.23
	BB	Accommodated	0.88	12.17	0.52	3.12
		ELL	0.66	15.25	0.28	3.01
		IEP	0.89	12.66	0.55	3.14
Geometry	A	Accommodated	0.95	9.96	0.81	5.42
		ELL	0.92	12.84	0.66	5.87
		IEP	0.95	10.01	0.81	5.40
	B	Accommodated	0.94	12.41	0.73	7.88
		ELL	0.89	16.75	0.59	8.07
		IEP	0.94	13.17	0.76	8.63
U.S. History	A	Accommodated	0.95	12.83	0.80	11.75
		ELL	0.87	11.16	0.58	9.07
		IEP	0.95	13.50	0.80	11.53
	B	Accommodated	0.90	11.49	0.65	9.43
		ELL	0.76	9.67	0.41	5.59
		IEP	0.90	10.97	0.67	9.59

Note: KMO = Kaiser's Measure of Sampling Adequacy; ELL = English Language Learners; IEP = Individualized Education Program.

Table 46. Spring 2013, Proficiency Level Cut Scores and Standard Error of Measurement (SEM)

Content	Form	Cut 1		Cut 2		Cut 3	
		Scale Score	SEM at Cut1	Scale Score	SEM at Cut2	Scale Score	SEM at Cut3
Algebra I	A	664	22	700	14	764	13
	B	665	22	700	14	764	13
Algebra II	A	657	32	702	22	787	19
	B	657	32	702	22	787	19
Biology I	A	654	22	701	20	774	26
	B	656	24	703	22	782	26
English II	AA	613	35	702	23	817	25
	AB	615	35	703	22	824	25
	BA	615	31	704	20	822	26
	BB	617	30	705	20	819	25
English III	AA	673	25	701	21	807	17
	AB	674	24	702	20	802	17
	BA	670	21	703	18	804	16
	BB	671	21	700	18	804	16
Geometry	A	637	27	703	19	781	16
	B	638	29	703	18	781	17
U.S. History	A	630	35	700	19	776	19
	B	634	36	700	21	775	19

Table 47. Spring 2013 Classification Consistency and Accuracy Statistics

Content	Form	Accuracy	Consistency	False Positive	False Negative	Kappa
Algebra I	A	0.80	0.73	0.09	0.11	0.59
	B	0.80	0.73	0.09	0.10	0.59
Algebra II	A	0.77	0.69	0.11	0.12	0.56
	B	0.77	0.69	0.11	0.12	0.56
Biology I	A	0.75	0.66	0.13	0.12	0.54
	B	0.73	0.64	0.14	0.13	0.51
English II	AA	0.82	0.75	0.09	0.09	0.57
	AB	0.83	0.75	0.09	0.09	0.57
	BA	0.83	0.76	0.08	0.09	0.57
	BB	0.83	0.77	0.08	0.08	0.58
English III	AA	0.80	0.73	0.09	0.10	0.56
	AB	0.82	0.75	0.09	0.10	0.57
	BA	0.83	0.77	0.08	0.09	0.58
	BB	0.84	0.78	0.08	0.08	0.60
Geometry	A	0.82	0.75	0.08	0.10	0.63
	B	0.82	0.75	0.08	0.09	0.62
U.S. History	A	0.79	0.71	0.10	0.11	0.59
	B	0.79	0.71	0.10	0.11	0.57

Table 48. Accuracy and Consistency Estimates by Cut Score

Content	Form	Accuracy			Consistency		
		U/L+P+A	U+L/P+A	U+L+P/A	U/L+P+A	U+L/P+A	U+L+P/A
Algebra I	A	0.97	0.93	0.91	0.95	0.90	0.88
	B	0.97	0.93	0.91	0.95	0.90	0.87
Algebra II	A	0.94	0.91	0.91	0.92	0.88	0.88
	B	0.95	0.91	0.91	0.93	0.88	0.87
Biology I	A	0.91	0.90	0.93	0.87	0.86	0.91
	B	0.91	0.89	0.93	0.87	0.85	0.90
English II	AA	0.99	0.92	0.91	0.99	0.89	0.87
	AB	0.99	0.93	0.91	0.99	0.90	0.87
	BA	0.99	0.93	0.90	0.99	0.91	0.86
	BB	0.99	0.94	0.90	0.99	0.91	0.87
English III	AA	0.96	0.93	0.90	0.94	0.90	0.87
	AB	0.97	0.94	0.91	0.95	0.91	0.87
	BA	0.98	0.95	0.90	0.97	0.93	0.86
	BB	0.98	0.96	0.90	0.97	0.94	0.86
Geometry	A	0.97	0.93	0.92	0.96	0.90	0.88
	B	0.98	0.94	0.91	0.97	0.91	0.87
U.S. History	A	0.97	0.92	0.91	0.95	0.88	0.88
	B	0.98	0.92	0.90	0.97	0.88	0.86

Note: U = Unsatisfactory; L = Limited Knowledge; P = Proficient; A = Advanced.

Table 49. Accuracy and Consistency Estimates by Cut Score: False Positive and False Negative Rates

Content	Form	U/L+P+A		U+L/P+A		U+L+P/A	
		False Positive	False Negative	False Positive	False Negative	False Positive	False Negative
Algebra I	A	0.01	0.02	0.03	0.04	0.05	0.04
	B	0.01	0.02	0.03	0.04	0.05	0.04
Algebra II	A	0.02	0.04	0.04	0.05	0.05	0.04
	B	0.02	0.03	0.04	0.05	0.05	0.04
Biology I	A	0.04	0.05	0.05	0.05	0.04	0.03
	B	0.04	0.05	0.05	0.06	0.04	0.03
English II	AA	0.00	0.01	0.03	0.05	0.05	0.04
	AB	0.00	0.01	0.03	0.05	0.05	0.04
	BA	0.00	0.00	0.02	0.04	0.06	0.04
	BB	0.00	0.00	0.02	0.04	0.06	0.04
English III	AA	0.01	0.03	0.03	0.04	0.05	0.04
	AB	0.01	0.02	0.02	0.04	0.05	0.04
	BA	0.01	0.01	0.02	0.03	0.06	0.05
	BB	0.01	0.01	0.01	0.03	0.06	0.05
Geometry	A	0.01	0.02	0.03	0.04	0.04	0.04
	B	0.01	0.01	0.03	0.04	0.05	0.04
U.S. History	A	0.01	0.02	0.04	0.05	0.05	0.04
	B	0.01	0.02	0.04	0.05	0.05	0.05

Note: U = Unsatisfactory; L = Limited Knowledge; P = Proficient; A = Advanced.

Figures

Figure 1. Spring 2013 Algebra I Form A operational scale score histogram

Kurtosis = 4.2209628918   Skewness = −0.256668171
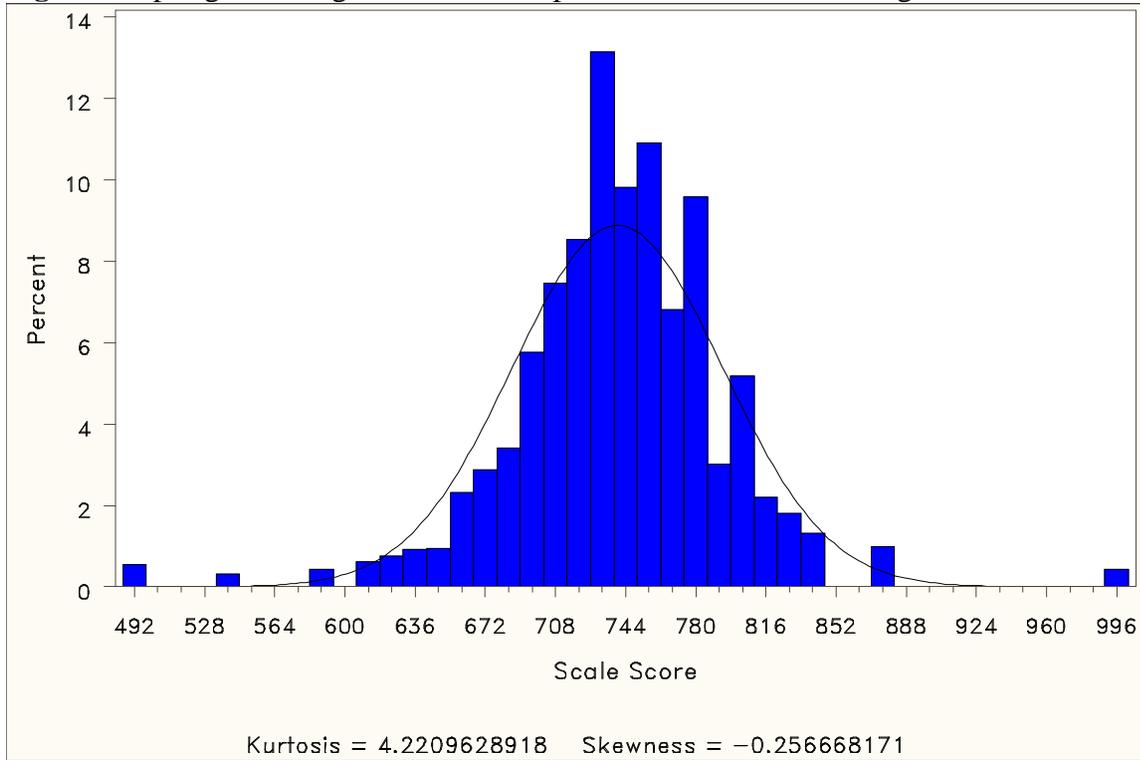
Figure 2. Spring 2013 Algebra I Form B operational scale score histogram

Kurtosis = 4.3801915904   Skewness = −0.062426429
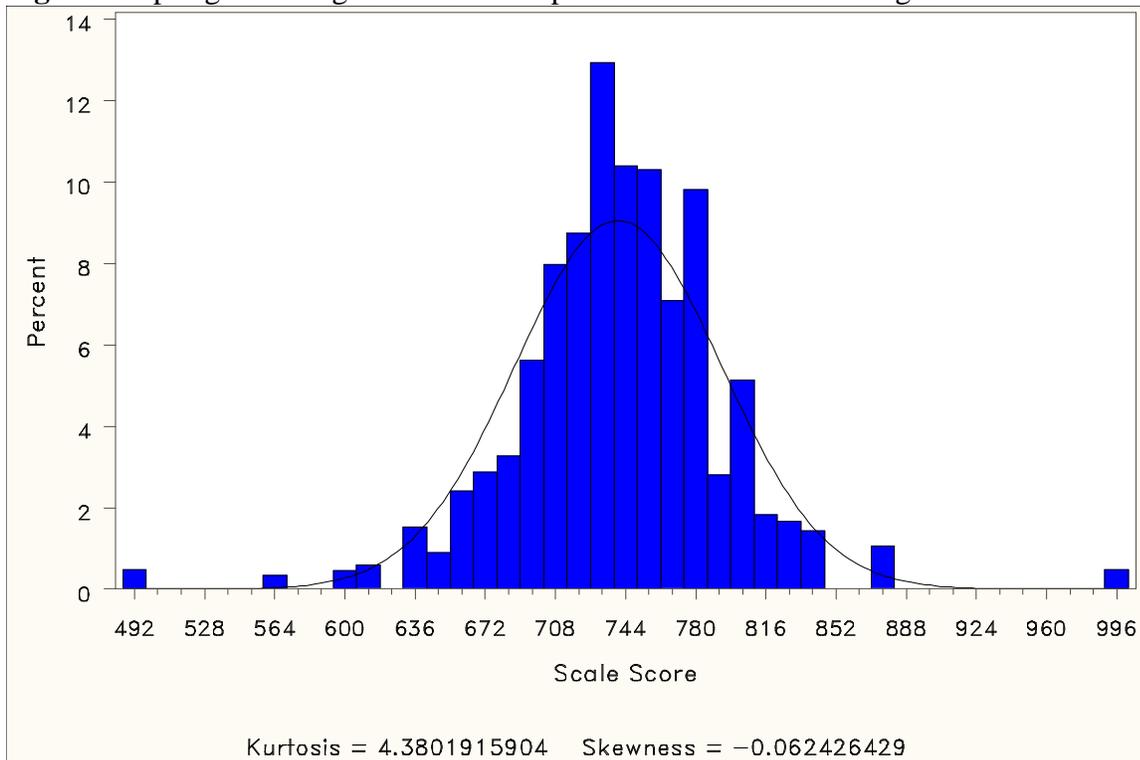
Figure 3. Spring 2013 Algebra II Form A operational scale score histogram

Kurtosis = 2.0616762546   Skewness = −0.566669132
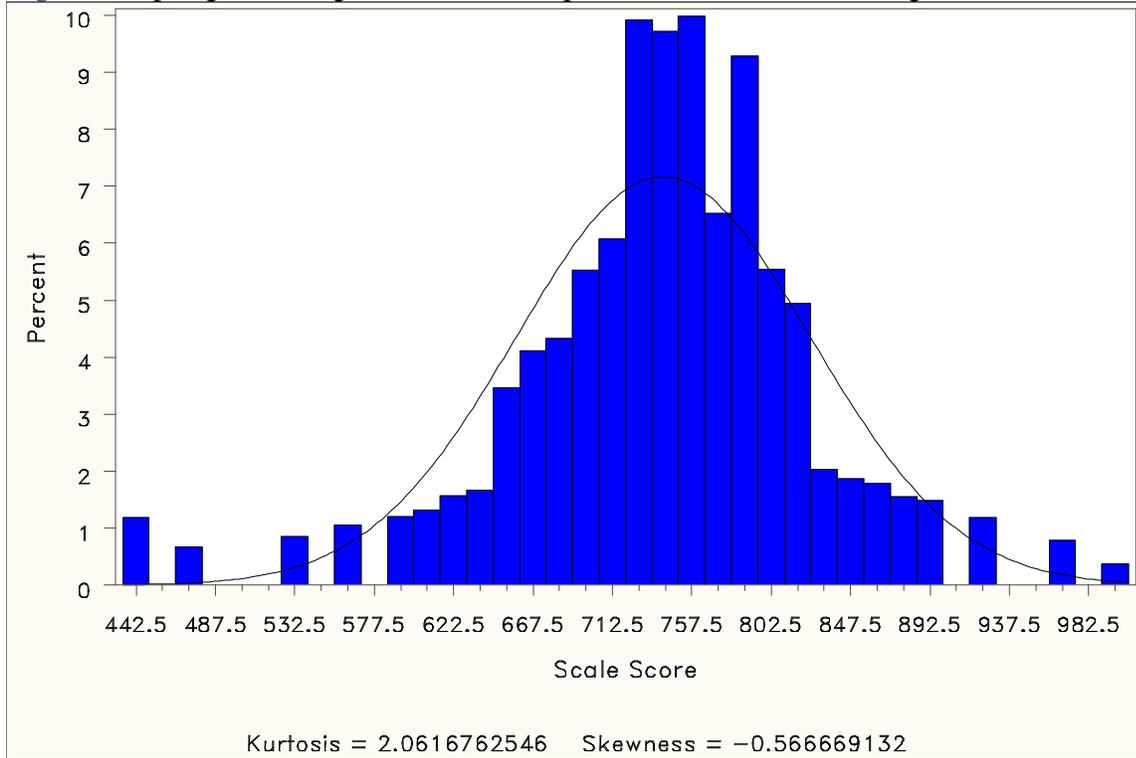
Figure 4. Spring 2013 Algebra II Form B operational scale score histogram

Kurtosis = 1.9132530633   Skewness = −0.3874274
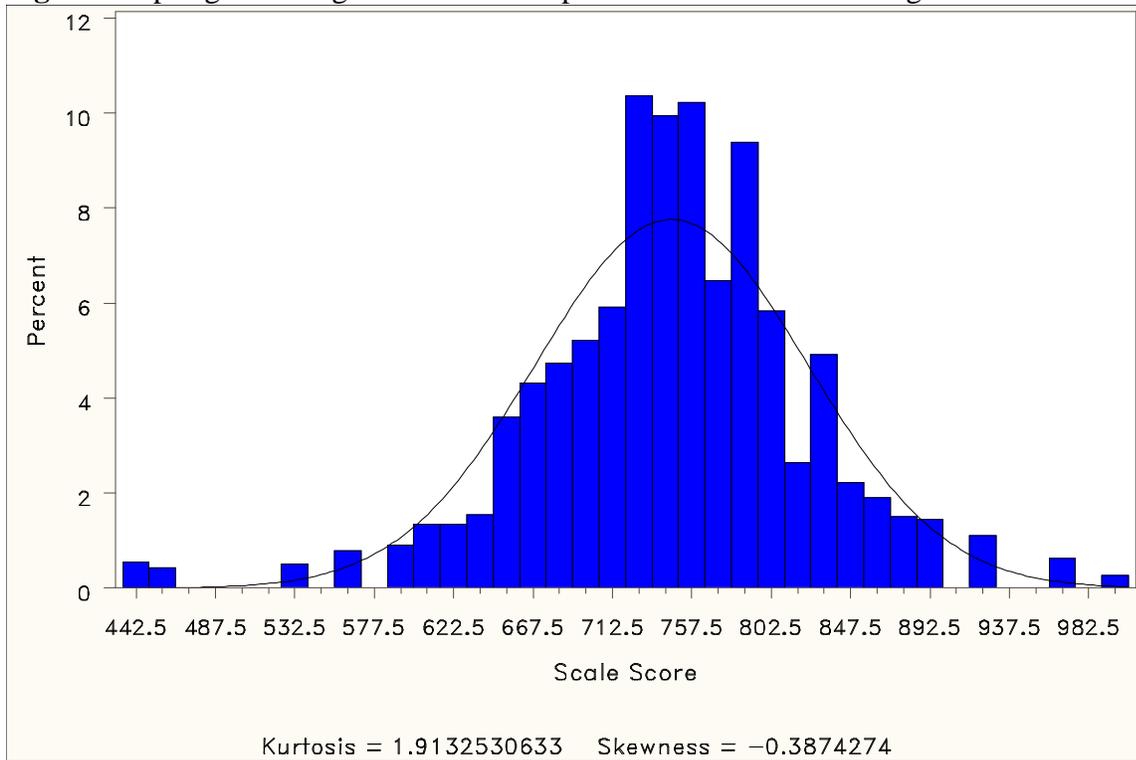
Figure 5. Spring 2013 Biology I Form A operational scale score histogram

Kurtosis = 1.255623594   Skewness = −0.428784464
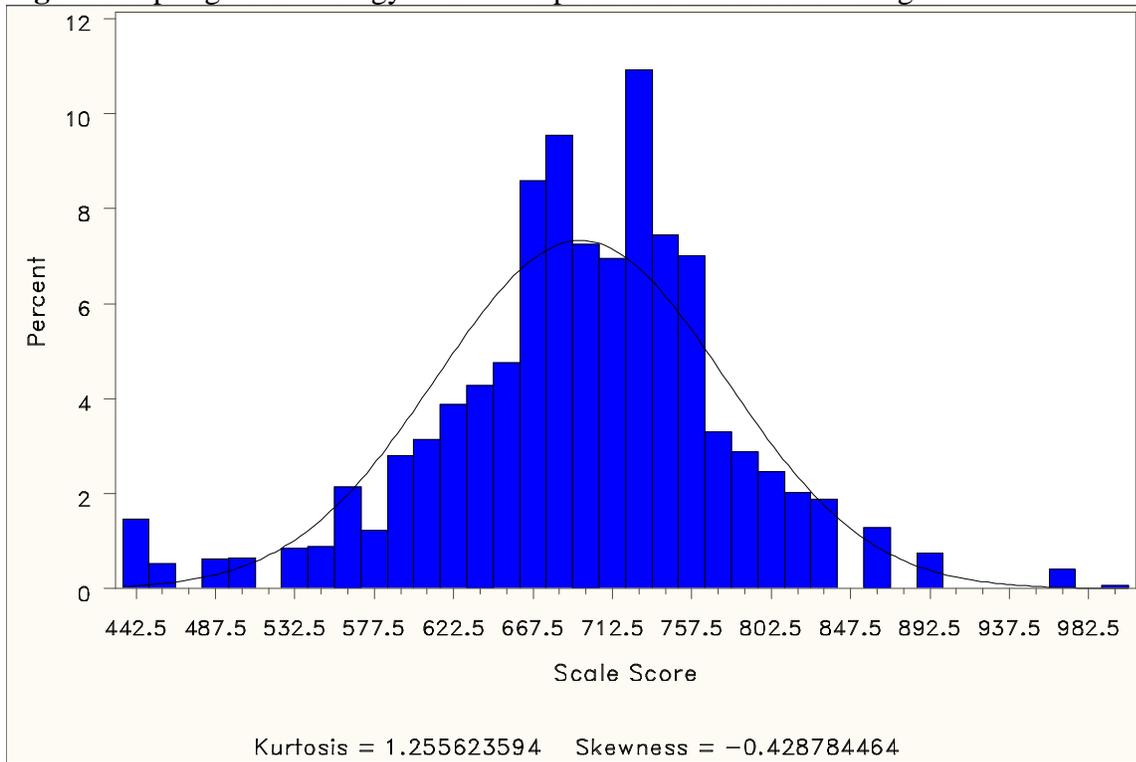
Figure 6. Spring 2013 Biology I Form B operational scale score histogram

Kurtosis = 1.0298933834   Skewness = −0.234225955
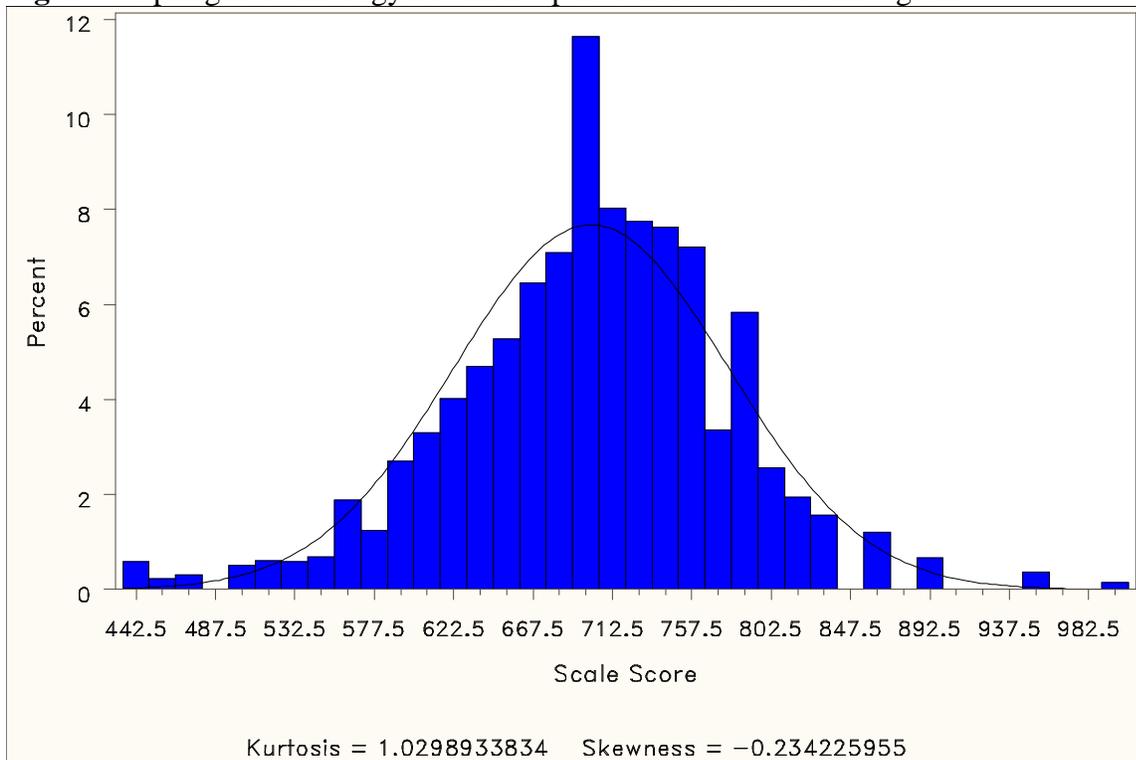
Figure 7. Spring 2013 English II Form AA operational scale score histogram

Kurtosis = 2.7906623146   Skewness = −0.762335875
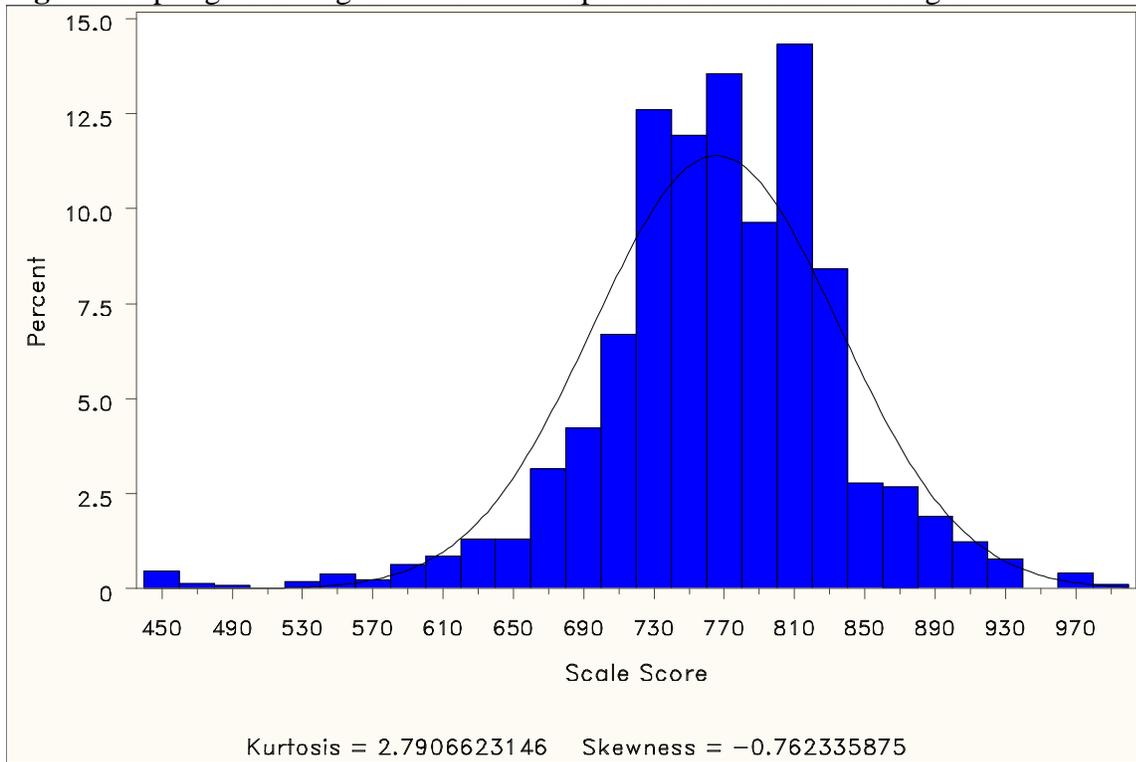
Figure 8. Spring 2013 English II Form AB operational scale score histogram

Kurtosis = 2.7517900719   Skewness = −0.684362972
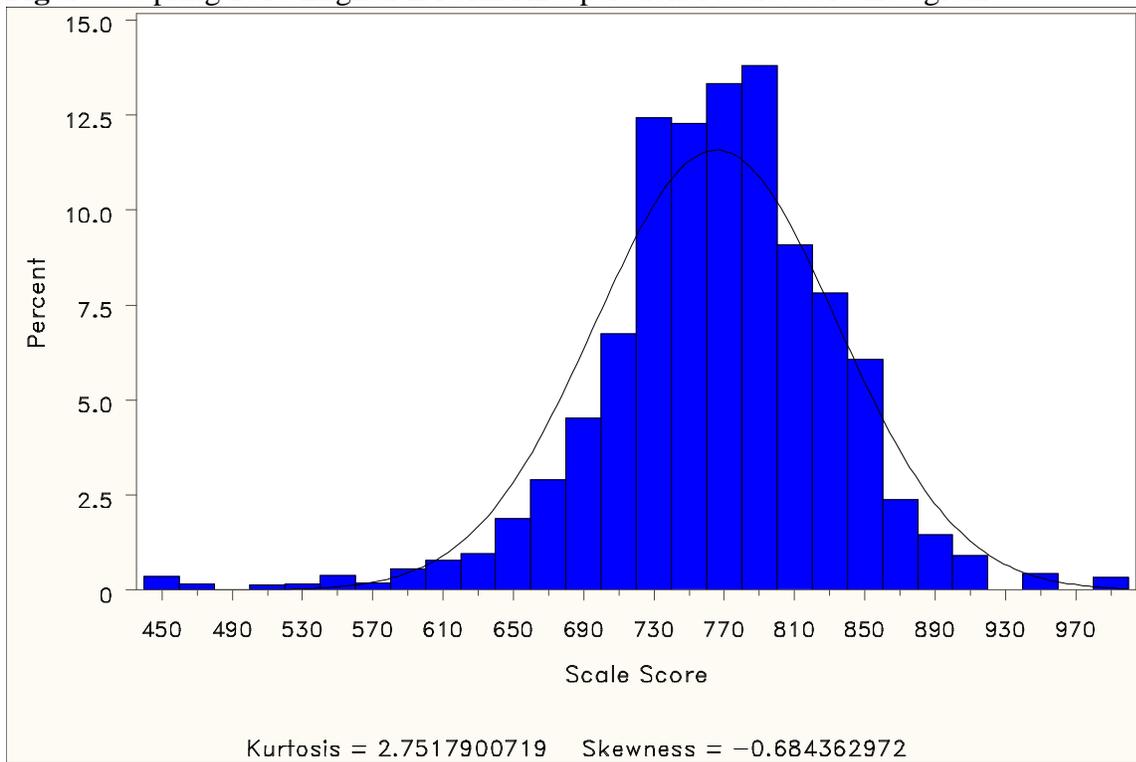
Figure 9. Spring 2013 English II Form BA operational scale score histogram

Kurtosis = 2.3356273929   Skewness = −0.377977227
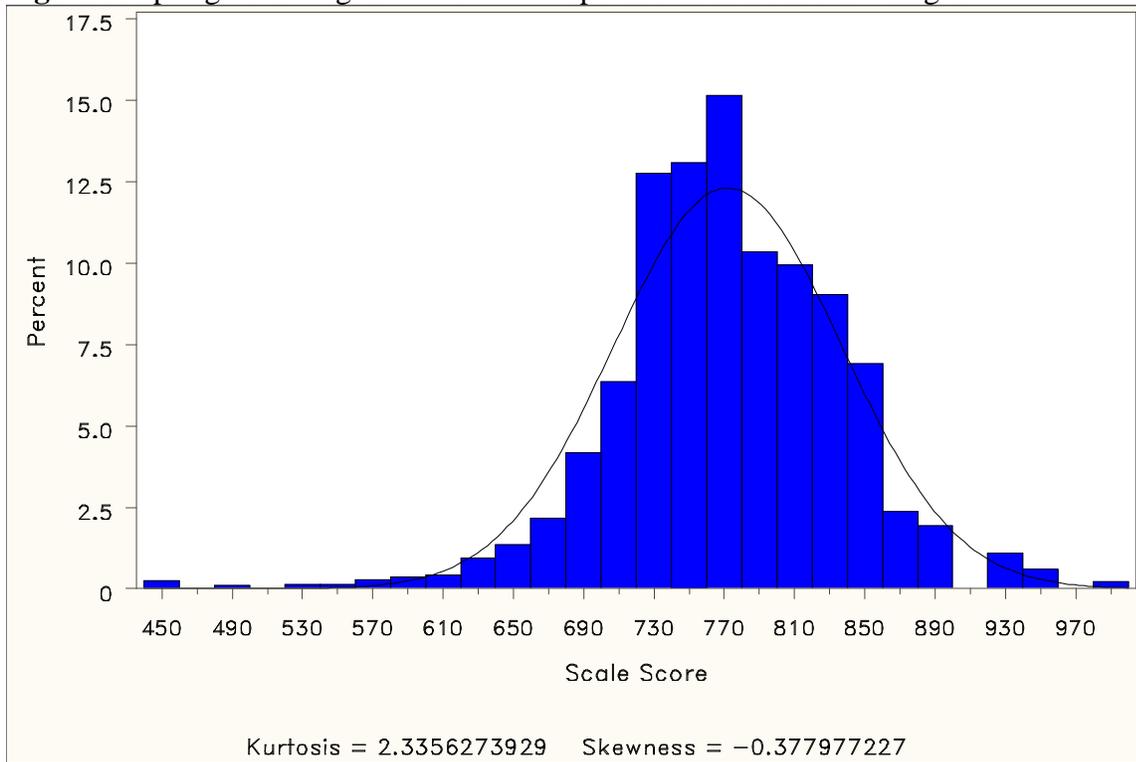
Figure 10. Spring 2013 English II Form BB operational scale score histogram

Kurtosis = 1.8364525217   Skewness = −0.237377437
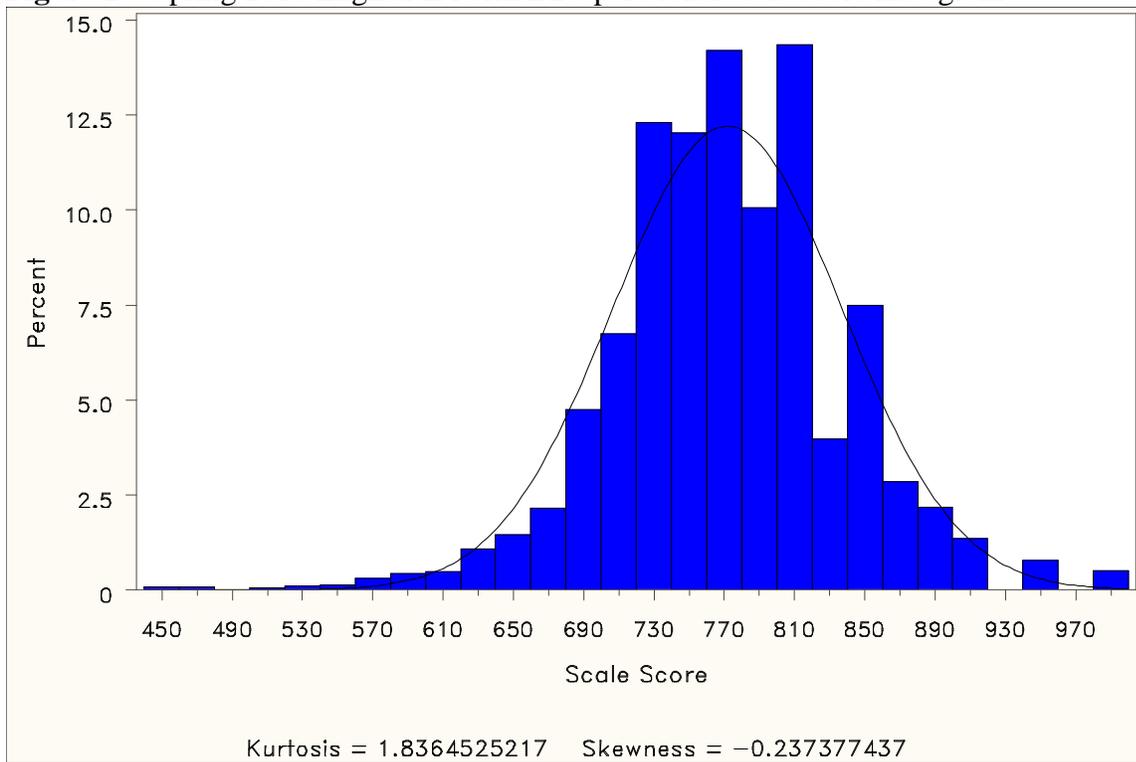
Figure 11. Spring 2013 English III Form AA operational scale score histogram

Kurtosis = 3.2916423959   Skewness = −1.078664483
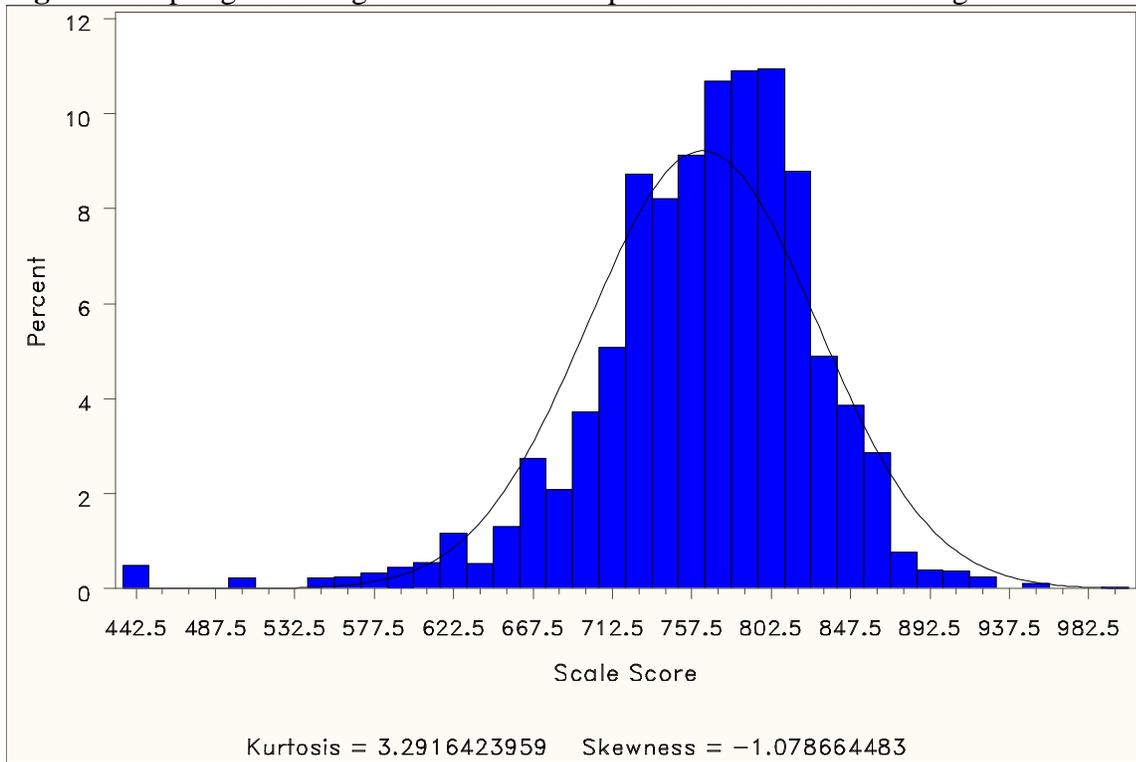
Figure 12. Spring 2013 English III Form AB operational scale score histogram

Kurtosis = 2.7004566821   Skewness = −0.868503542
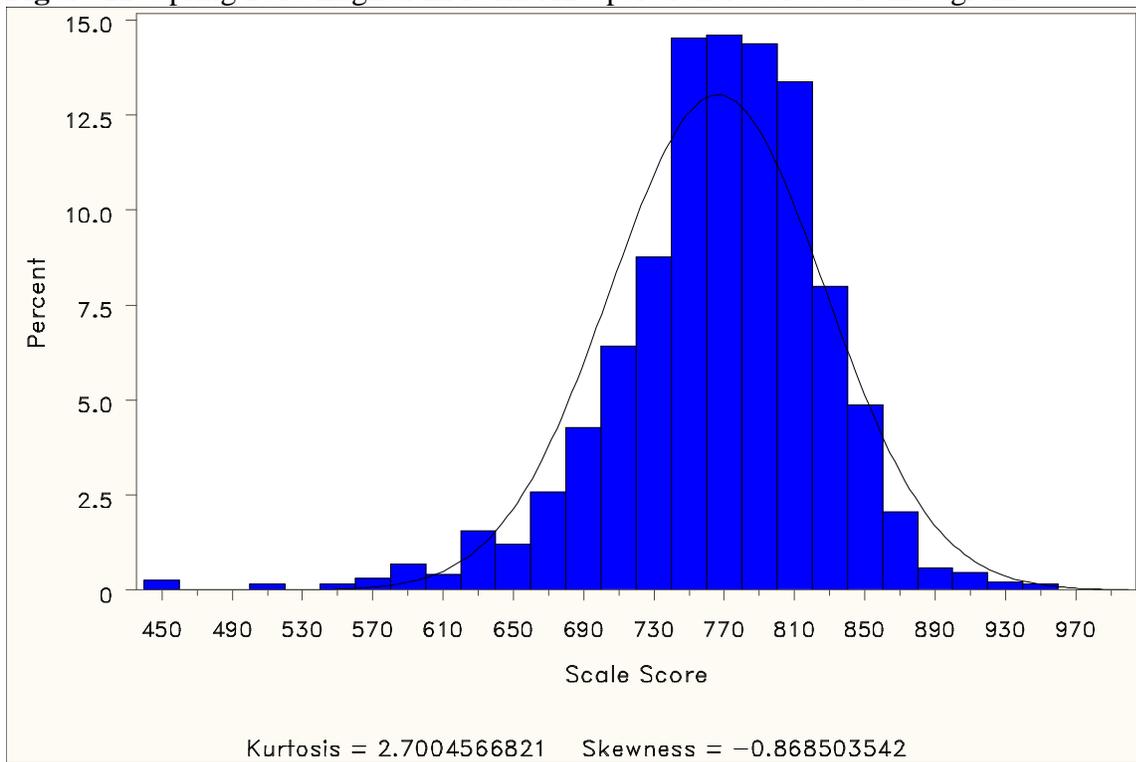
Figure 13. Spring 2013 English III Form BA operational scale score histogram

Kurtosis = 2.9375627726   Skewness = −0.567744748
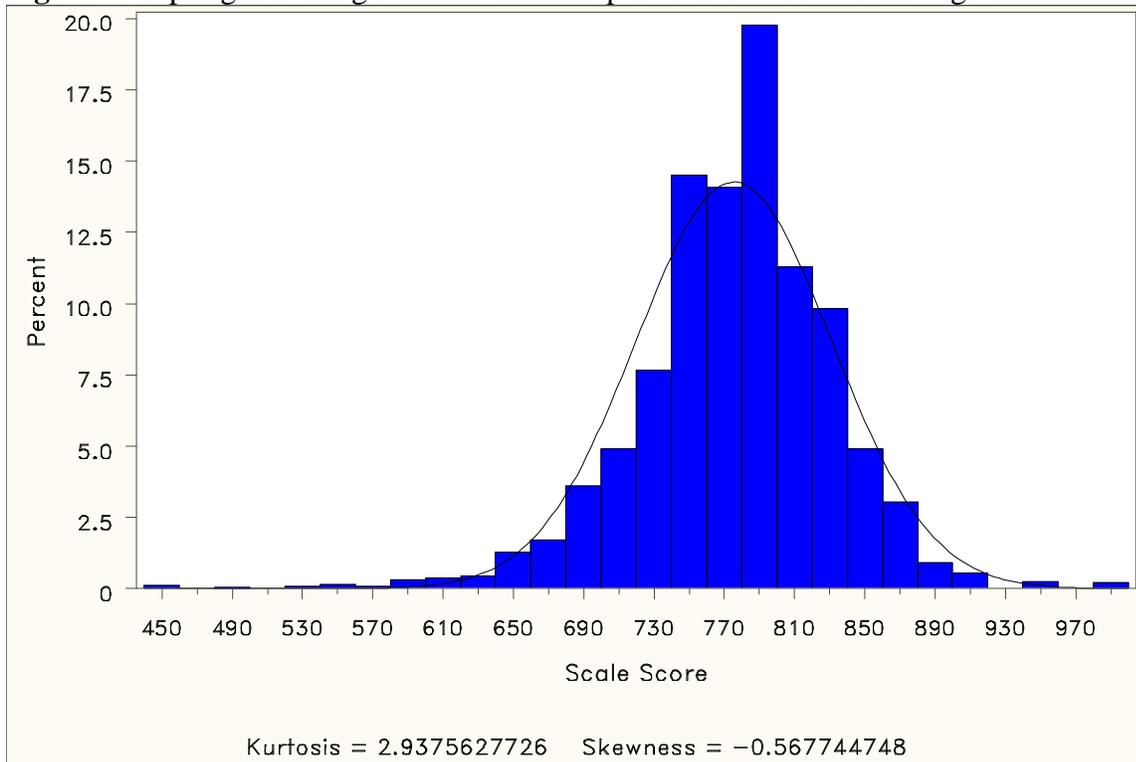
Figure 14. Spring 2013 English III Form BB operational scale score histogram

Kurtosis = 2.6133003632   Skewness = −0.512989137
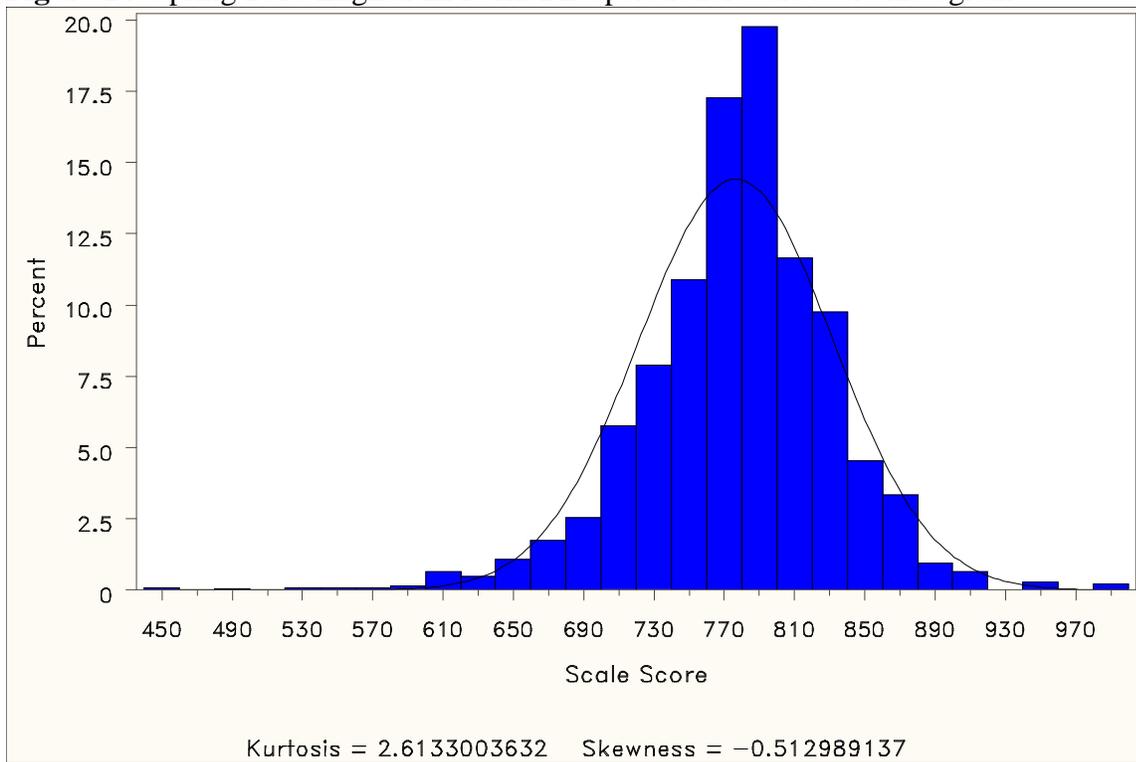
Figure 15. Spring 2013 Geometry Form A operational scale score histogram

Kurtosis = 1.6912539625   Skewness = −0.242248519
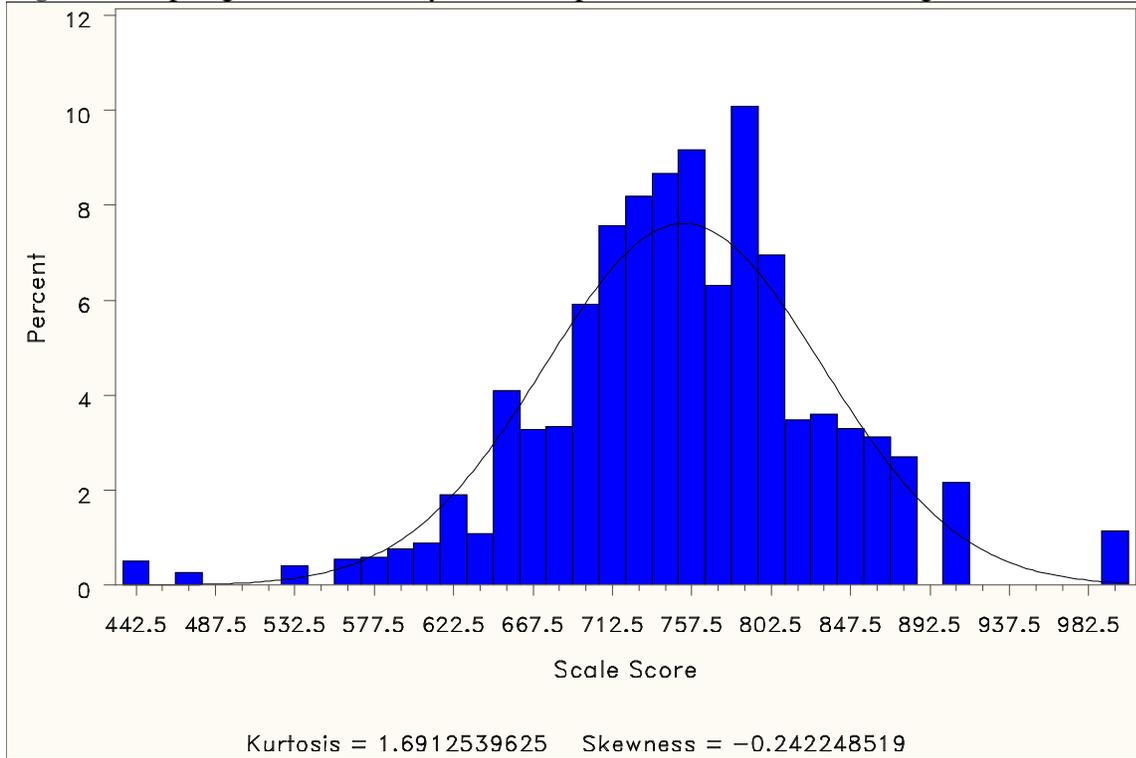
Figure 16. Spring 2013 Geometry Form B operational scale score histogram

Kurtosis = 1.6719673091   Skewness = −0.043042486
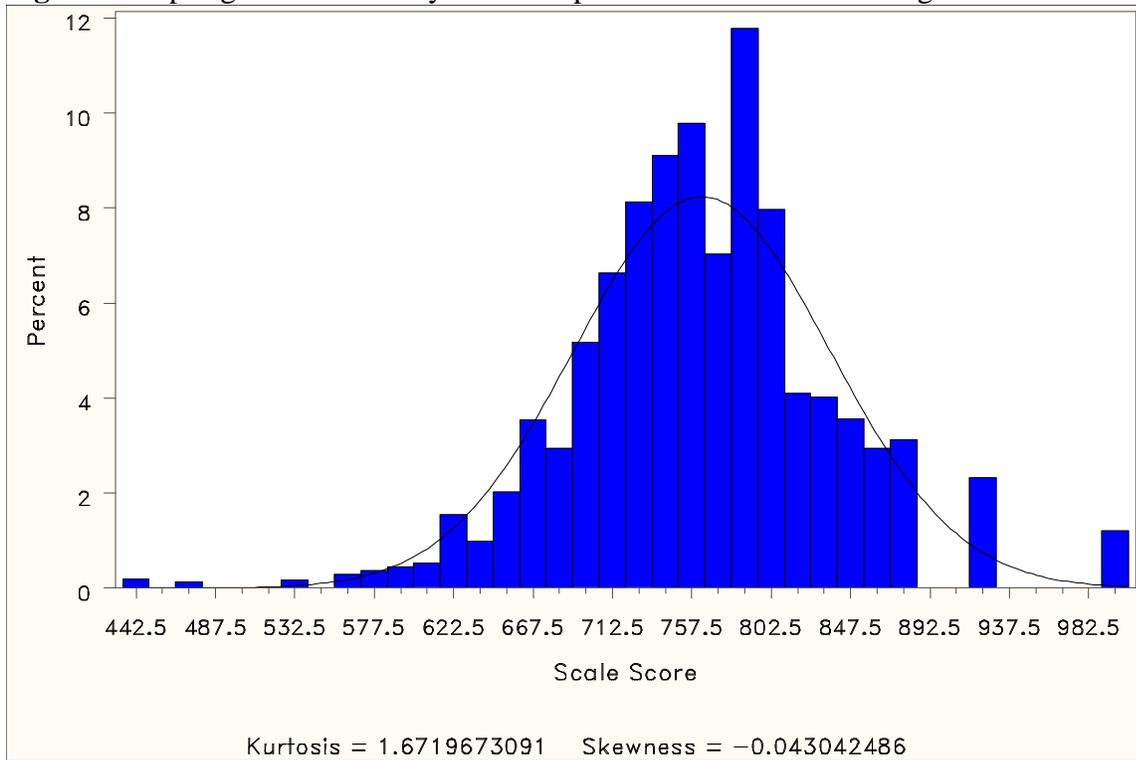
Figure 17. Spring 2013 U.S. History Form A operational scale score histogram

Kurtosis = 1.693651205   Skewness = −0.491453002
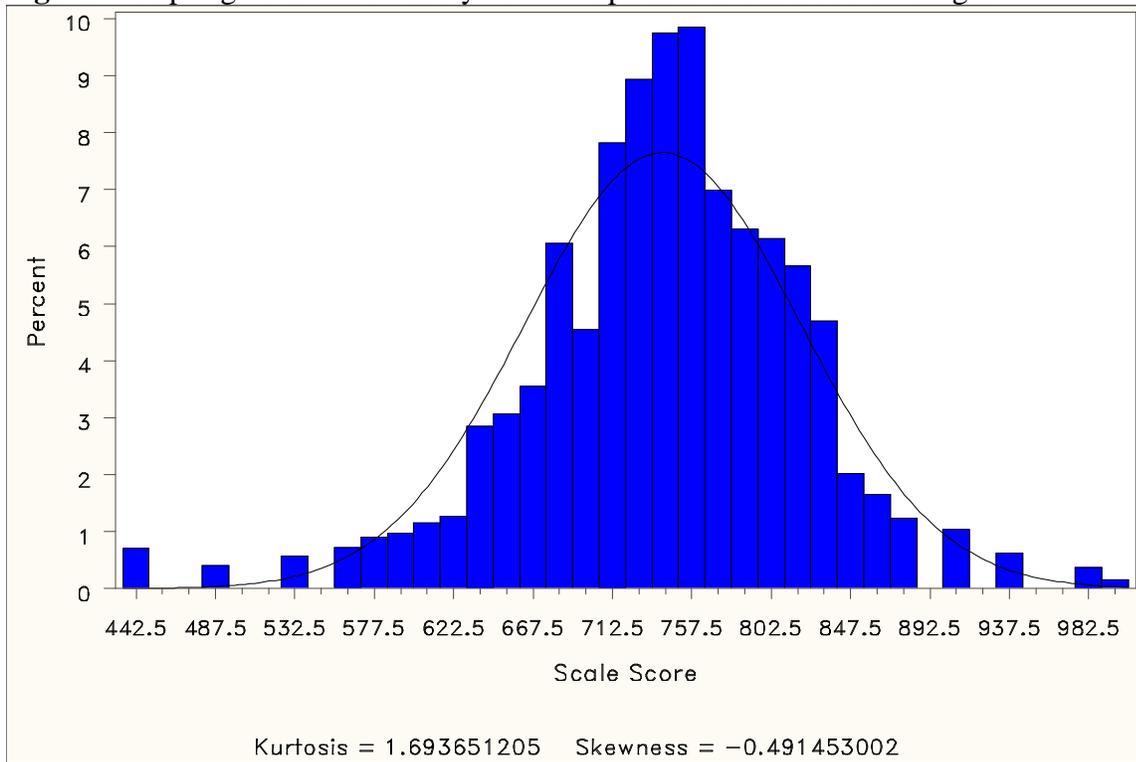
Figure 18. Spring 2013 U.S. History Form B operational scale score histogram

Kurtosis = 1.7444118576   Skewness = −0.426314717
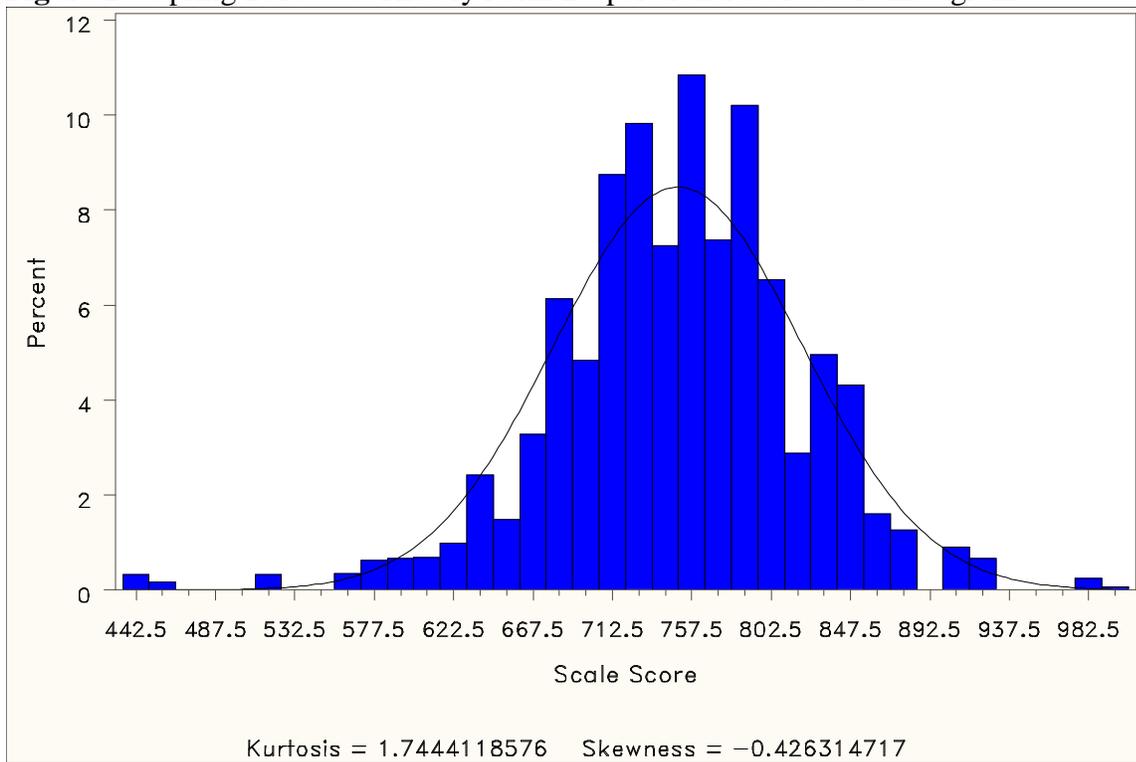
Figure 19. Spring 2013 Algebra I Form A operational test characteristic curve and standard error of measurement curve
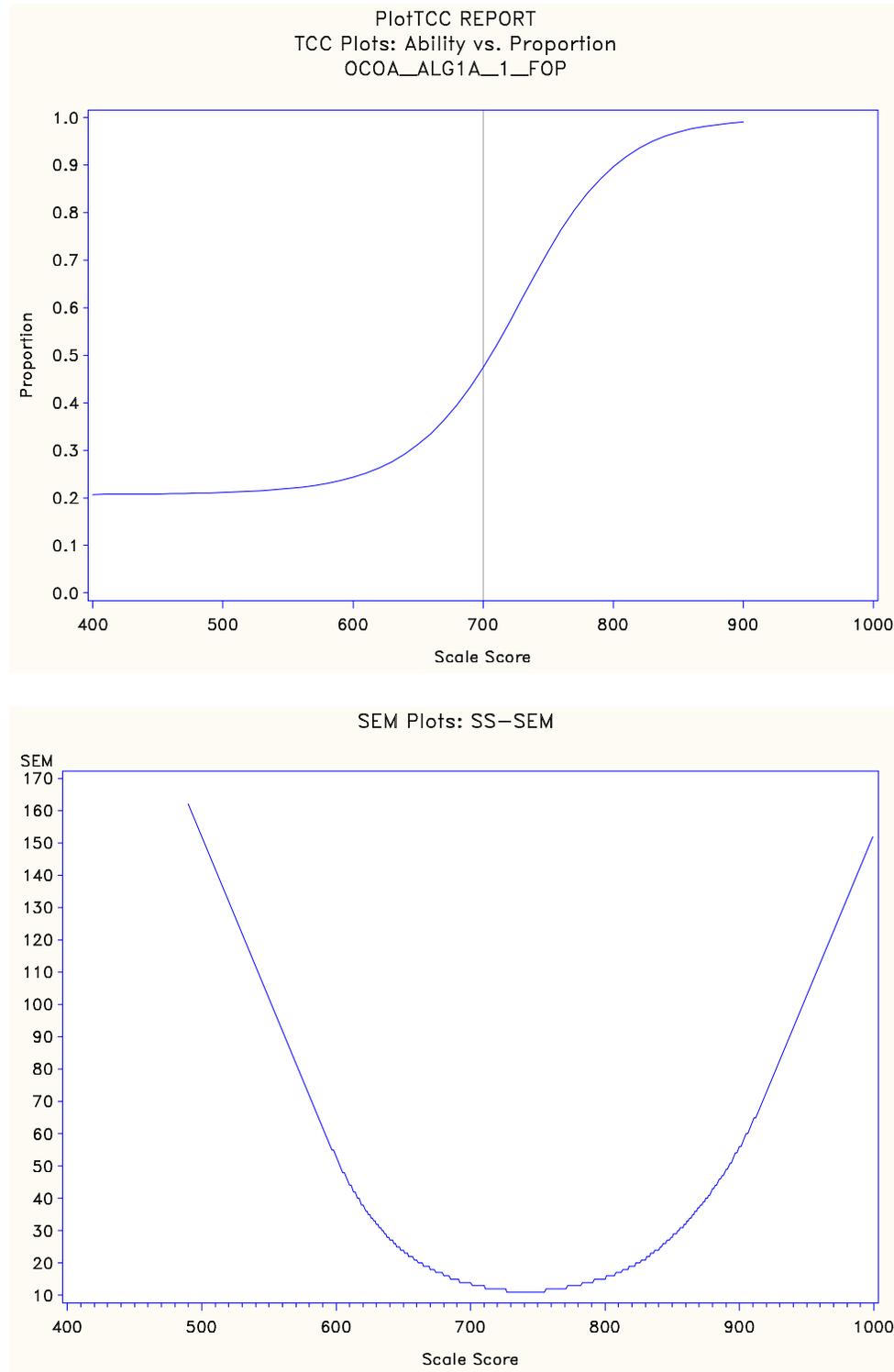
Figure 20. Spring 2013 Algebra I Form B operational test characteristic curve and standard error of measurement curve
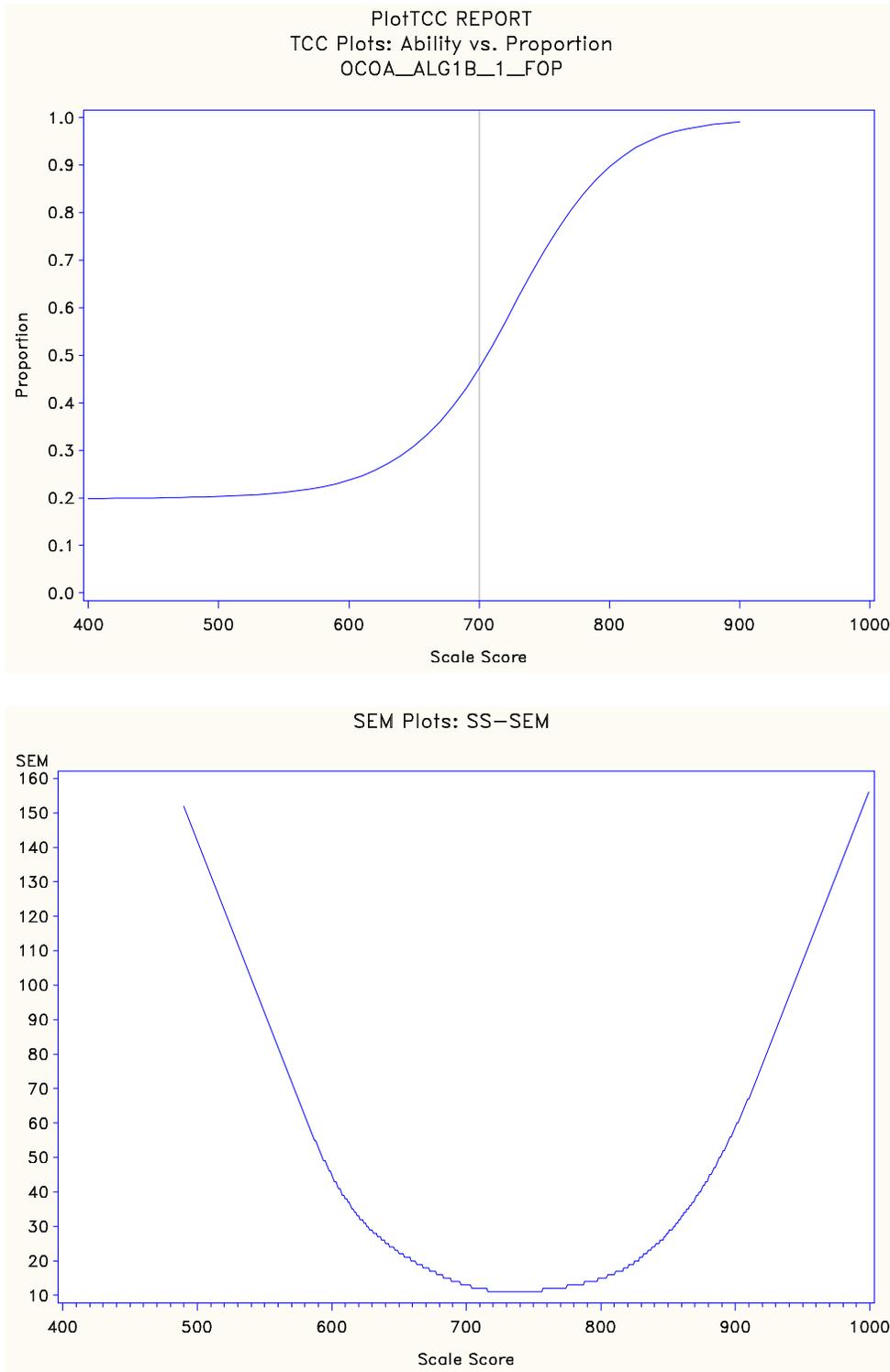
Figure 21. Spring 2013 Algebra II Form A operational test characteristic curve and standard error of measurement curve
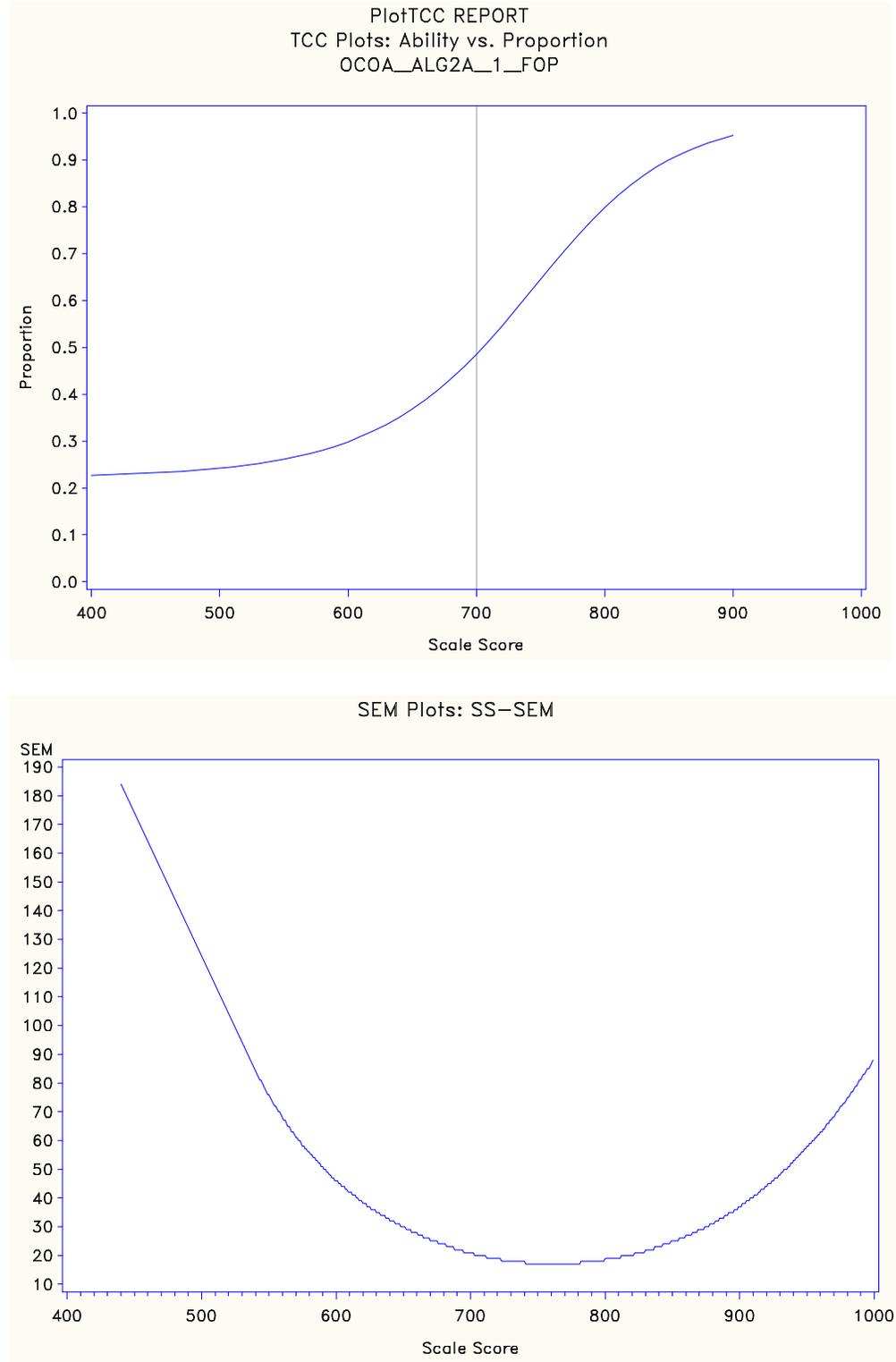
Figure 22. Spring 2013 Algebra II Form B operational test characteristic curve and standard error of measurement curve
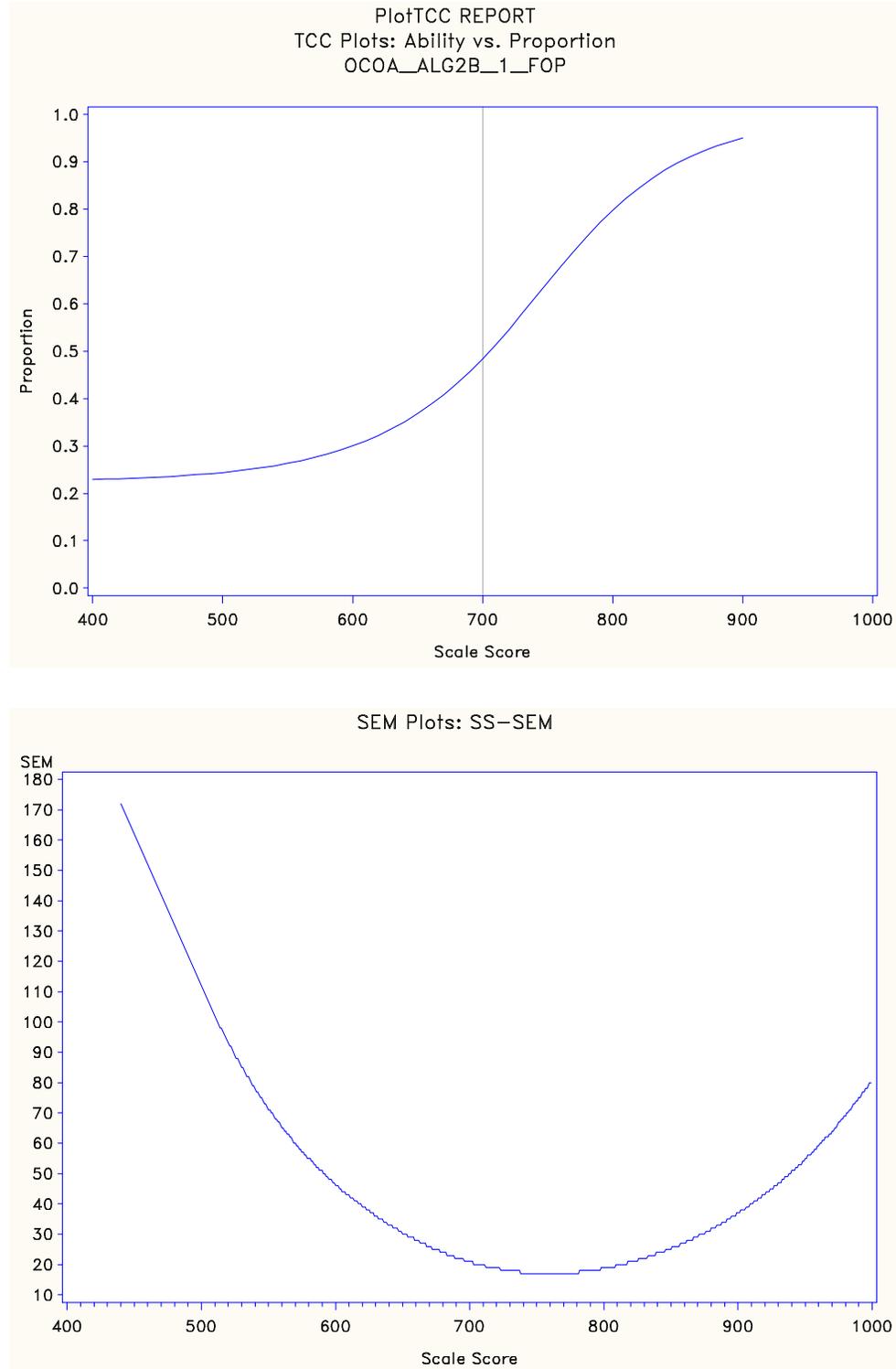
Figure 23. Spring 2013 Biology I Form A operational test characteristic curve and standard error of measurement curve
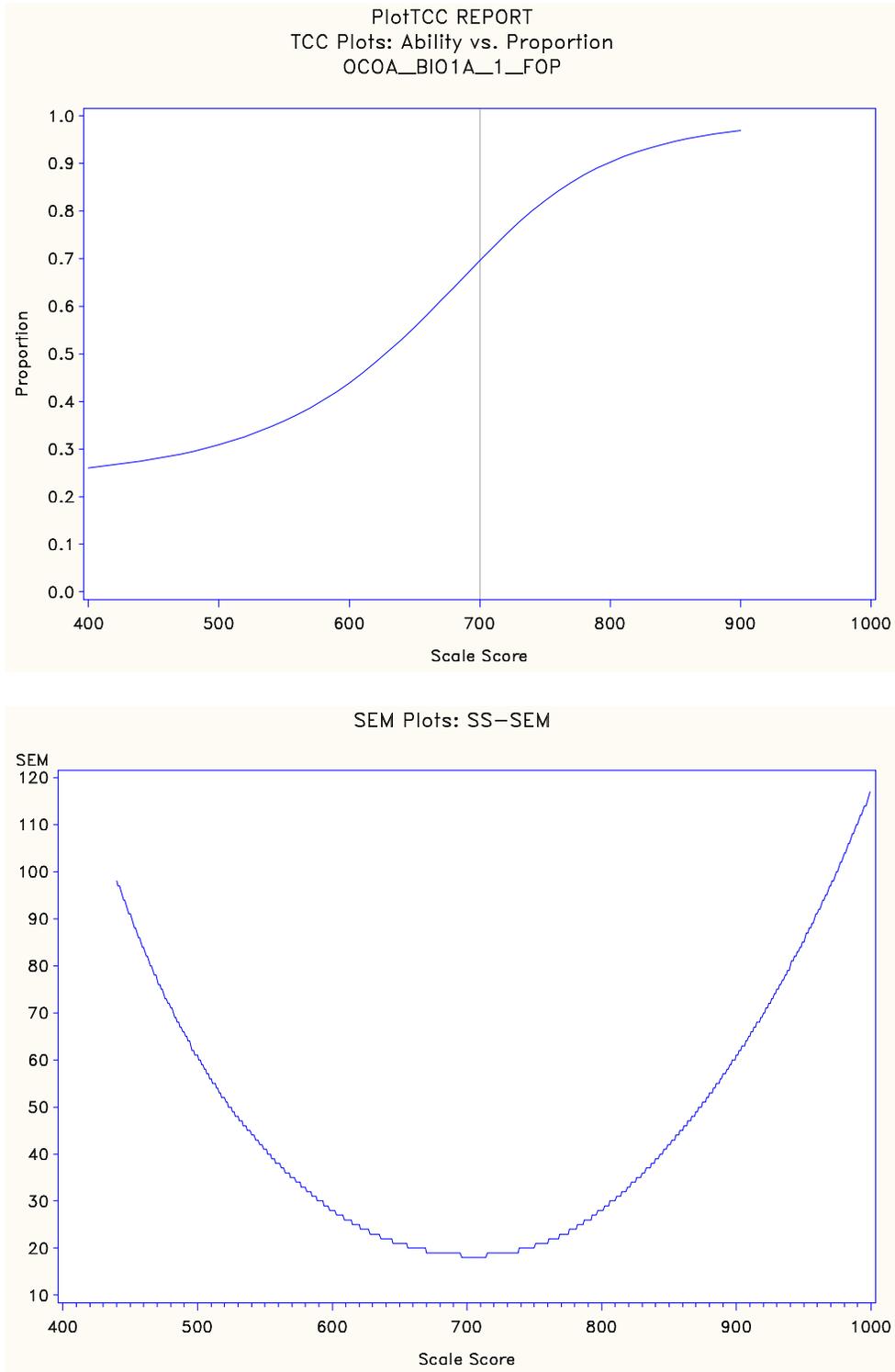
Figure 24. Spring 2013 Biology I Form B operational test characteristic curve and standard error of measurement curve
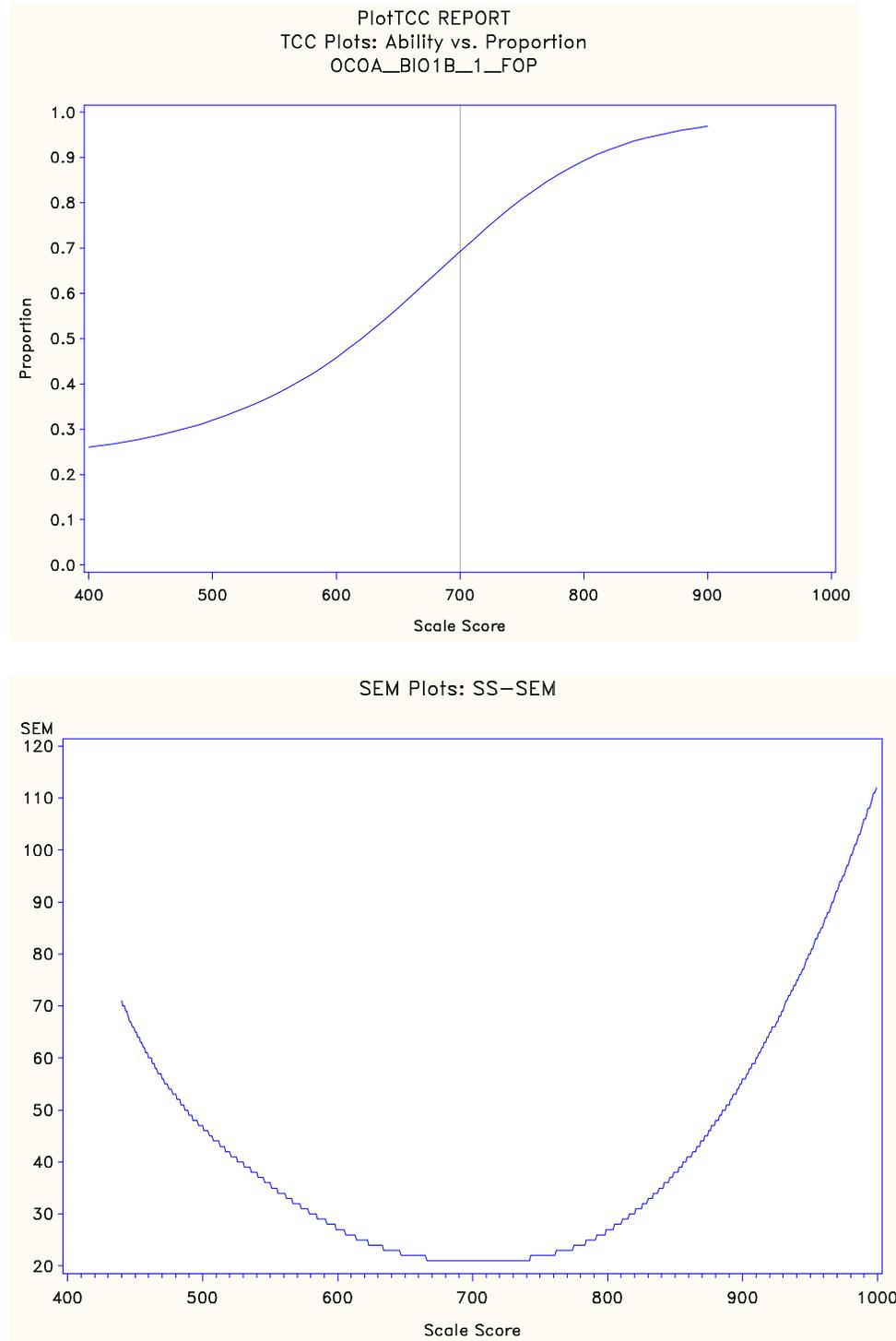
Figure 25. Spring 2013 English II Form AA operational test characteristic curve and standard error of measurement curve
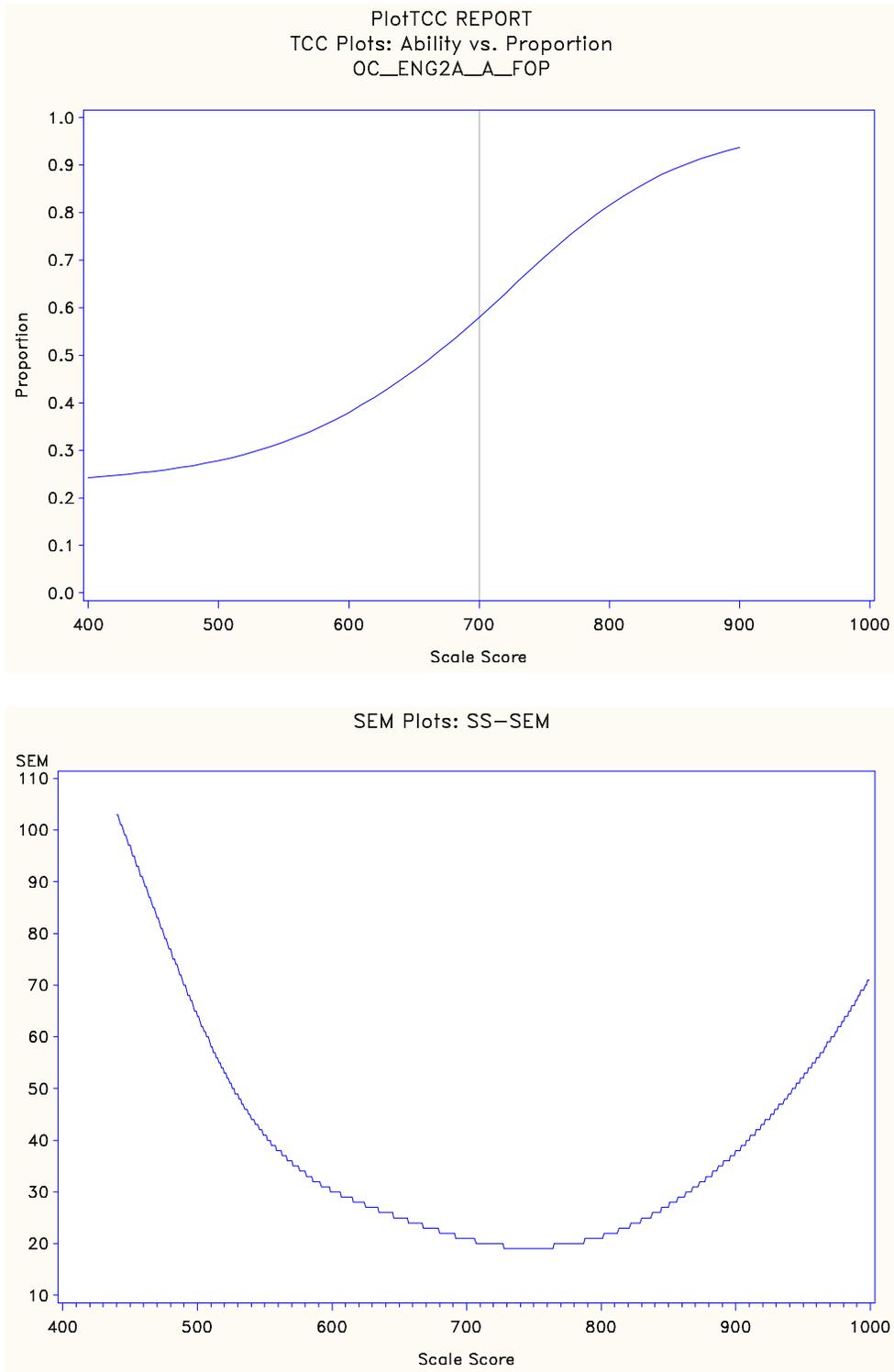
Figure 26. Spring 2013 English II Form AB operational test characteristic curve and standard error of measurement curve
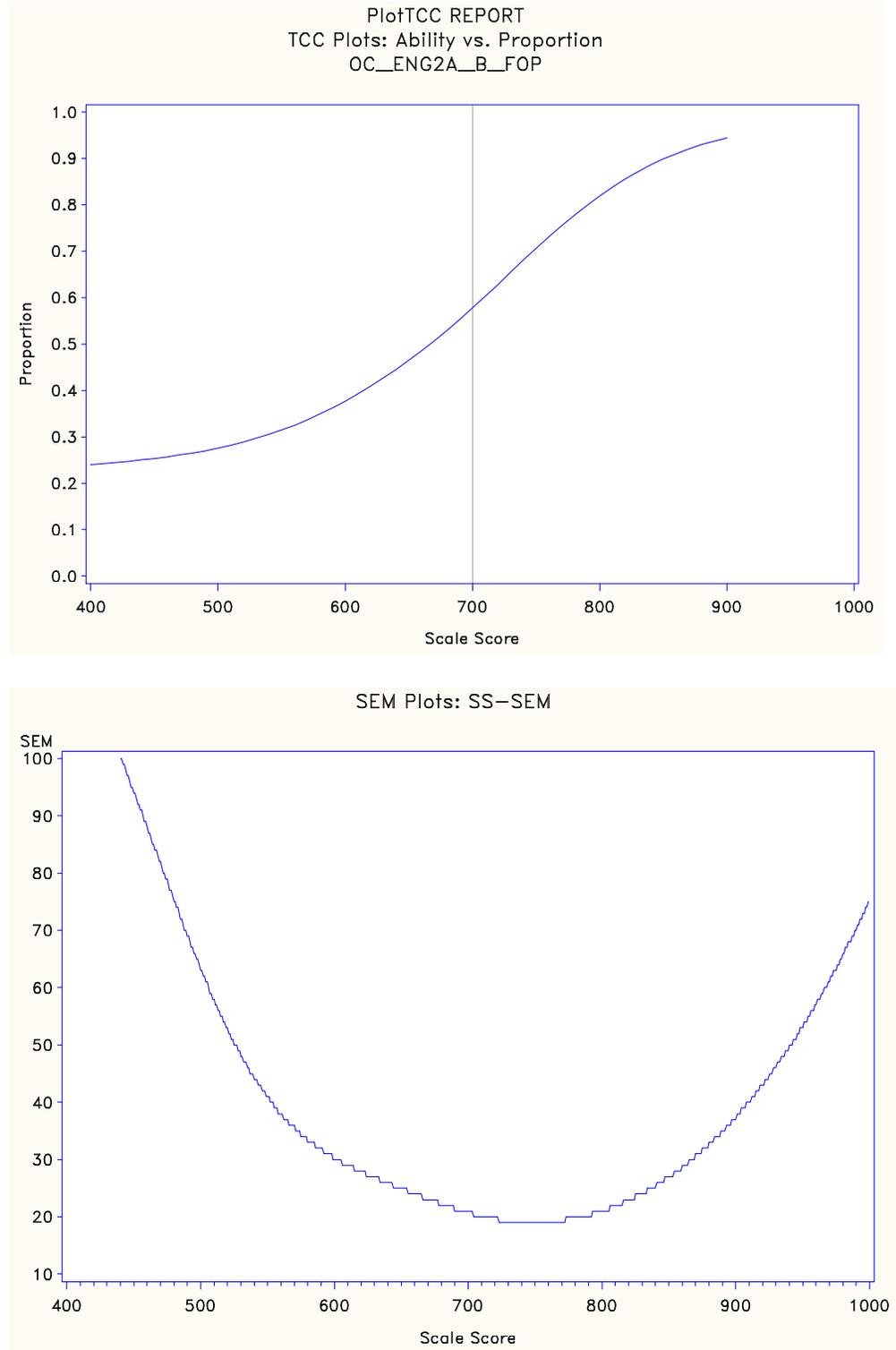
Figure 27. Spring 2013 English II Form BA operational test characteristic curve and standard error of measurement curve
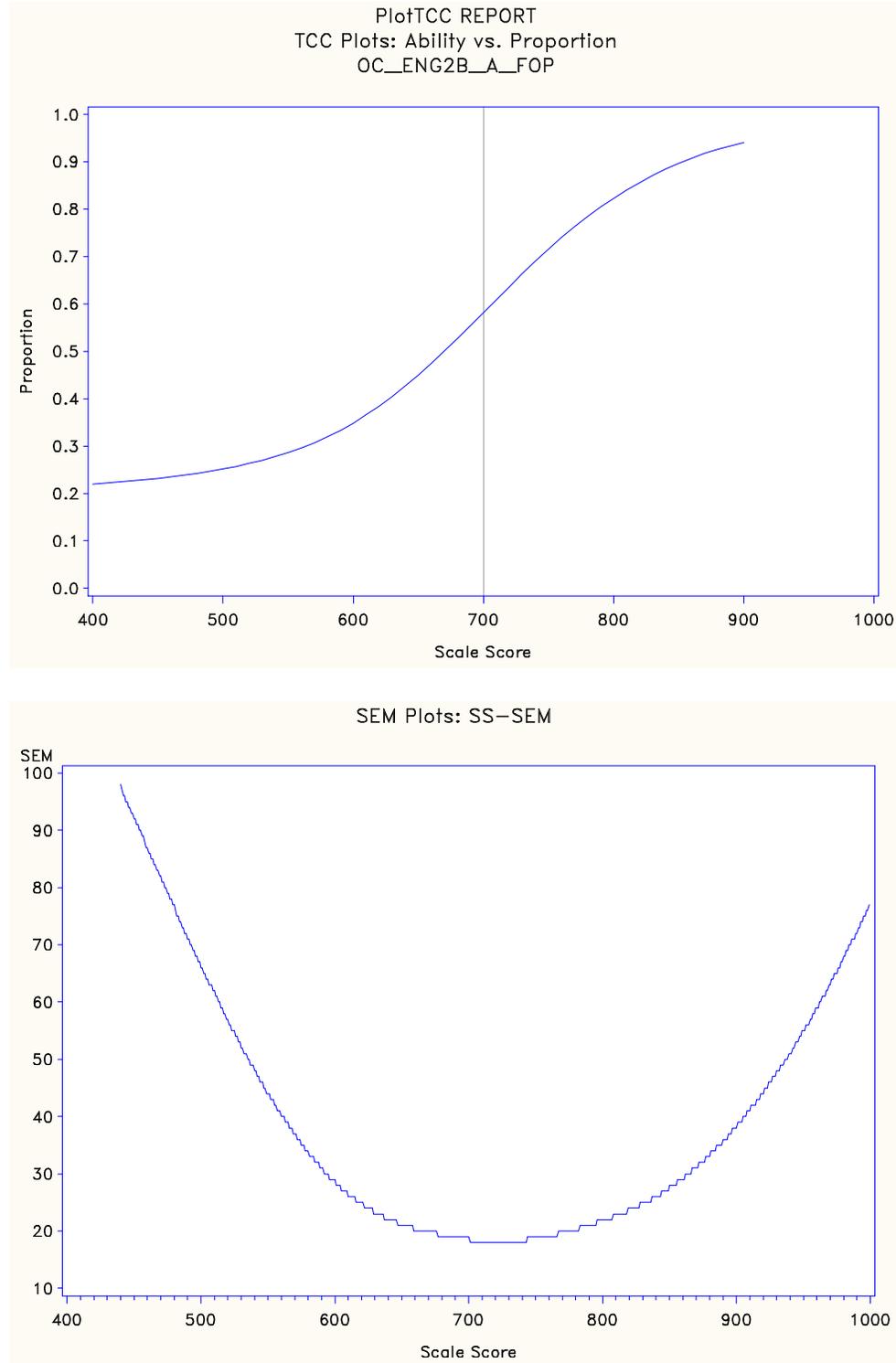
Figure 28. Spring 2013 English II Form BB operational test characteristic curve and standard error of measurement curve

PlotTCC REPORT
TCC Plots: Ability vs. Proportion
OC__ENG2B__B__FOP

SEM Plots: SS−SEM
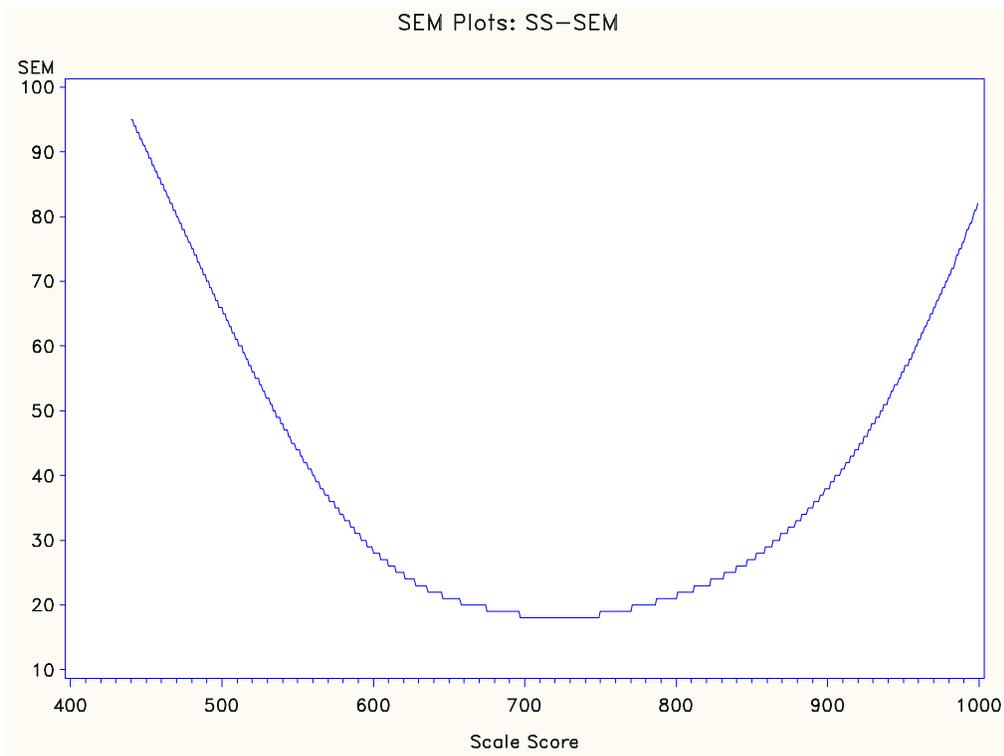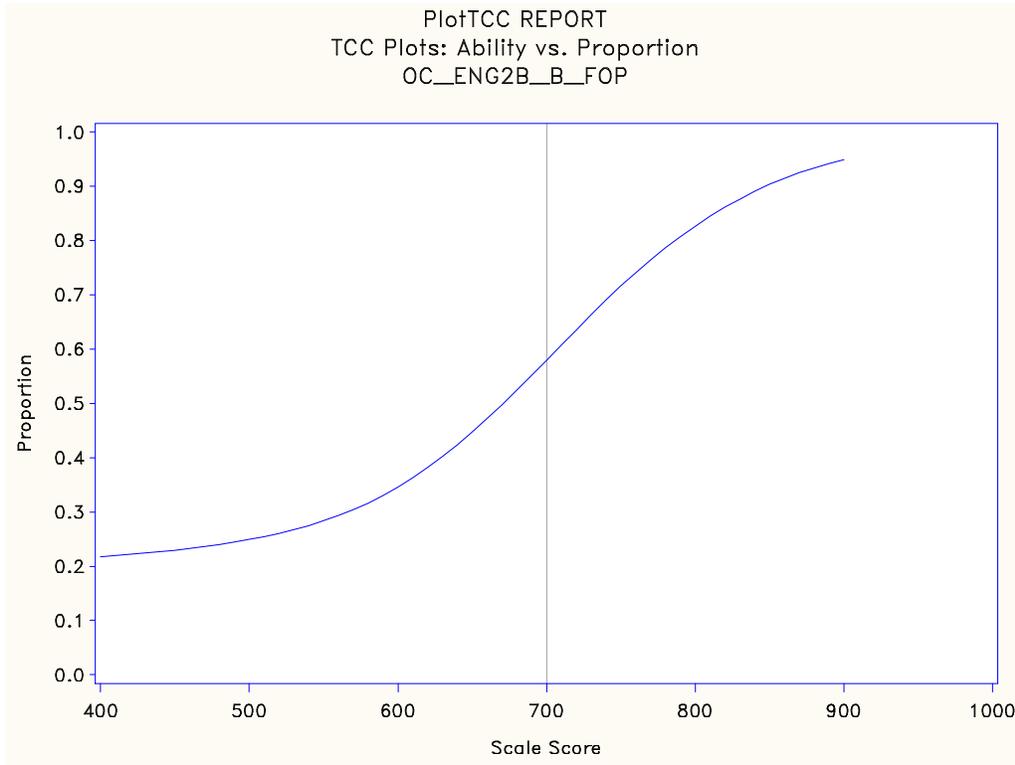
Figure 29. Spring 2013 English III Form AA operational test characteristic curve and standard error of measurement curve
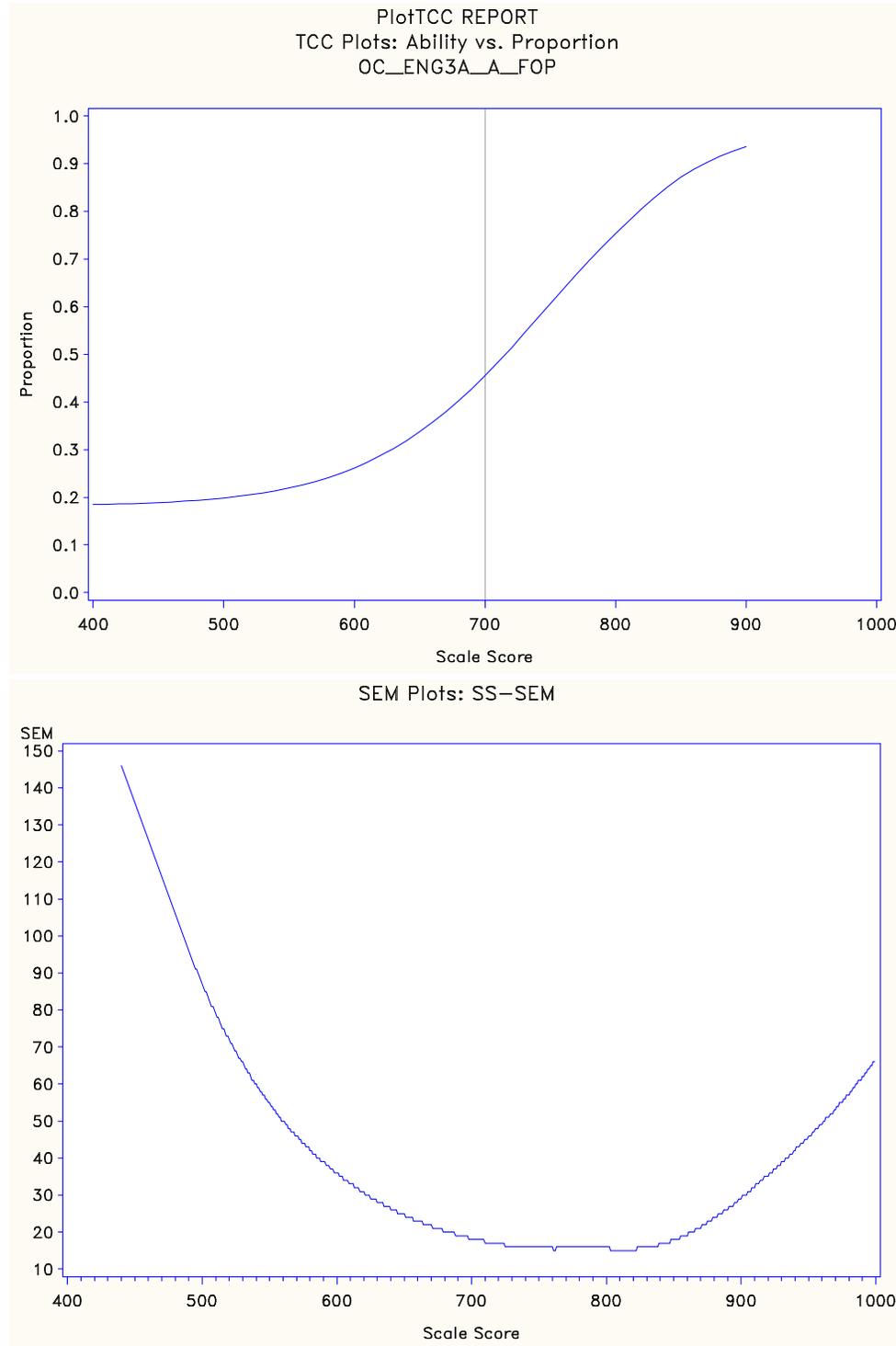
Figure 30. Spring 2013 English III Form AB operational test characteristic curve and standard error of measurement curve
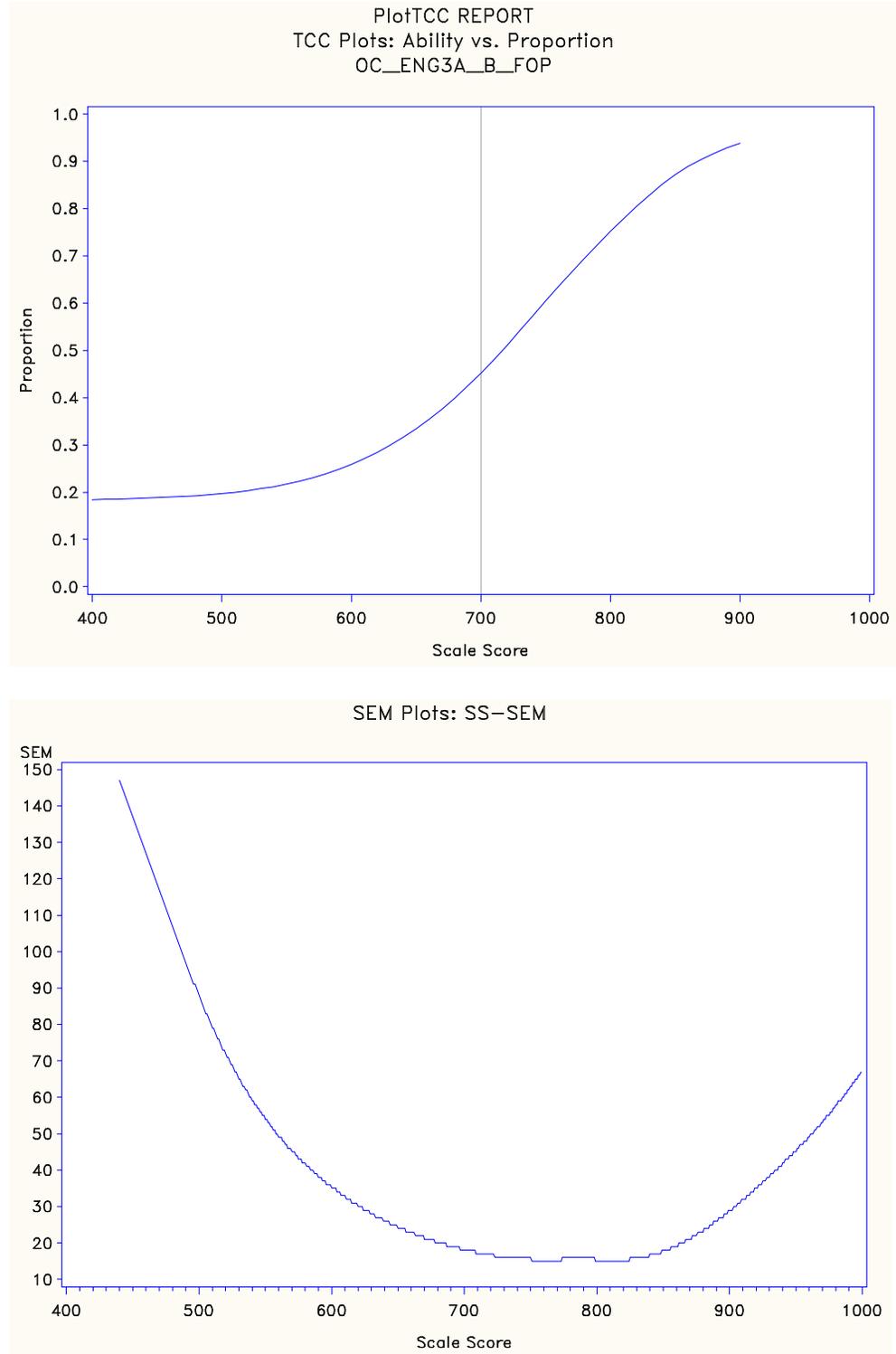
Figure 31. Spring 2013 English III Form BA operational test characteristic curve and standard error of measurement curve
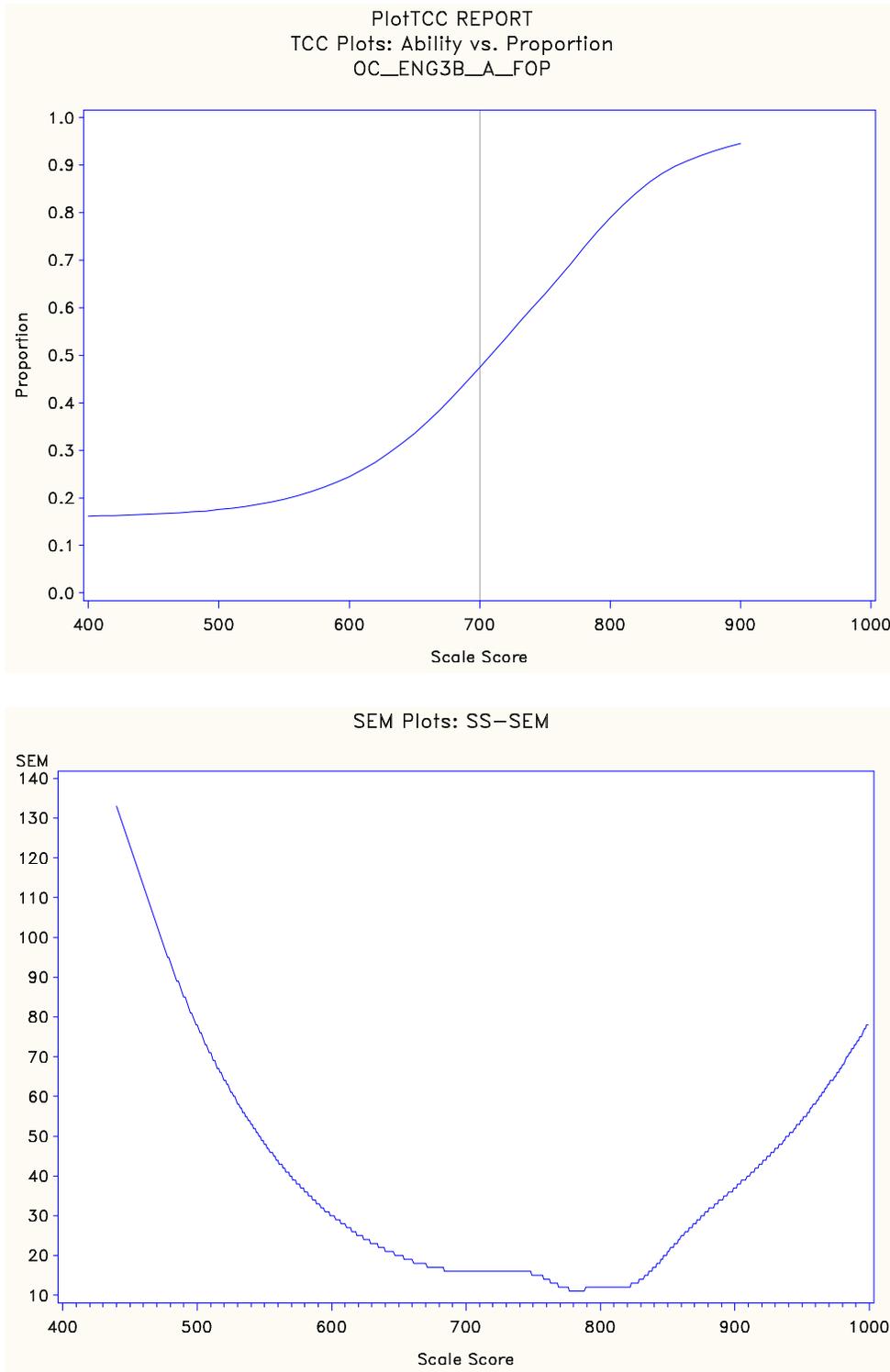
Figure 32. Spring 2013 English III Form BB operational test characteristic curve and standard error of measurement curve
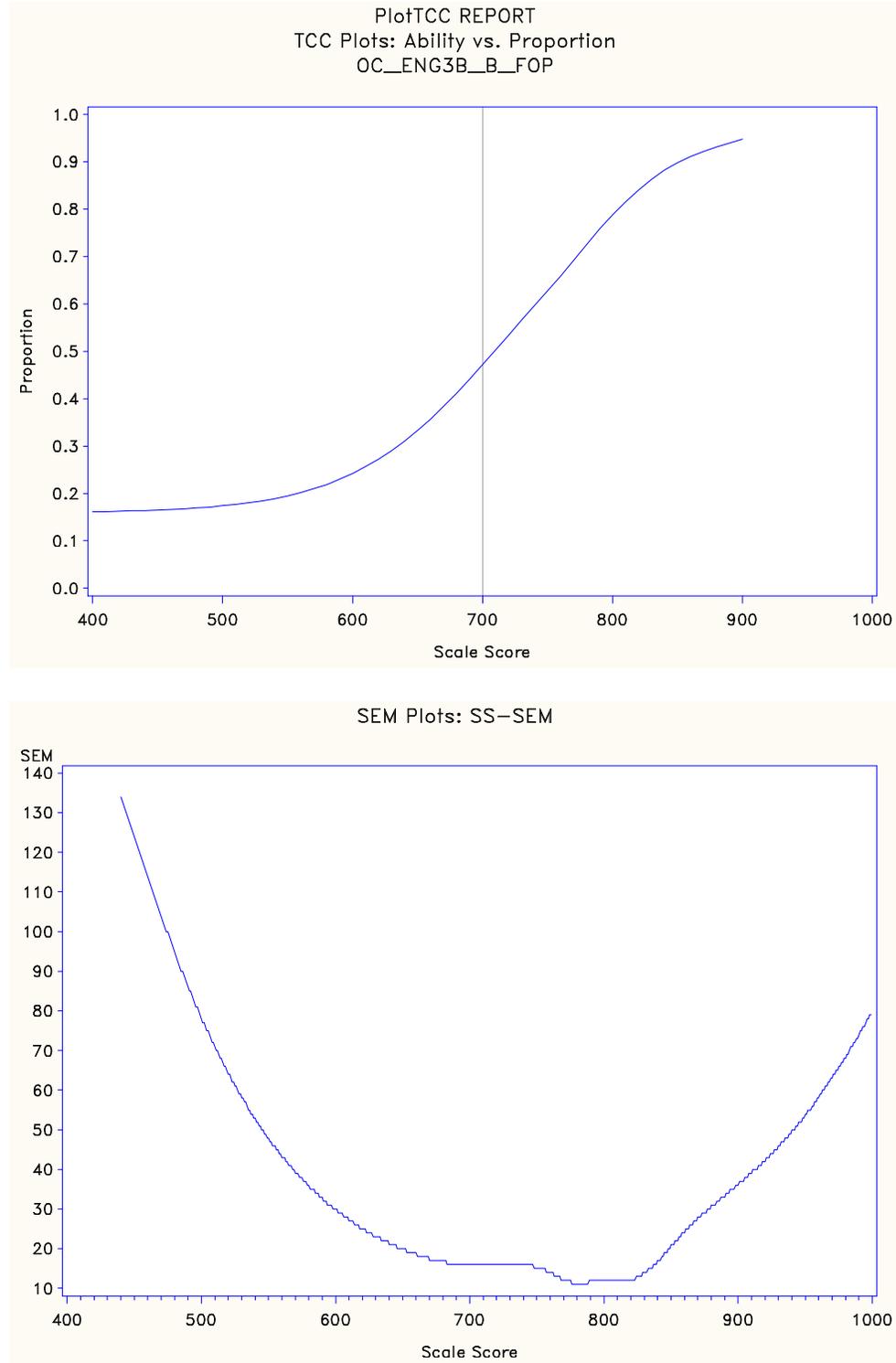
Figure 33. Spring 2013 Geometry Form A operational test characteristic curve and standard error of measurement curve
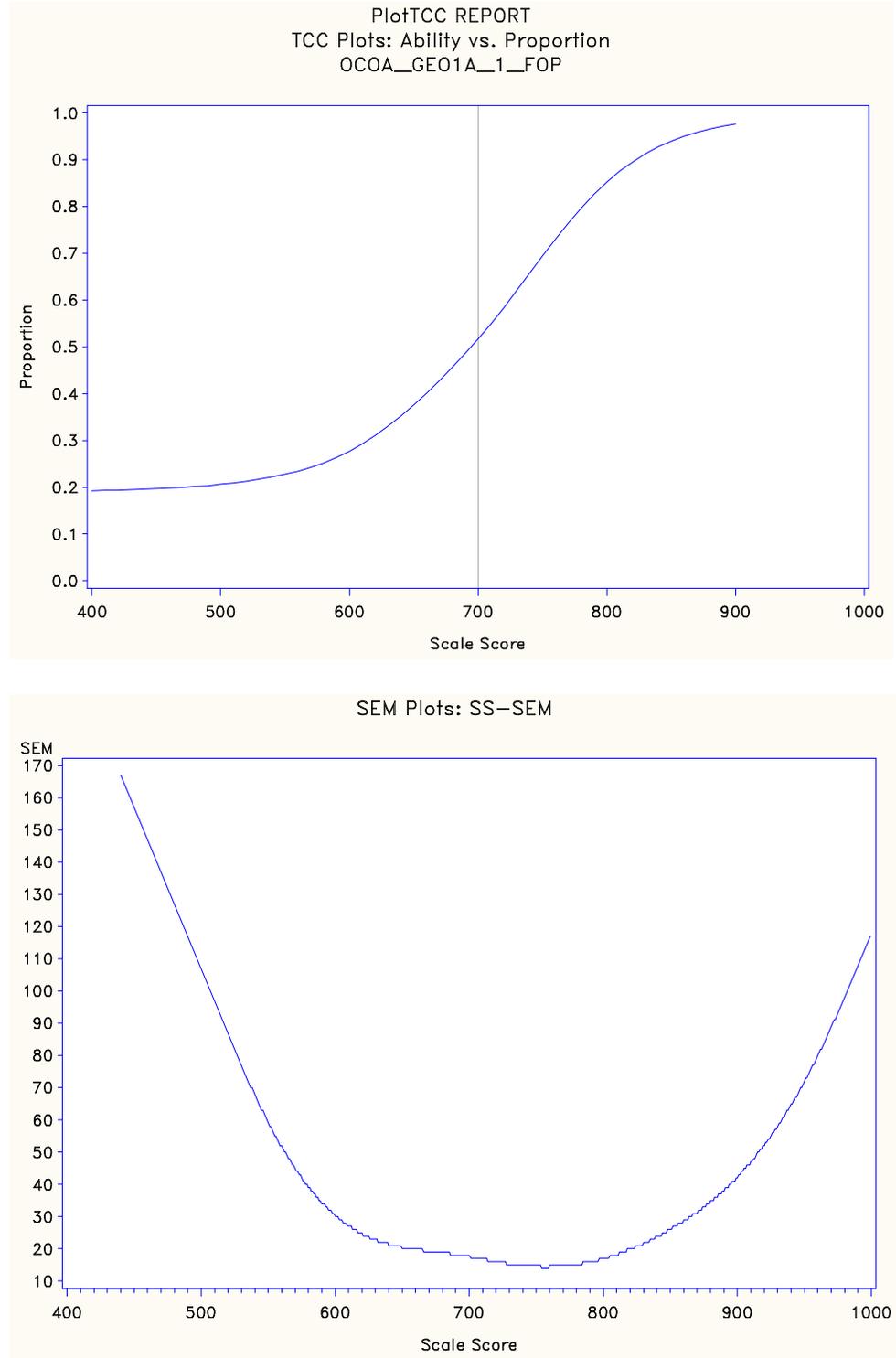
Figure 34. Spring 2013 Geometry Form B operational test characteristic curve and standard error of measurement curve
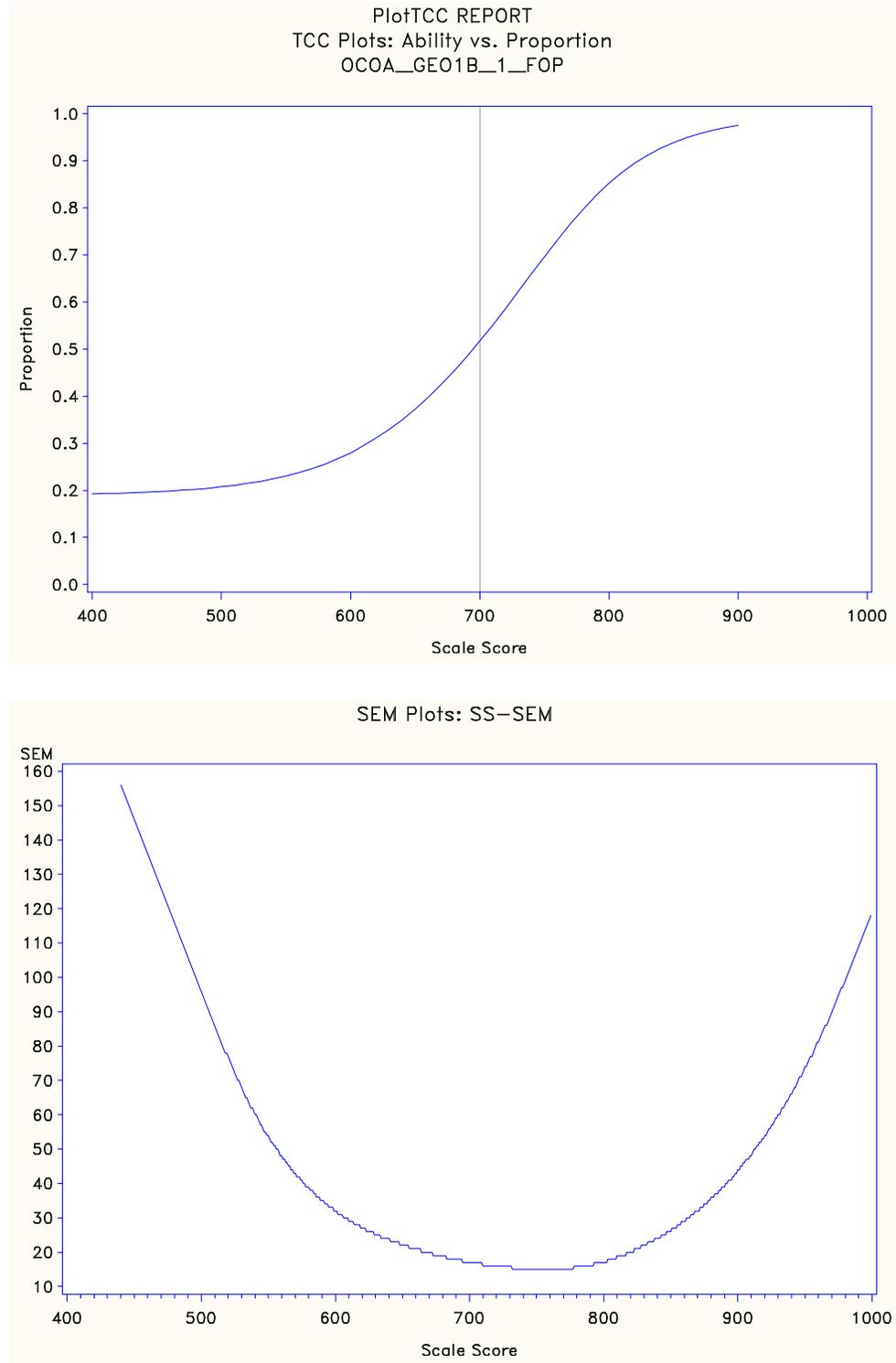
Figure 35. Spring 2013 U.S. History Form A operational test characteristic curve and standard error of measurement curve
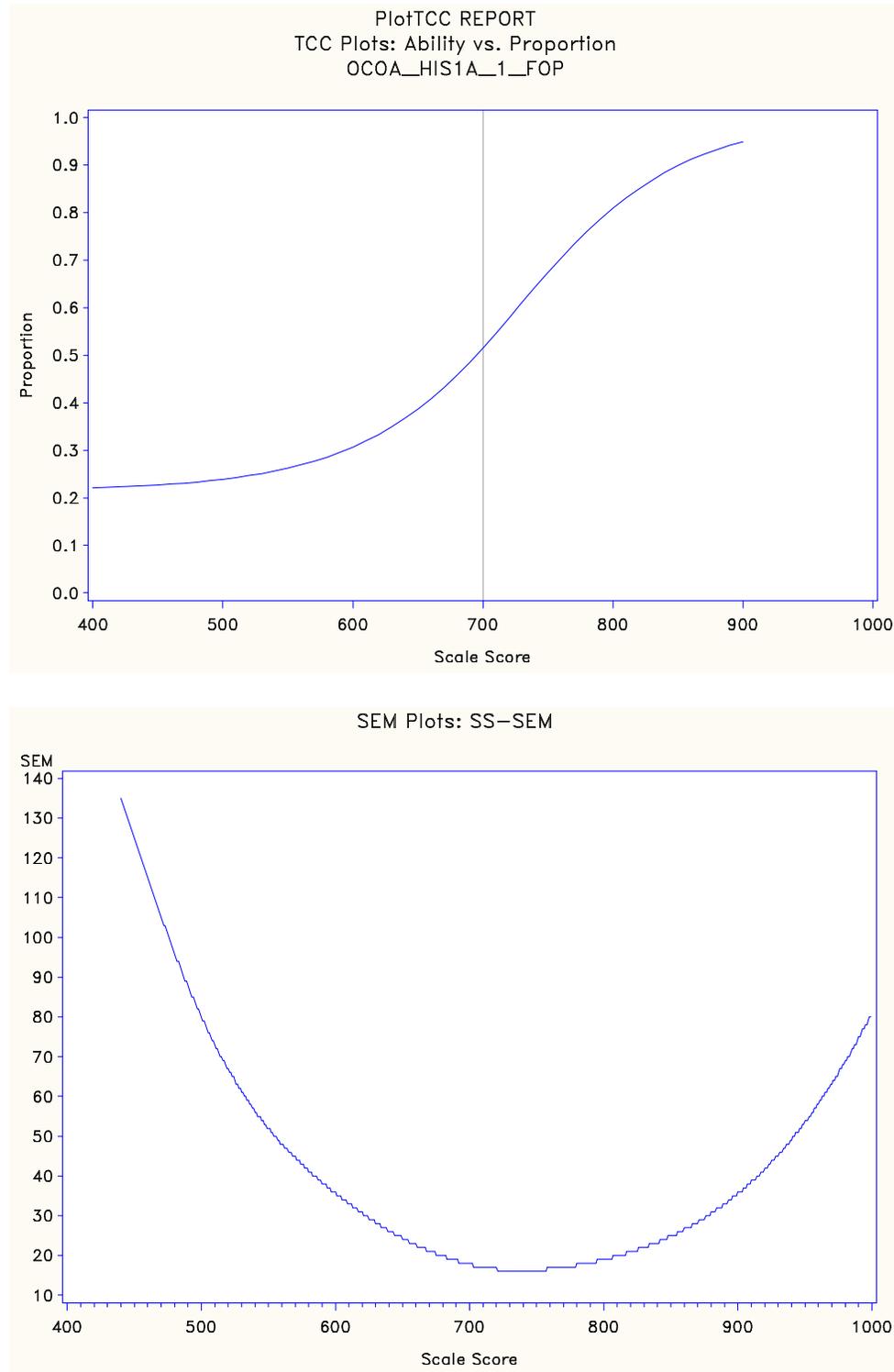
Figure 36. Spring 2013 U.S. History Form B operational test characteristic curve and standard error of measurement curve
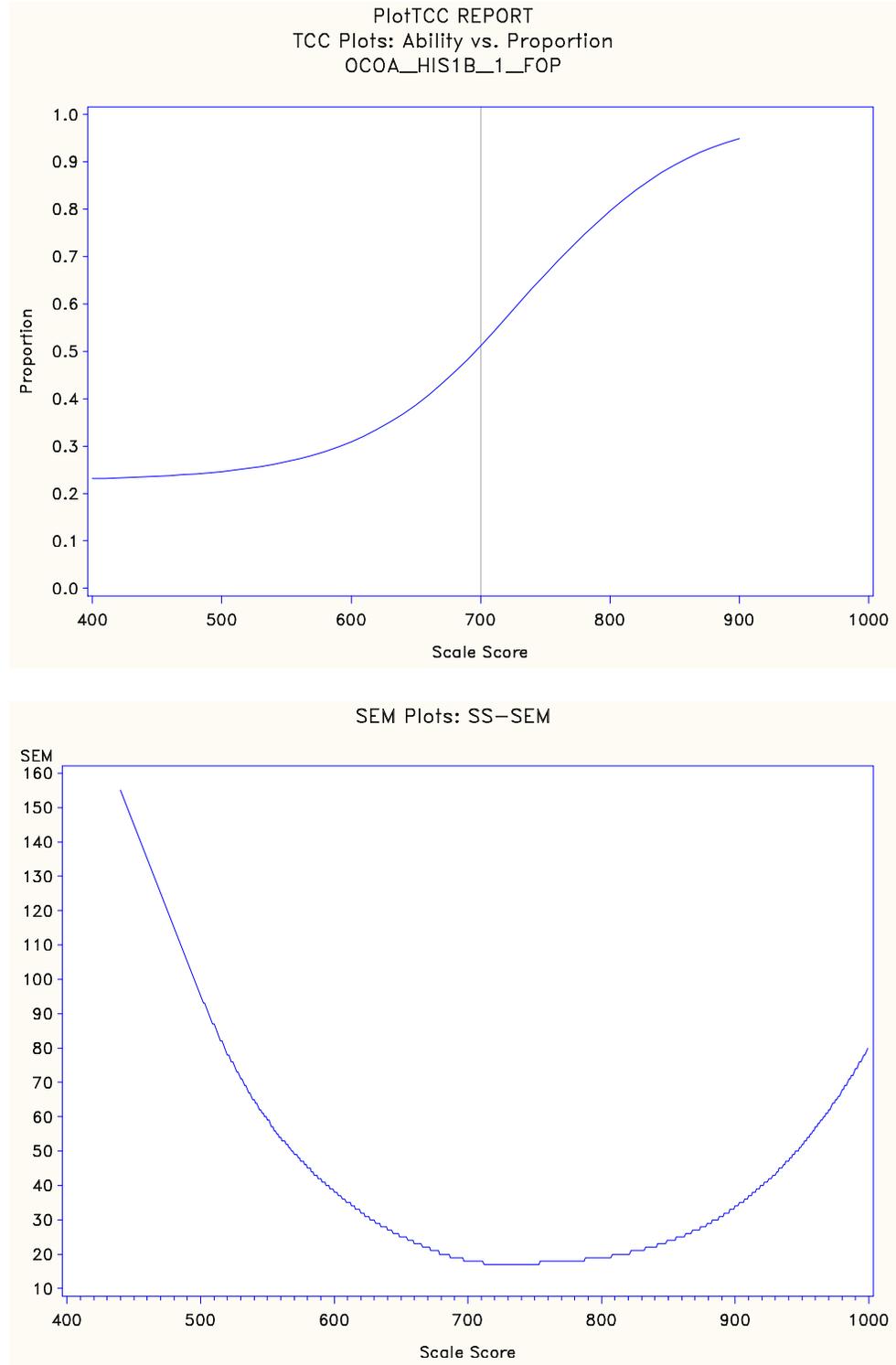
Figure 37. Spring 2013 Algebra I Form A scree plot: All
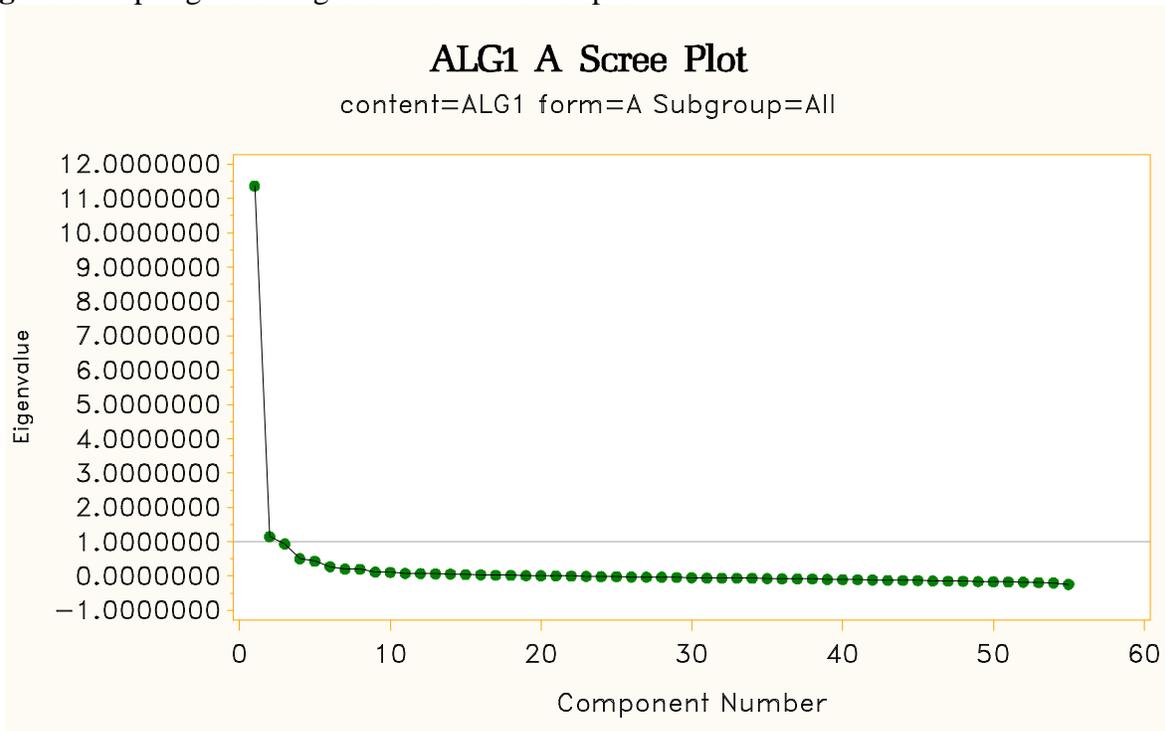
Figure 38. Spring 2013 Algebra I Form A scree plot: Accommodated
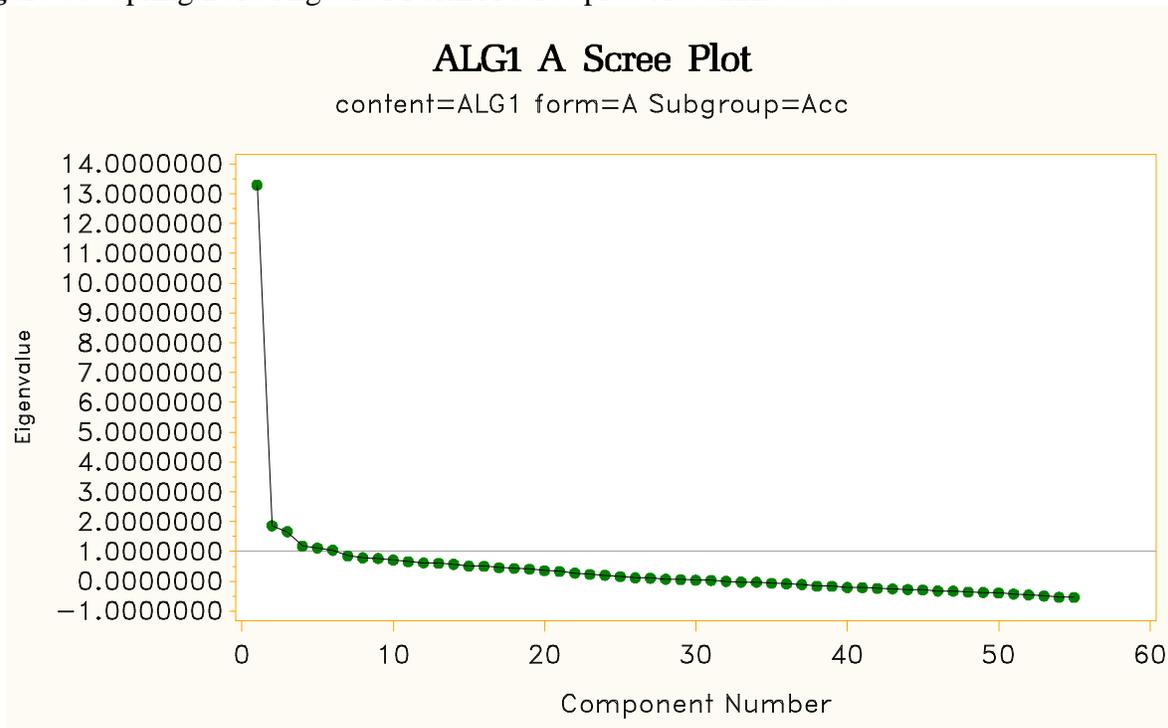
Figure 39. Spring 2013 Algebra I Form A scree plot: English Language Learner
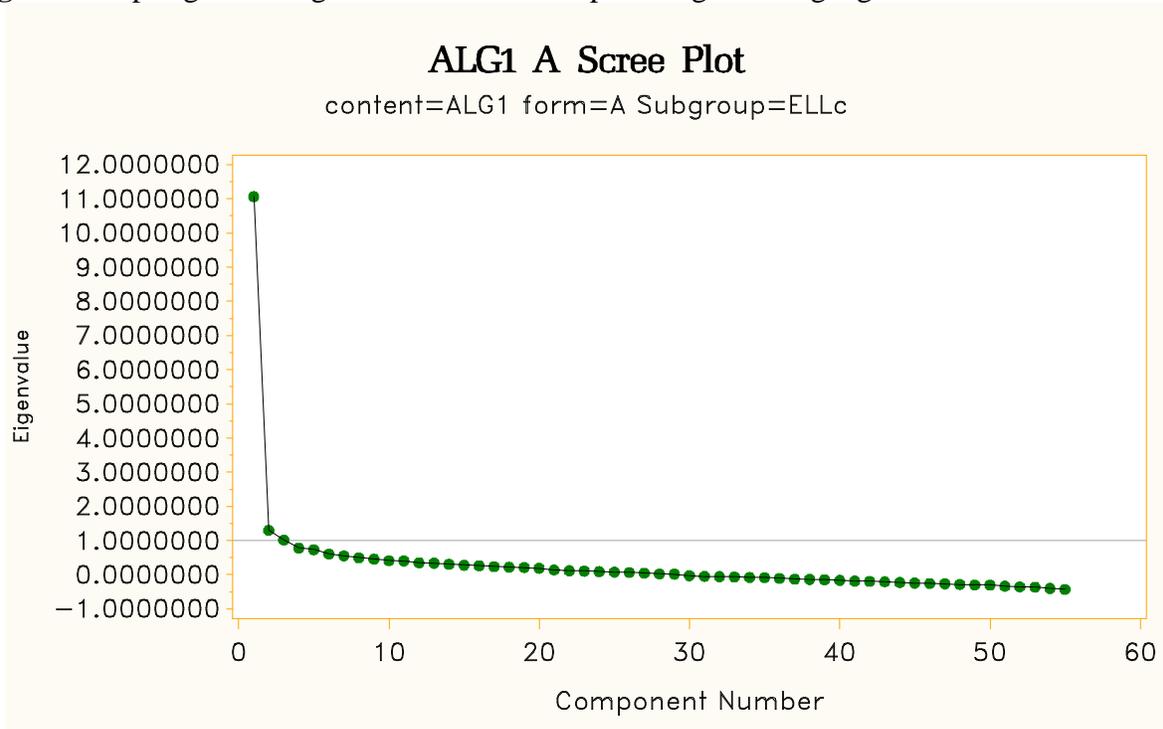
**Figure 40.** Spring 2013 Algebra I Form A scree plot: Individualized Education Program
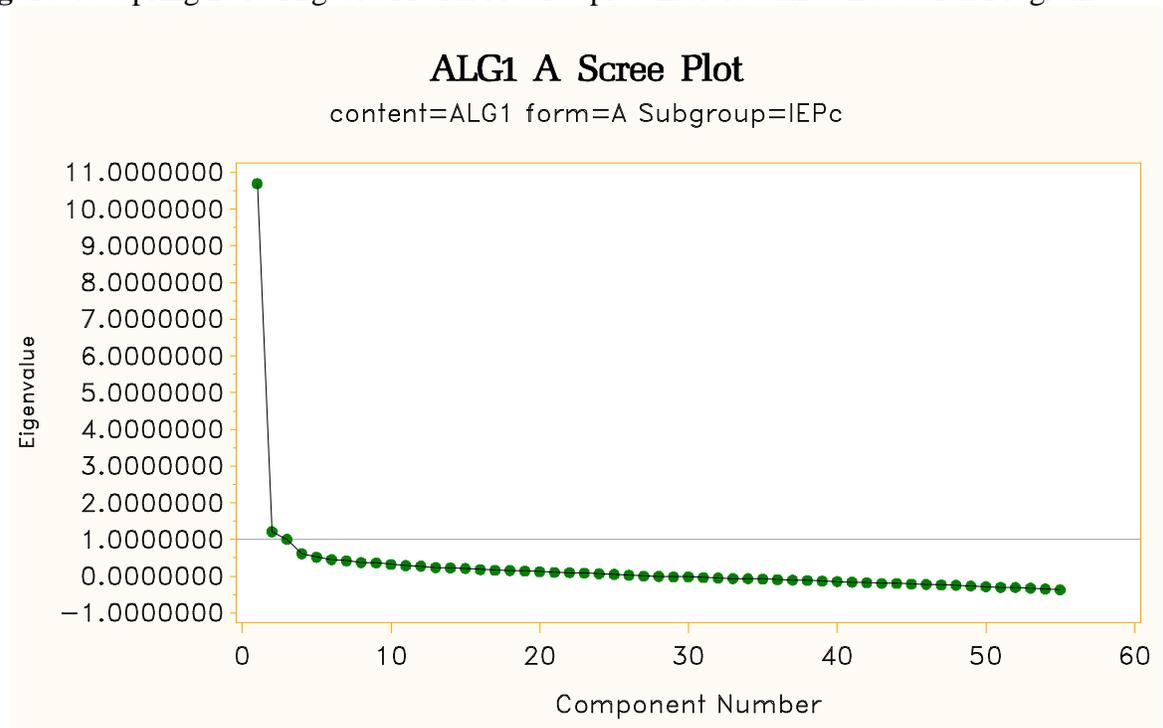
Figure 41. Spring 2013 Algebra I Form B scree plot: All
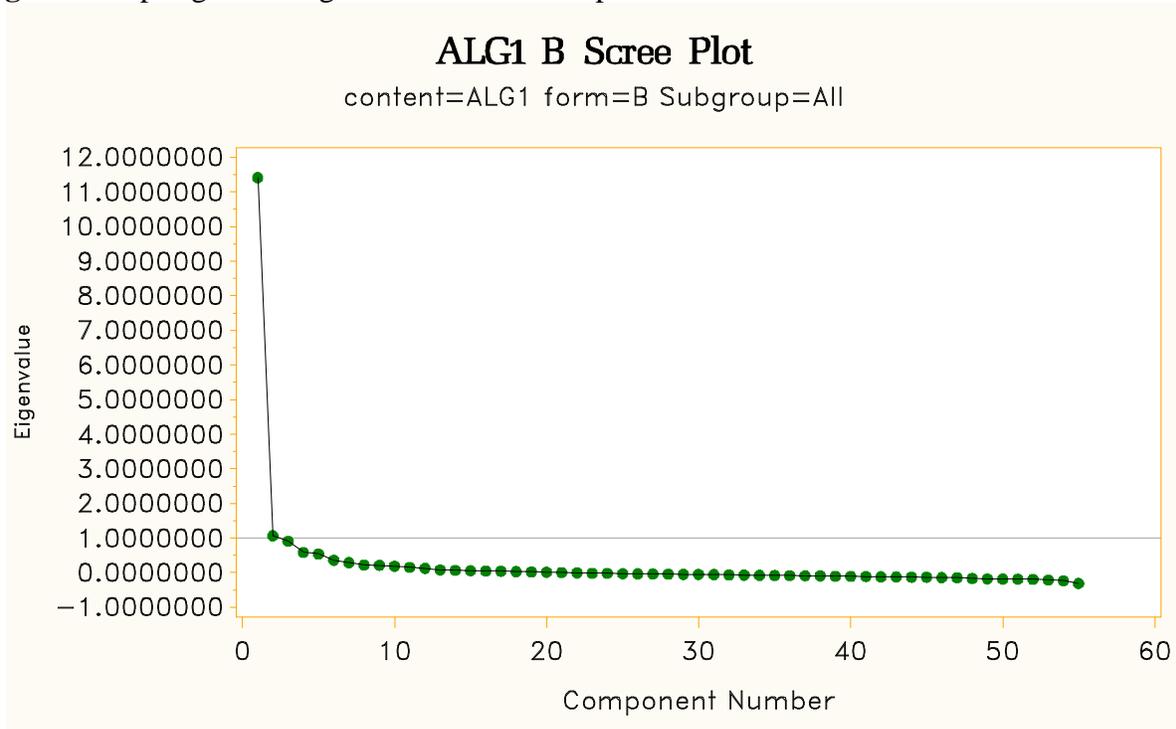
Figure 42. Spring 2013 Algebra I Form B scree plot: Accommodated
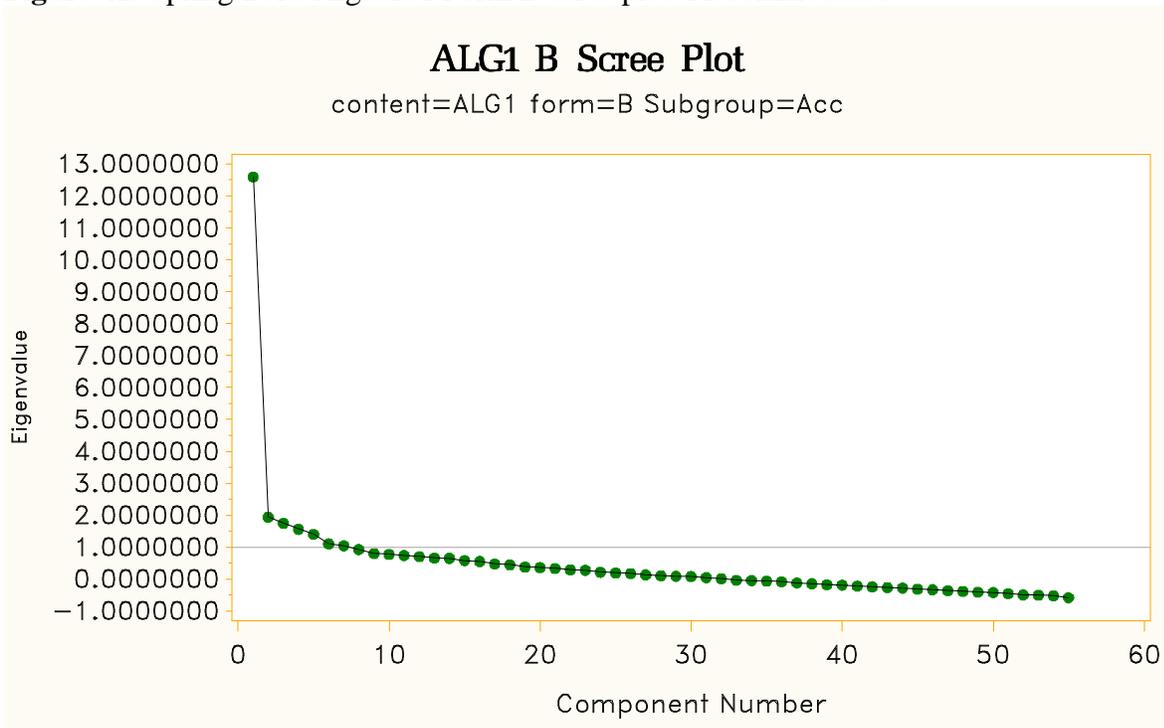
Figure 43. Spring 2013 Algebra I Form B scree plot: English Language Learner
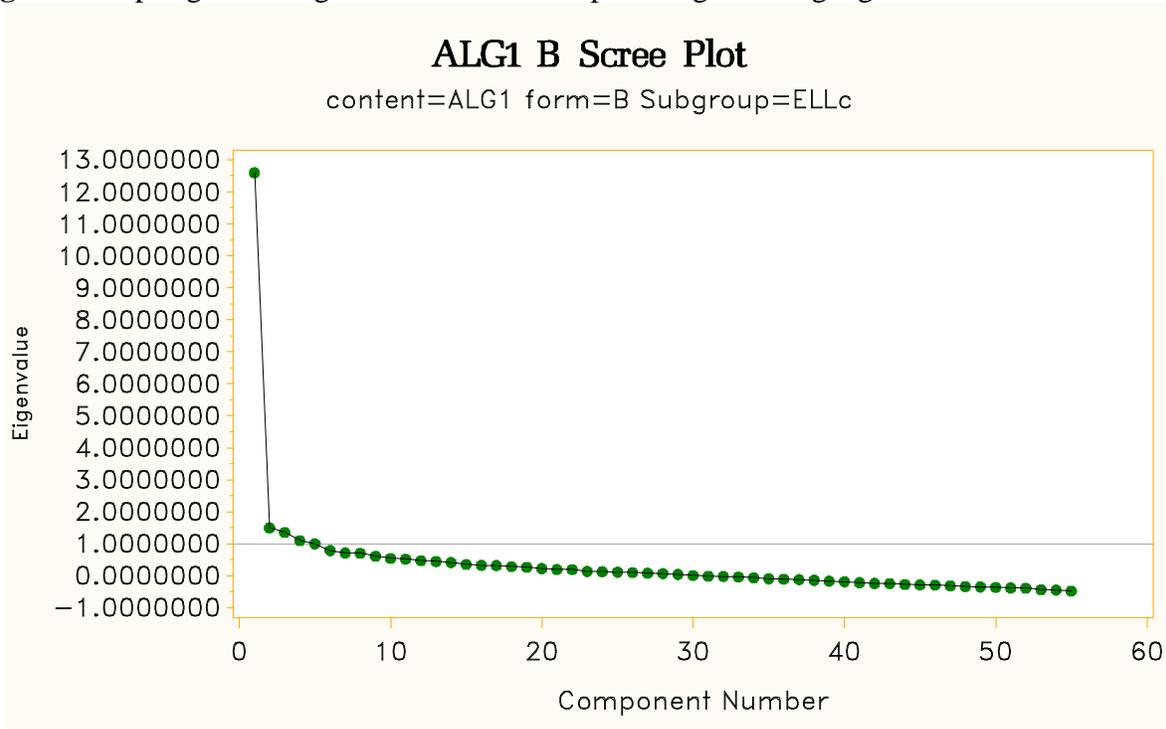
Figure 44. Spring 2013 Algebra I Form B scree plot: Individualized Education Program
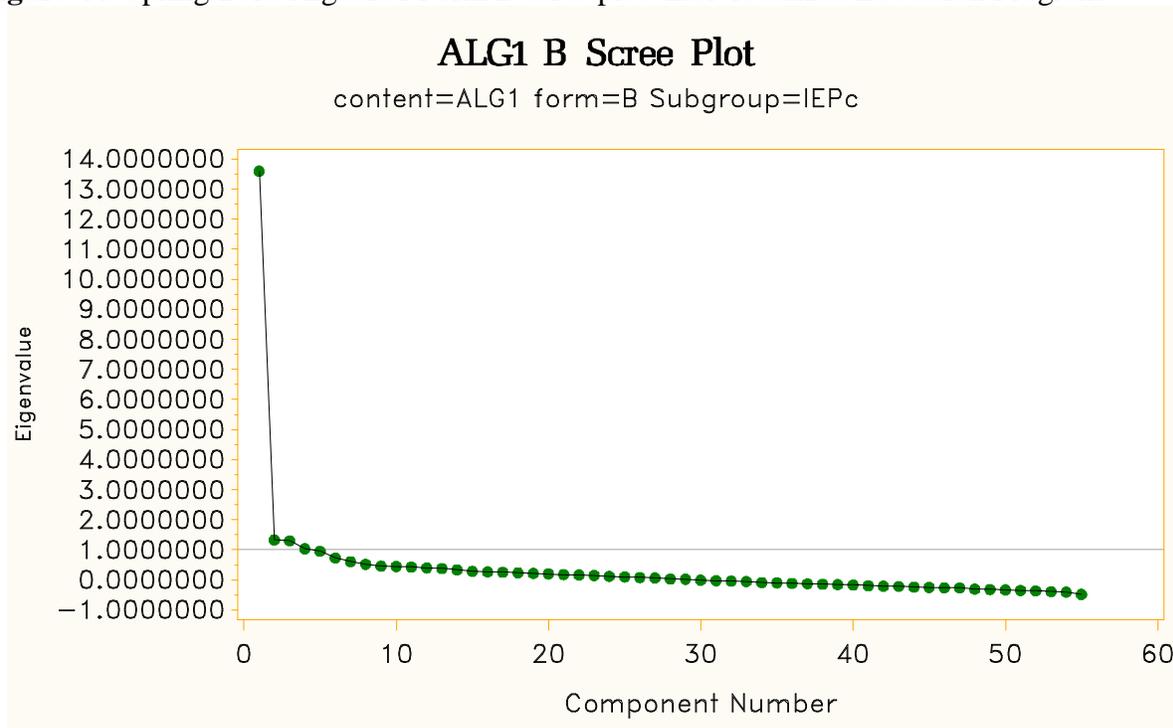
**Figure 45.** Spring 2013 Algebra II Form A scree plot: All
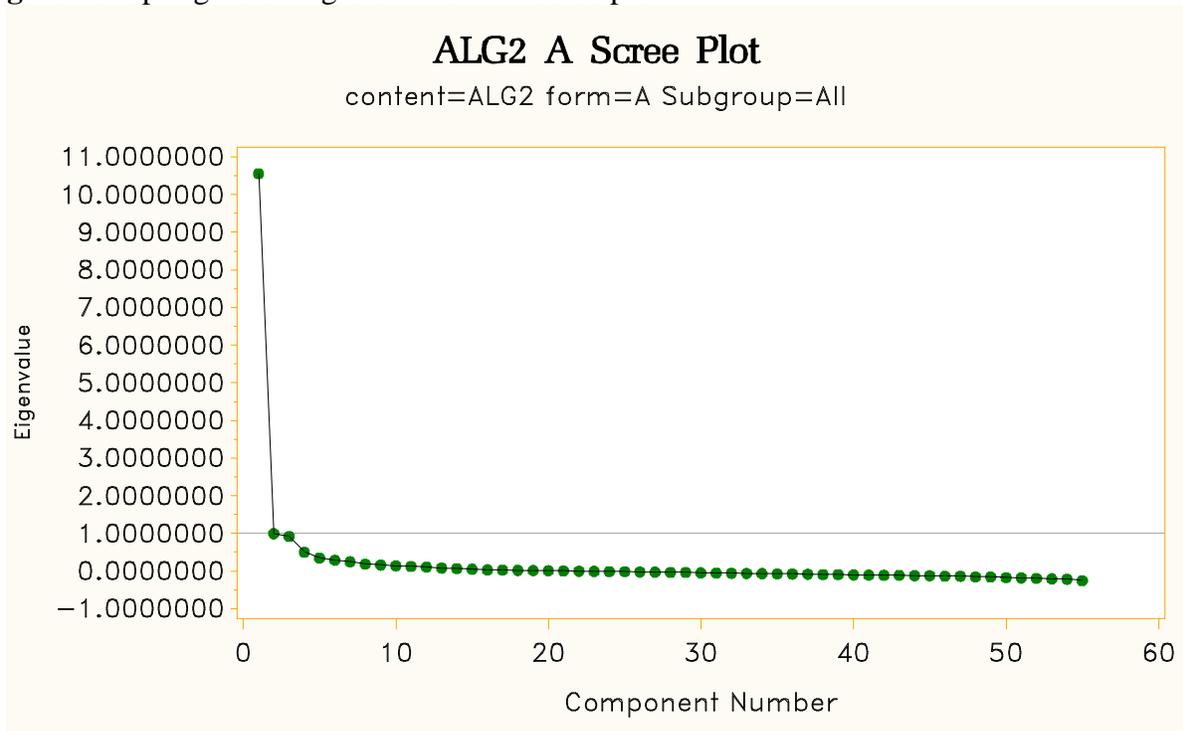
Figure 46. Spring 2013 Algebra II Form A scree plot: Accommodated
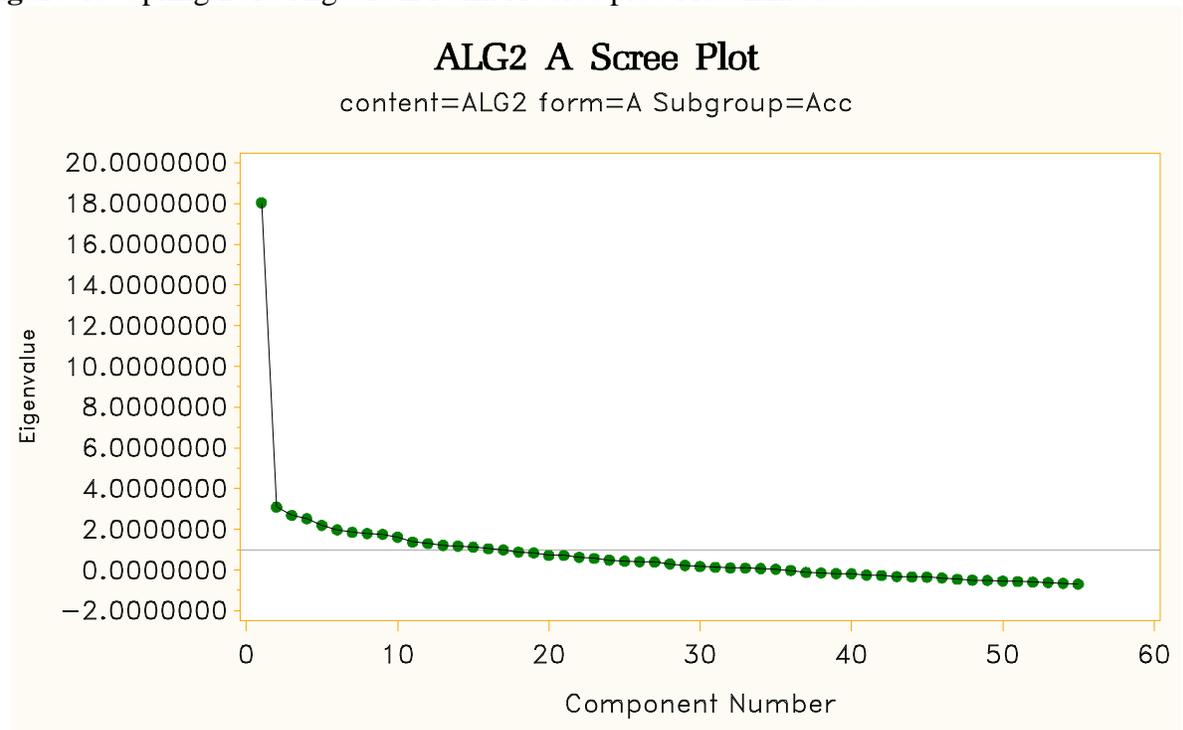
Figure 47. Spring 2013 Algebra II Form A scree plot: English Language Learner
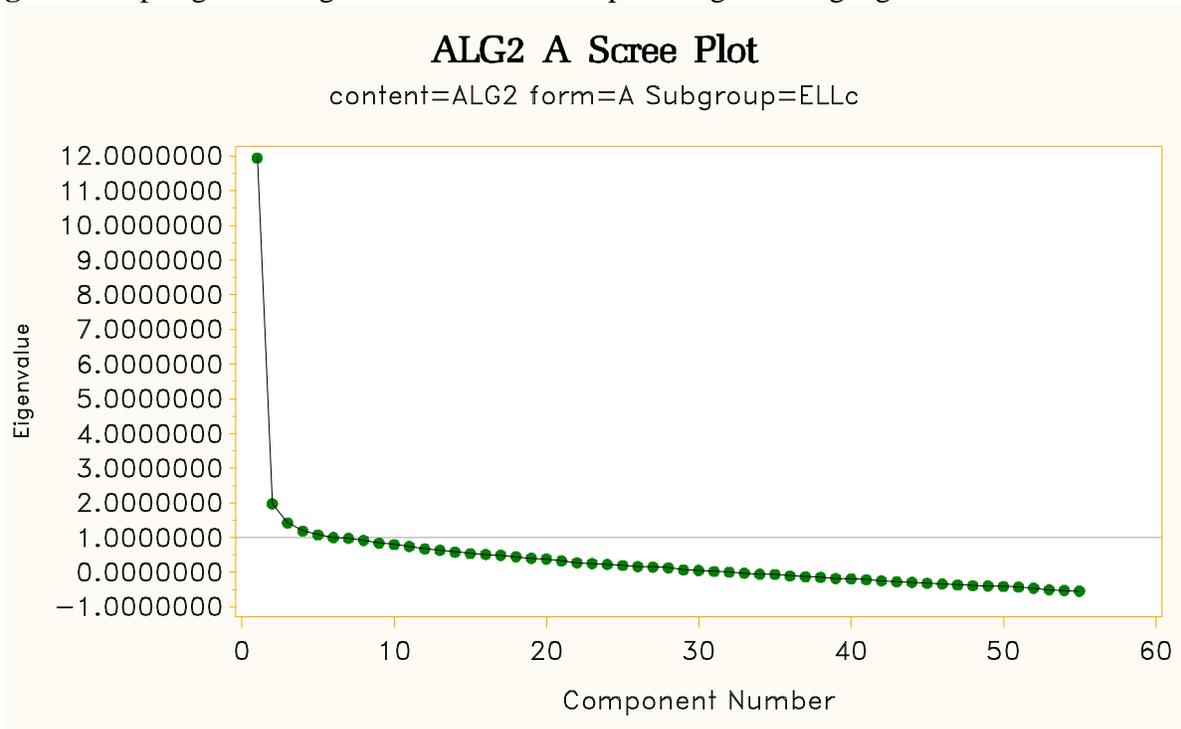
Figure 48. Spring 2013 Algebra II Form A scree plot: Individualized Education Program
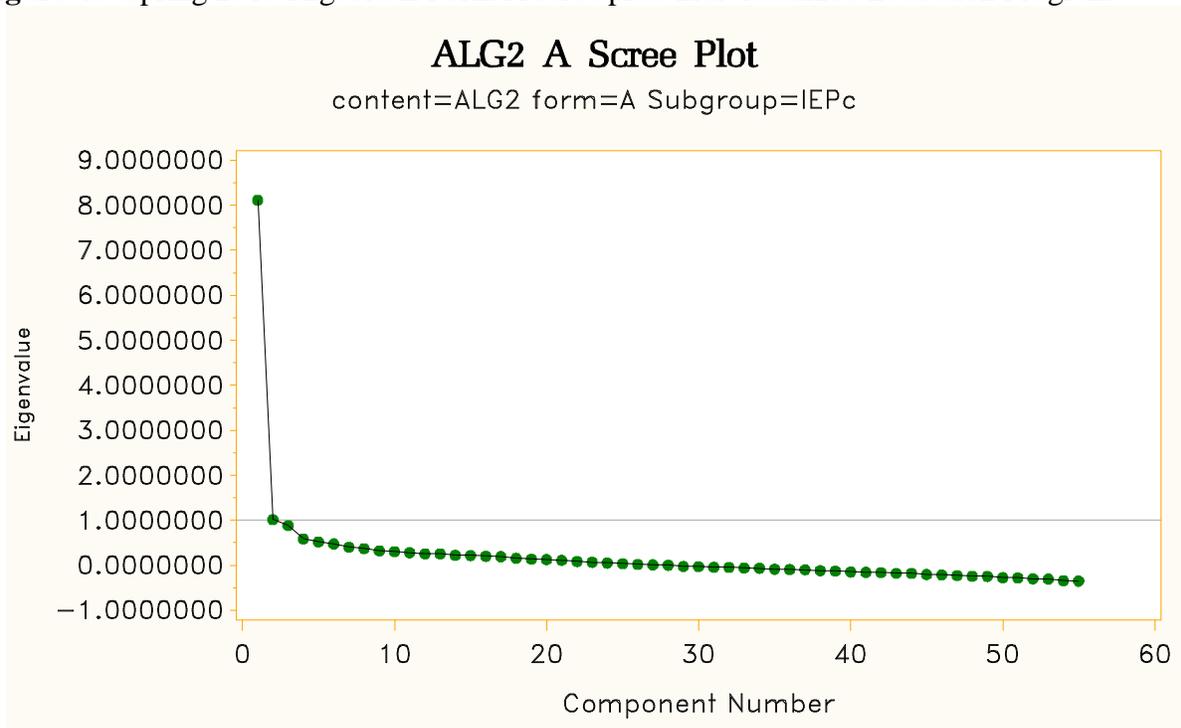
**Figure 49.** Spring 2013 Algebra II Form B scree plot: All
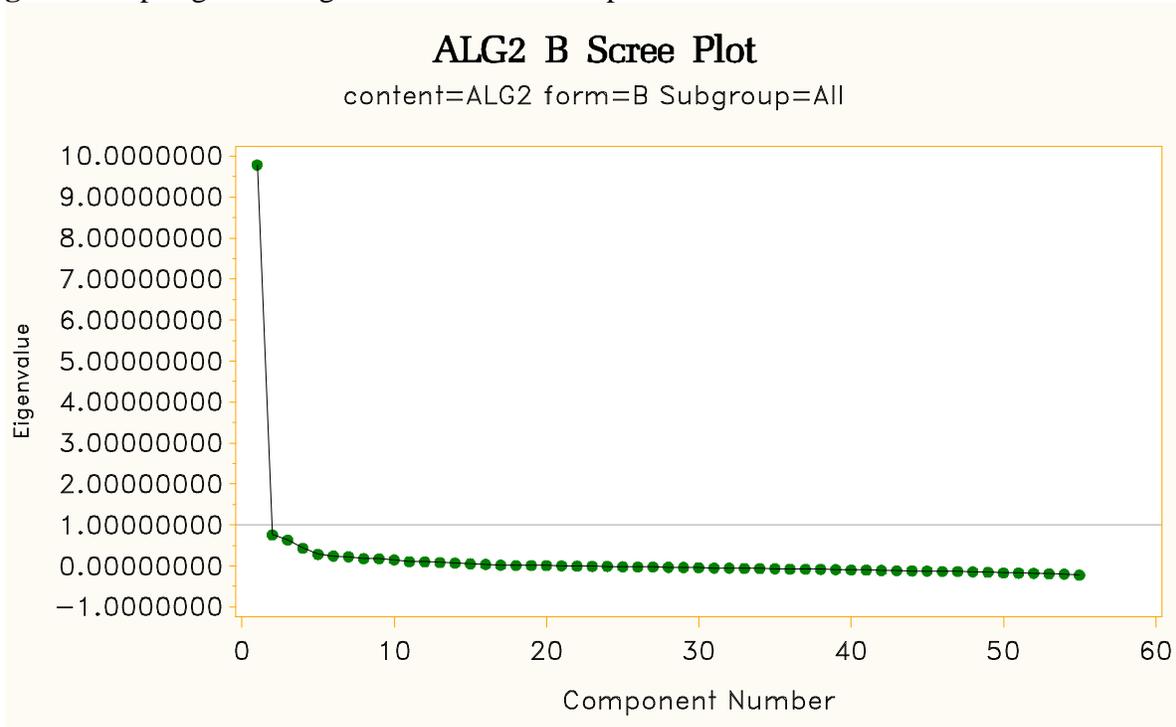
Figure 50. Spring 2013 Algebra II Form B scree plot: Accommodated
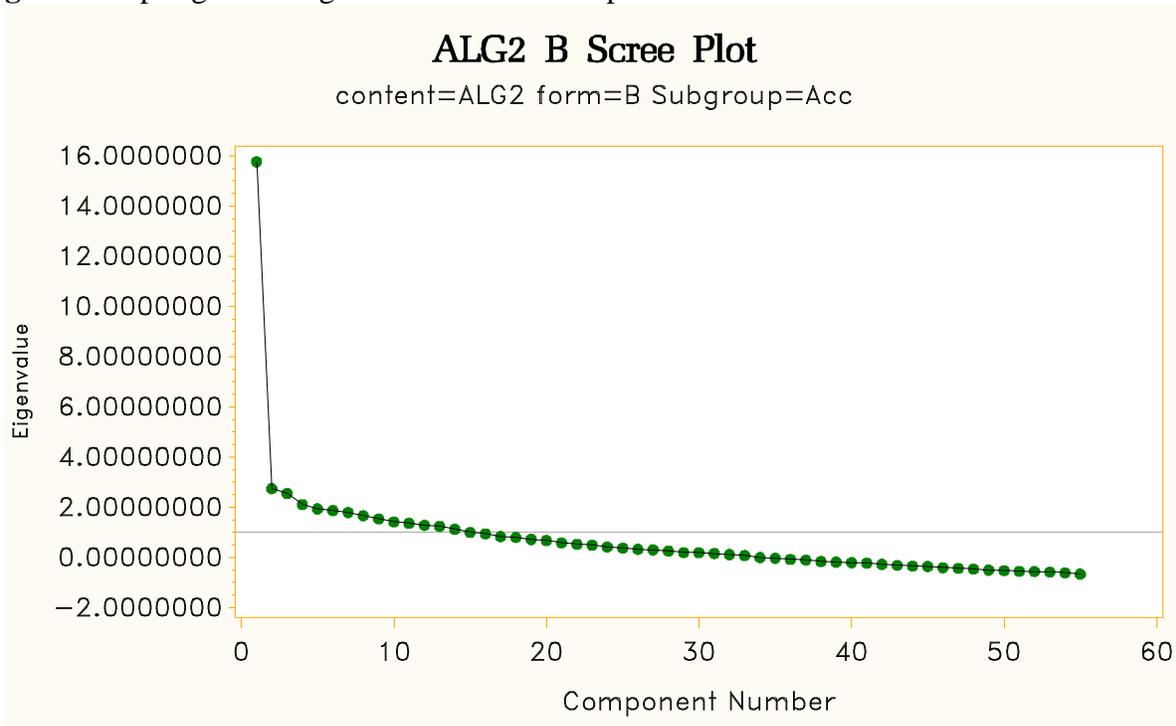
Figure 51. Spring 2013 Algebra II Form B scree plot: English Language Learner
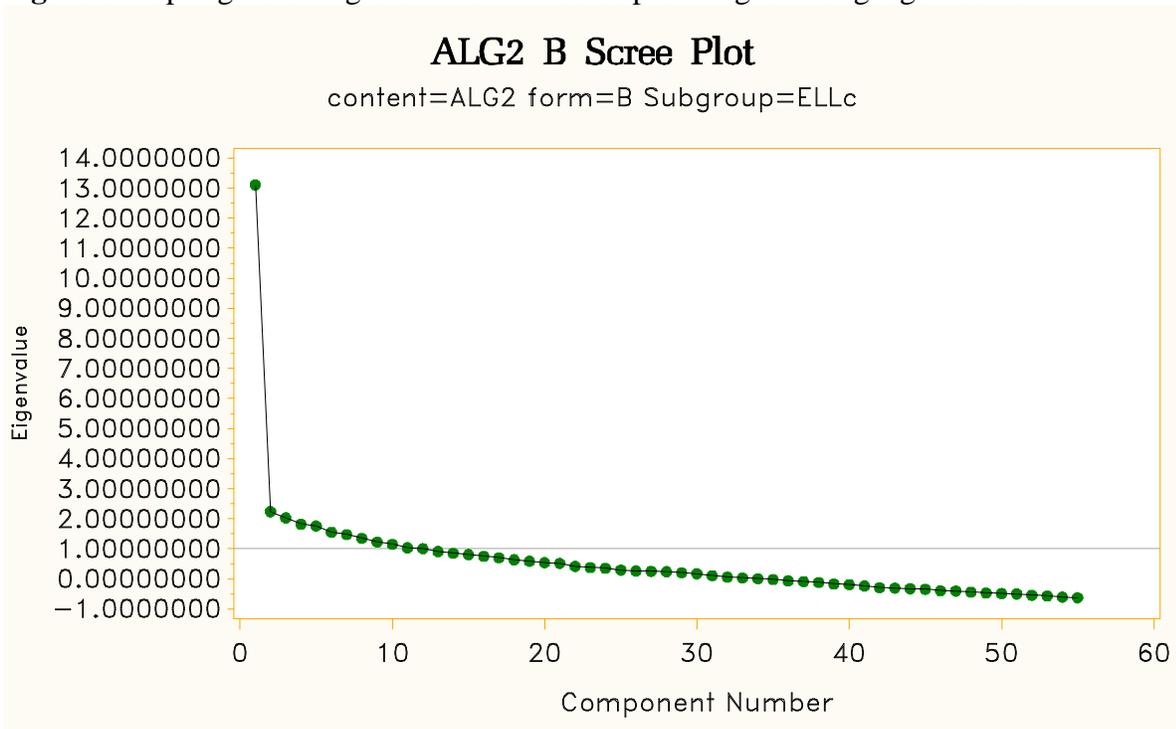
Figure 52. Spring 2013 Algebra II Form B scree plot: Individualized Education Program
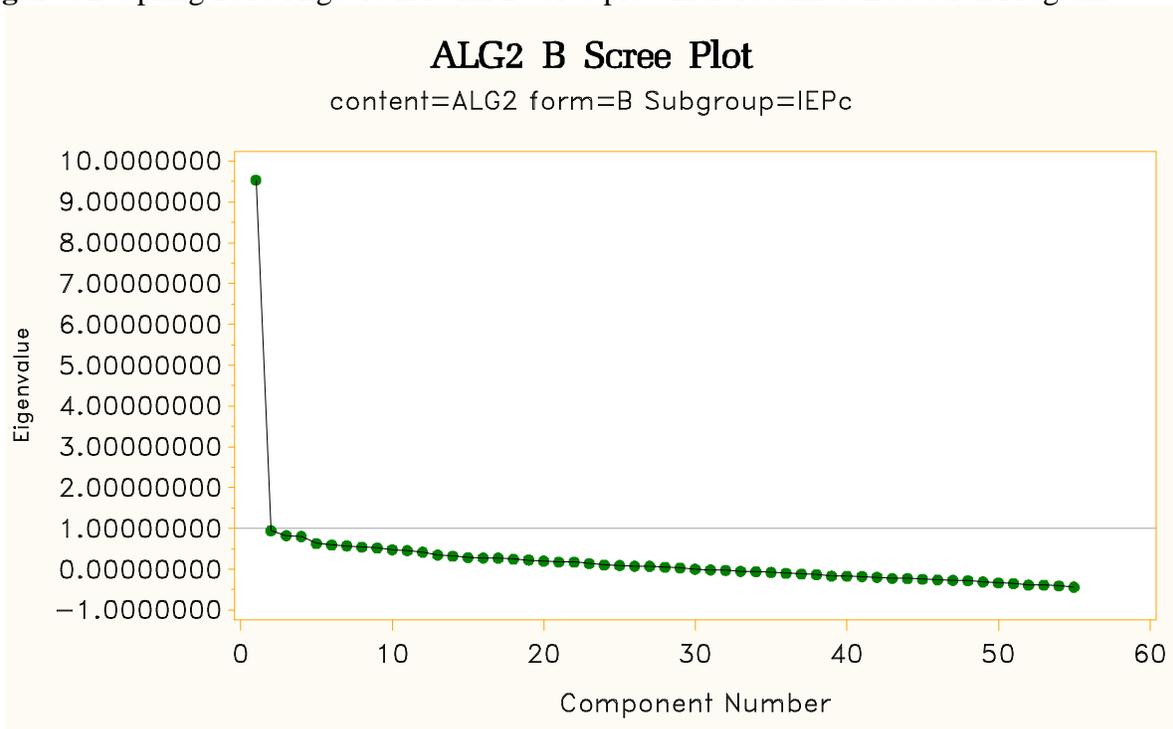
Figure 53. Spring 2013 Biology I Form A scree plot: All
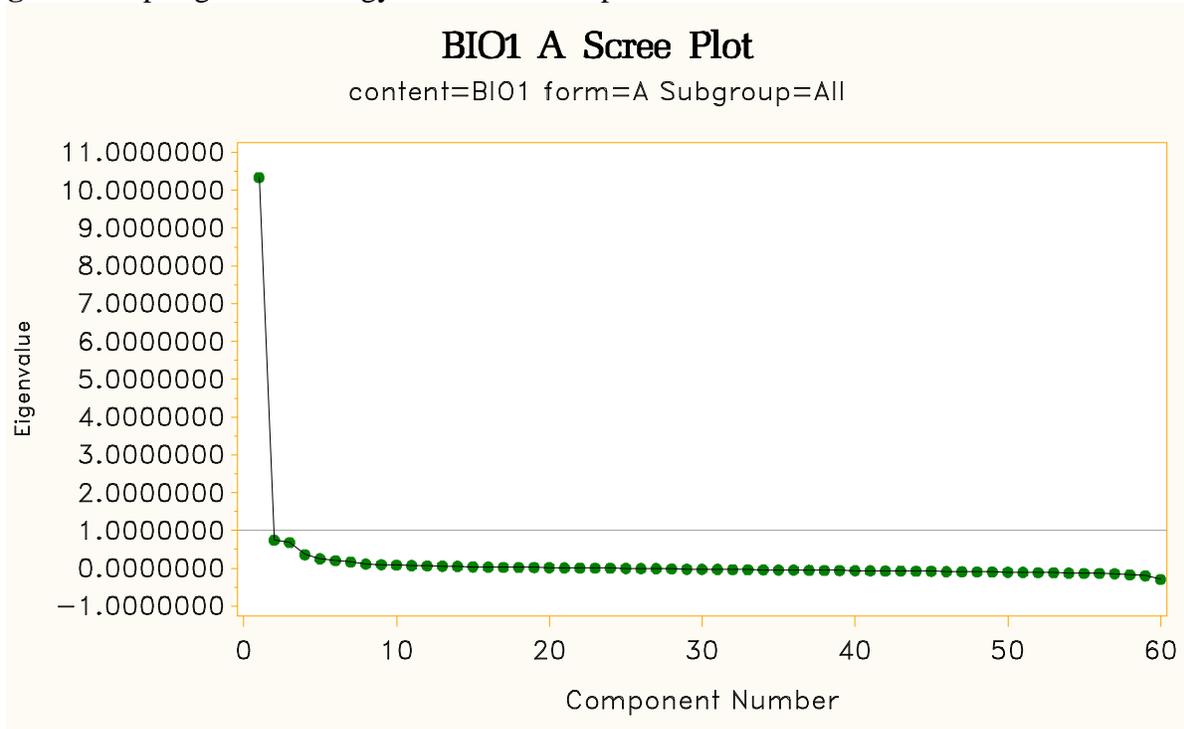
**Figure 54.** Spring 2013 Biology I Form A scree plot: Accommodated
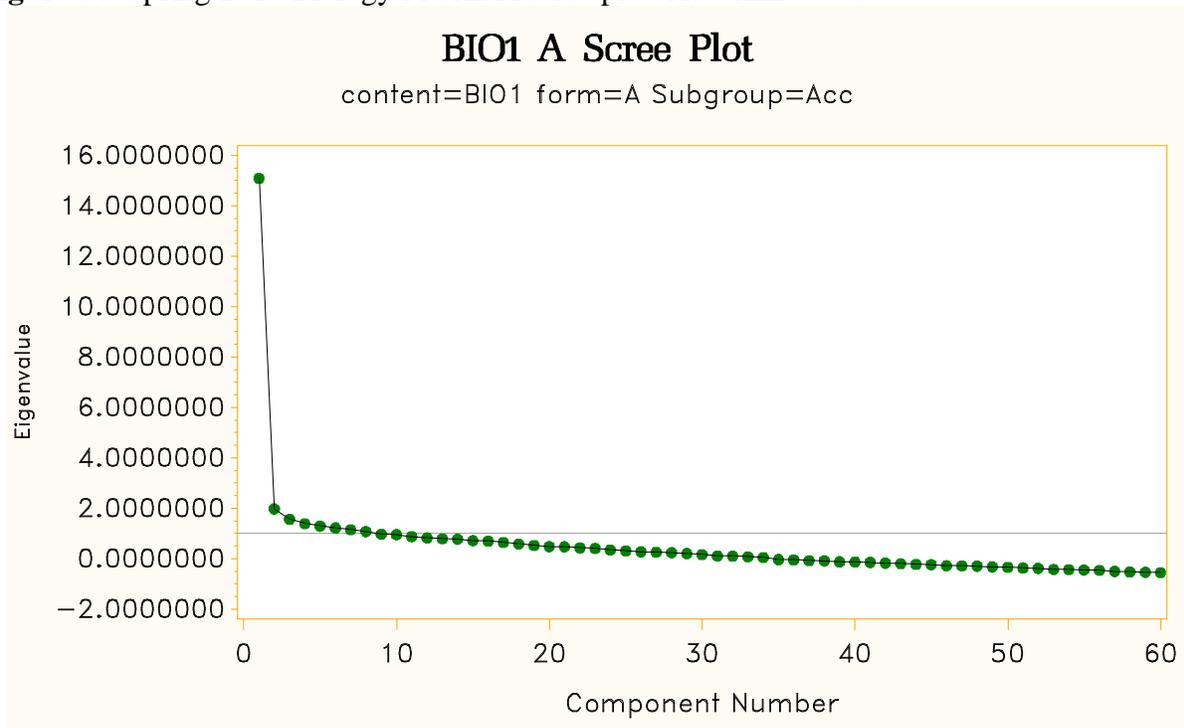
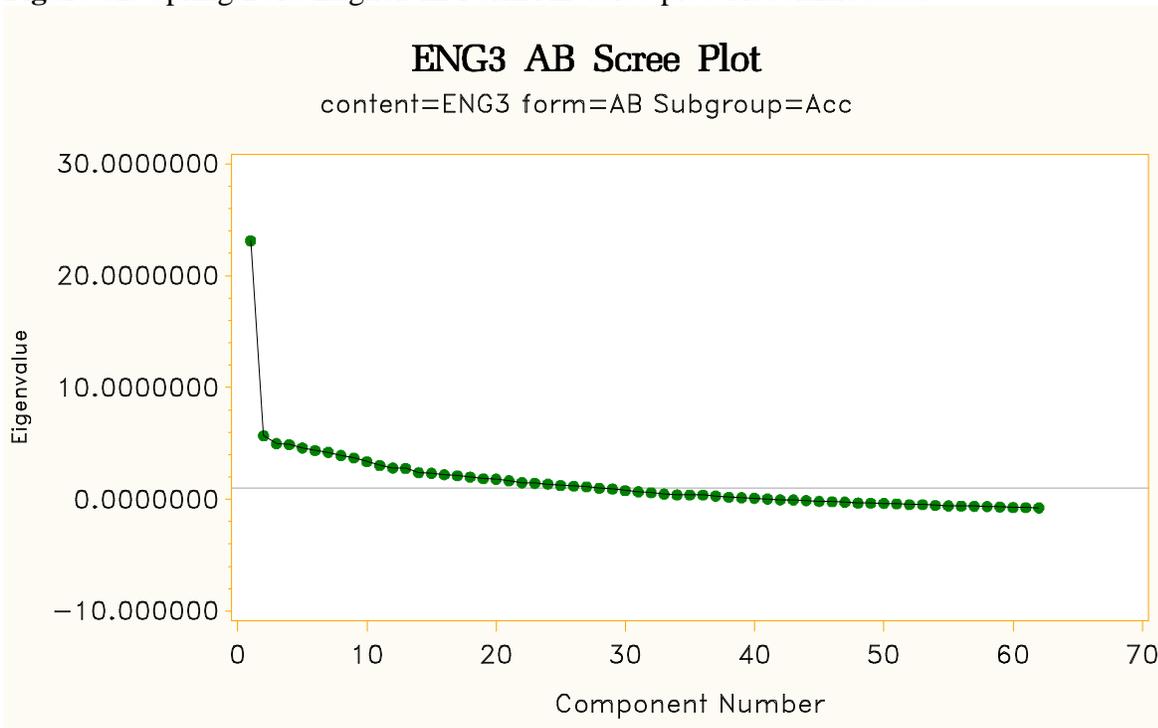**Figure 55.** Spring 2013 Biology I Form A scree plot: English Language Learner
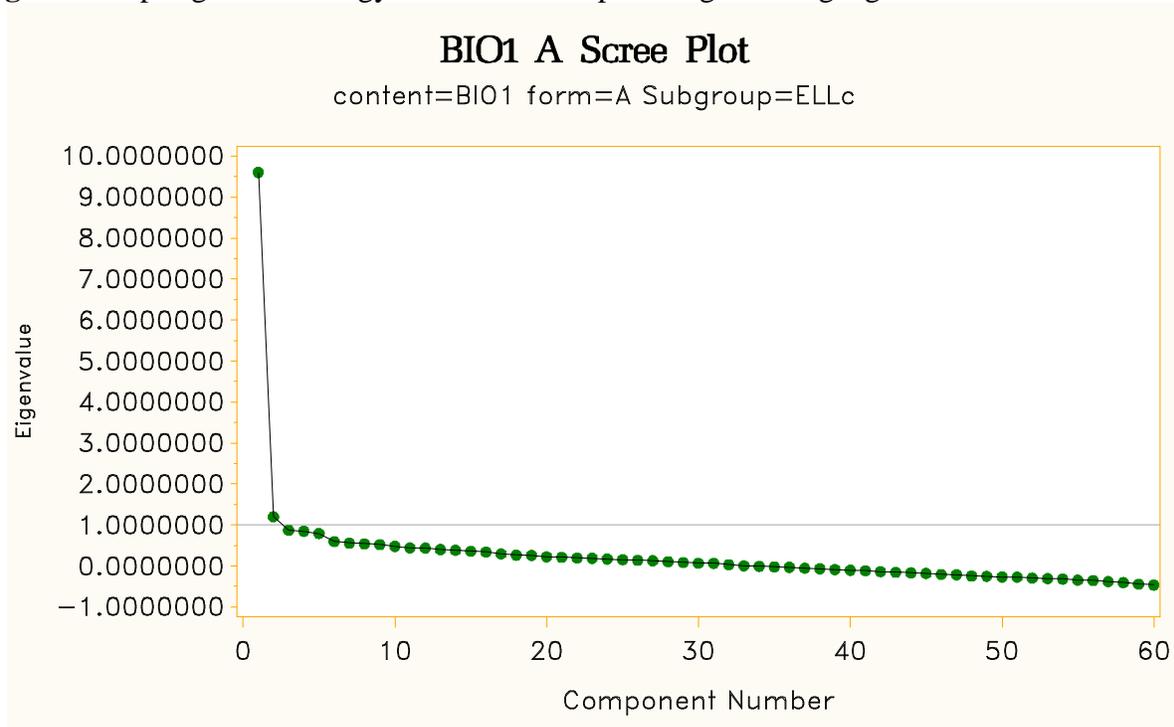
Figure 56. Spring 2013 Biology I Form A scree plot: Individualized Education Program
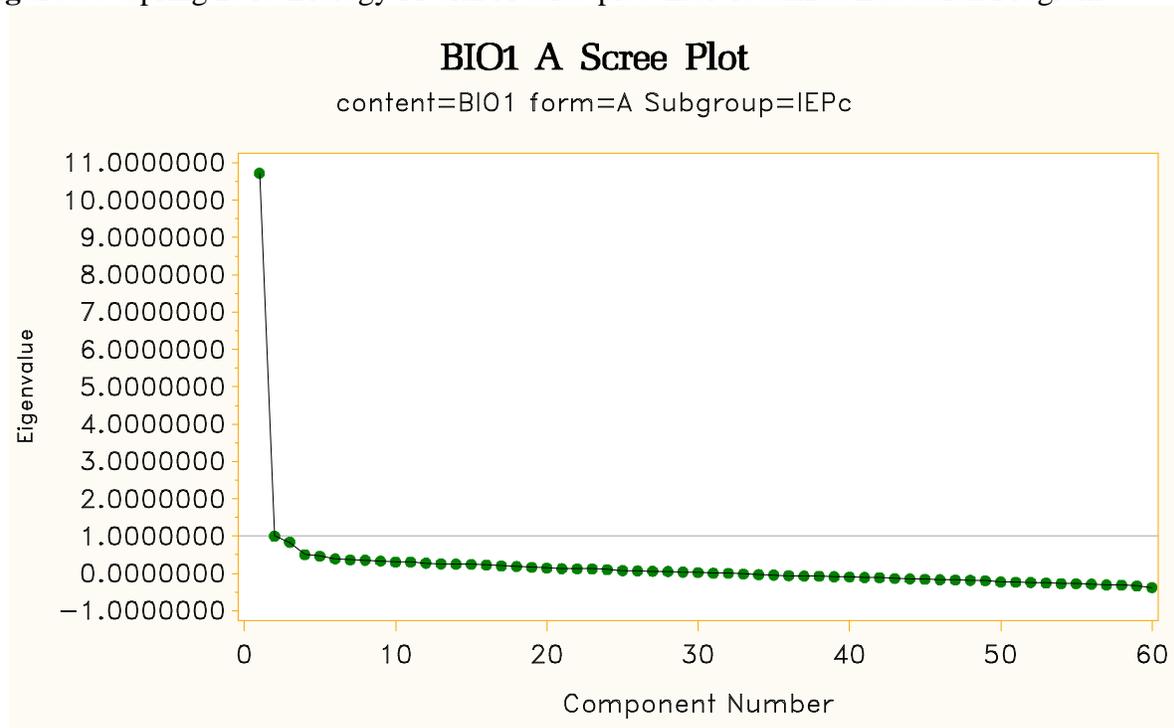
Figure 57. Spring 2013 Biology I Form B scree plot: All
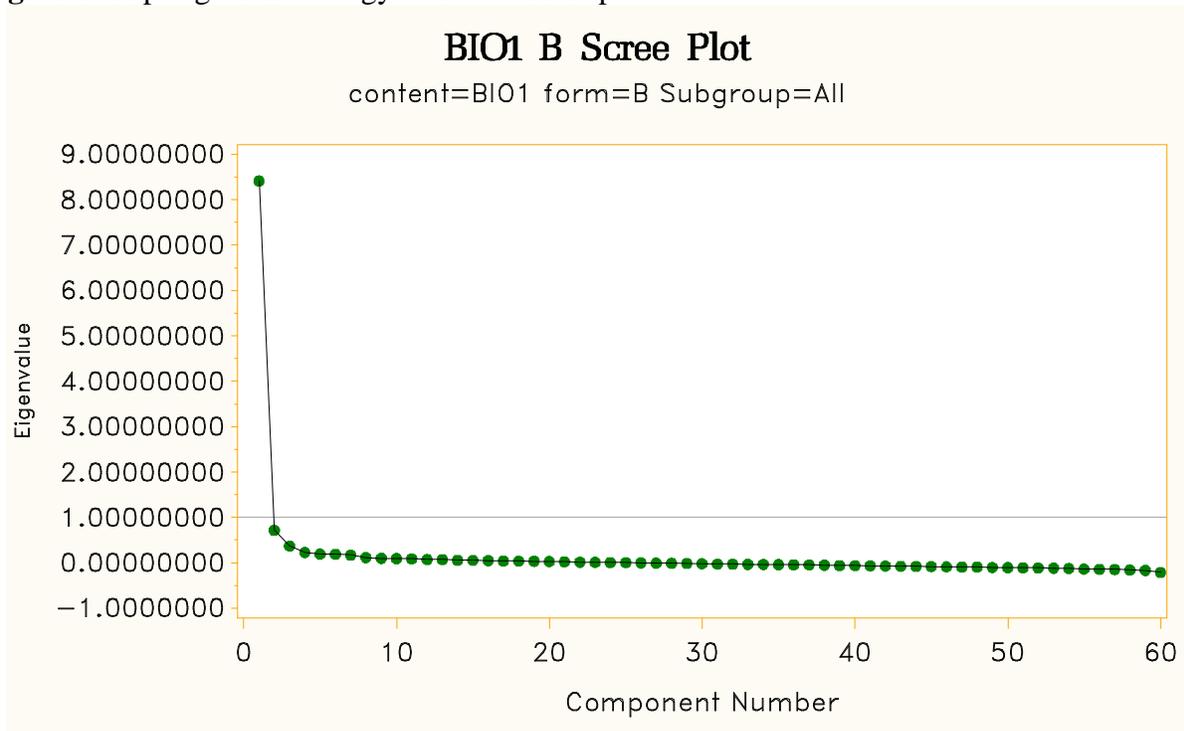
Figure 58. Spring 2013 Biology I Form B scree plot: Accommodated
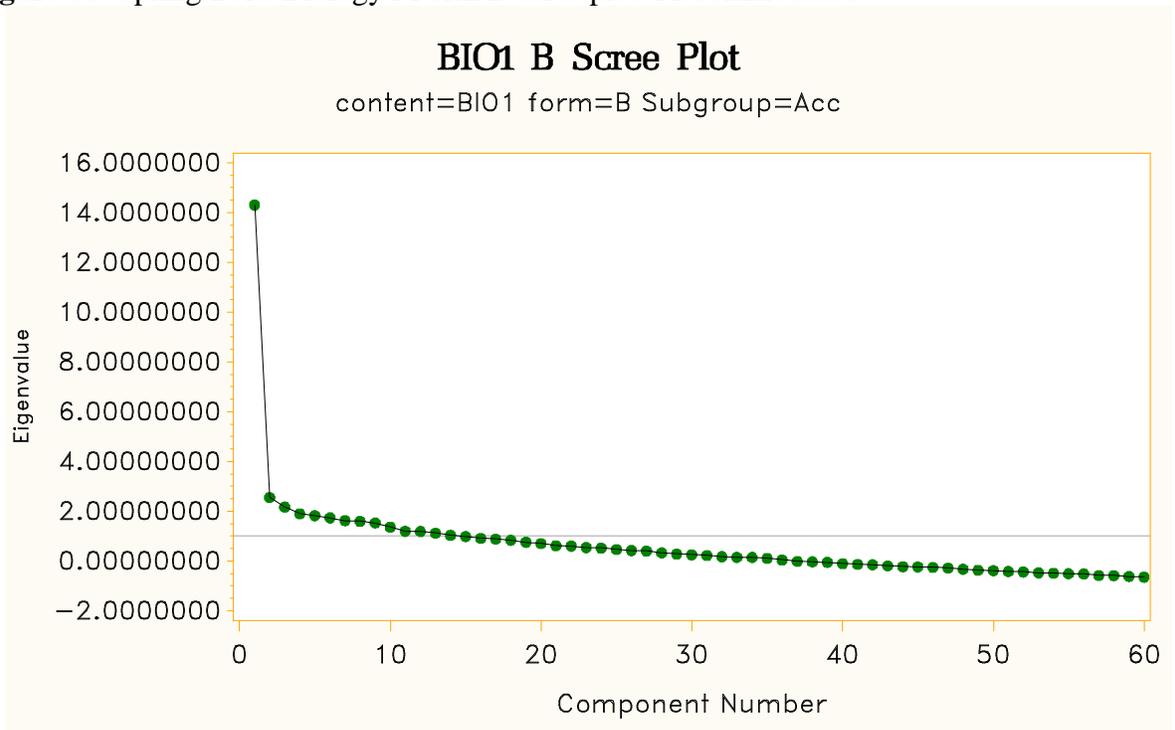
Figure 59. Spring 2013 Biology I Form B scree plot: English Language Learner
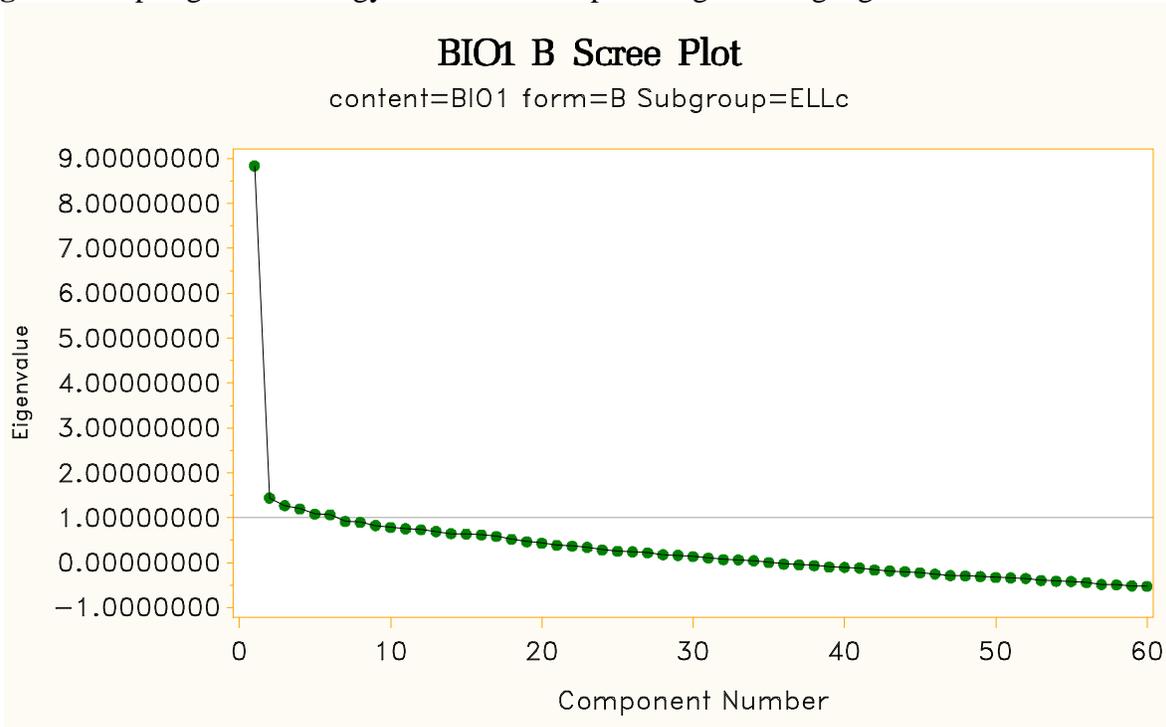
Figure 60. Spring 2013 Biology I Form B scree plot: Individualized Education Program
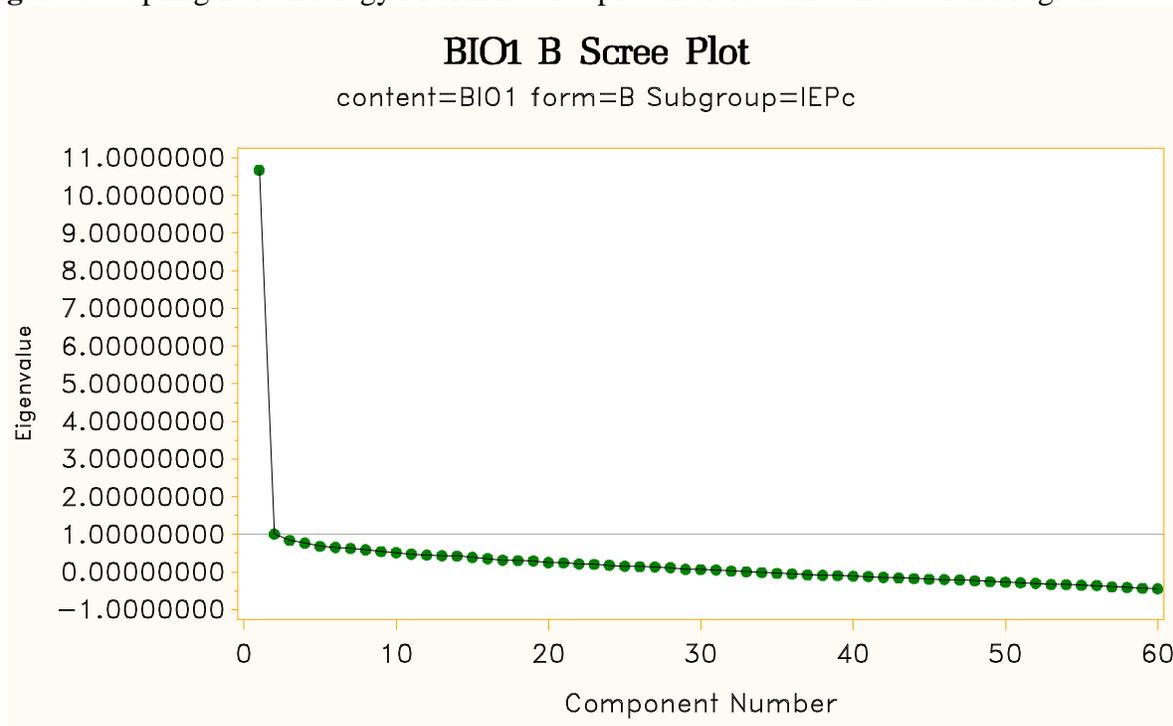
Figure 61. Spring 2013 English II Form AA scree plot: All
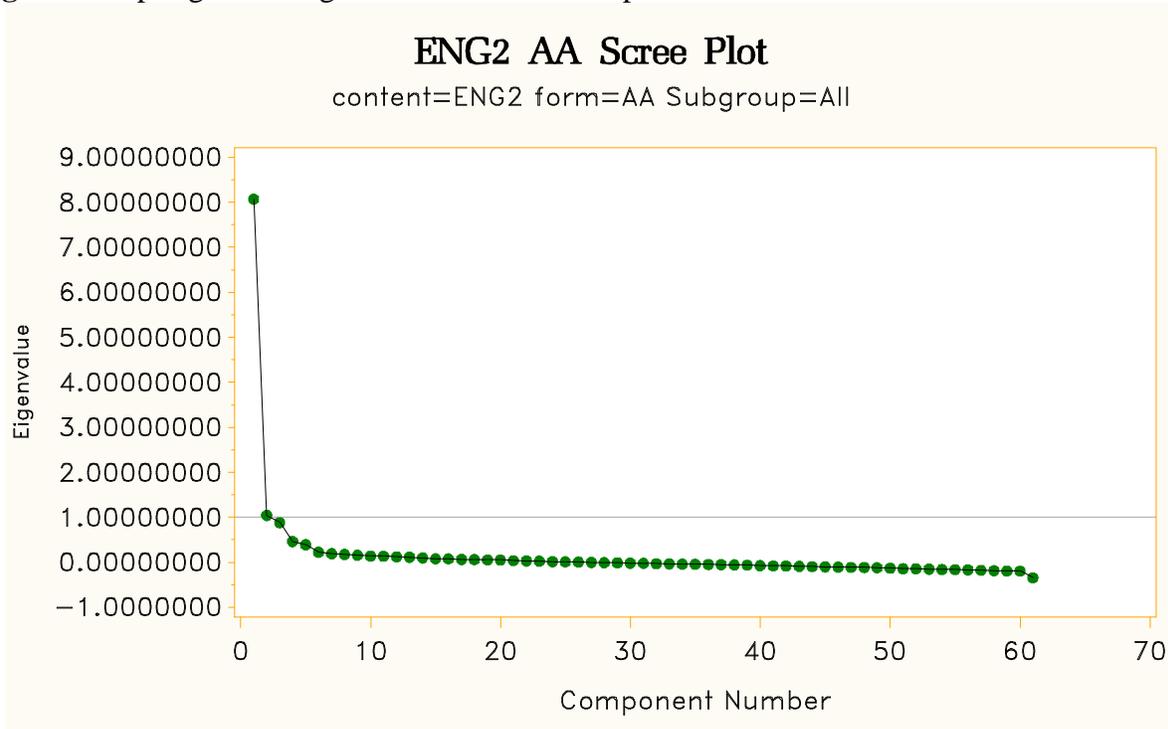
Figure 62. Spring 2013 English II Form AA scree plot: Accommodated
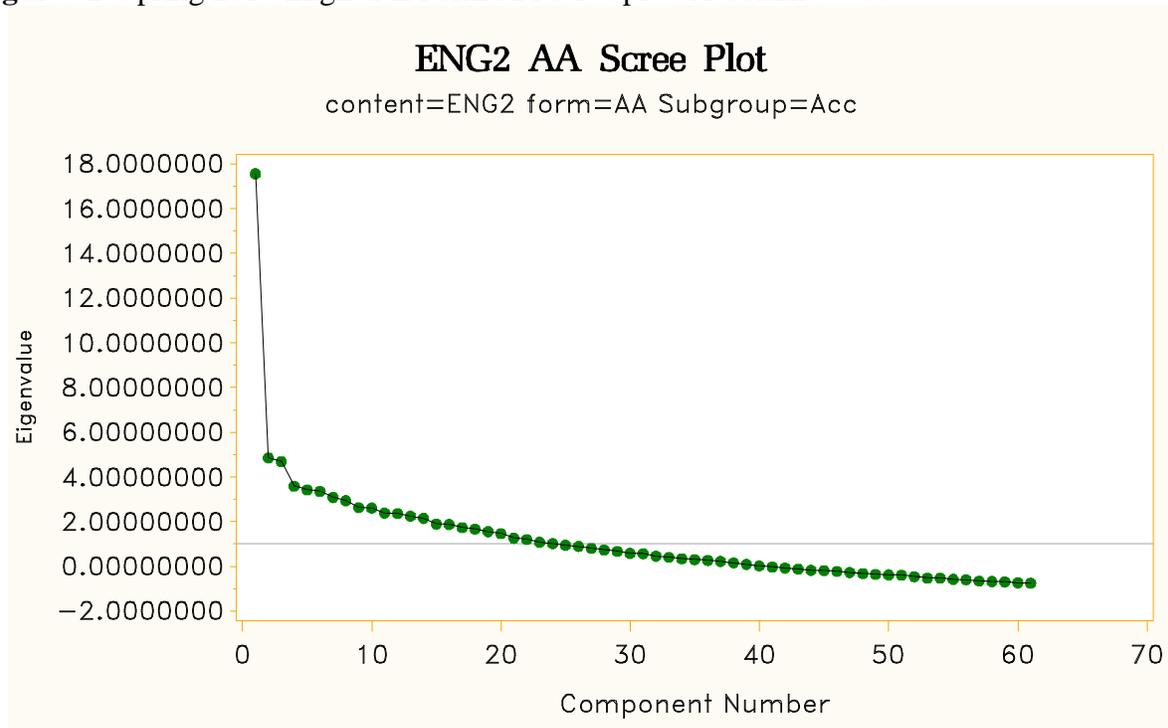
Figure 63. Spring 2013 English II Form AA scree plot: English Language Learner
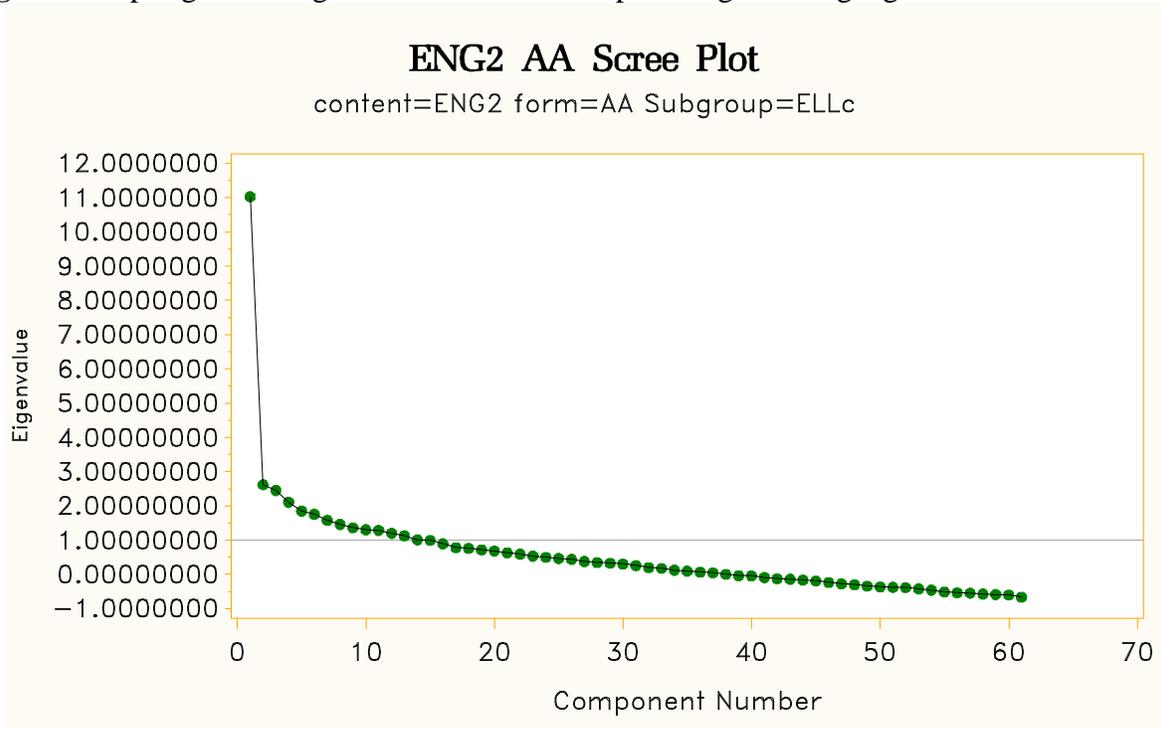
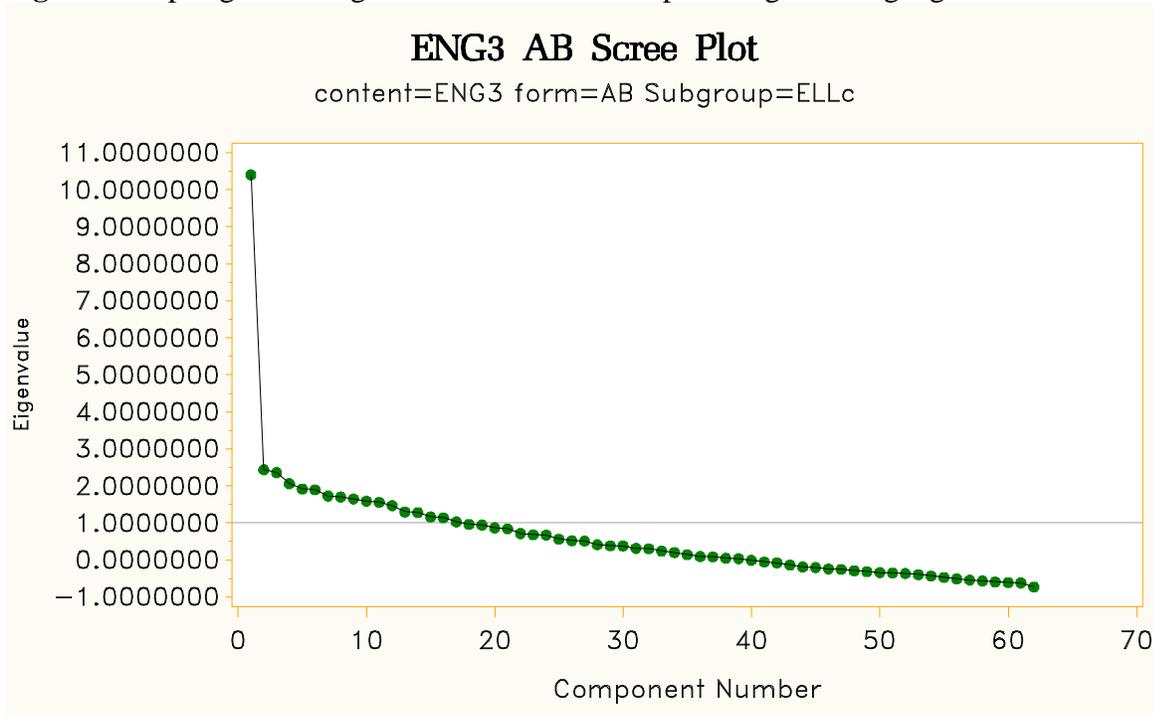**Figure 64.** Spring 2013 English II Form AA scree plot: Individualized Education Program
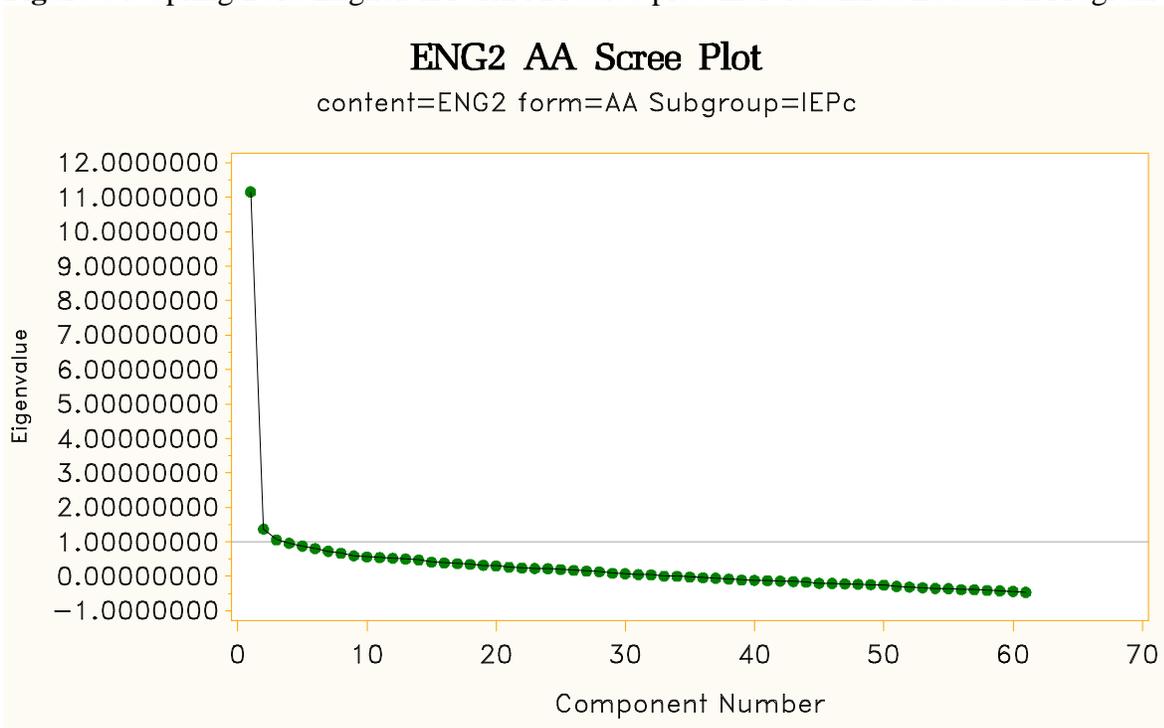
Figure 65. Spring 2013 English II Form AB scree plot: All
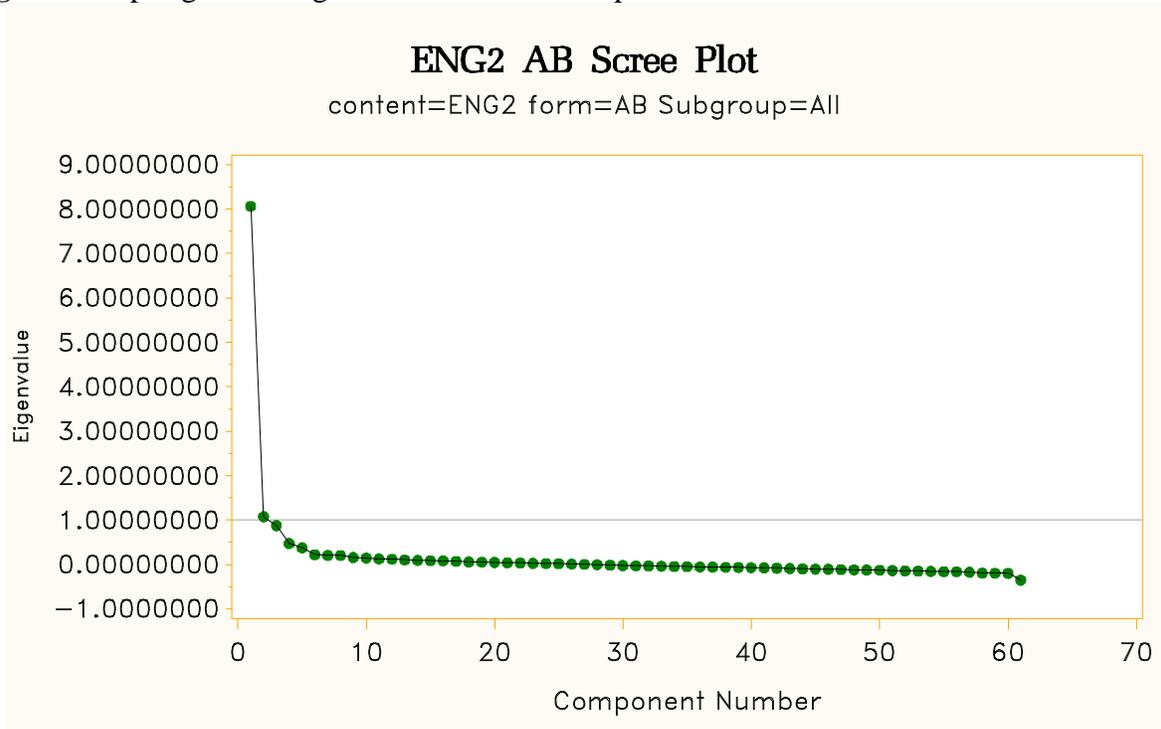
Figure 66. Spring 2013 English II Form AB scree plot: Accommodated
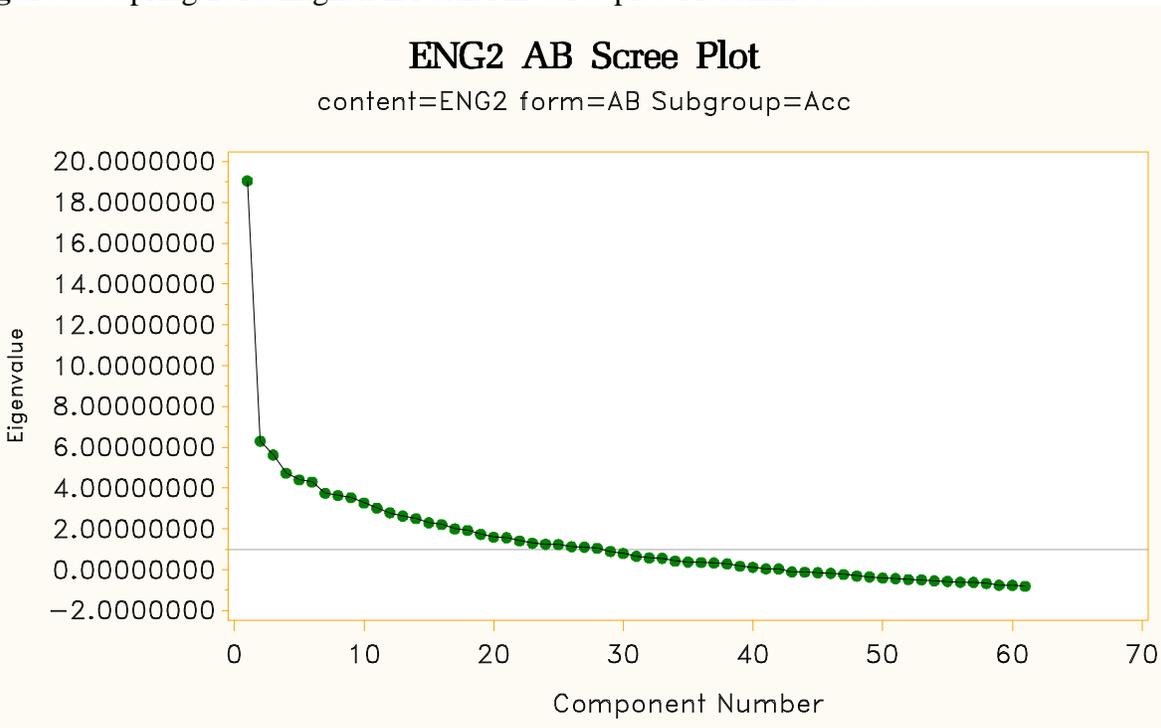
Figure 67. Spring 2013 English II Form AB scree plot: English Language Learner
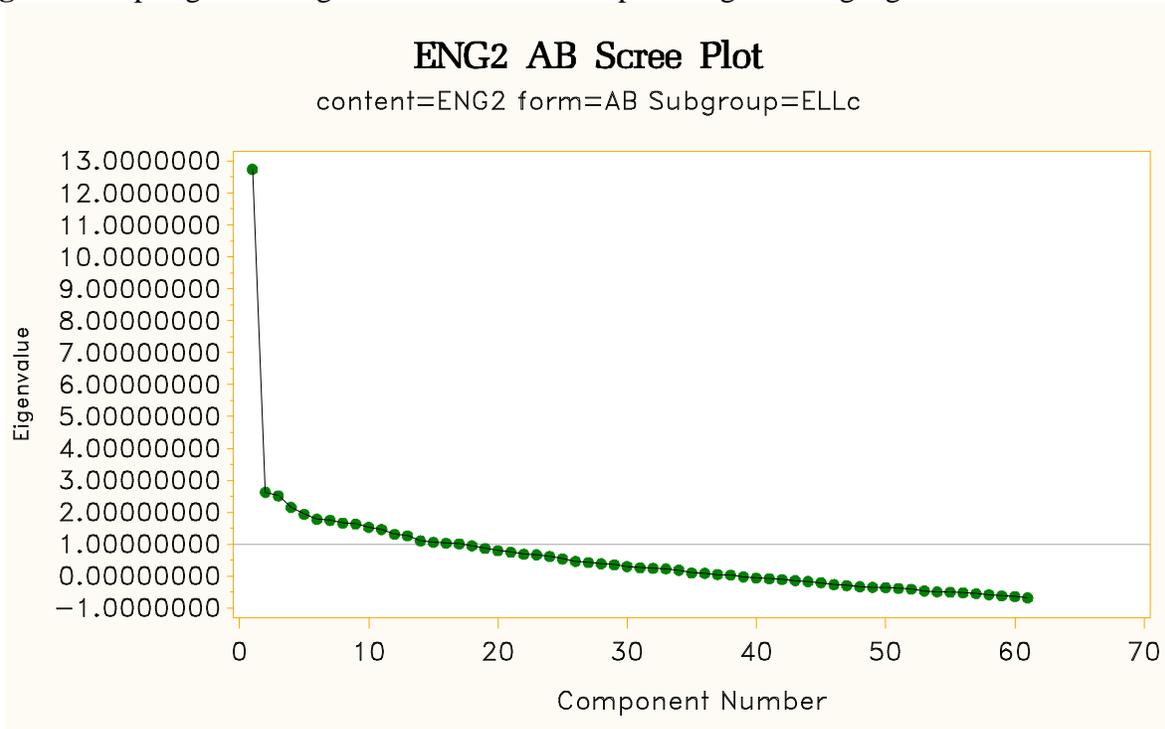
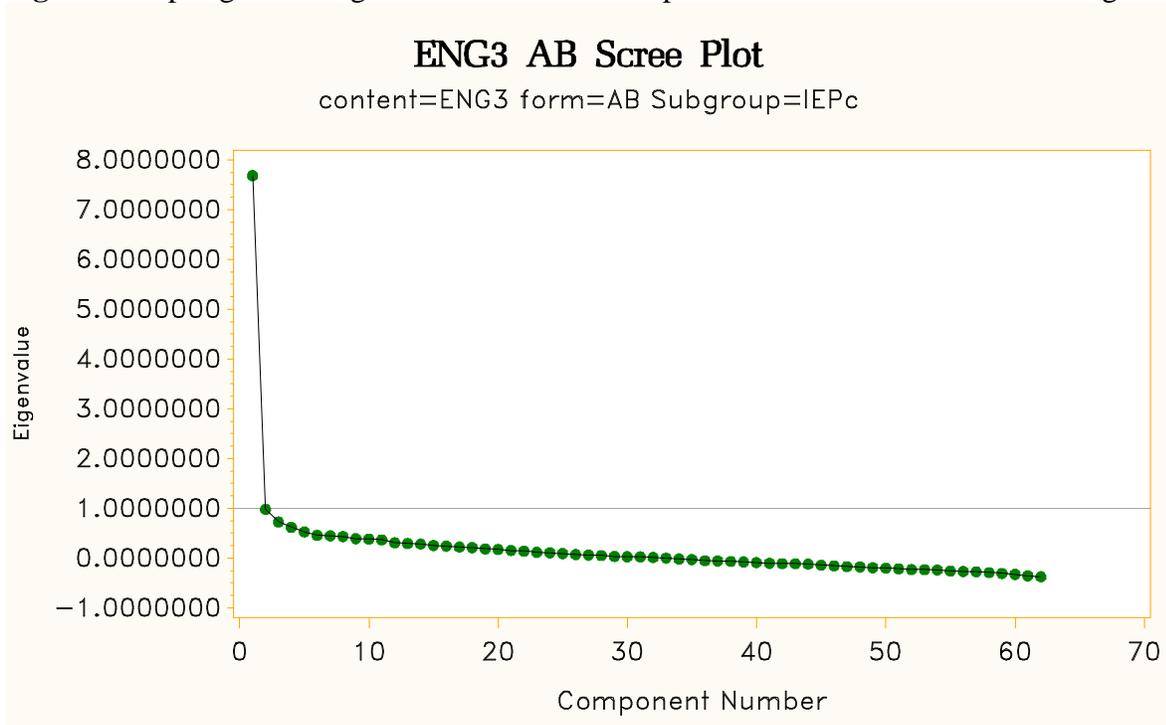**Figure 68.** Spring 2013 English II Form AB scree plot: Individualized Education Program
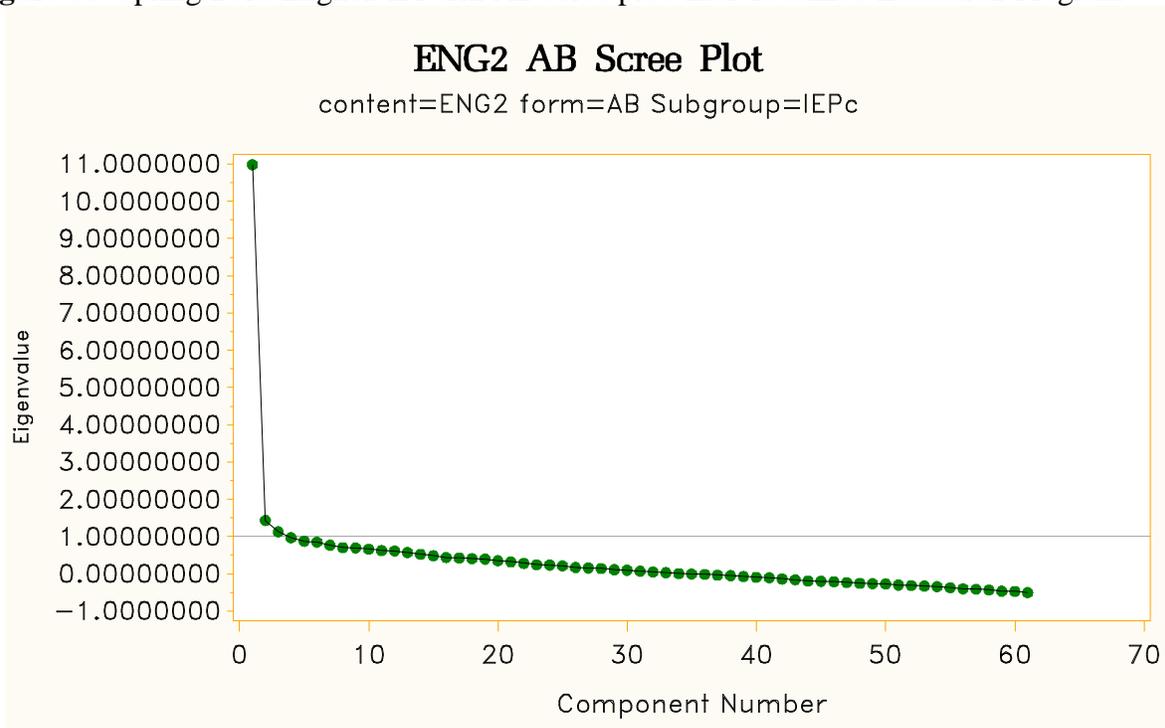
Figure 69. Spring 2013 English II Form BA scree plot: All
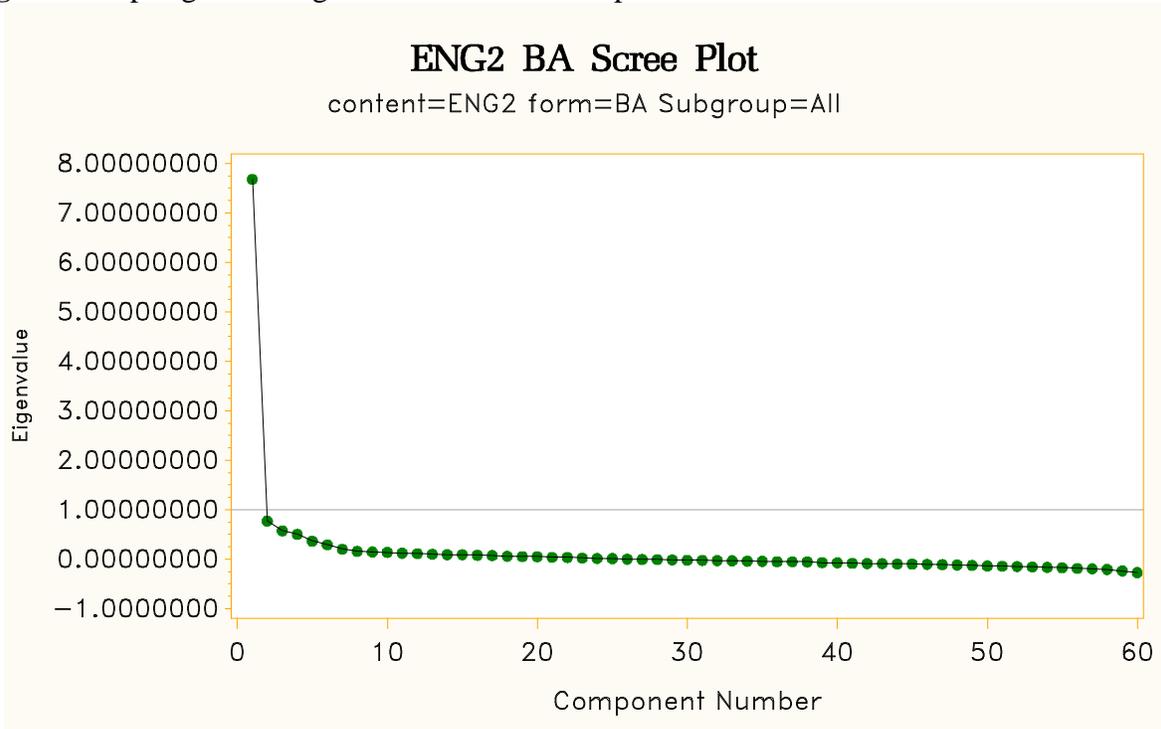
Figure 70. Spring 2013 English II Form BA scree plot: Accommodated
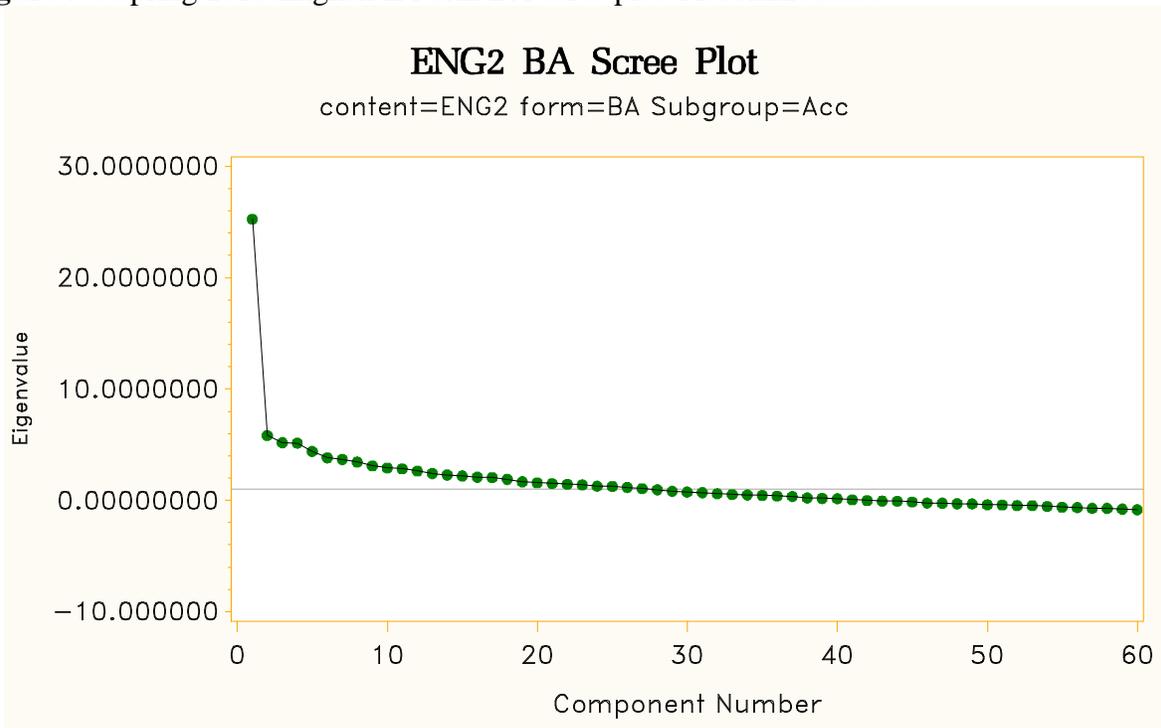
Figure 71. Spring 2013 English II Form BA scree plot: English Language Learner
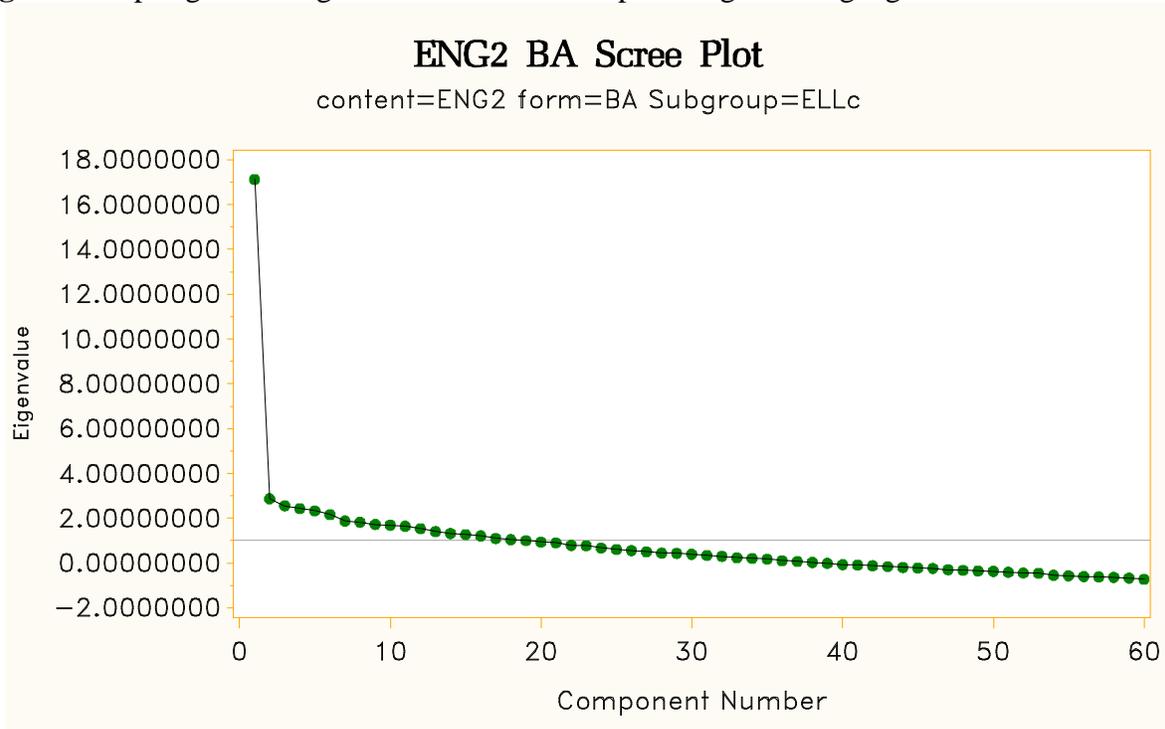
Figure 72. Spring 2013 English II Form BA scree plot: Individualized Education Program
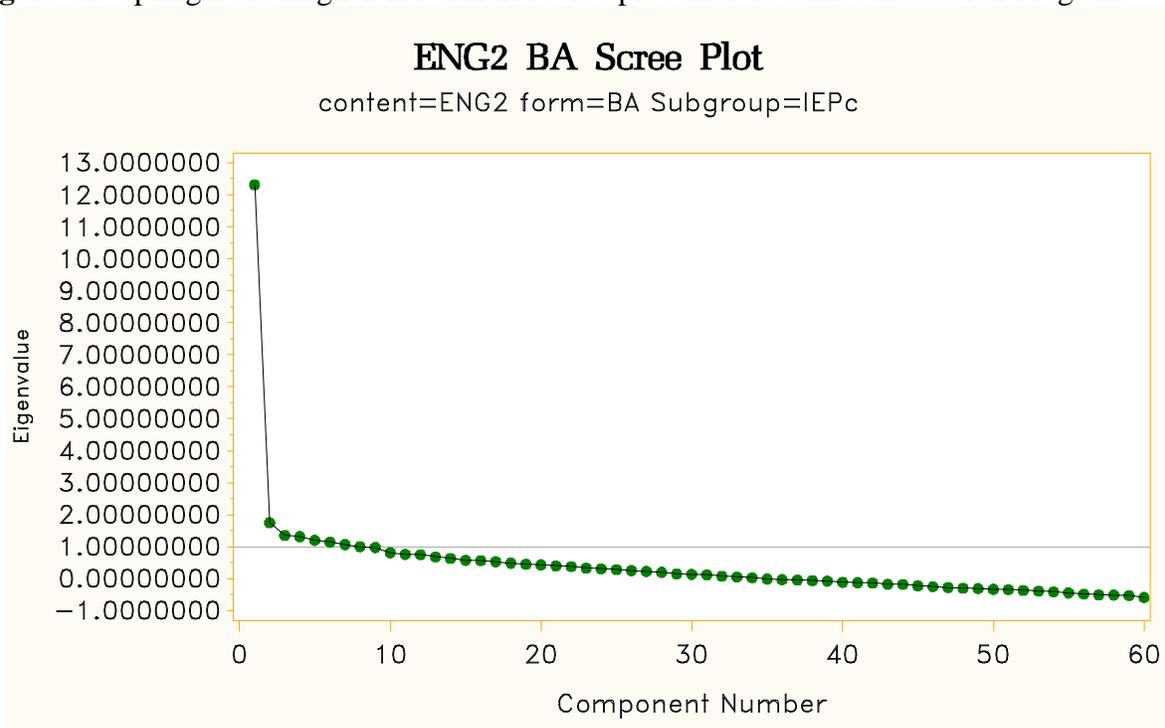
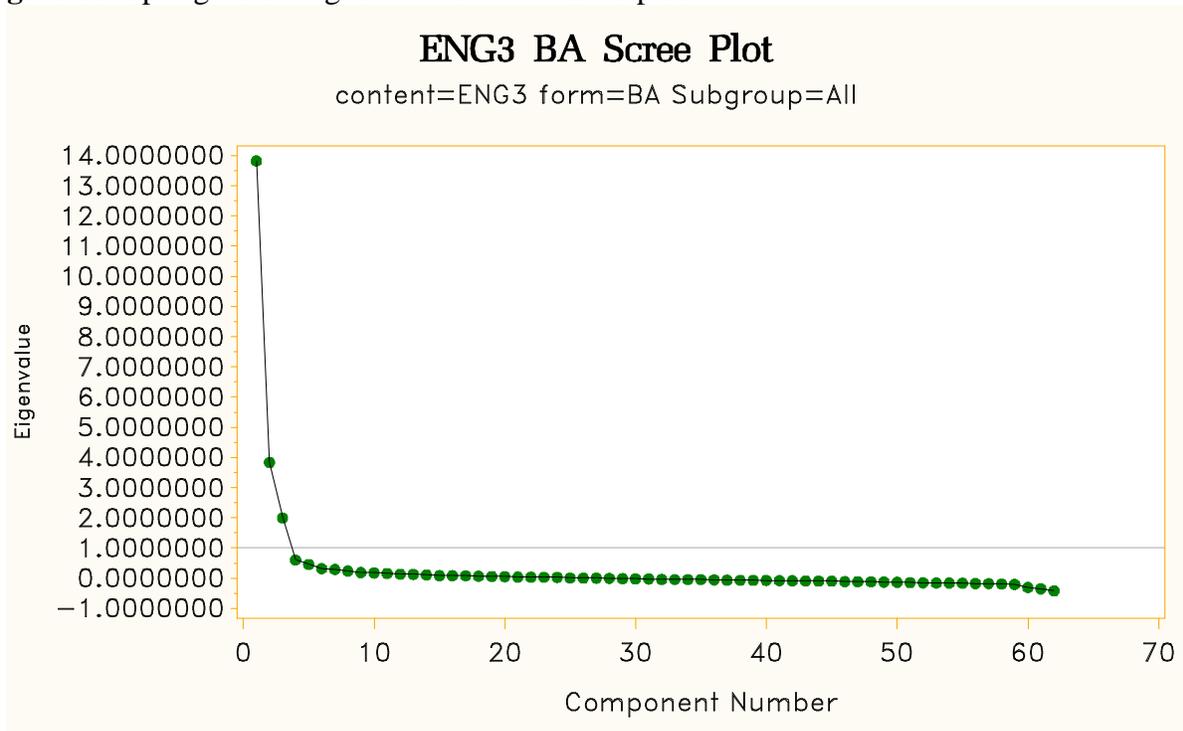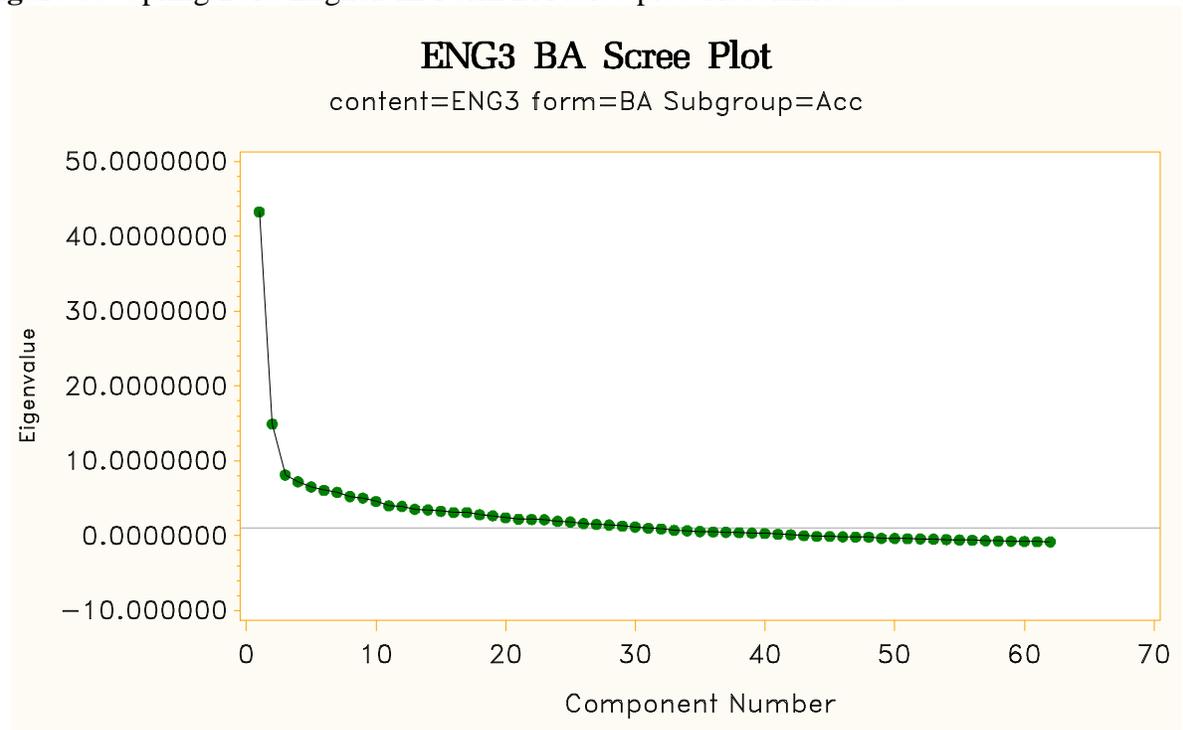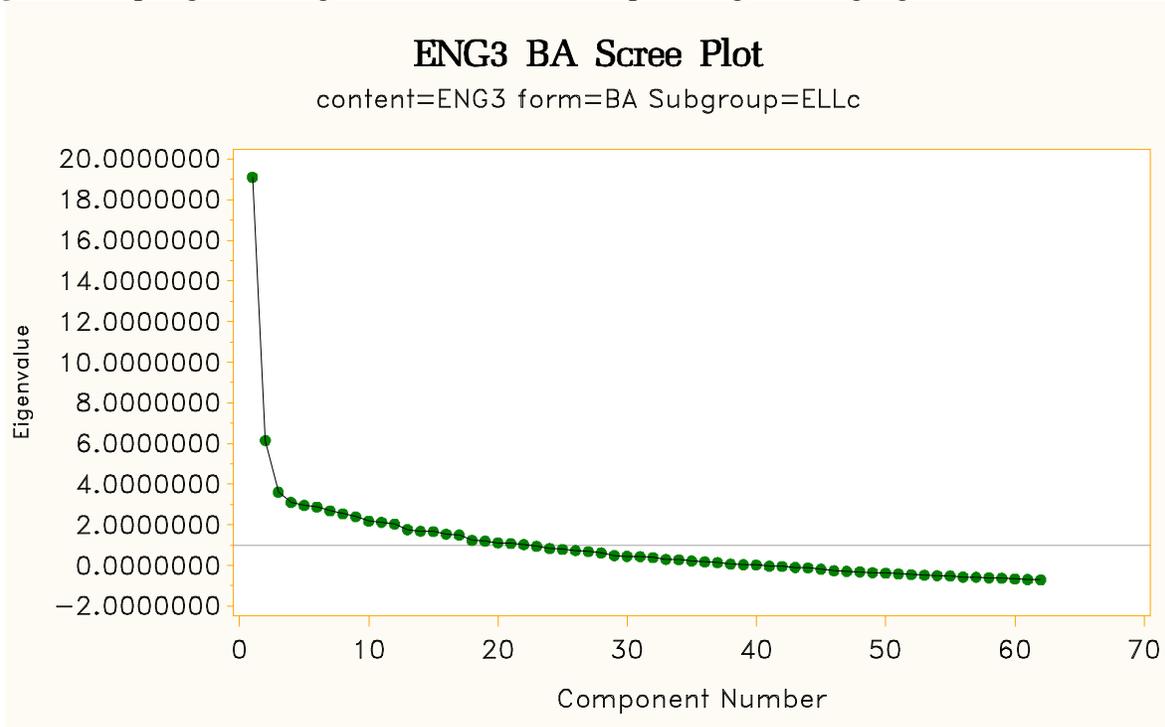**Figure 73.** Spring 2013 English II Form BB scree plot: All
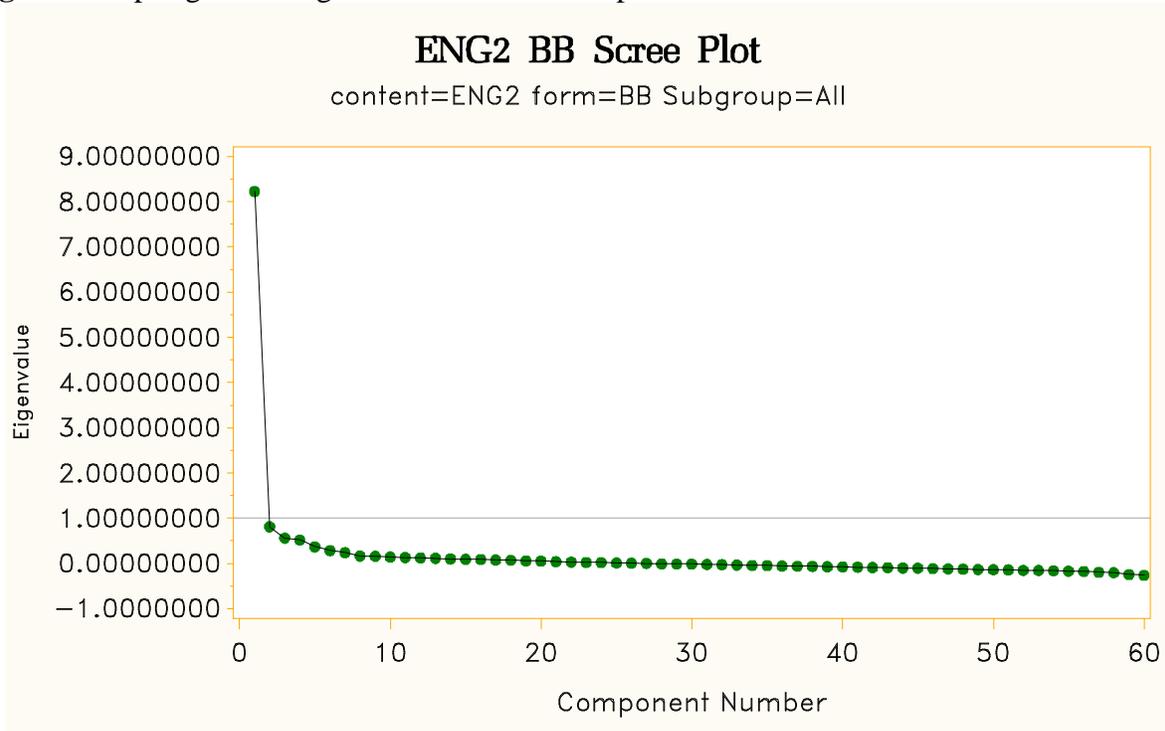
Figure 74. Spring 2013 English II Form BB scree plot: Accommodated
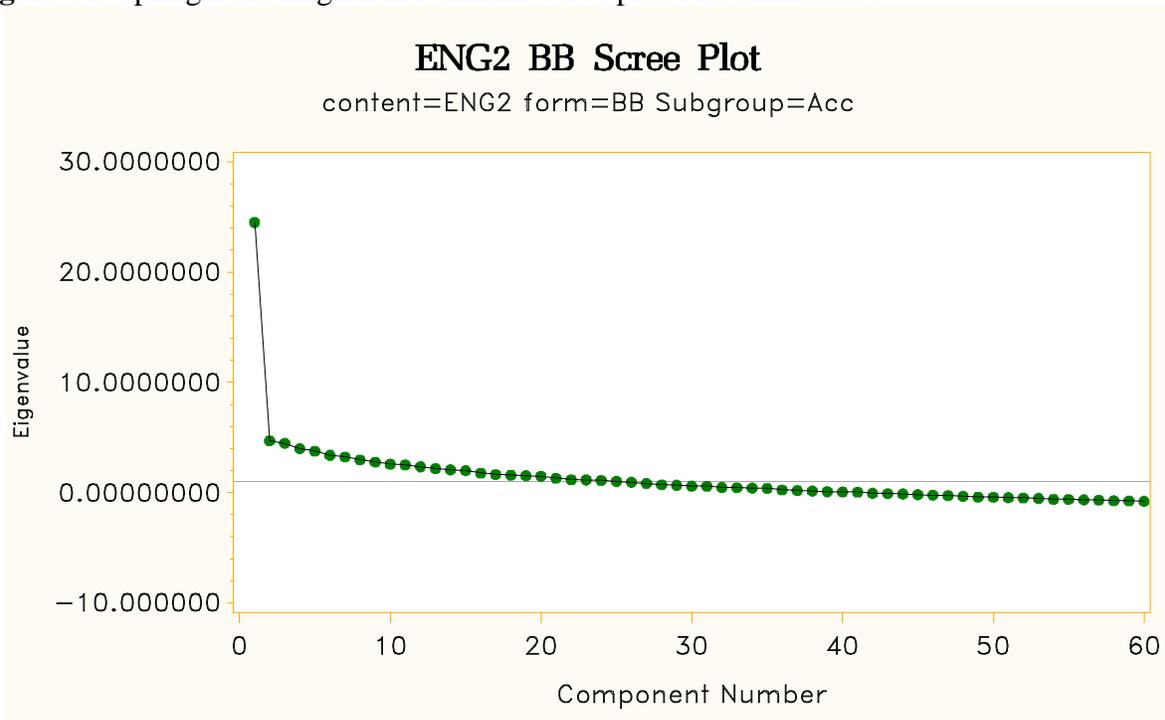
Figure 75. Spring 2013 English II Form BB scree plot: English Language Learner
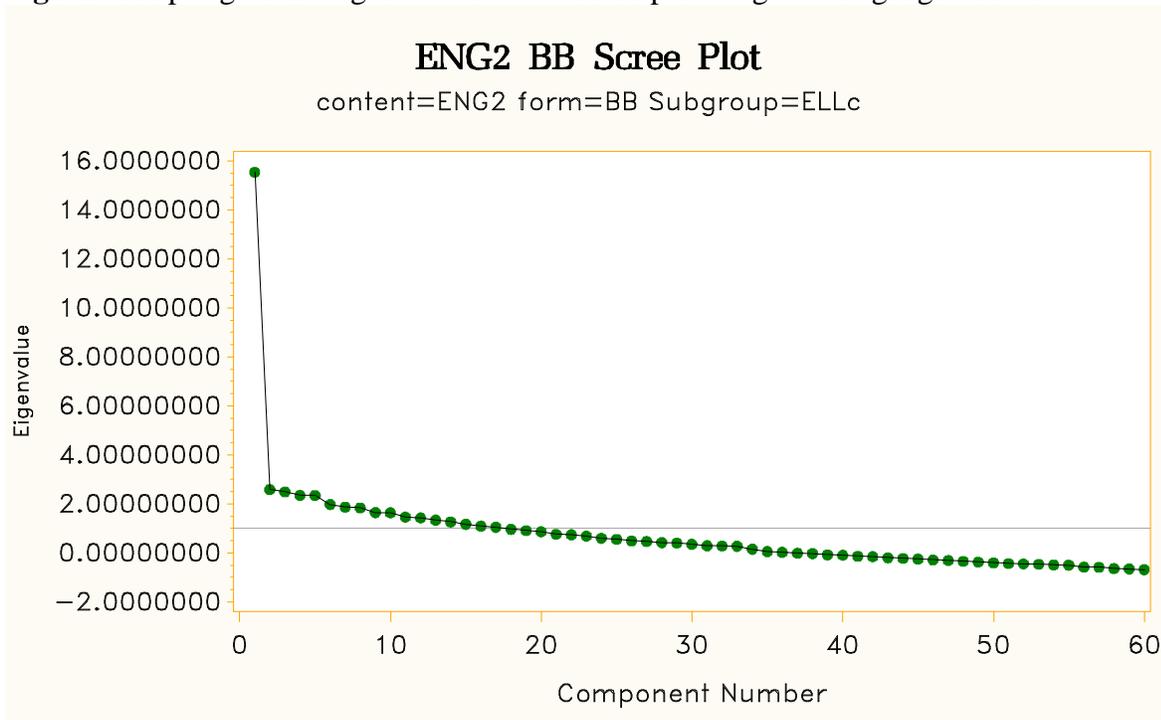
Figure 76. Spring 2013 English II Form BB scree plot: Individualized Education Program
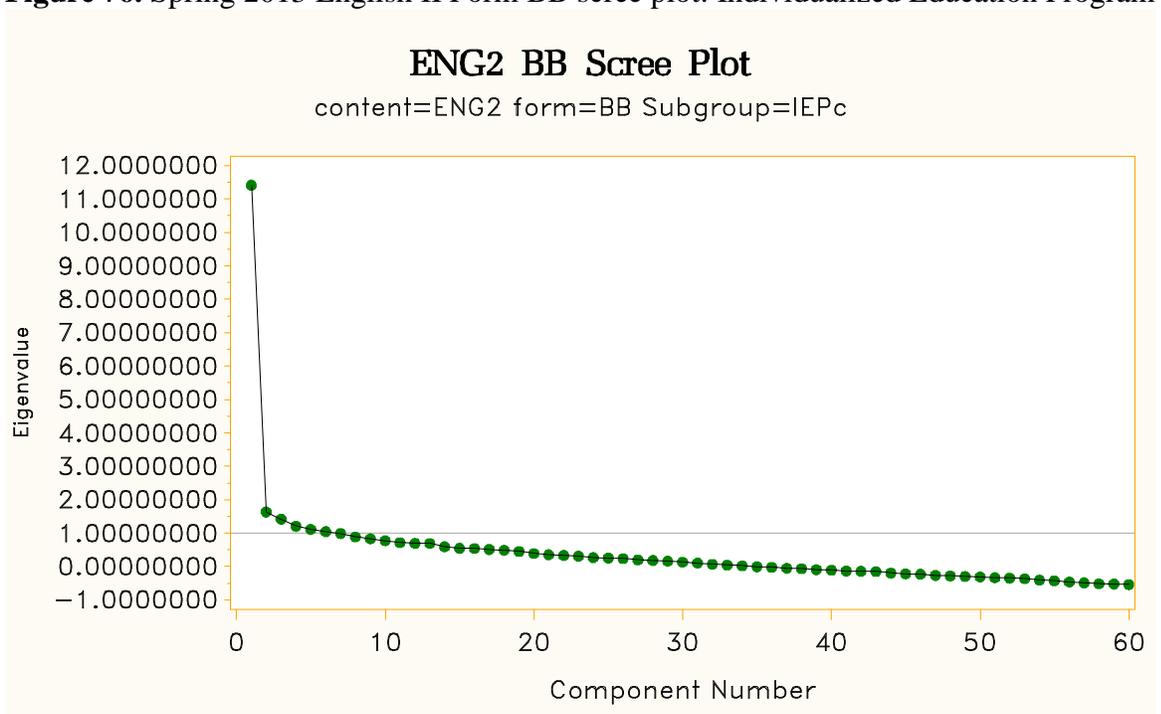
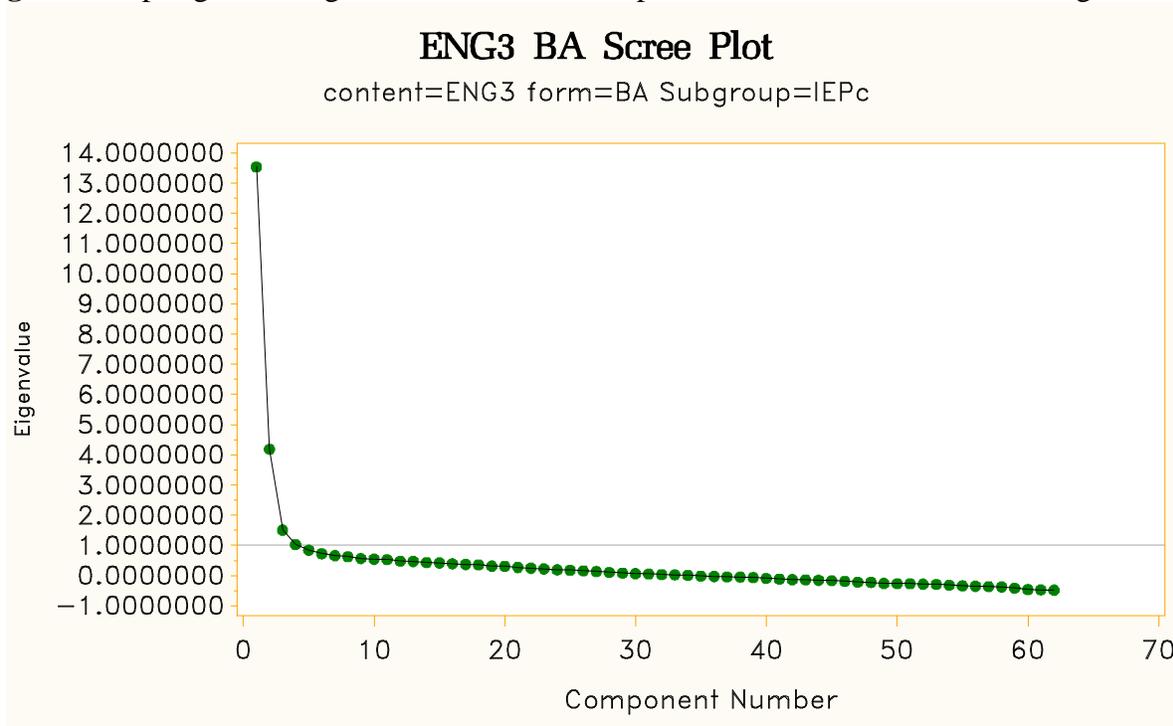**Figure 77.** Spring 2013 English III Form AA scree plot: All
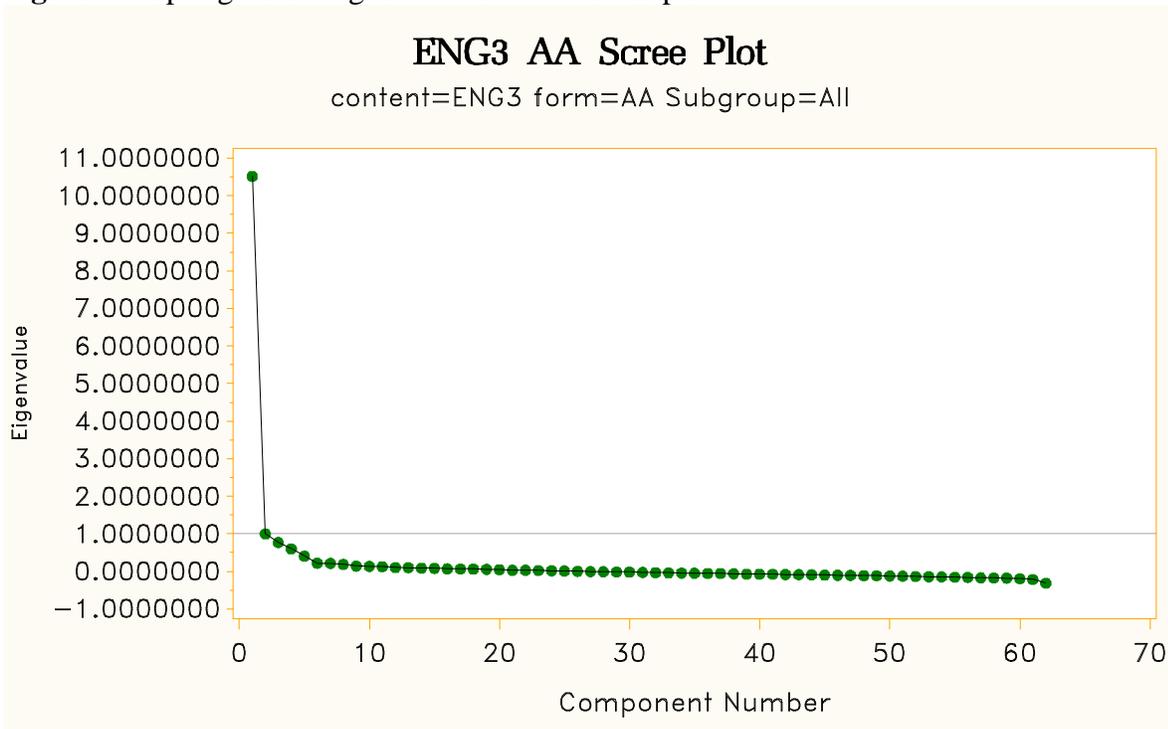
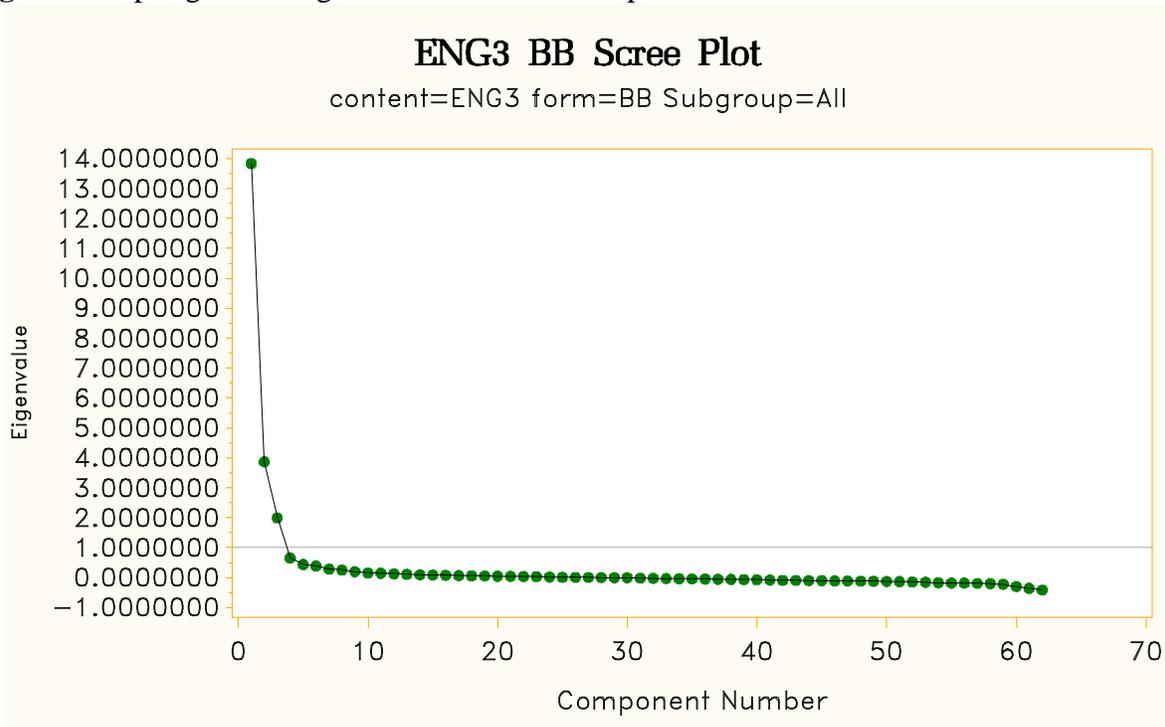**Figure 78.** Spring 2013 English III Form AA scree plot: Accommodated
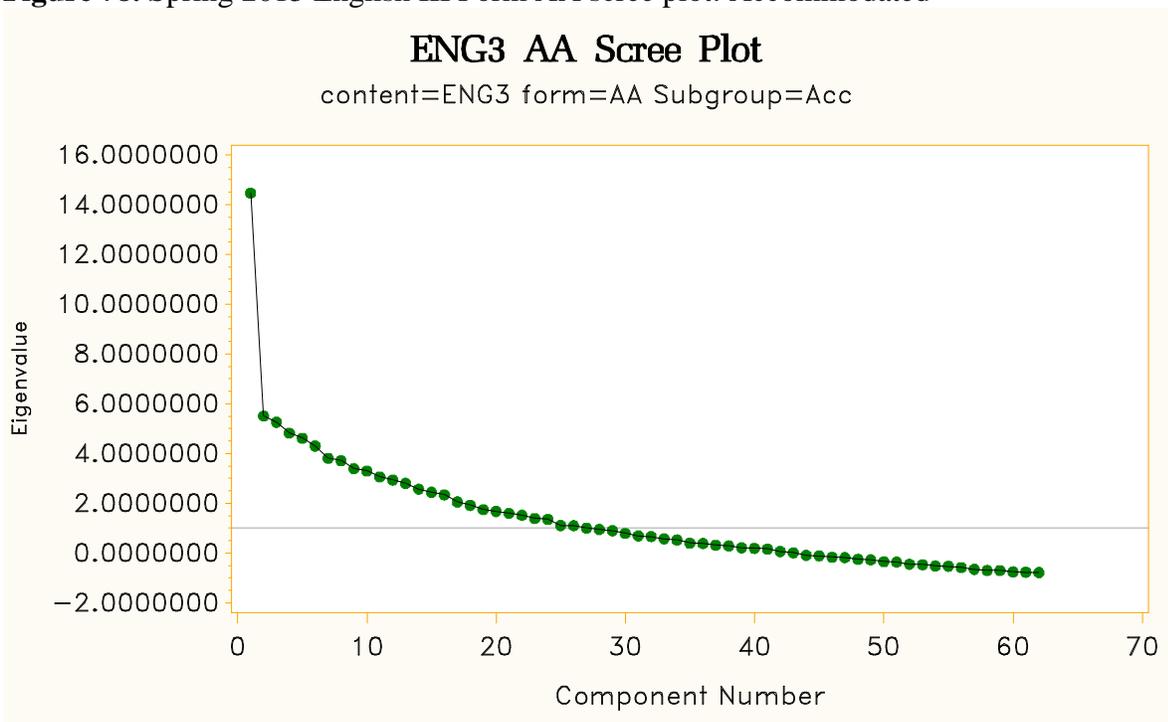
Figure 79. Spring 2013 English III Form AA scree plot: English Language Learner
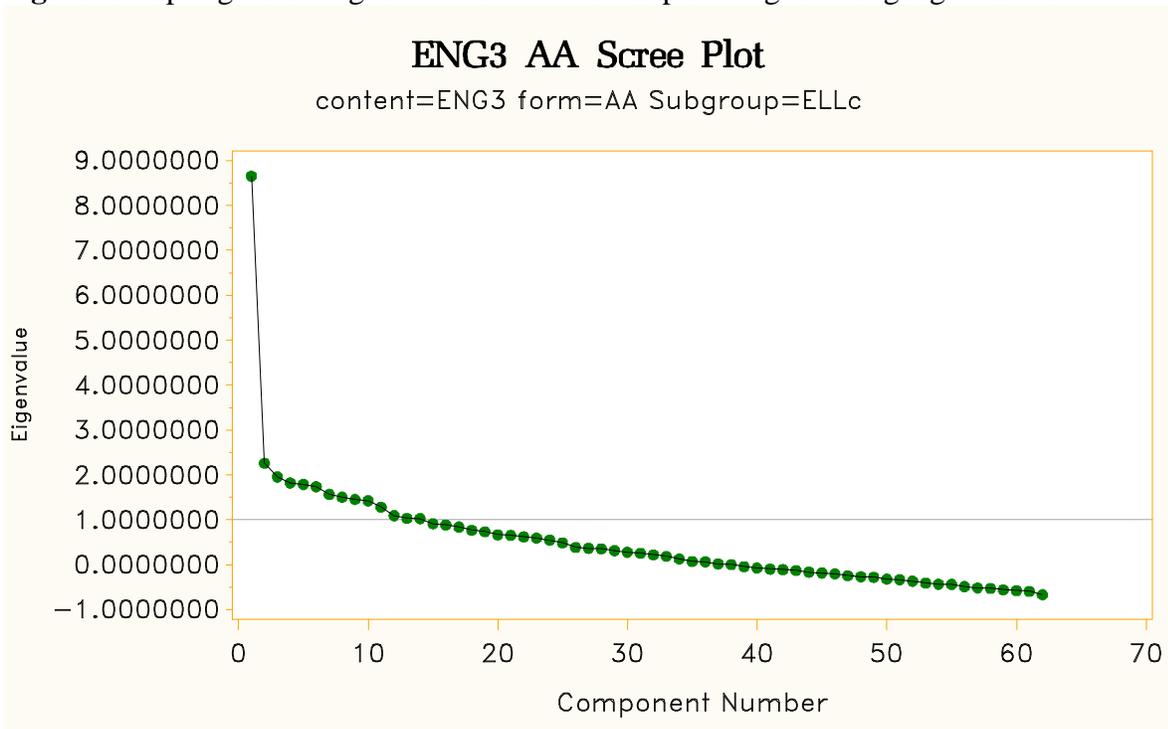
Figure 80. Spring 2013 English III Form AA scree plot: Individualized Education Program
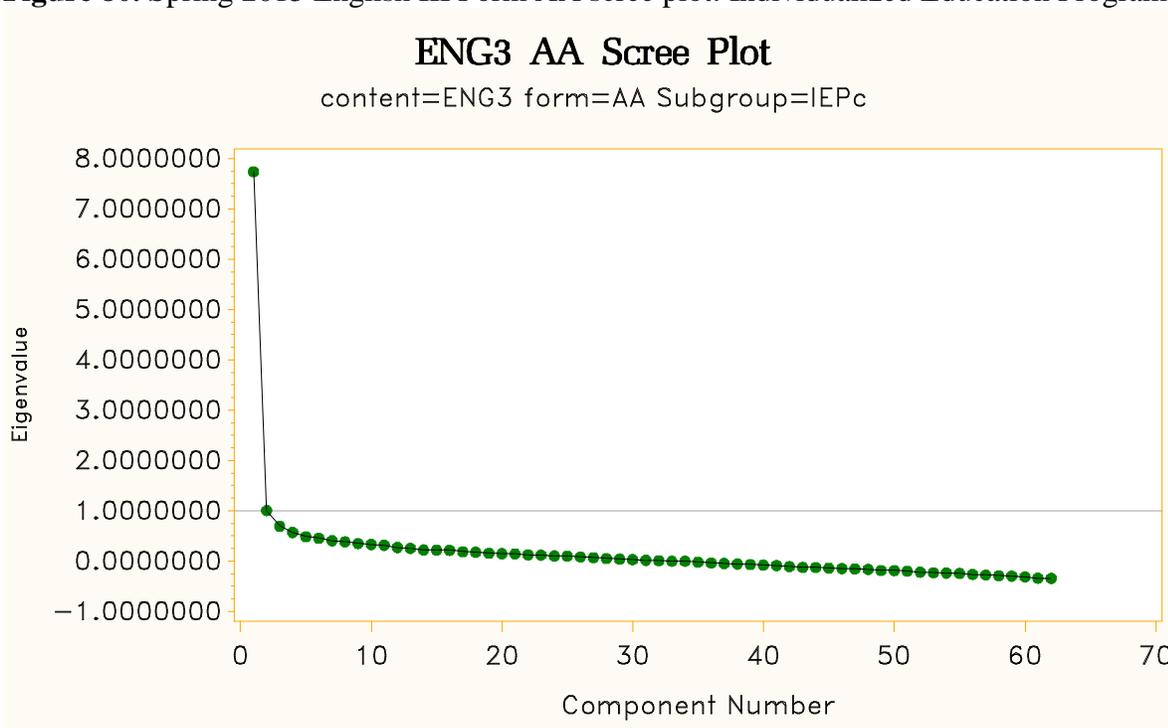
Figure 81. Spring 2013 English III Form AB scree plot: All
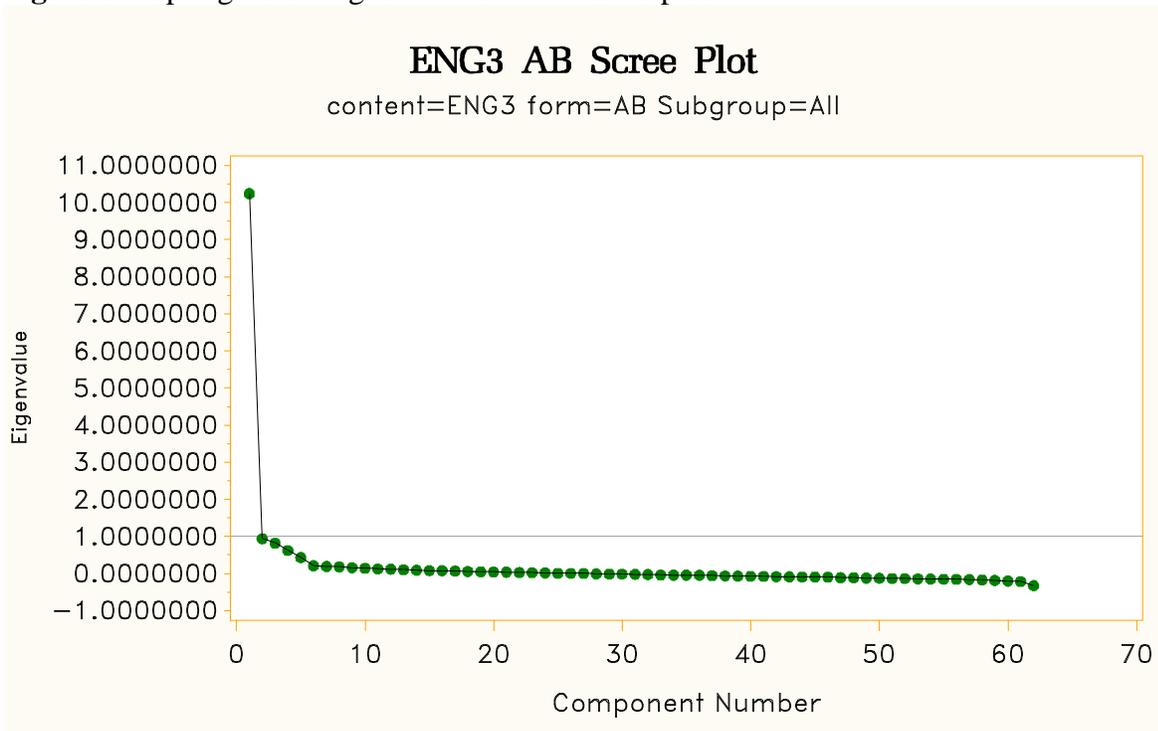
Figure 82. Spring 2013 English III Form AB scree plot: Accommodated

Figure 83. Spring 2013 English III Form AB scree plot: English Language Learner

Figure 84. Spring 2013 English III Form AB scree plot: Individualized Education Program

Figure 85. Spring 2013 English III Form BA scree plot: All

Figure 86. Spring 2013 English III Form BA scree plot: Accommodated

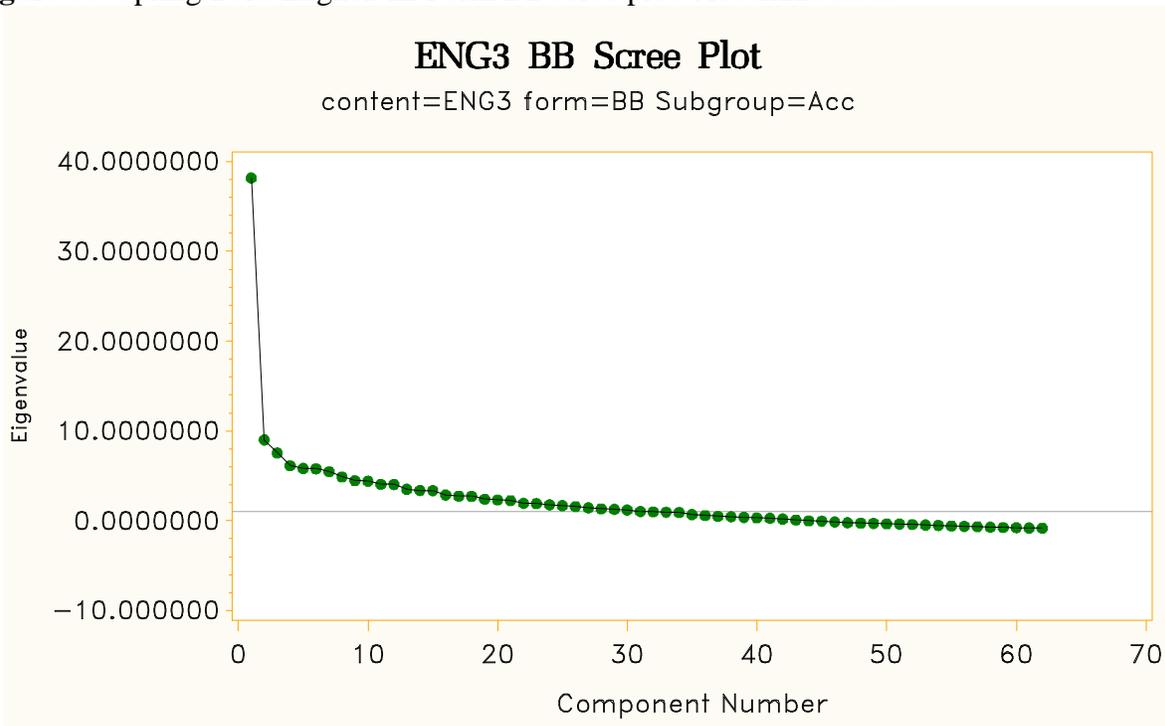**Figure 87.** Spring 2013 English III Form BA scree plot: English Language Learner

Figure 88. Spring 2013 English III Form BA scree plot: Individualized Education Program
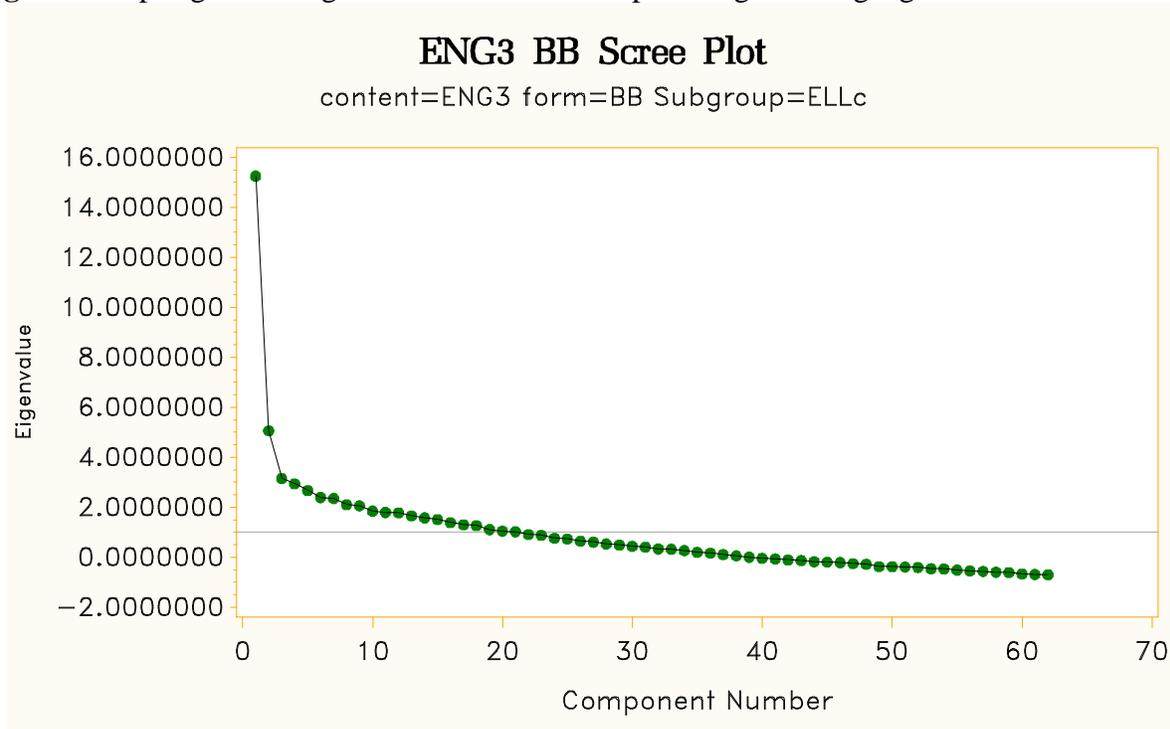
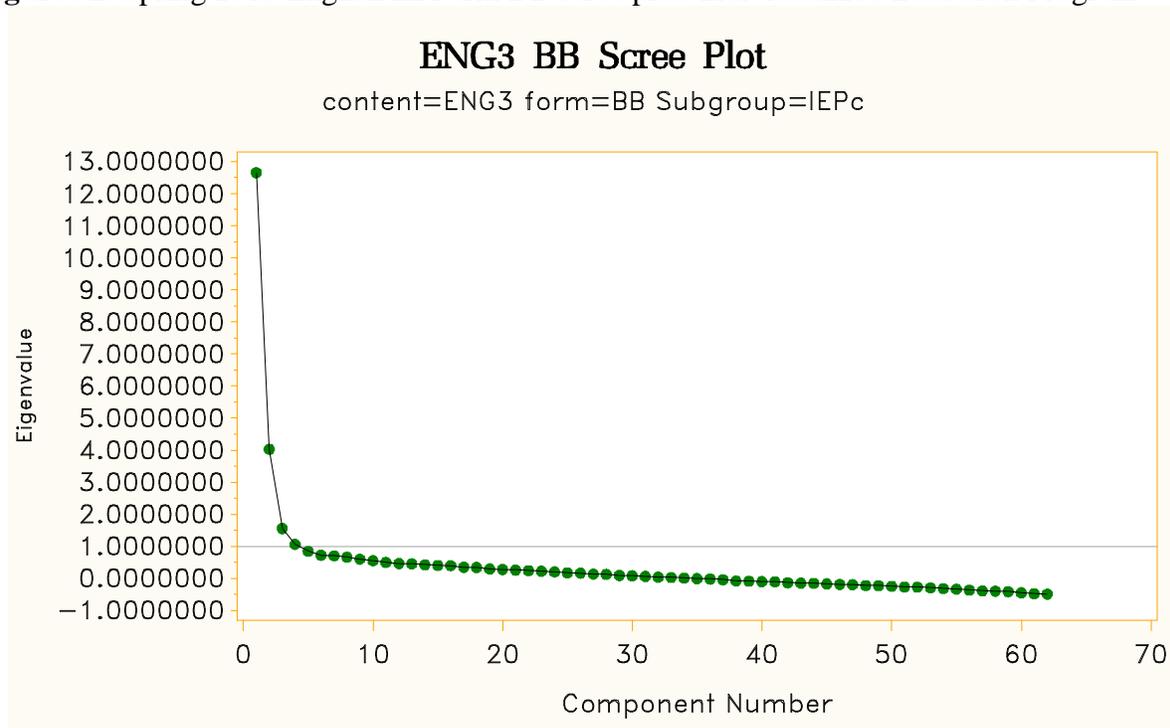**Figure 89.** Spring 2013 English III Form BB scree plot: All

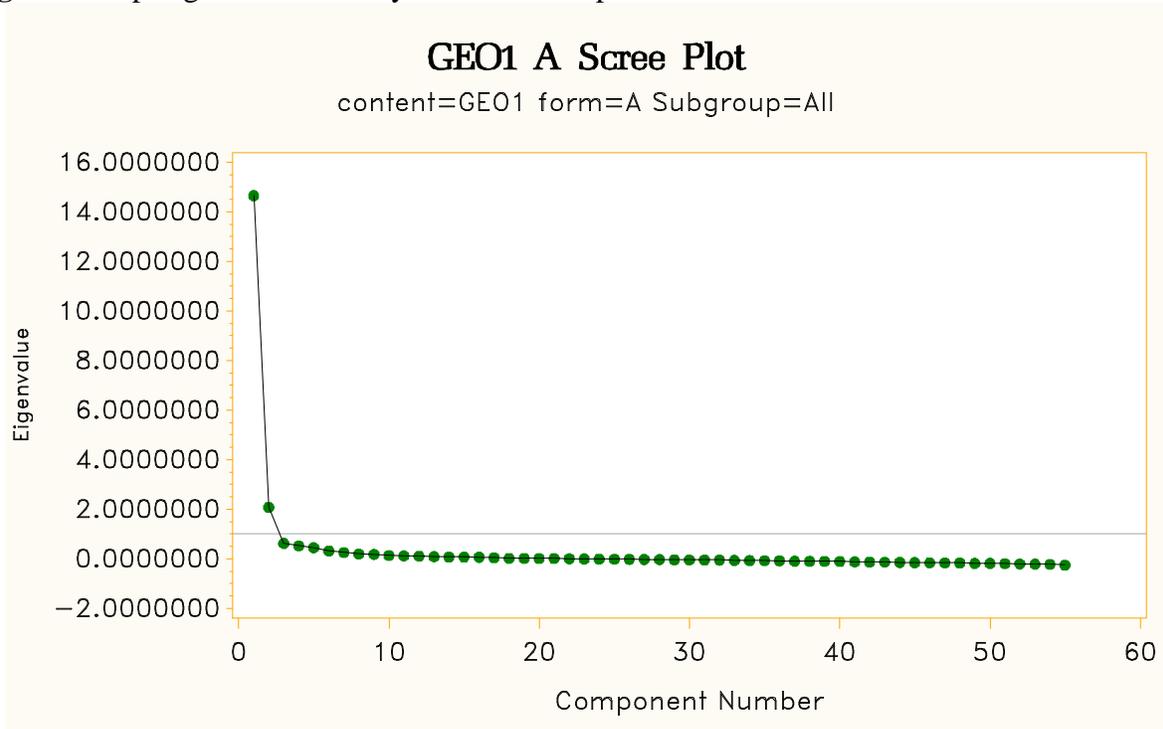**Figure 90.** Spring 2013 English III Form BB scree plot: Accommodated

Figure 91. Spring 2013 English III Form BB scree plot: English Language Learner
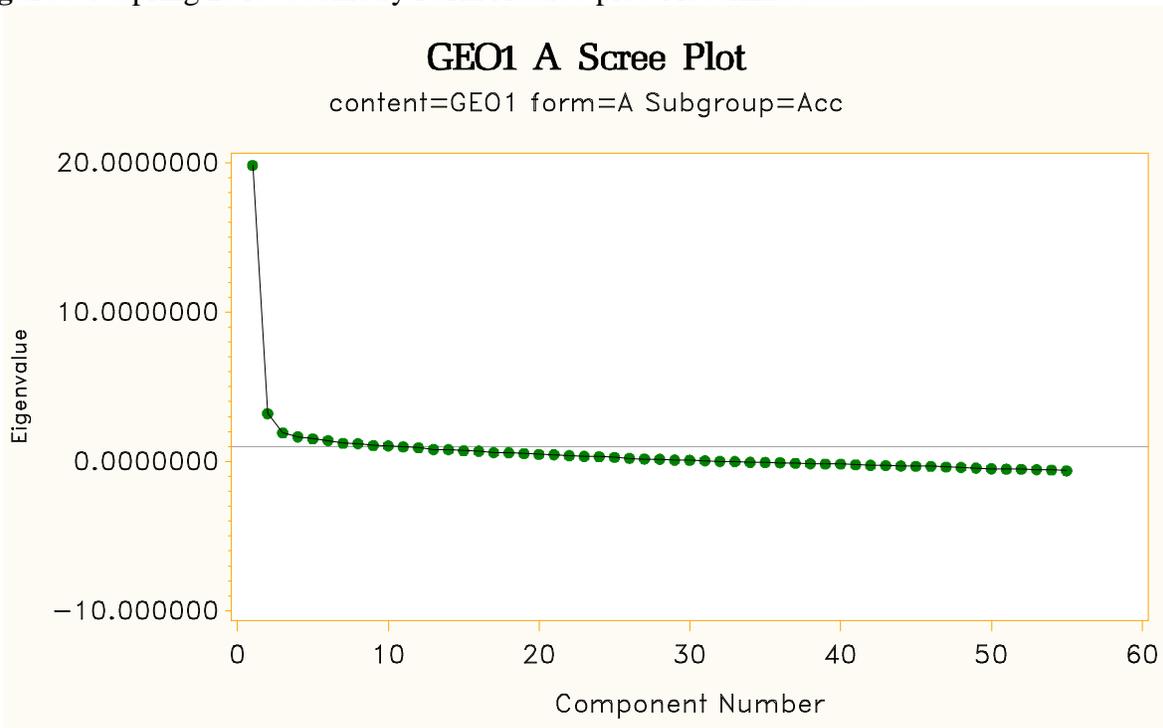

Figure 92. Spring 2013 English III Form BB scree plot: Individualized Education Program

Figure 93. Spring 2013 Geometry Form A scree plot: All

Figure 94. Spring 2013 Geometry Form A scree plot: Accommodated

Figure 95. Spring 2013 Geometry Form A scree plot: English Language Learner

Figure 96. Spring 2013 Geometry Form A scree plot: Individualized Education Program
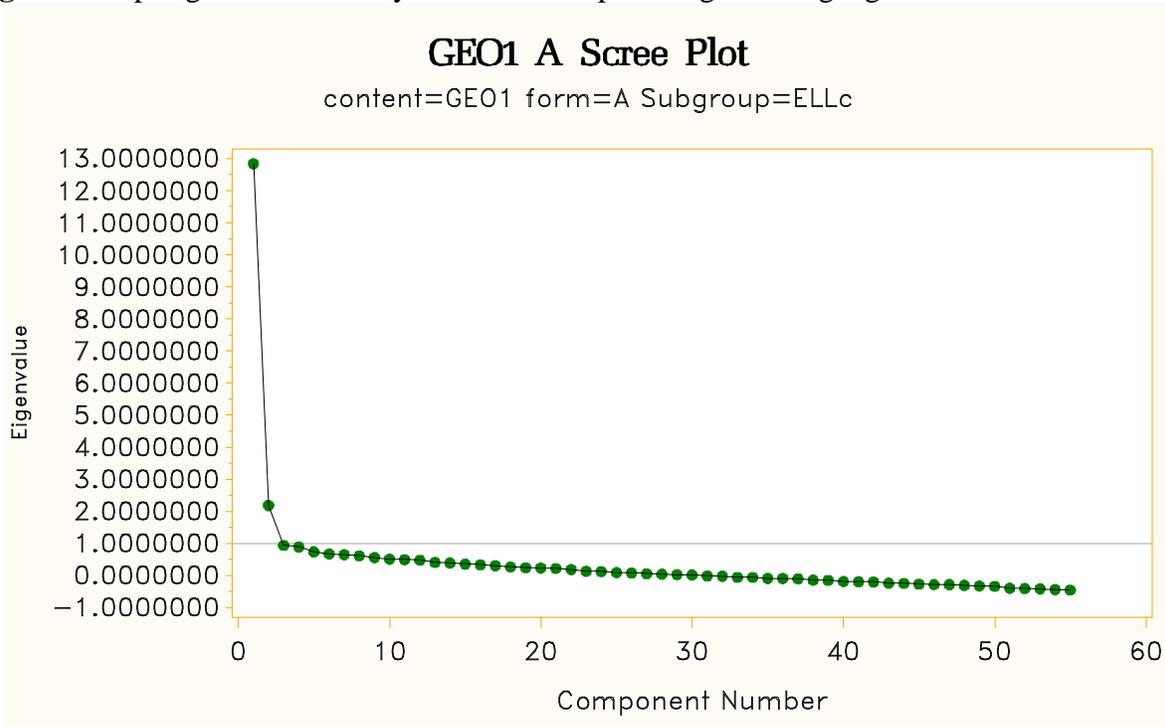
Figure 97. Spring 2013 Geometry Form B scree plot: All
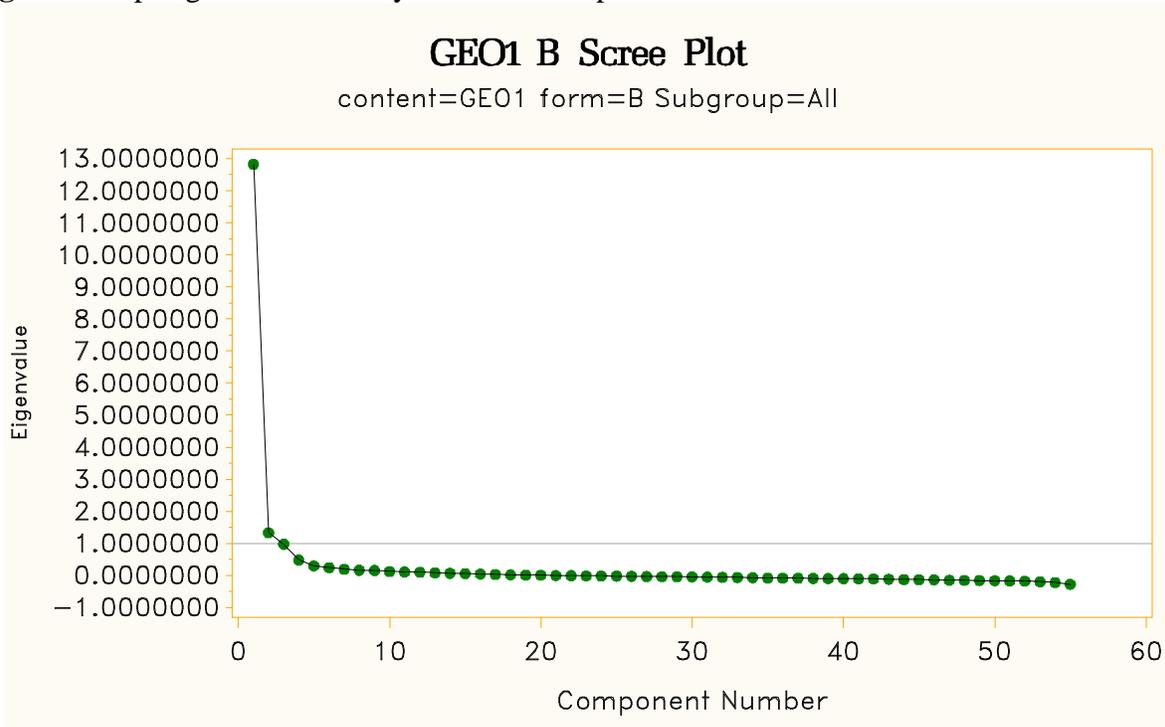
Figure 98. Spring 2013 Geometry Form B scree plot: Accommodated
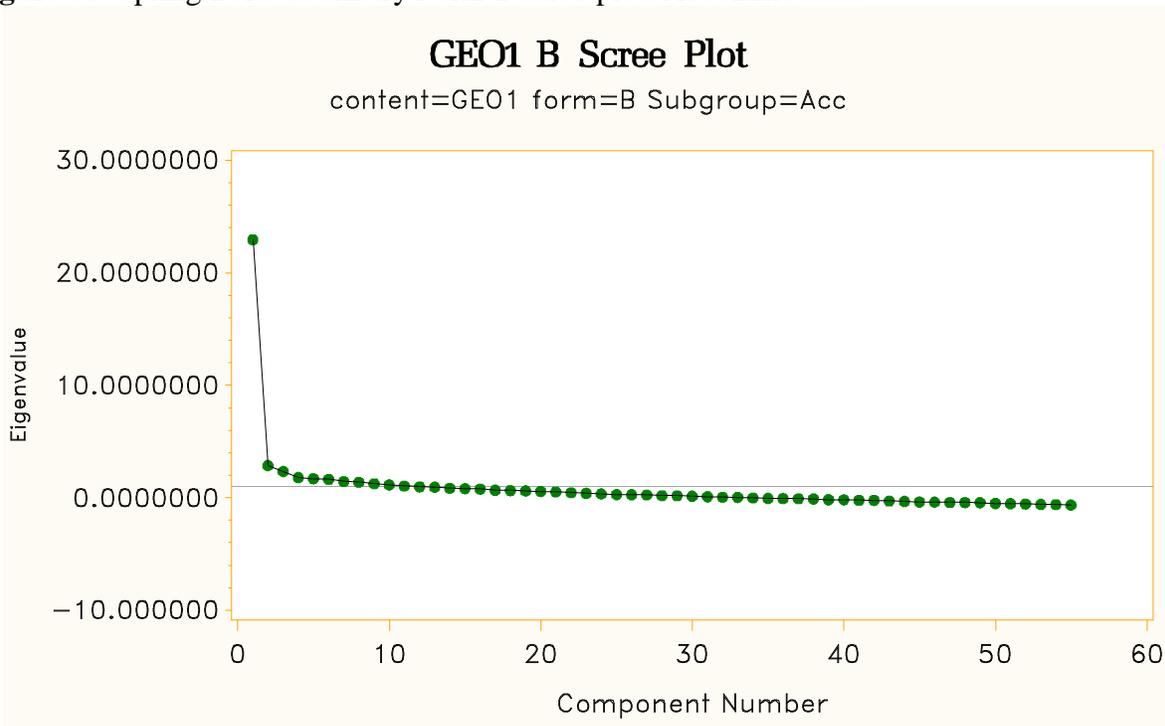
Figure 99. Spring 2013 Geometry Form B scree plot: English Language Learner
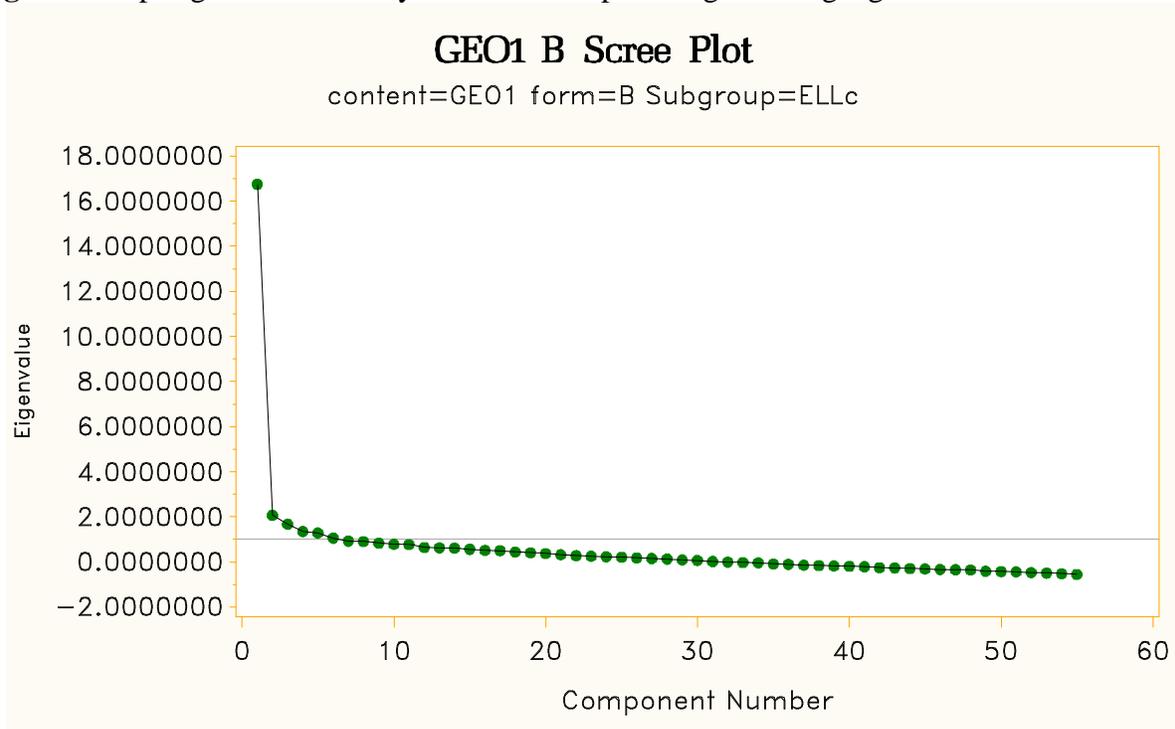
Figure 100. Spring 2013 Geometry Form B scree plot: Individualized Education Program

Figure 101. Spring 2013 U.S. History Form A scree plot: All

Figure 102. Spring 2013 U.S. History Form A scree plot: Accommodated
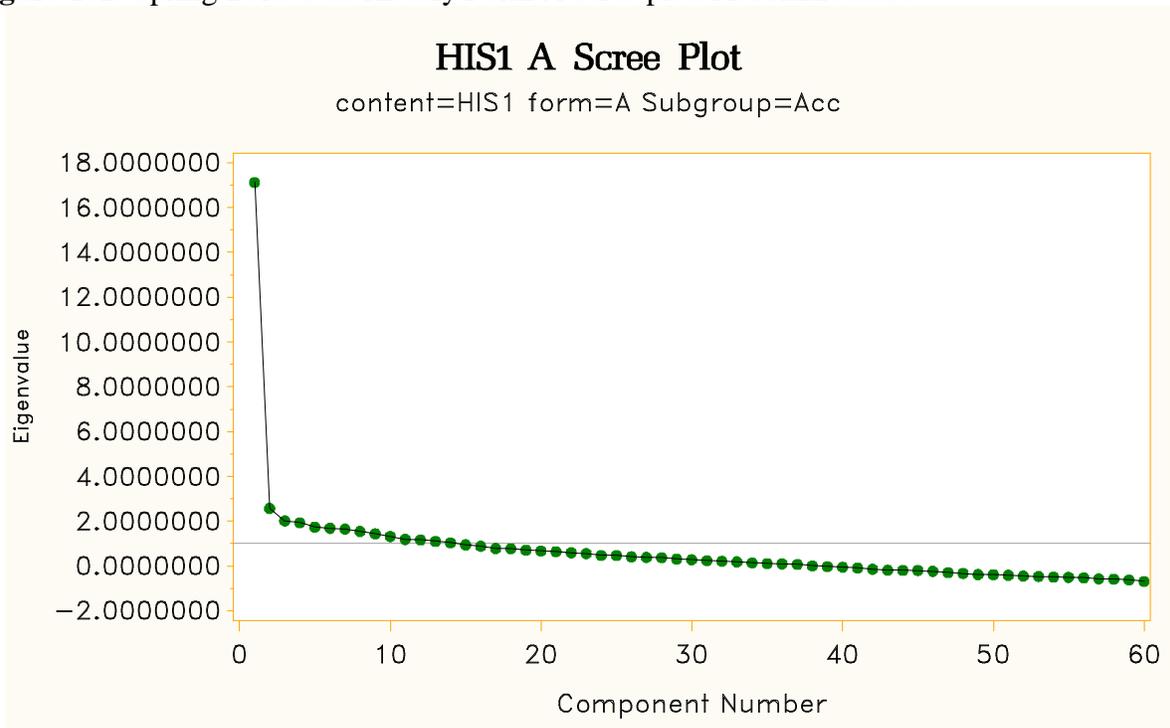
Figure 103. Spring 2013 U.S. History Form A scree plot: English Language Learner
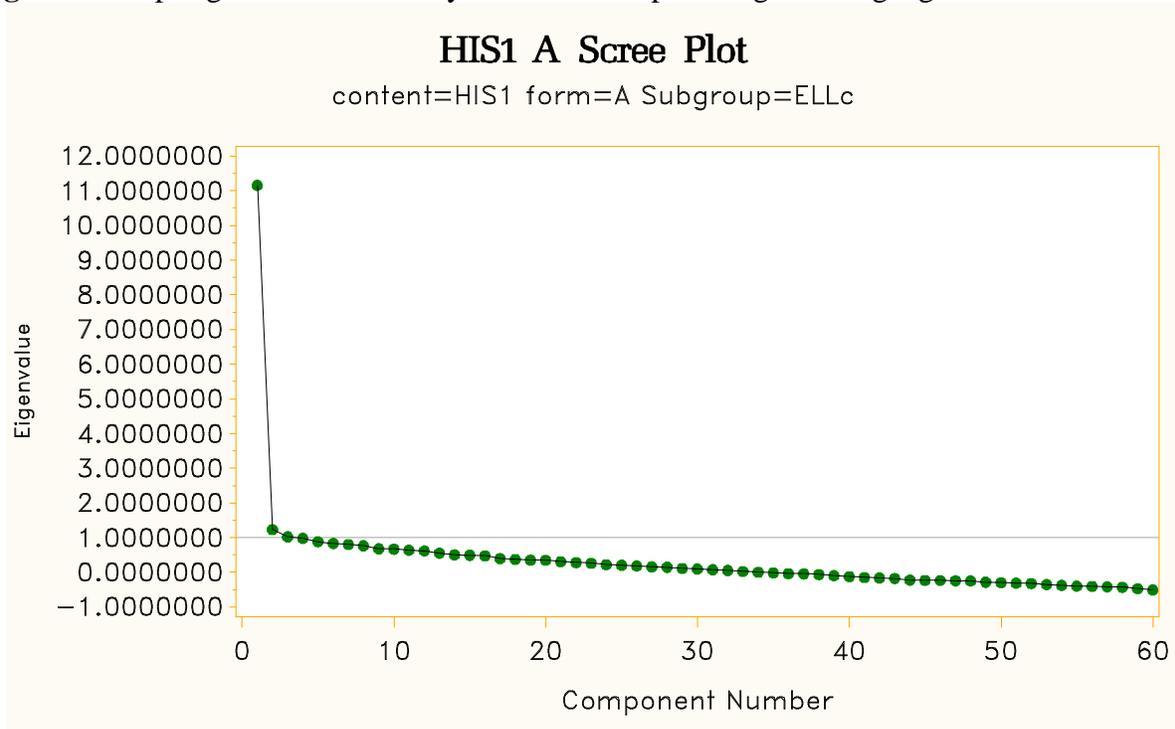
Figure 104. Spring 2013 U.S. History Form A scree plot: Individualized Education Program
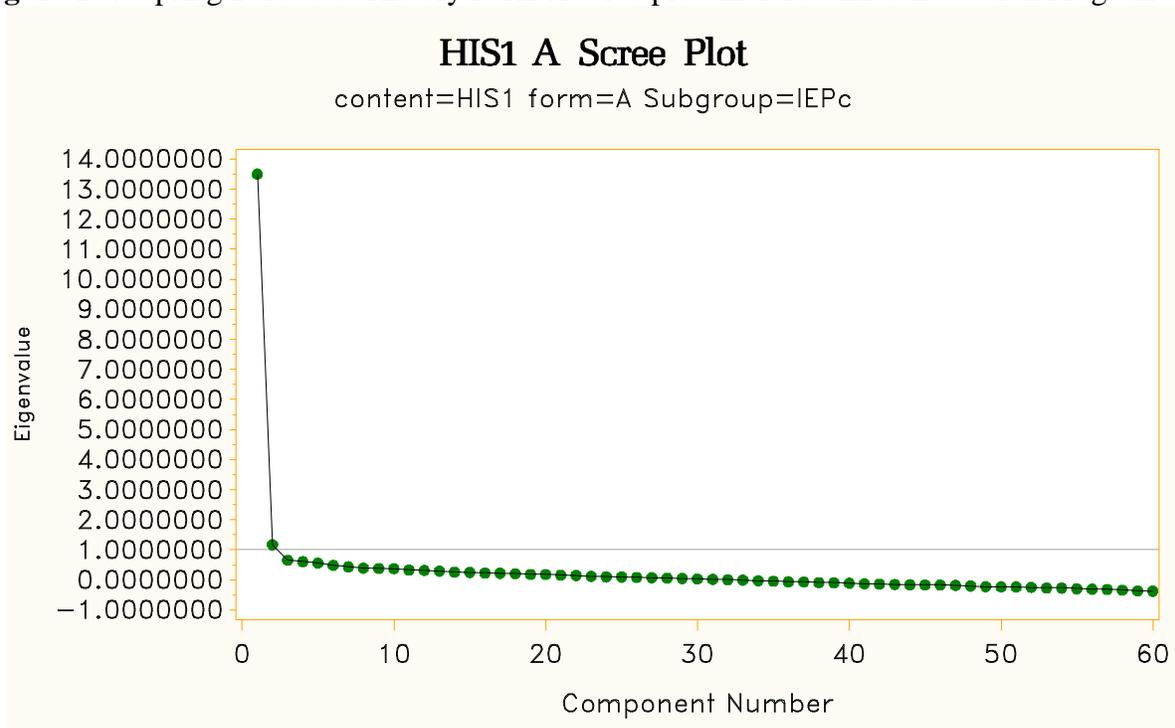
Figure 105. Spring 2013 U.S. History Form B scree plot: All
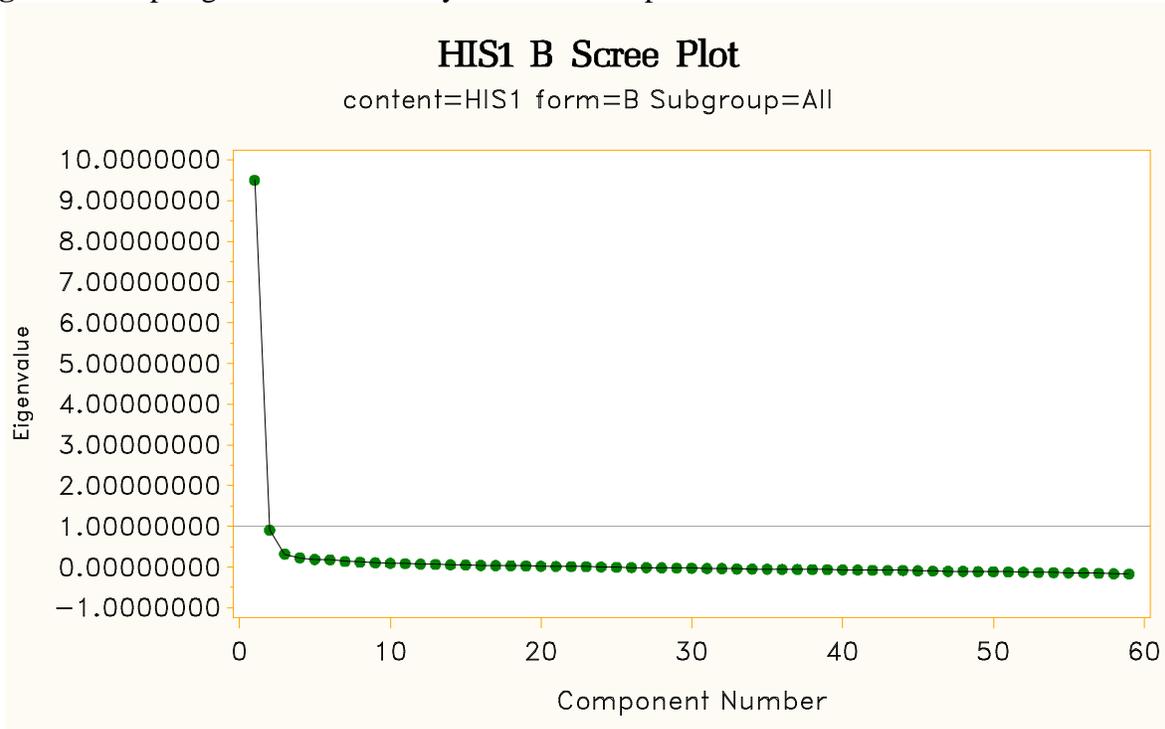
Figure 106. Spring 2013 U.S. History Form B scree plot: Accommodated
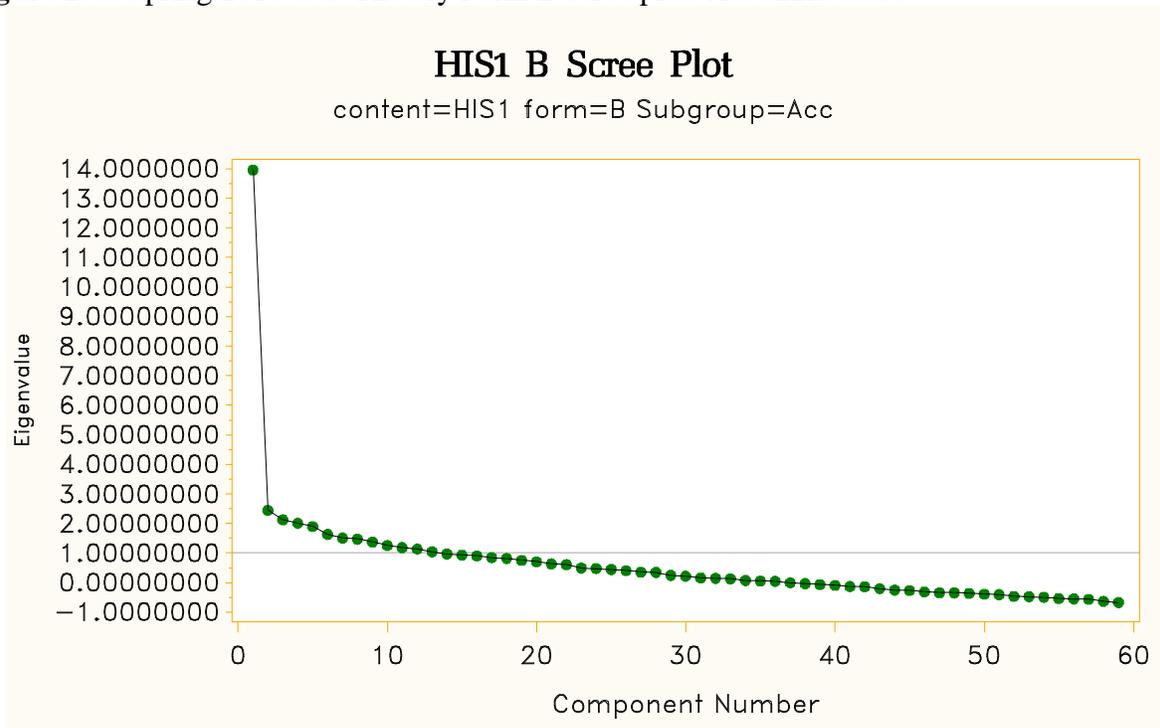
Figure 107. Spring 2013 U.S. History Form B scree plot: English Language Learner

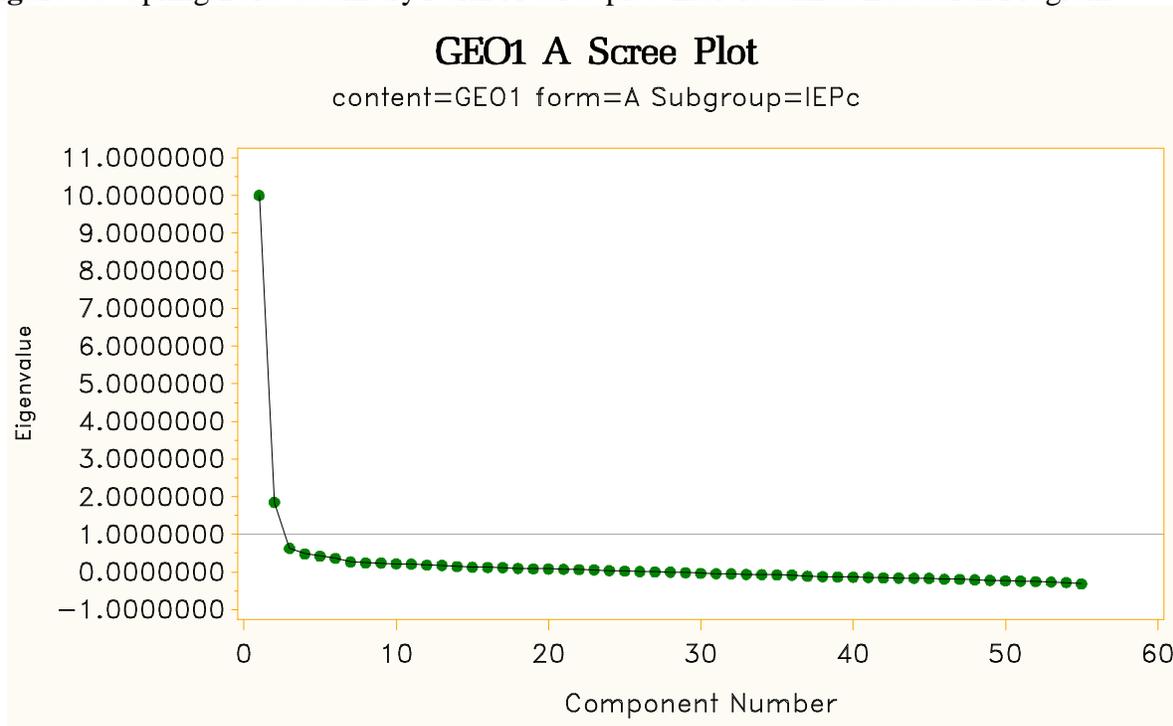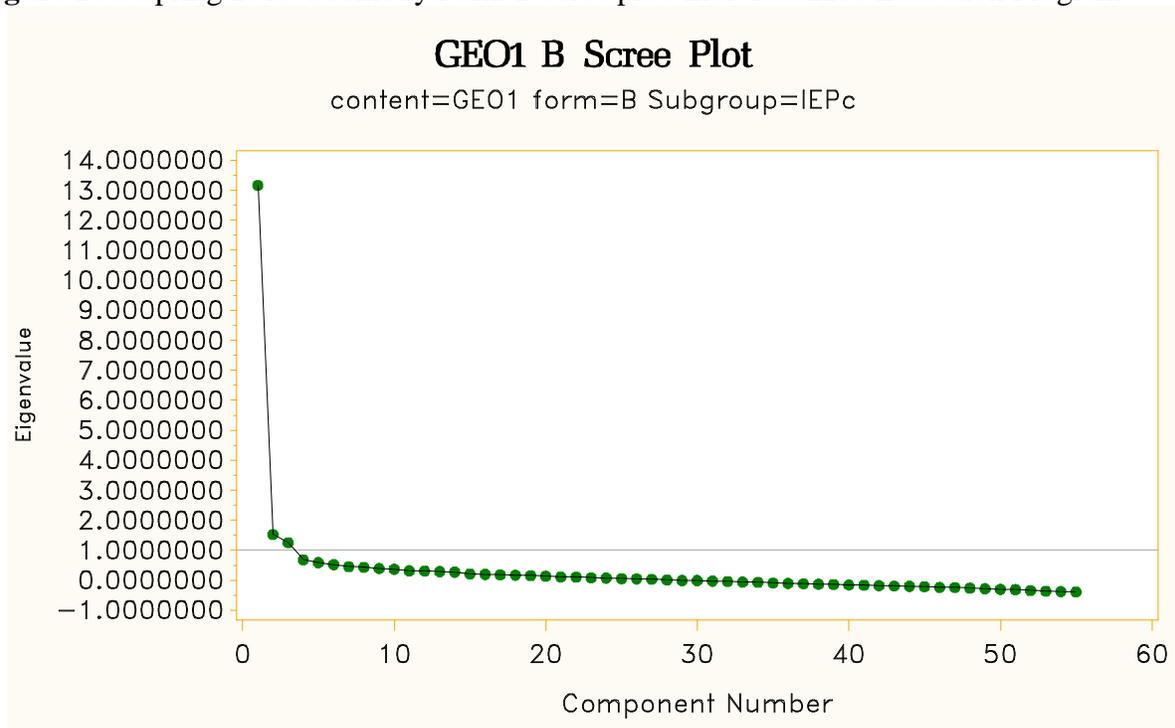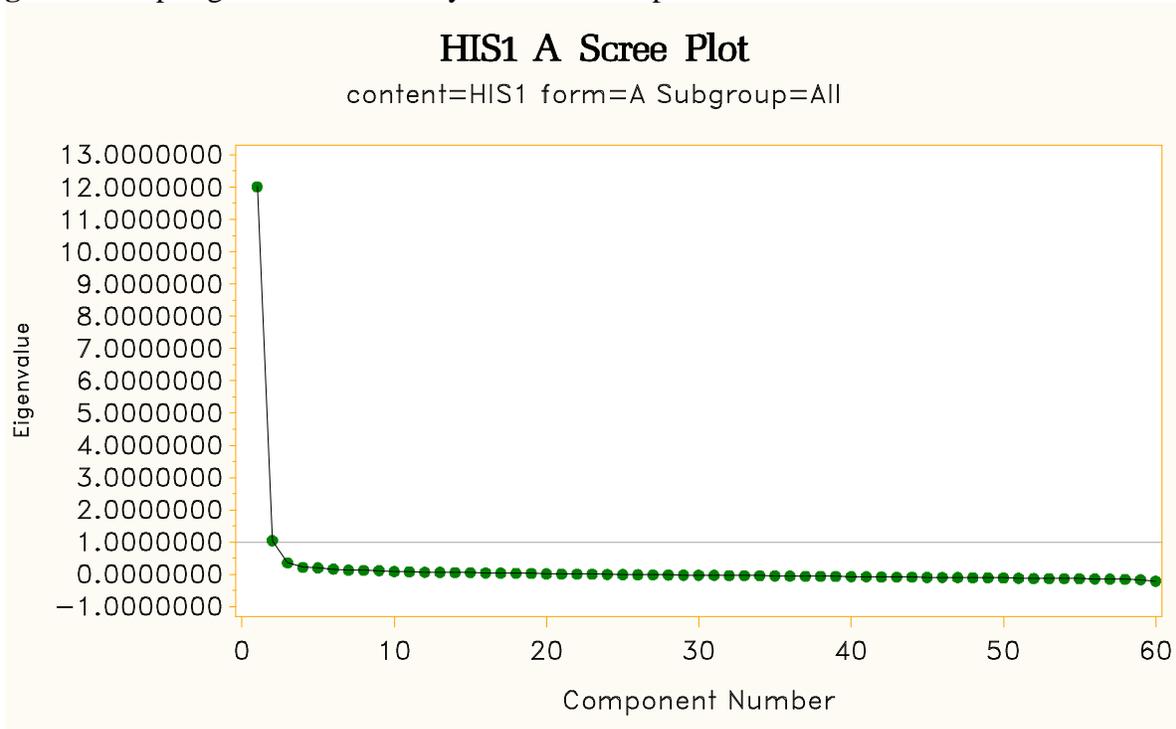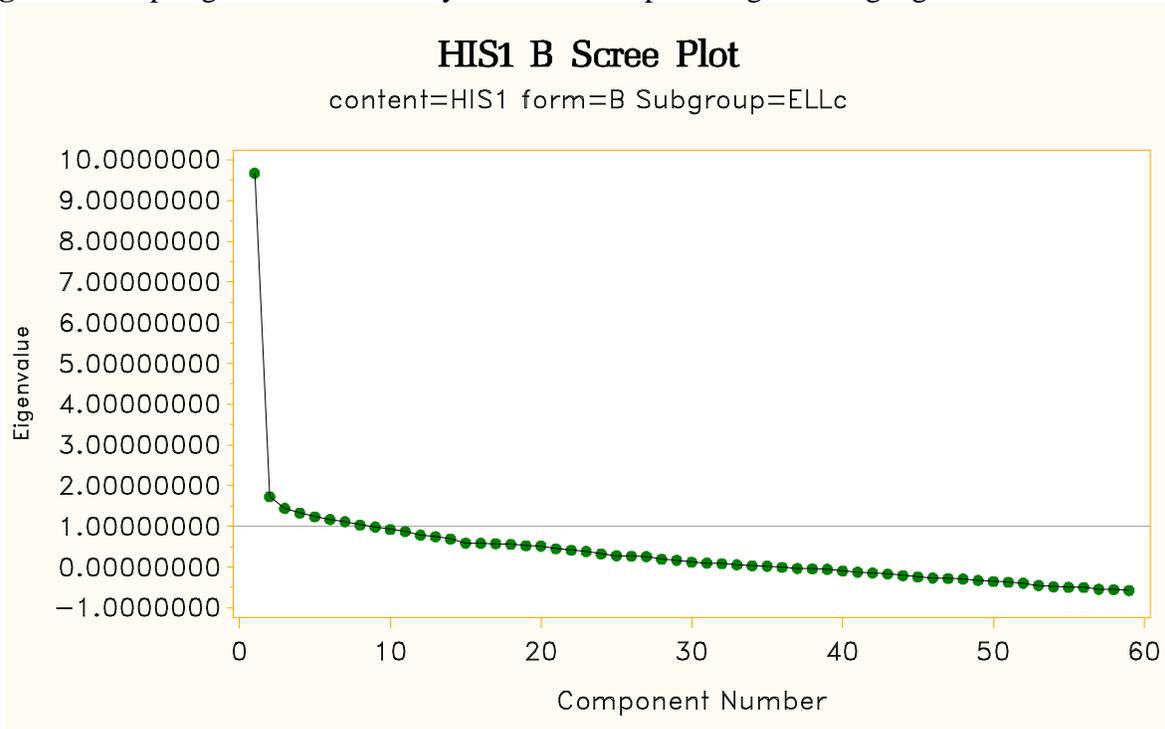**Figure 108.** Spring 2013 U.S. History Form B scree plot: Individualized Education Program
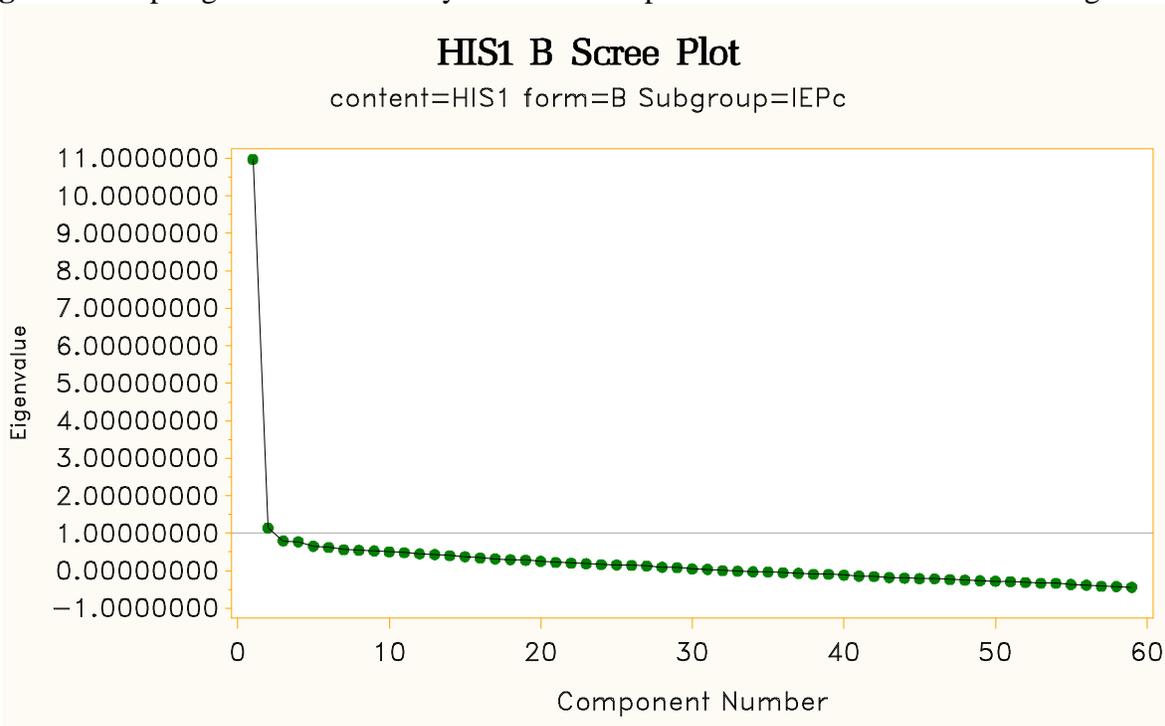
Appendices

Appendix A
Standards, Objectives/Skills, and Processes Assessed by Subject

Table A1. OCCT Test Blueprint and Actual Item Counts: Algebra I

OKC³ Standard and Objective	Ideal Number of Items for Alignment to *OKC³**	Actual Number of Items on 2013 Test Form A	Actual Number of Items on 2013 Test Form B
Number Sense and Algebraic Operations	**15**		
Equations and Formulas (1.1)	6	6	6
Expressions (1.2)	9	9	9
Relations and Functions	**31**		
Relations and Functions (2.1)	6	6	6
Linear Equations and Graphs (2.2)	15	15	15
Linear Inequalities and Graphs (2.3)	6	6	6
Systems of Equations (2.4)	4	4	4
Data Analysis, Probability, and Statistics	**9**		
Data Analysis (3.1)	5	5	5
Line of Best Fit (3.3)	4	4	4
Total Test	**55**	**55**	**55**

*Suppressed item in this reporting category.

Table A2. OCCT Test Blueprint and Actual Item Counts: Algebra II

OKC^3 Standard and Objective	Ideal Number of Items for Alignment to OKC^3	Actual Number of Items on 2013 Test Form A	Actual Number of Items on 2013 Test Form B
Number Sense and Algebraic Operations	**15**		
Rational Exponents (1.1)	5-6	6	5
Polynomial and Rational Expressions (1.2)	5-6	5	6
Complex Numbers (1.3)	4	4	4
Relations and Functions	**31**		
Functions and Function Notation (2.1)	5	5	5
Systems of Equations (2.2)	5	5	5
Quadratic Equations and Functions (2.3)	5	5	5
Conic Sections (2.4)	4	4	4
Exponential and Logarithmic Functions (2.5)	4	4	4
Polynomial Equations and Functions (2.6)	4	4	4
Rational Equations and Functions (2.7)	4	4	4
Data Analysis, Probability, and Statistics	**9**		
Analysis of Collected Data (3.1)	5	5	5
Arithmetic and Geometric Sequences (3.2)	4	4	4
Total Test	**55**	**55**	**55**

Table A3. OCCT Test Blueprint and Actual Item Counts: Geometry

OKC³ Standard and Objective	Ideal Number of Items for Alignment to *OKC³*	Actual Number of Items on 2013 Test Form A	Actual Number of Items on 2013 Test Form B
Logical Reasoning	**6**		
Inductive and Deductive Reasoning (1.1)	4	4	4
Conditional Statements (1.2)	2	2	2
Properties of 2-Dimensional Figures	**20**		
Line and Angle Relationships (2.1)	4	4	4
Polygons and Other Plane Figures (2.2)	4	4	4
Similarity (2.3)	4	4	4
Congruence (2.4)	4	4	4
Circles (2.5)	4	4	4
Triangles and Trigonometric Ratios	**12**		
Pythagorean Theorem (3.1)	4	4	4
Right Triangle Relationships (3.2)	4	4	4
Trigonometric Functions (3.3)	4	4	4
Properties of 3-Dimensional Figures	**10**		
Polyhedra and Other Solids (4.1)	6	6	6
Similarity (4.2)	2	2	2
Models and Perspectives (4.3)	2	2	2
Coordinate Geometry	**7**		
Properties of Points, Segments, and Lines (5.1)	4	4	4
Properties of Figures (5.2)	3	3	3
Total Test	**55**	**55**	**55**

Table A4. OCCT Test Blueprint and Actual Item Counts: English II

OKC³ Standard and Objective	Ideal Number of Items for Alignment to *OKC³**	Actual Number of Items on 2013 Test Form A	Actual Number of Items on 2013 Test Form B
Reading/Literature			
Vocabulary	**6-8**	**6**	**6**
Comprehension	**16-20**		
Literal Understanding (2.1)	4-5	4	4
Inferences and Interpretation (2.2)	4-5	5	4
Summary and Generalization (2.3)	4-5	4	5
Analysis and Examination (2.4)	4-5	5	4
Literature	**17-20**		
Literary Genres (3.1)	4-5	4	4
Literary Elements (3.2)	5-6	6	6
Figurative Language and Sound Devices (3.3)	4-5	4	5
Literary Works (3.4)	4-5	4	4
Research and Information	**6**	**6**	**5**
Writing/Grammar/Usage and Mechanics			
Writing (1.0, 2.0)	**1 (6 points)**		
Writing Prompt	1	1	1
Grammar/Usage and Mechanics	**12**		
Standard English Usage (3.1)	4	4	4
Mechanics and Spelling (3.2)	4	4	4
Sentence Structure (3.3)	4	4	4*
Total Test	**61 (66 Points)**	**61**	**60**

*Suppressed Item

Table A5. OCCT Test Blueprint and Actual Item Counts: English III

OKC³ Standard and Objective	Ideal Number of Items for Alignment to *OKC³**	Actual Number of Items on 2013 Test Form A	Actual Number of Items on 2013 Test Form B
Reading/Literature			
Vocabulary (1.0)	**6-8**	**7**	**7**
Comprehension	**16-20**		
Literal Understanding (2.1)	4-5	5	5
Inferences and Interpretation (2.2)	4-5	3	3
Summary and Generalization (2.3)	4-5	6	4
Analysis and Examination (2.4)	4-5	4	4
Literature	**17-20**		
Literary Genres (3.1)	4-5	3	3
Literary Elements (3.2)	5-6	7	8
Figurative Language and Sound Devices (3.3)	4-5	3	4
Literary Works (3.4)	4-5	4	4
Research and Informatiom	**6-7**	**6**	**6**
Writing/Grammar/Usage and Mechanics			
Writing (1.0, 2.0)	**1 (10 points)**		
Writing Prompt	1	1	1
Grammar/Usage and Mechanics	**12**		
Standard English Usage (3.1)	4-5	3*	2*
Mechanics and Spelling (3.2)	0-2	1	1
Sentence Structure (3.3)	4-5	5	2
Total Test	**61 (66 Points)**	**61**	**60**

*Suppressed Item

Table A6. OCCT Test Blueprint and Actual Item Counts: Biology I

OKC³ Standard and Objective	Ideal Number of Items for Alignment to *OKC³**	Actual Number of Items on 2013 Test Form A	Actual Number of Items on 2013 Test Form B
Process Standards			
Observe and Measure	**6**		
Qualitative/quantitative observations and changes (P1.1)	4	4	4
Use appropriate tools & (P1.2) and	0	0	0
Use appropriate SI units (P1.3)	2	2	2
Classify	**7-8**		
Use observable properties to classify (P2.1)	4	4	3
Identify properties of a classification system (P2.2)	3-4	4	4
Experimental Design	**16-19**		
Evaluate the design of investigations (P3.1)	4-5	5	7
Hazards/practice safety (P3.2) & Identify a testable hypothesis in a biology investigation (P3.4)	5-6	4	2
Use mathematics to show relationships (P3.3)	4-6	5	5
Identify potential hazards and practice safety procedures in all science activities (P3.5)	3	3	3
Interpret and Communicate	**20-24**		
Select predictions based on observed patterns of evidence (P4.1)	4-5	4	4
Interpret line, bar, trend, and circle graphs (P4.3)	4-5	4	4
Accept or reject a hypothesis (P4.4)	4-5	5	4
Make logical conclusions based on experimental data (4.5)	4-5	4	6
Identify an appropriate graph or chart (4.8)	4	4	4
Translate quantitative information expressed in words into visual form (4.8a)	0	0	0
Translate information expressed visually or mathematically (4.8b)	0	0	0
Model	**8**		
Interpret a model which explains a given set of observations (5.1)	4	4	4
Select predictions based on models, using mathematics when appropriate (5.2)	4	4	4
Total Test	**60**	**60**	**60**

Table A6. OCCT Test Blueprint and Actual Item Counts: Biology I (*continued*)

OKC³ Standard and Objective	Ideal Number of Items for Alignment to *OKC³**	Actual Number of Items on 2013 Test Form A	Actual Number of Items on 2013 Test Form B
Content Standards			
The Cell	**12-15**		
Cells structures and functions (C1.1)	4-6	4	4
Differentiation of cells (C1.2)	4-6	4	4
Specialized cells (C1.3)	4	4	4
The Molecular Basis of Heredity	**12-15**		
DNA structure and function in heredity (C2.1)	6-8	6	5
Sorting and recombination of genes (C2.2)	6-7	6	7
Experimental Design	**16-19**		
Evaluate the design of investigations (P3.1)	4-5	5	7
Hazards/practice safety (P3.2) & Identify a testable hypothesis in a biology investigation (P3.4)	5-6	4	2
Use mathematics to show relationships (P3.3)	4-6	5	5
Identify potential hazards and practice safety procedures in all science activities (P3.5)	3	3	3
Biological Diversity	**12-15**		
Variation among organisms (C3.1)	4-6	4	4
Natural selection and biological adaptations (C3.2)	4-6	4	5
Behavior patterns can be used to ensure reproductive success (C3.3)	4	4	4
The Interdependence of Organisms	**8-10**		
Organisms both cooperate and compete (C4.1)	4-6	4	4
Population dynamics (C4.2)	4-6	4	4
Matter/Energy/Organization in Living Systems	**12-15**		
Complexity and organization used for survival (C5.1)	4	5	4
Matter and energy flow in living and nonliving systems (C5.2)	4	4	4
Earth cycles including abiotic and biotic factors (C5.3)	4	4	4
Total Test	**57**	**57**	**57**

** Items from the Safety Objective (P3.5) are not dual aligned to a content standard

Table A7. OCCT Test Blueprint and Actual Item Counts: U.S. History

OKC³ Standard and Objective	Ideal Number of Items for Alignment to *OKC³**	Actual Number of Items on 2013 Test Form A	Actual Number of Items on 2013 Test Form B
Post-Reconstruction to the Progressive Era, 1878-1900	**8**		
Post Reconstruction Amendments (1.1)	2-4	0	2
Immigration, Westward Movement, and Native American Experiences (1.2)	2-4	3	2
Impact of Industrialization on Society, Economics, and Politics (1.3)	2-4	5	3
Expanding Role of the United States in International Affairs	**6**	**9**	**8**
Cycles of Economic Boom and Bust in the 1920s and 1930s	**8**		
Economic, Political, & Social Transformation Between the World Wars (3.1)	3-5	5	5
Economic Destabilization and the Great Depression/New Deal (3.2, 3.3)	3-5	6	6
Foreign and Domestic Policies during the Cold War, 1945-1975	**18**		
The Cold War - Foreign and Domestic (5.1, 5.2)	4-5	5	8
The Vietnam War Era (5.3)	4-5	5	5
The African American Civil Rights Movement (5.4)	4-6	5	4
Social Political Transformation (5.5)	4-5	6	3*
Total Test	**60**	**60**	**59**

Appendix B
A Statistical Investigation of Oklahoma Computer Disruptions

A Statistical Investigation of Oklahoma Computer Disruptions

Final Report

Prepared for: **Oklahoma Department of Education**
2500 N. Lincoln Blvd.
Oklahoma City, OK 73105-4599

Authors: **Matthew S. Swain**
Arthur A. Thacker
Bethany H. Bynum

Date: **August 6, 2013**

A Statistical Investigation of Oklahoma Computer Disruptions

Final Report

Prepared for: **Oklahoma Department of Education**
2500 N. Lincoln Blvd.
Oklahoma City, OK 73105-4599

Authors: **Matthew S. Swain**
Arthur A. Thacker
Bethany H. Bynum

Date: **August 6, 2013**

A Statistical Investigation of Oklahoma Computer Disruptions

Table of Contents

List of Tables

A Statistical Investigation of Oklahoma Computer Disruptions

Abstract

On April 29th and April 30th, 2013, students in Oklahoma experienced technological delays or "interruptions" while completing their statewide Math and Reading assessments. The potential impact of those disruptions on test scores is the focus of this investigation. Multiple analytical approaches were conducted to determine the presence and magnitude of any effects. Additionally, analyses were conducted on multiple "cohorts" of students to examine if the effects of disruption on these two days were disparate from typical delays or interruptions during testing.

The testing vendor, CTB McGraw/Hill (CTB), conducted their own investigation of the impact of the test interruptions. The Oklahoma Department of Education (ODE) requested that an independent investigation of the interruption also be conducted. The Human Resources Research Organization (HumRRO) was selected to conduct the independent investigation[1]. Data were provided by CTB (for the 2013 cohort of students) and ODE (test data from the prior vendor). CTB also shared their investigation methodology and results. This investigation verified CTB's initial data (e.g. percent of students impacted) but used independently derived methodology to investigate the impact of the interruptions and draw conclusions. HumRRO primarily compared students' predicted scores to their actual interrupted scores to investigate the presence and magnitude of the interruption effect.

HumRRO's analyses detected a small effect of the interruption. This effect was not consistent across grades and/or subjects, however. Students in some grade/subjects appear to have been disadvantaged by the interruption, while others performed better than expected. Even when statistically significant differences in mean scores were found within a grade/subject, the difference was not consistent across the distribution of students taking the test. Therefore, HumRRO does not recommend a mathematical adjustment of test scores.

Methodology

The primary goal of this investigation was to determine if the score a disrupted student received would have been different if the student had not been disrupted. That is, if two students who were very similar on all available variables that predicted 2013 test scores differed only in whether or not they were disrupted, any differences in those test scores could be more confidently stated to be an effect of disruption. Propensity score matching provides the ability to match students with similar distributions on a set of variables (Connelly, Sackett, & Waters, 2013). Each student in the disrupted group for each grade and subject test was matched with a student from the non-disrupted group. This matching was done using all available variables except for pertinent 2013 scores (i.e., the dependent variable). Multiple grades and subject tests were investigated leading to formation of four "cohorts."

Description of Four Cohorts

Students in grades 6 through 8 completing their Math and Reading tests and students in high school completing their end of instruction (EOI) exams were included in the investigation. The

---

[1] A similar study was carried out by Hill (2013) to investigate this same issue in Indiana.

OK testing window for the grades 3-8 tests runs from April 10 to April 30. Cohort A was defined first as all students in grades 6 through 8 who were interrupted during their Math and Reading tests. Because the interruptions on 4/29 and 4/30 were of particular interest, Cohort B was further refined from Cohort A to only include students who were interrupted on those two days. Tables 1 and 2 provide the frequencies of interruption by day and grade for Math and Reading tests, respectively. Clearly, interruptions on 4/29 and 4/30 were atypical.

Table 1. Number of Students Interrupted by Day and Grade on Math Test

Day	Grade			
	6	7	8	Total
4/10/2013	139	23	12	174
4/11/2013	192	96	40	328
4/12/2013	122	55	26	203
4/15/2013	97	92	14	203
4/16/2013	95	75	45	215
4/17/2013	275	158	43	476
4/18/2013	108	69	61	238
4/19/2013	45	38	41	124
4/22/2013	48	23	31	102
4/23/2013	59	27	154	240
4/24/2013	66	46	167	279
4/25/2013	35	24	43	102
4/26/2013	14	12	12	38
4/29/2013	**263**	**244**	**427**	**934**
4/30/2013	**490**	**344**	**676**	**1510**
5/1/2013	10	5	15	30
5/2/2013	9	12	8	29
5/3/2013	1	4	1	6
5/6/2013	0	0	3	3
5/7/2013	0	1	0	1
Total	**2068**	**1348**	**1819**	**5235**

A Statistical Investigation of Oklahoma Computer Disruptions

Table 2. Number of Students Interrupted by Day and Grade on Reading Test

Day	Grade			
	6	7	8	Total
4/10/2013	174	86	78	338
4/11/2013	161	160	265	586
4/12/2013	276	146	191	613
4/15/2013	142	338	115	595
4/16/2013	286	216	124	626
4/17/2013	177	77	122	376
4/18/2013	84	127	93	304
4/19/2013	27	30	62	119
4/22/2013	45	52	71	168
4/23/2013	109	95	169	373
4/24/2013	116	68	106	290
4/25/2013	95	19	92	206
4/26/2013	9	20	82	111
4/29/2013	**220**	**173**	**384**	**777**
4/30/2013	**357**	**269**	**264**	**890**
5/1/2013	52	10	11	73
5/2/2013	7	16	11	34
5/3/2013	1	3	1	5
5/6/2013	1	0	3	4
5/7/2013	0	1	11	12
Total	**2339**	**1906**	**2255**	**6500**

Interruptions also occurred during the EOI exams completed primarily by high school students. Using the available data, two additional cohorts were constructed to explore the presence of disruption effects on EOI exams. These cohorts were chosen primarily based on what data were available. More specifically, the goal was to choose exams that had sufficient sample size to detect disruption effects, if any. Cohort C focused on students interrupted during their Algebra I test whereas Cohort D focused on students interrupted during their English II exam.

The report that follows will focus first on Cohorts A and B followed by Cohorts C and D.

Computer Disruption Data Overview

Interruption data provided to HumRRO included the test subject and grade level that each student was completing at the time of interruption for Cohorts A and B. Archival data regarding the student's prior year performance, demographic information, as well as current year performance on other exams was also provided. A total of 130,429 valid, unique IDs with complete Math test scores were recorded and 130,373 valid, unique IDs with complete Reading test scores were also recorded. After integrating the interruption data provided to HumRRO, it appears that 5,235 students were interrupted during their Math test *across all days* out of 130,429 students. For the Reading test, 6,500 students were interrupted *across all days* out of 130,373 students. These students formed Cohort A. An interruption rate was calculated to

examine the prevalence of interruption. The numbers of disrupted and non-disrupted students before matching, separated by grade level, are provided in Table 3.

Table 3. Interruption Data by Subject and Grade

	Math			Reading		
Grade	All Students	Disrupted	Interruption Rate	All Students	Disrupted	Interruption Rate
6	43,976	2,068	4.6%	43,957	2,339	5.3%
7	43,728	1,348	3.0%	43,716	1,906	4.3%
8	42,725	1,819	4.2%	42,700	2,255	5.2%
Total	**130,429**	**5,235**	**4.0%**	**130,373**	**6,500**	**5.0%**

The numbers of students in Table 3 are fairly consistent with those provided by CTB, for both non-disrupted and disrupted students. For those students who were disrupted, some were interrupted or delayed during their testing session multiple times. Although most students who were disrupted only experienced one delay, some students experienced many delays. Table 2 provides the frequency of disruption for each student in the disrupted group by test.

Table 4. Frequency of Interruption by Test

Interruption Number	Math	Reading
1	4,455	5,446
2	423	681
3	186	171
4	80	88
5	42	52
6	25	26
7	13	19
8	8	7
9	1	3
10	2	5
11	0	1
12	0	0
13	0	1

Propensity Score Matching

In an attempt to isolate the potential effect of disruption, propensity score matching was employed to select a sample of students from the non-disrupted group that closely resembles the disrupted group on all variables available that relate to 2013 scores. In one sense, propensity score matching attempts to "control for" these variables. With these groups established, differences in 2013 scores are more likely to be due to an effect of disruption.

The following variables were used as matching variables for Cohorts A and B:

- 2012 Math and Reading scaled scores
- Other subject 2013 scale scores (Reading score used if interrupted on 2013 Math test)
- Gender (0 = Female, 1 = Male)
- Ethnicity (Dummy coded)
- Limited English proficiency
- Student-level free/reduced lunch
- School-level proportion of free/reduced lunch students
- School-level achievement (Average of 2012 student-level Reading and Math scores)

Differences between disrupted and non-disrupted students on these matching variables were examined prior to matching. These differences were calculated using Cohen's d with "rule of thumb" cutoff values suggested by Harder, Stuart, & Anthony (2010). Across grades and subjects, the disrupted students seemed to be disproportionately African American and from schools with more free and reduced lunch students and lower achievement means. Given the relationships between these variables and the test scores on the disrupted test, it is clear why propensity score matching is a necessary method.

Next, using logistic regression, we regressed group membership (disrupted or not disrupted) on to the matching variables. Generally, the pseudo R-square values of the logistic regression were small, ranging from .0042 to .0242. Overall, the small values suggest that the combination of prior year student achievement, demographics, SES, school-level achievement or school-level SES had little relationship to the likelihood that a student experienced disruptions.

To match the two samples, the predicted disruption probabilities from the logistic regression analyses were saved. The predicted probabilities represent the probability that a student was in the disrupted sample. We used the *nearest neighbor method* to match the two samples. That is, the predicted disruption probability for each student in the disrupted sample was matched to the student with the closest predicted disruption probability in the non-disrupted sample. The sampling was done without replacement so that each student in the disrupted sample was matched with a unique student in the non-disrupted sample. The average difference between a disrupted student's predicted probability and the matched non-disrupted student's predicted probability was .000003 for Reading and .000003 for Math. The largest difference was .0016. Differences no larger than .20 have been shown to reduce bias and produce accurate group difference estimates (Austin, 2009; Connelly, Sackett, & Waters, 2013). The results suggest that everyone in the disrupted sample was matched with a student in the non-disrupted sample with a very similar predicted disruption probability. To further evaluate the closeness of the matched sample, we examined the mean difference of the matching variables. The average Cohen's d between the two samples was .002 and ranged from -.060 to .075 after matching. All of the differences greater than .10 prior to matching were reduced. Near zero effect sizes suggest that

our samples have been effectively balanced on prior year achievement, school-level achievement, gender, ethnicity, and school-level free-or-reduced percentage. A summary of the mean, standard deviations, and effect sizes before and after matching can be found in the Appendix.

2013 Score Differences

Using the matched samples, we examined whether students' scores for tests impacted by computer interruptions differed from students' scores that were not impacted by computer disruptions. By matching the samples on variables that are likely to predict student scores, any difference between the two samples can be better attributed to the computer disruptions. We used a number of analyses to examine differences in scores. Scale scores for Math and Reading for the matched groups were examined four ways:

1. Mean differences on 2013 scores
2. R^2 change when combing groups and adding dichotomous disruption variable
3. R^2 differences when predicting 2013 scores separately
4. Applying Non-Disrupted regression equation from step 3 onto Disrupted group as well as 5th, 10th, 90th, and 95th percentile cuts

Cohort A Analyses

Differences in average 2013 test scores

Because these two matched groups are alike on variables that are typically predictive of test scores, if testing disruptions had no overall impact, then the averages of the 2013 test scores for the two groups would be expected to be nearly identical. Differences in the average 2013 test scores would be evidence that computer disruption impacted test performance. We examined mean differences between the two samples using a t-test and Cohen's *d* effect size. Table 5 presents the results by grade and subject for Cohort A.

Table 5. Descriptive and Inferential Statistics of 2013 Test Scores After Matching – Cohort A

	N	Disrupted		Non-Disrupted		t	d
		Mean	SD	Mean	SD		
Grade 6 Reading	2,002	733.7	77.87	731.9	73.84	-0.75	0.024
Grade 7 Reading	1,625	724.6	66.99	726.4	68.69	0.76	-0.027
Grade 8 Reading	1,900	757.8	84.37	758.6	78.62	0.32	-0.010
Grade 6 Math	1,710	744.9	77.77	743.6	75.24	-0.49	0.017
Grade 7 Math	1,085	726.5	84.46	730.8	80.98	1.22	-0.052
Grade 8 Math	1,533	727.8	79.72	732.0	86.09	1.42	-0.051

Note. No *p*-values below .05.

For Reading, the effect sizes ranged from -.027 to .024, suggesting no overall directional effect. For Math, the effect sizes ranged from -.052 to .017. All effect sizes indicate very small mean differences between scale scores of the matched samples.

A Statistical Investigation of Oklahoma Computer Disruptions

Examining the predictability of 2013 test scores

First, to examine predictability of 2013 test scores, we examined the incremental variance that disruption accounted for when added to the prediction of the 2013 test scores beyond other known indicators of performance (including prior year achievement, ethnicity, gender, school-level achievement, and school-level percentage of free-or-reduced lunch). If the inclusion of disruption in the multiple regression models adds to the estimation of 2013 scores, then this would supply evidence that disruption impacted 2013 test scores. For this model, the two groups were included in the regression models together.

Table 6 reports the R-square values for each model and the R-square change between the two models. With the inclusion of disruption in the model, the R-square value changed very little or not at all, indicating that disruption did not add to the prediction of 2013 scores.

Table 6. Combined Regression Equation Adding Disruption – Cohort A

	N	Covariates Only R^2	Covariates + Disrupted R^2	ΔR^2
Grade 6 Reading	4,004	0.5759	0.5760	<0.001
Grade 7 Reading	3,250	0.5331	0.5339	0.001
Grade 8 Reading	3,800	0.5105	0.5105	<0.001
Grade 6 Math	3,420	0.5788	0.5791	<0.001
Grade 7 Math	2,170	0.5550	0.5552	<0.001
Grade 8 Math	3,066	0.5162	0.5167	0.001

Next, using all of the data we have available to create equations that predict students' 2013 test scores, we can statistically determine whether disrupted students scored differently than expected. Prediction equations were estimated for the disrupted students and separately for the matched sample of non-disrupted students. In addition to prediction equations, this technique gives us multiple regression coefficients (R-square) that can be interpreted like a correlation coefficient to tell us how well 2013 test scores can actually be predicted from our available data. If students' performance was affected, the strength of the prediction for the disrupted students should be less than the non-disrupted students as shown by lower multiple regression coefficients. A lower coefficient means that students' performance in the disrupted group was not as predictable as students' performance in the non-disrupted group. This would supply another piece of evidence about the impact of the computer disruptions. Table 7 presents the multiple regression coefficients (R-square) for the disrupted and non-disrupted groups. Overall, 2013 test scores were well predicted for both samples, with 52% to 59% of the variance accounted for by the predictor variables for Reading, and 55% to 59% for Math. Generally, there were slightly higher R-square values for the non-disrupted group, although the difference in variance accounted for was practically small ranging from 0.9% to 6%.

Table 7. Separate Regression Equations for Disrupted and Non-Disrupted Groups – Cohort A

	N	Disrupted R^2	Non-Disrupted R^2	ΔR^2
Grade 6 Reading	2,002	0.574	0.589	-0.015
Grade 7 Reading	1,625	0.521	0.551	-0.030
Grade 8 Reading	1,900	0.533	0.494	0.039
Grade 6 Math	1,710	0.578	0.587	-0.009
Grade 7 Math	1,085	0.548	0.577	-0.029
Grade 8 Math	1,533	0.488	0.547	-0.059

Examining distributions of predicted student scores

The prediction equations for the non-disrupted students give us a statistical statement about what to expect normally for students testing under non-disrupted conditions. We know that the prediction is not perfect, but given the high R-square values we can use the prediction equation to calculate how disrupted students might have scored had they not been disrupted. For each disrupted student, we computed their predicted score using the regression equation computed for the non-disrupted students. Next, we took the difference between the observed score and predicted score, where positive values indicate higher observed scores than predicted and negative values indicate higher predicted scores than observed--imagine a distribution of observed minus predicted score differences. "*Large numbers*" of students with "*notable differences*" between obtained and predicted scores provides another piece of evidence about the impact of the computer disruptions. Table 8 presents the distribution of observed and predicted scores for the disrupted and non-disrupted sample for Reading and Math. The difference between the observed and predicted score is also reported.

Finally, we assessed "*large number*" and a "*notable difference*" between obtained and predicted scores for interrupted students by comparing the difference in observed and predicted scores for the non-disrupted group to the difference for the disrupted group. Since our non-disrupted group of students represents what would be expected under normal testing conditions, on average, the difference between the observed and predicted score can be considered a baseline. That is, the average difference for non-disrupted students is what is expected to naturally occur based on the fact that our prediction, inherently, can never be perfect.

We evaluated differences in two ways. First, we examined the difference in the standard deviation of differences for the disrupted and non-disrupted group using an *F*-test. This provides information on whether the spread of differences in predicted and observed scores is statistically different for the two groups (i.e., students are more affected, on average, in the disrupted group than the non-disrupted group). Results suggest that the standard deviation of the difference scores is larger in the disrupted group for grade 6 Reading and grades 6 and 7 Math (see bolded SDs in Table 8).

Next, we took the difference in observed and predicted scores at the 5th, 10th, 90th and 95th percentile for the non-disrupted group and determined the number of students in the disrupted group that were at or below the same cut point for the 5th and 10th percentile and those that were at or above the cut point for the 90th and 95th percentile. If there are a large number of students below the 5th and 10th percentile of the non-disrupted group than what would be expected (5% and 10%, respectively) then more students in the disrupted group scored lower than expected. On the other hand, if there are a large number of students above the 90th and

A Statistical Investigation of Oklahoma Computer Disruptions

95th percentile of the non-disrupted group (10% and 5%, respectively) then more students in the disrupted group scored higher than expected. Either case would provide evidence that the computer disruption had an impact on scores. Table 9 presents the percent of students in the disrupted group with score differences less than the 5th and 10th percentile and higher than the 90th and 95th percentile of the non-disrupted group.

Table 8. Distribution of the Difference between Predicted and Observed Scores for Non-disrupted and Disrupted Groups – Cohort A

		Disrupted							Non-Disrupted					
	N	Mean	SD	5th	10th	90th	95th	N	Mean	SD	5th	10th	90th	95th
Grade 6 Reading														
Difference	2002	1.81	**52.02**	-73.94	-55.14	59.77	84.53	2002	0.00	47.36	-71.02	-52.14	56.69	78.04
Predicted Score	2002	731.94	**61.38**	634.72	660.52	805.18	826.03	2002	731.95	56.65	642.62	665.18	801.51	821.08
Observed Score	2002	733.75	**77.87**	613	639	833	860	2002	731.95	73.84	613	639	813	860
Grade 7 Reading														
Difference	1625	-3.87	46.91	-77.33	-58.60	49.90	67.02	1625	-0.07	46.01	-63.31	-48.98	50.19	74.55
Predicted Score	1625	728.44	51.25	645.67	667.32	788.63	806.22	1625	726.45	51.02	644.46	666.61	784.50	801.89
Observed Score	1625	724.56	66.99	621	650	797	849	1625	726.38	68.69	621	644	797	849
Grade 8 Reading														
Difference	1900	-0.64	58.57	-88.05	-68.67	68.58	98.49	1900	0.08	55.94	-86.26	-65.72	67.76	104.63
Predicted Score	1900	758.45	55.38	670.25	690.90	823.10	845.99	1900	758.57	55.23	670.35	692.53	822.76	843.49
Observed Score	1900	757.81	**84.37**	623	652.5	842	909	1900	758.65	78.62	630	658	842	909
Grade 6 Math														
Difference	1710	2.97	**51.22**	-72.99	-54.18	58.81	83.88	1710	0.04	48.37	-73.80	-53.87	59.80	80.12
Predicted Score	1710	741.94	**61.18**	649.31	671.46	815.01	841.66	1710	743.58	57.64	649.82	676.08	813.13	834.56
Observed Score	1710	744.91	77.77	622	651	833	865	1710	743.62	75.24	622	651	830	865
Grade 7 Math														
Difference	1085	-2.92	**58.45**	-89.32	-61.70	59.89	76.58	1085	0.00	52.70	-76.44	-57.62	56.04	73.67
Predicted Score	1085	729.39	64.43	623.49	654.87	802.54	828.87	1085	730.80	61.49	636.73	655.83	800.52	822.44
Observed Score	1085	726.46	84.46	605	629	827	852	1085	730.80	80.98	605	639	827	852
Grade 8 Math														
Difference	1533	-3.94	57.72	-89.46	-69.96	62.50	87.98	1533	0.01	57.97	-86.19	-66.31	68.29	95.63
Predicted Score	1533	731.72	57.56	636.86	666.24	800.67	824.81	1533	732.02	63.64	627.90	656.37	809.02	831.84
Observed Score	1533	727.78	79.72	602	632	816	854	1533	732.02	86.09	602	632	830	866

Note. Bolded values indicate significantly larger standard deviation for Disrupted group at $p < .05$.

Compare predictions for interrupted students to non-interrupted students

Table 9. Percent of Disrupted Students with Predicted and Observed Score Differences at the 5th, 10th, 90th and 95th Percentile of Non-disrupted Students – Cohort A

	N	5th	10th	90th	95th
		Larger than 5% or 10% = Disadvantaged		Larger than 5% or 10% = Advantaged	
Grade 6 Reading	2,002	5.29%	11.24%	10.94%	5.89%
Grade 7 Reading	1,625	8.74%	13.48%	9.97%	3.75%
Grade 8 Reading	1,900	5.32%	11.00%	10.05%	4.63%
Grade 6 Math	1,710	4.85%	10.23%	9.82%	5.79%
Grade 7 Math	1,085	6.54%	11.43%	11.24%	5.81%
Grade 8 Math	1,533	5.54%	11.02%	8.61%	3.78%

For Reading, the results show a slightly higher percent of students above the 5th and 10th cut point for grade 7. These results suggest that for grade 7 Reading, approximately 3.74% of students had lower observed scores than predicted scores based on the non-disrupted sample. In Math, the differences around the 5th and 10th percentiles were less pronounced, although the actual percentages did bounce around what would be expected.

On the upper end of the distribution, it can be noted that some students in the disrupted group scored higher than what would be expected. For example, some students in the disrupted sample scored high in Grade 7 Math than what was expected, but there were other students in this same group that were disadvantaged.

Cohort B Analyses

Because the disruptions on 4/29 and 4/30 appeared to be more pronounced, the above analyses were replicated using *only students disrupted on 4/29 and 4/30* as the disrupted sample. In this way, any effect of disruption, on average, may be more pronounced on this sample than by looking at students disrupted on any day.

Table 10. Cohort A and Cohort B Comparison

Grade	Math			Reading		
	Cohort A	Cohort B	Proportion of Cohort A	Cohort A	Cohort B	Proportion of Cohort A
6	2,068	753	0.36	2,339	577	0.25
7	1,348	588	0.44	1,906	442	0.23
8	1,819	1,103	0.61	2,255	648	0.29
Total	**5,235**	**2,444**		**6,500**	**1,667**	

Interestingly, the proportions of students in Cohort A who are also in Cohort B vary by grade and subject. Grade 8 Math students appear to have primarily been interrupted on 4/29 or 4/30.

Propensity score matching was again used to match the reduced sample of disrupted students to similar students in the non-disrupted group. The same four analyses conducted on Cohort A were also conducted on Cohort B.

Table 11. Descriptive and Inferential Statistics of 2013 Test Scores After Matching – Cohort B

		Disrupted		Non-Disrupted			
	N	Mean	SD	Mean	SD	*t*	*d*
Grade 6 Reading	468	756.3	82.39	750.2	77.78	-1.16	0.08
Grade 7 Reading	617	743.2	78.27	744.9	74.07	0.39	-0.02
Grade 8 Reading	369	726.7	73.87	726.1	67.22	-0.12	0.01
Grade 6 Math	455	728.5	89.21	726.8	72.02	-0.33	0.02
Grade 7 Math	543	755.5	93.14	762.5	86.29	1.28	-0.08
Grade 8 Math	932	719.5	81.63	729.0	86.01	2.44*	-0.11

* p < .05.

Mean scores between the matched disrupted and non-disrupted students were neither statistically nor practically significant except for Grade 8 Math. Although the mean score of disrupted students was lower than the non-disrupted students at a statistical level, the practical significance was still small. Additionally, with 932 students in both samples, the power of this statistical test is very high so even a small difference may be statistically significant.

Examining the predictability of 2013 test scores

Regression equations were formed using the same predictor variables as in Cohort A. Again, one regression combined the disrupted and non-disrupted groups to examine the increase in predictive validity when adding "disruption" to the equation. Overall, the changes in R^2 values were very small.

Table 12. Combined Regression Equation Adding Disruption – Cohort B

	N	Covariates Only R^2	Covariates + Disrupted R^2	ΔR^2
Grade 6 Reading	936	0.6005	0.6009	<0.001
Grade 7 Reading	738	0.5319	0.5321	<0.001
Grade 8 Reading	1,086	0.5444	0.5444	<0.001
Grade 6 Math	1,234	0.5292	0.5293	<0.001
Grade 7 Math	910	0.4892	0.4894	<0.001
Grade 8 Math	1,864	0.5051	0.5080	0.003

Like Cohort A, two more regression equations were built to predict 2013 scores from the same predictor variables in the previous model. These models were run on both groups separately to examine if the difference in R^2 was substantively smaller in the disrupted group than the non-disrupted group. This may signal unexplained variance possibly attributable to a disruption effect. The R^2 values were smaller only in the Grade 8 Math exam, although this change is small, accounting for about 1% of variance explained.

Table 13. Separate Regression Equations for Disrupted and Non-Disrupted Groups – Cohort B

	N	Disrupted R^2	Non-Disrupted R^2	Δ R^2
Grade 6 Reading	468	0.619	0.604	0.015
Grade 7 Reading	369	0.555	0.532	0.024
Grade 8 Reading	543	0.574	0.530	0.045
Grade 6 Math	617	0.534	0.534	<0.000
Grade 7 Math	455	0.504	0.498	0.006
Grade 8 Math	932	0.461	0.560	-0.099

Examining distributions of predicted student scores

By looking at the distribution of difference scores (i.e., observed minus predicted scores), we can determine, as best as our models predict, the scores students who were disrupted would have received if they were not disrupted. Given that Cohort B is a refined sample of Cohort A, any "*notable difference*" or effect of disruption could be expected to be more extreme in this sample.

An *F*-test was conducted comparing the variance of observed, predicted, and difference scores, of both groups. The significantly larger standard deviations for the disrupted group are bolded in Table 14. Interestingly, the Grade 8 Math disrupted group did not have statistically significantly larger variance than its matched non-disrupted group.

Table 14. Distribution of the Difference between Predicted and Observed Scores for Non-disrupted and Disrupted Groups – Cohort B

		Disrupted							Non-Disrupted					
	N	Mean	SD	5th	10th	90th	95th	N	Mean	SD	5th	10th	90th	95th
Grade 6 Reading														
Difference	468	3.28	53.38	-87.02	-58.63	60.40	89.60	468	0.00	48.97	-69.55	-53.25	62.78	81.71
Predicted Score	468	753.03	53.73	659.27	684.55	820.58	838.81	468	750.23	60.43	653.23	678.31	829.04	849.02
Observed Score	468	756.31	82.39	619	659	862	894	468	750.23	77.78	633	652	833	860
Grade 7 Reading														
Difference	369	-1.86	**51.87**	-86.94	-63.94	51.56	71.41	369	-0.01	46.01	-65.14	-53.04	55.97	79.64
Predicted Score	369	728.55	47.56	649.69	668.17	785.31	799.94	369	726.09	49.01	646.52	665.73	781.42	797.76
Observed Score	369	726.69	73.87	593	633	818	849	369	726.08	67.22	621	650	797	849
Grade 8 Reading														
Difference	543	-1.92	63.31	-100.42	-70.16	71.13	98.62	543	-0.03	59.18	-86.42	-65.67	75.25	108.75
Predicted Score	543	757.46	66.52	639.76	672.18	837.02	865.69	543	762.54	62.79	666.74	691.77	837.14	858.30
Observed Score	543	755.53	93.14	591	637	870	909	543	762.51	86.29	623	658	870	909
Grade 6 Math														
Difference	617	0.04	50.59	-75.67	-57.21	57.35	81.22	617	1.69	54.40	-73.58	-60.63	60.76	84.91
Predicted Score	617	744.82	54.10	654.18	679.22	809.00	825.66	617	741.46	59.81	655.87	672.37	807.55	839.04
Observed Score	617	744.86	74.07	622	659	830	865	617	743.15	78.27	632	651	830	865
Grade 7 Math														
Difference	455	-0.94	**65.38**	-110.29	-66.03	71.21	81.95	455	-0.04	51.00	-79.98	-59.48	53.93	78.50
Predicted Score	455	729.48	53.50	637.79	659.40	791.01	812.24	455	726.81	50.85	646.49	670.06	784.81	799.42
Observed Score	455	728.54	**89.21**	597	639	827	852	455	726.76	72.02	618	648	807	838
Grade 8 Math														
Difference	932	-8.80	61.36	-99.62	-77.95	60.03	87.78	932	0.02	57.07	-86.13	-61.29	67.03	93.85
Predicted Score	932	728.32	55.36	638.45	667.91	790.78	811.54	932	728.98	64.35	622.63	652.67	802.95	825.13
Observed Score	932	719.52	81.63	588	623	816	846	932	729.00	86.01	589	623	830	866

Note. Bolded values indicate significantly larger standard deviation for Disrupted group at $p < .05$.

Compare predictions for interrupted students to non-interrupted students

Table 15. Percent of Disrupted Students with Predicted and Observed Score Differences at the 5th, 10th, 90th and 95th Percentile of Non-disrupted Students – Cohort B

	N	5th	10th	90th	95th
		Larger than 5% or 10% = Disadvantaged		Larger than 5% or 10% = Advantaged	
Grade 6 Reading	468	8.55%	11.97%	9.40%	6.62%
Grade 7 Reading	369	9.76%	13.55%	8.13%	4.07%
Grade 8 Reading	543	6.81%	12.52%	8.84%	4.05%
Grade 6 Math	617	4.70%	10.86%	11.02%	5.67%
Grade 7 Math	455	7.47%	11.43%	16.04%	6.15%
Grade 8 Math	932	6.97%	14.81%	8.80%	4.08%

According to Table 15, some disrupted students scored lower than expected whereas others scored higher than expected. The largest group of disadvantaged students appears to be Grade 8 Math at the 10th percentile mark; however, other groups have larger proportions of students at the 5th percentile mark.

Cohort C Analyses

In addition to the interruptions in Grades 6 through 8 exams, HumRRO was provided student scores and interruption data for EOI exams. These exams are primarily completed by high school students and are often taken predominantly in one grade level. This is relevant in selecting variables to predict the current year scores. Since prior performance is the best predictor of future performance, this fact limited our investigation to certain exams with available prior performance data.

Algebra I test scores were selected for Cohort C because most students who complete that exam are 9th graders and their Grade 8 Math exam is available as a sufficient indicator of prior performance. However, some 7th and 8th graders also completed the Algebra 2013 test. These students also completed their Math exams in 2013 and 2012, which can be used as prior performance indicators. Given that these students are likely advanced in this subject (i.e., upper end of the ability distribution), we can examine these students separately to determine if disruption had a differential effect on higher ability students.

To be clear, Cohort C is comprised of students who completed the Algebra I test in 2013 matched with 2012 Math scores; therefore, this sample should consist primarily of students who were 9th graders in spring 2013. The same 2013 Algebra I scores were matched to 2013 Math scores to result in a sample that primarily consists of students who were 7th and 8th graders in spring 2013. The same interruption data on the Algebra I test was used for both groups. The number of valid, unique Algebra I score records is 31,021; however, grade information was only provided in the interruption file. The number of interrupted cases during the 2013 Algebra I test by grade is displayed in Table 16.

Table 16. Interrupted Cases on 2013 Algebra I
Test by Grade

	N	Percentage
Grade 7	59	2.07%
Grade 8	602	21.08%
Grade 9	1,759	61.59%
Grade 10	347	12.15%
Grade 11	72	2.52%
Grade 12	17	0.60%
Total	**2,856**	**100%**

Note that only the Grades 7 and 8 records from the interruption file were matched with the 2013 Algebra I score file to form the subset "2013 Algebra (7th & 8th)." All analyses were conducted on all groups after propensity score matching. The matching formulas were identical for Cohort C; however, Math 2013 scores were used for the grades 7 and 8 subset.

Differences in average 2013 test scores

After propensity score matching, mean difference tests were performed on the matched samples. Note that the performance of the 7th and 8th grade group is higher than the group overall. Additionally, both Disrupted samples seem to have lower scores than the Non-Disrupted matched samples, overall.

Table 17. Descriptive and Inferential Statistics of 2013 Test Scores After Matching – Cohort C

		Disrupted		Non-Disrupted			
	N	Mean	SD	Mean	SD	t	d
2013 Algebra	2,179	739.5	50.17	744.7	49.41	3.40*	0.10
2013 Algebra (7th & 8th)	621	763.5	55.45	771.5	48.81	2.70*	0.15

*$p < .05$.

Examining the predictability of 2013 test scores

The regression equations were performed on both groups in Cohort C. R^2 values were lower for the 2013 Algebra group likely due to only Math 2012 scores being available for a prior performance indicator. Including Math and Reading 2013 scores would have reduced the sample to the 7th and 8th grade group since 9th graders likely did not complete the 2013 Math and Reading tests meant for 8th graders. Importantly, the change in R^2 values were not practically large, indicating a near-zero effect of disruption on Algebra score variance.

Table 18. Combined Regression Equation Adding Disruption – Cohort C

	N	Covariates Only R^2	Covariates + Disrupted R^2	Δ R^2
2013 Algebra	4,358	0.3445	0.3460	-0.002
2013 Algebra (7th & 8th)	621	0.5553	0.5587	-0.003

A Statistical Investigation of Oklahoma Computer Disruptions

The separate regression equations indicated less variance explained in the Disrupted group than the Non-Disrupted group by about 2 or 3%. Although this difference is likely practically small, it is in the direction of concern. Further analyses may signal a consistent disadvantage in one of the Cohort C groups.

Table 19. Separate Regression Equations for Disrupted and Non-Disrupted Groups – Cohort C

	N	Disrupted R^2	Non-Disrupted R^2	ΔR^2
2013 Algebra	2,179	0.332	0.360	-0.029
2013 Algebra (7th & 8th)	621	0.557	0.577	-0.021

Examining distributions of predicted student scores

The distributions of observed, predicted, and difference scores were examined for Cohort C focusing on the 7th and 8th grade students who were at the higher end of the ability distribution. The *F*-tests to compare the variances of the scores between the groups was again conducted. Those results may be found in Table 20.

The 7th and 8th grade group had a statistically significantly larger variance in difference scores for the Disrupted group than the Non-Disrupted group. This result foreshadows larger proportions of students below (and perhaps above) the percentile cuts.

Table 20. Distribution of the Difference between Predicted and Observed Scores for Non-disrupted and Disrupted Groups – Cohort C

		Disrupted							Non-Disrupted					
	N	Mean	SD	5th	10th	90th	95th	*N*	Mean	SD	5th	10th	90th	95th
Algebra														
Difference	2179	-3.77	41.22	-58.05	-44.34	37.99	54.30	2179	-0.03	39.51	-56.61	-42.93	42.08	54.70
Predicted Score	2179	743.32	29.03	706.32	717.52	775.22	785.03	2179	744.71	29.66	704.66	716.15	777.28	787.29
Observed Score	2179	739.55	50.17	670	687	793	815	2179	744.67	49.41	670	687	798	826
Algebra (7th & 8th)														
Difference	621	-6.02	**38.20**	-52.27	-39.94	32.08	51.54	621	0.01	31.74	-41.43	-32.17	32.84	44.01
Predicted Score	621	769.52	38.77	705.60	723.70	816.70	843.30	621	771.49	37.08	714.00	730.40	818.90	838.10
Observed Score	621	763.50	**55.45**	691	704	826	843	621	771.51	48.81	708	722	826	842

Note. Bolded values indicate significantly larger standard deviation for Disrupted group at *p* < .05.

Compare predictions for interrupted students to non-interrupted students

Table 21. Percent of Disrupted Students with Predicted and Observed Score Differences at the 5th, 10th, 90th and 95th Percentile of Non-disrupted Students – Cohort C

	N	5th	10th	90th	95th
		Larger than 5% or 10% = Disadvantaged		Larger than 5% or 10% = Advantaged	
2013 Algebra	2,179	5.42	10.65	8.08	4.91
2013 Algebra (7th & 8th)	621	9.34	17.07	9.66	6.60

Table 21 contains the percentages of students in the Disrupted group above or below the percentile cuts of the Non-Disrupted Students. The 7th and 8th grade students who were interrupted while completing the Algebra I test in 2013 were proportionately more disadvantaged than advantaged. In other words, it appears that 7th and 8th grade students were differentially affected by interruption, on average, than others completing the same exam.

Cohort D Analyses

Another set of students completing an EOI test was examined. The English II test was chosen due to the available data in the EOI dataset. A measure of prior performance was desired for creating the matched sample. However, because students typically take the English II test in Grade 10, the nearest other assessment where most students also had scores was the US History EOI. Using the US History EOI was not a perfect solution, however. The content was obviously different, but highly correlated. The bigger issue was that not all English II EOI interrupted student also had a US History score.

After examining the students who were disrupted on the English II test who had US History test scores, another approach was developed to increase the sample size. More data was provided to HumRRO that enabled us to use 2011 reading test scores (from when the Grade 10 student were in Grade 8) as a prior performance indicator. This approach increased the sample size of non-disrupted students and provided an improved prediction model for the 2013 English II scores. Therefore, both groups are presented as Cohort D, much like Cohort C had two groups focused on the same exam.

For clarity, "2013 English Group 1" is all students who had US History exam scores as a "prior performance" indicator and "2013 English Group 2" is all students who had 2011 Reading scores as a predictor of performance. The prediction models differed primarily in these variables. Group 2 also included 2011 Math scores since the data was available for nearly every student.

Differences in average 2013 test scores

After propensity score matching, mean difference tests were performed on the matched samples. Disrupted students in Group 1 did better than matched Non-Disrupted students; however, this difference was neither statistically nor practically different. Group 2 mean difference was in the opposite direction, but this was also non-significant and small.

Table 22. Descriptive and Inferential Statistics of 2013 Test Scores After Matching – Cohort D

		Disrupted		Non-Disrupted			
	N	Mean	SD	Mean	SD	t	d
2013 English Group 1	427	748.4	74.80	746.1	73.47	-0.44	-0.03
2013 English Group 2	2,688	763.6	64.10	765.2	67.92	0.93	0.03

Note. No *p*-values below .05.

Examining the predictability of 2013 test scores

The regression equations were performed on both groups in Cohort D. R^2 values were lower for the Group 2 equations because the 2011 scores were not a recent test score, limiting the predictive power. However, for both groups, the "Disrupted" dichotomous variable does not additionally explain a practically large amount of English II test score variance.

Table 23. Combined Regression Equation Adding Disruption – Cohort D

	N	Covariates Only R^2	Covariates + Disrupted R^2	ΔR^2
2013 English Group 1	854	0.5187	0.5188	<0.001
2013 English Group 2	5,376	0.3503	0.3504	<0.001

The separate regression equations indicated *more* variance explained in the Disrupted group than the Non-Disrupted group in the Group 1 analyses, with a difference of about 4%. This is not a large difference and is of less concern that if Disrupted students had less variance explained. The difference in Group 2 equations is practically very small.

Table 24. Separate Regression Equations for Disrupted and Non-Disrupted Groups – Cohort D

	N	Disrupted R^2	Non-Disrupted R^2	ΔR^2
2013 English Group 1	427	0.545	0.503	0.042
2013 English Group 2	2,688	0.351	0.354	-0.003

Examining distributions of predicted student scores

The distributions of observed, predicted, and difference scores were examined for Cohort D and the *F*-tests to compare the variances of the scores between the groups were again conducted. Those results may be found in Table 25. Only one variable had larger variance for the Disrupted group at a significant level.

Table 25. Distribution of the Difference between Predicted and Observed Scores for Non-disrupted and Disrupted Groups – Cohort D

		Disrupted							Non-Disrupted					
	N	Mean	SD	5th	10th	90th	95th	N	Mean	SD	5th	10th	90th	95th
English Group 1														
Difference	427	1.26	51.41	-77.54	-54.29	57.77	74.86	427	-0.04	51.82	-87.26	-59.88	60.36	79.94
Predicted Score	427	747.09	52.14	653.61	683.34	814.11	834.58	427	746.15	52.09	655.99	680.26	817.72	834.51
Observed Score	427	748.36	74.80	615	656	838	859	427	746.11	73.47	622	656	838	864
English Group 2														
Difference	2688	-1.78	52.00	-80.89	-61.72	60.27	80.57	2688	-0.05	55.00	-82.57	-60.67	59.01	80.11
Predicted Score	2688	765.34	**43.00**	704.70	725.90	808.70	822.30	2688	765.29	40.00	709.20	724.60	807.50	821.90
Observed Score	2688	763.56	64.00	662	693	838	864	2688	765.24	68.00	654	687	840	859

Note. Bolded values indicate significantly larger standard deviation for Disrupted group at $p < .05$.

Compare predictions for interrupted students to non-interrupted students

Table 26. Percent of Disrupted Students with Predicted and Observed Score Differences at the 5th, 10th, 90th and 95th Percentile of Non-disrupted Students – Cohort D

	N	5th	10th	90th	95th
		Larger than 5% or 10% = Disadvantaged		Larger than 5% or 10% = Advantaged	
2013 English Group 1	427	3.75	8.20	9.84	3.75
2013 English Group 2	2,688	4.84	10.27	10.42	5.21

The percentages of students in the Disrupted group beyond the cutoff of the Non-Disrupted group are provided in Table 26. It seems that fewer students were both advantaged and disadvantaged in Group 1, but these differences do not appear large. The percentages for Group 2 are very close to the expected values and there appear to be no strong directional concerns.

Summary

This report provides a statistical investigation of computer disruptions on student test scores. A number of analyses were conducted on many groups of students completing different tests to investigate whether computer disruptions impacted scores. The only group that seemed to be disadvantaged was the 7th and 8th graders completing the Algebra I test, and the disadvantage was not consistent throughout the distribution of student scores. These differences rely on the predictability of the model, which is inherently imperfect. While there is some evidence to suggest that there were effects from the disruption, nothing emerged in a systematic way across grades, subjects, or methodologies that merit a statistical adjustment. While we cannot know for an individual student that the computer disruption did not impact his or her test score, we conclude that for the overall population of Oklahoma students the computer disruptions had little impact on student test scores.

References

Austin, P. C. (2009). Some methods of propensity-score matching had superior performance to others: Results of an empirical investigation and Monte Carlo simulation. *Biometrical Journal, 51*, 171-184. doi: 10.1002/bimj.200810488

Connelly, B. S., Sackett, P.R., & Waters, S. D. (2013). Balancing treatment and control groups in quasi-experiments: An Introduction to propensity scoring. *Personnel Psychology, 66*, 407-442. doi: 10.1111/peps.12020

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods, 15*(3), 234-249. doi: 10.1037/a0019623

Hill, R. (2013). An analysis of the Impact of Interruptions on the 2013 administration of the Indiana Statewide Testing for Educational Progress—Plus (ISTEP+). The National Center for the Improvement of Educational Assessment, Inc. Downloaded July 28, 2013 from http://www.doe.in.gov/sites/default/files/news/hill-report.pdf.

Mean Differences between Non-Disrupted and Disrupted Samples before and after Propensity Matching

Table A1. Mean Covariate Differences for Grade 6 Reading – Cohort A

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	737.90	78.11	40872	734.47	82.55	2283	0.043
2012 Math Scale	742.12	99.21	38109	736.55	97.65	2145	0.057
2012 Reading Scale	732.91	95.92	38012	726.85	101.34	2136	0.062
School Free-Lunch	0.56	0.20	38896	0.58	0.19	2161	-0.088
School Achievement	738.61	26.44	38896	736.18	27.10	2161	0.091
African American	0.11	0.31	41444	0.15	0.35	2317	**-0.117**
American Indian	0.23	0.42	41444	0.21	0.41	2317	0.043
Asian	0.02	0.16	41444	0.03	0.17	2317	-0.034
Hispanic	0.14	0.34	41442	0.14	0.34	2317	0.005
Pacific Islander	0.01	0.07	41444	0.01	0.07	2317	0.000
Caucasian	0.70	0.46	41444	0.68	0.47	2317	0.044
Male	0.50	0.50	41428	0.50	0.50	2314	0.011
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	737.68	75.80	2002	739.20	79.82	2002	-0.020
2012 Math Scale	740.15	93.20	2002	739.21	95.92	2002	0.010
2012 Reading Scale	731.14	87.70	2002	729.53	98.76	2002	0.017
School Free-Lunch	0.59	0.20	2002	0.58	0.19	2002	0.023
School Achievement	736.03	26.64	2002	736.34	26.95	2002	-0.012
African American	0.15	0.36	2002	0.15	0.36	2002	0.011
American Indian	0.21	0.41	2002	0.21	0.41	2002	-0.007
Asian	0.02	0.15	2002	0.03	0.17	2002	-0.042
Hispanic	0.13	0.34	2002	0.14	0.34	2002	-0.016
Pacific Islander	0.01	0.08	2002	0.01	0.07	2002	0.013
Caucasian	0.68	0.47	2002	0.68	0.47	2002	-0.014
Male	0.48	0.50	2002	0.49	0.50	2002	-0.015

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A2. Mean Covariate Differences for Grade 6 Math – Cohort A

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Reading Scale	731.49	77.29	41178	732.79	82.16	1977	-0.016
2012 Math Scale	741.98	98.90	38635	736.42	94.67	1857	0.057
2012 Reading Scale	732.71	96.48	38153	731.60	90.90	1833	0.012
School Free-Lunch	0.57	0.20	39992	0.57	0.17	1896	-0.044
School Achievement	737.76	26.56	39992	738.05	24.20	1896	-0.011
African American	0.11	0.31	41965	0.16	0.36	2011	**-0.152**
American Indian	0.22	0.42	41965	0.22	0.41	2011	0.018
Asian	0.02	0.16	41965	0.04	0.19	2011	-0.067
Hispanic	0.14	0.35	41964	0.14	0.35	2011	-0.014
Pacific Islander	0.01	0.07	41965	0.00	0.07	2011	0.013
Caucasian	0.70	0.46	41965	0.66	0.47	2011	0.083
Male	0.50	0.50	41864	0.48	0.50	2011	0.050
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Reading Scale	739.53	75.52	1710	738.15	79.91	1710	0.018
2012 Math Scale	742.62	92.83	1710	739.51	94.26	1710	0.033
2012 Reading Scale	736.69	89.18	1710	733.03	91.75	1710	0.041
School Free-Lunch	0.57	0.20	1710	0.57	0.17	1710	-0.002
School Achievement	738.20	25.85	1710	738.25	23.42	1710	-0.002
African American	0.16	0.37	1710	0.16	0.37	1710	0.000
American Indian	0.22	0.41	1710	0.21	0.41	1710	0.017
Asian	0.04	0.19	1710	0.04	0.19	1710	0.000
Hispanic	0.15	0.36	1710	0.15	0.35	1710	0.020
Pacific Islander	0.00	0.03	1710	0.00	0.06	1710	-0.060
Caucasian	0.66	0.47	1710	0.66	0.47	1710	-0.006
Male	0.49	0.50	1710	0.48	0.50	1710	0.033

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A3. Mean Covariate Differences for Grade 7 Reading – Cohort A

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	733.18	80.19	41134	723.82	85.05	1847	**0.113**
2012 Math Scale	733.69	95.85	38422	723.88	102.50	1732	0.099
2012 Reading Scale	729.60	96.57	38486	719.00	102.50	1731	**0.106**
School Free-Lunch	0.55	0.19	39075	0.56	0.18	1765	-0.043
School Achievement	732.65	23.81	39056	728.23	25.35	1764	**0.180**
African American	0.11	0.31	41815	0.16	0.36	1887	**-0.129**
American Indian	0.23	0.42	41815	0.21	0.41	1887	0.030
Asian	0.02	0.15	41815	0.02	0.15	1887	0.000
Hispanic	0.13	0.34	41813	0.12	0.32	1887	0.046
Pacific Islander	0.00	0.06	41815	0.01	0.08	1887	-0.029
Caucasian	0.69	0.46	41815	0.67	0.47	1887	0.048
Male	0.50	0.50	41799	0.50	0.50	1886	0.008
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	724.80	77.89	1625	728.01	82.05	1625	-0.040
2012 Math Scale	721.00	102.66	1625	724.99	102.34	1625	-0.039
2012 Reading Scale	716.65	109.24	1625	719.01	103.90	1625	-0.022
School Free-Lunch	0.56	0.19	1625	0.55	0.18	1625	0.038
School Achievement	728.14	24.37	1625	728.30	25.57	1625	-0.006
African American	0.16	0.37	1625	0.15	0.36	1625	0.024
American Indian	0.22	0.41	1625	0.21	0.41	1625	0.019
Asian	0.03	0.17	1625	0.02	0.15	1625	0.027
Hispanic	0.12	0.32	1625	0.12	0.32	1625	-0.004
Pacific Islander	0.01	0.07	1625	0.01	0.08	1625	-0.016
Caucasian	0.66	0.47	1625	0.67	0.47	1625	-0.030
Male	0.50	0.50	1625	0.49	0.50	1625	0.033

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A4. Mean Covariate Differences for Grade 7 Math – Cohort A

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Reading Scale	730.70	67.27	41702	721.07	74.39	1276	**0.136**
2012 Math Scale	733.52	95.92	39121	721.55	103.24	1191	**0.120**
2012 Reading Scale	729.45	97.02	38842	720.16	99.00	1169	0.095
School Free-Lunch	0.55	0.19	40363	0.60	0.16	1227	**-0.290**
School Achievement	731.88	23.93	40343	725.46	25.27	1227	**0.261**
African American	0.11	0.31	42421	0.16	0.36	1307	**-0.136**
American Indian	0.23	0.42	42421	0.23	0.42	1307	-0.003
Asian	0.02	0.15	42421	0.02	0.13	1307	0.042
Hispanic	0.13	0.34	42421	0.17	0.37	1307	-0.099
Pacific Islander	0.00	0.07	42421	0.01	0.07	1307	-0.016
Caucasian	0.69	0.46	42421	0.66	0.47	1307	0.060
Male	0.50	0.50	42408	0.51	0.50	1306	-0.018
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Reading Scale	729.50	68.49	1085	727.81	71.04	1085	0.024
2012 Math Scale	727.52	92.31	1085	724.73	100.98	1085	0.029
2012 Reading Scale	725.59	97.65	1085	722.12	96.18	1085	0.036
School Free-Lunch	0.60	0.18	1085	0.61	0.16	1085	-0.054
School Achievement	724.77	24.25	1085	725.15	25.27	1085	-0.015
African American	0.18	0.38	1085	0.17	0.37	1085	0.022
American Indian	0.20	0.40	1085	0.22	0.41	1085	-0.043
Asian	0.02	0.15	1085	0.02	0.14	1085	0.032
Hispanic	0.20	0.40	1085	0.18	0.38	1085	0.075
Pacific Islander	0.00	0.06	1085	0.00	0.07	1085	-0.014
Caucasian	0.66	0.47	1085	0.65	0.48	1085	0.021
Male	0.49	0.50	1085	0.50	0.50	1085	-0.011

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A5. Mean Covariate Differences for Grade 8 Reading – Cohort A

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Math Scale	731.00	83.07	39830	732.46	87.45	2189	-0.017
2012 Math Scale	734.49	97.70	37151	732.62	94.76	2041	0.019
2012 Reading Scale	737.40	91.04	37350	733.56	95.21	2048	0.041
School Free-Lunch	0.52	0.20	37910	0.54	0.18	2070	**-0.102**
School Achievement	736.91	24.41	37910	734.38	23.52	2070	**0.106**
African American	0.11	0.31	40681	0.15	0.35	2229	**-0.105**
American Indian	0.22	0.42	40681	0.21	0.41	2229	0.034
Asian	0.02	0.15	40681	0.02	0.15	2229	0.006
Hispanic	0.13	0.33	40679	0.14	0.34	2228	-0.037
Pacific Islander	0.00	0.07	40681	0.01	0.08	2229	-0.027
Caucasian	0.70	0.46	40681	0.68	0.47	2229	0.041
Male	0.50	0.50	40654	0.50	0.50	2228	-0.001
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Math Scale	738.54	83.13	1900	737.43	86.79	1900	0.013
2012 Math Scale	734.74	106.92	1900	733.75	94.06	1900	0.010
2012 Reading Scale	733.97	103.24	1900	735.68	93.53	1900	-0.017
School Free-Lunch	0.54	0.20	1900	0.54	0.18	1900	0.010
School Achievement	734.67	24.33	1900	734.68	23.68	1900	0.000
African American	0.14	0.35	1900	0.15	0.36	1900	-0.025
American Indian	0.21	0.41	1900	0.21	0.41	1900	0.013
Asian	0.03	0.16	1900	0.02	0.15	1900	0.014
Hispanic	0.15	0.36	1900	0.14	0.34	1900	0.040
Pacific Islander	0.01	0.08	1900	0.01	0.08	1900	-0.007
Caucasian	0.68	0.47	1900	0.67	0.47	1900	0.002
Male	0.49	0.50	1900	0.49	0.50	1900	-0.009

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A6. Mean Covariate Differences for Grade 8 Math – Cohort A

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Reading Scale	750.65	82.08	40262	749.63	79.88	1760	0.013
2012 Math Scale	734.59	97.70	37628	727.36	96.68	1644	0.074
2012 Reading Scale	737.67	91.72	37406	731.69	84.76	1632	0.068
School Free-Lunch	0.53	0.20	38969	0.58	0.19	1694	**-0.280**
School Achievement	736.40	24.65	38969	727.38	22.13	1694	**0.386**
African American	0.11	0.31	40939	0.14	0.34	1786	-0.083
American Indian	0.22	0.42	40939	0.19	0.40	1786	0.068
Asian	0.02	0.15	40939	0.01	0.12	1786	0.071
Hispanic	0.13	0.33	40939	0.17	0.38	1786	**-0.132**
Pacific Islander	0.00	0.07	40939	0.01	0.08	1786	-0.018
Caucasian	0.69	0.46	40939	0.70	0.46	1786	-0.004
Male	0.50	0.50	40929	0.49	0.50	1784	0.013
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Reading Scale	756.06	83.99	1533	755.45	78.36	1533	0.008
2012 Math Scale	728.61	109.66	1533	728.40	94.36	1533	0.002
2012 Reading Scale	733.60	104.04	1533	733.15	83.63	1533	0.005
School Free-Lunch	0.58	0.19	1533	0.58	0.19	1533	0.022
School Achievement	727.34	23.99	1533	727.22	22.16	1533	0.005
African American	0.15	0.35	1533	0.14	0.35	1533	0.017
American Indian	0.20	0.40	1533	0.19	0.40	1533	0.008
Asian	0.02	0.12	1533	0.02	0.12	1533	-0.005
Hispanic	0.17	0.38	1533	0.18	0.38	1533	-0.014
Pacific Islander	0.01	0.07	1533	0.01	0.07	1533	0.000
Caucasian	0.69	0.46	1533	0.69	0.46	1533	-0.016
Male	0.49	0.50	1533	0.48	0.50	1533	0.014

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A7. Mean Covariate Differences for Grade 6 Reading – Cohort B

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	737.52	78.25	42606	752.91	84.37	549	**-0.189**
2012 Math Scale	741.58	99.22	39749	761.13	89.66	505	**-0.207**
2012 Reading Scale	732.38	96.37	39644	749.29	83.11	504	**-0.188**
School Free-Lunch	0.56	0.20	40538	0.48	0.22	519	**0.419**
School Achievement	738.31	26.38	40538	751.51	30.66	519	**-0.463**
African American	0.11	0.31	43201	0.13	0.34	560	-0.062
American Indian	0.23	0.42	43201	0.17	0.38	560	**0.140**
Asian	0.02	0.16	43201	0.05	0.21	560	**-0.119**
Hispanic	0.14	0.34	43199	0.13	0.34	560	0.021
Pacific Islander	0.01	0.07	43201	0.00	0.06	560	0.025
Caucasian	0.69	0.46	43201	0.70	0.46	560	-0.013
Male	0.50	0.50	43185	0.49	0.50	557	0.023
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	757.31	82.91	468	758.79	79.65	468	-0.018
2012 Math Scale	762.20	100.46	468	766.50	83.36	468	-0.047
2012 Reading Scale	746.76	105.49	468	753.35	75.84	468	-0.073
School Free-Lunch	0.48	0.22	468	0.47	0.22	468	0.011
School Achievement	751.82	28.10	468	752.24	30.34	468	-0.014
African American	0.12	0.33	468	0.12	0.33	468	0.000
American Indian	0.17	0.38	468	0.17	0.38	468	0.006
Asian	0.07	0.26	468	0.05	0.22	468	0.081
Hispanic	0.15	0.36	468	0.13	0.34	468	0.067
Pacific Islander	0.00	0.07	468	0.00	0.07	468	0.000
Caucasian	0.71	0.46	468	0.72	0.45	468	-0.028
Male	0.46	0.50	468	0.49	0.50	468	-0.051

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

A Statistical Investigation of Oklahoma Computer Disruptions

Table A8. Mean Covariate Differences for Grade 6 Math – Cohort B

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Reading Scale	731.54	77.39	42443	732.05	84.63	712	-0.006
2012 Math Scale	741.81	98.78	39827	736.71	94.85	665	0.053
2012 Reading Scale	732.67	96.39	39329	731.86	86.03	657	0.009
School Free-Lunch	0.57	0.20	41208	0.57	0.17	680	-0.023
School Achievement	737.74	26.49	41208	739.33	24.83	680	-0.062
African American	0.11	0.31	43257	0.19	0.40	719	**-0.246**
American Indian	0.22	0.42	43257	0.21	0.41	719	0.030
Asian	0.03	0.16	43257	0.04	0.19	719	-0.064
Hispanic	0.14	0.35	43256	0.15	0.35	719	-0.021
Pacific Islander	0.01	0.07	43257	0.00	0.05	719	0.041
Caucasian	0.70	0.46	43257	0.63	0.48	719	**0.143**
Male	0.50	0.50	43156	0.48	0.50	719	0.046
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Reading Scale	742.25	71.63	617	740.30	81.34	617	0.026
2012 Math Scale	745.58	98.78	617	739.02	96.46	617	0.067
2012 Reading Scale	736.91	99.64	617	733.21	85.08	617	0.040
School Free-Lunch	0.56	0.20	617	0.57	0.17	617	-0.046
School Achievement	740.61	24.75	617	739.55	24.79	617	0.043
African American	0.21	0.41	617	0.20	0.40	617	0.016
American Indian	0.19	0.40	617	0.21	0.41	617	-0.040
Asian	0.04	0.20	617	0.04	0.19	617	0.008
Hispanic	0.15	0.36	617	0.15	0.36	617	-0.014
Pacific Islander	0.01	0.08	617	0.00	0.06	617	0.047
Caucasian	0.64	0.48	617	0.63	0.48	617	0.034
Male	0.49	0.50	617	0.47	0.50	617	0.023

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A9. Mean Covariate Differences for Grade 7 Reading – Cohort B

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	732.88	80.36	42557	722.22	86.01	424	**0.128**
2012 Math Scale	733.40	96.02	39757	720.21	109.55	397	**0.128**
2012 Reading Scale	729.22	96.80	39822	721.26	102.03	395	0.080
School Free-Lunch	0.55	0.19	40436	0.52	0.18	404	**0.150**
School Achievement	732.46	23.85	40416	732.63	28.40	404	-0.007
African American	0.11	0.32	43268	0.10	0.30	434	0.055
American Indian	0.23	0.42	43268	0.21	0.40	434	0.051
Asian	0.02	0.15	43268	0.02	0.13	434	0.052
Hispanic	0.13	0.34	43266	0.12	0.33	434	0.026
Pacific Islander	0.00	0.07	43268	0.01	0.08	434	-0.035
Caucasian	0.69	0.46	43268	0.76	0.43	434	**-0.156**
Male	0.50	0.50	43252	0.50	0.50	433	-0.007
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	721.38	84.57	369	728.33	82.75	369	-0.083
2012 Math Scale	721.03	100.50	369	722.94	110.59	369	-0.018
2012 Reading Scale	721.72	110.46	369	721.48	102.69	369	0.002
School Free-Lunch	0.51	0.20	369	0.52	0.18	369	-0.031
School Achievement	732.72	23.51	369	732.57	28.85	369	0.006
African American	0.10	0.30	369	0.09	0.29	369	0.028
American Indian	0.19	0.39	369	0.21	0.40	369	-0.041
Asian	0.01	0.09	369	0.02	0.13	369	-0.075
Hispanic	0.14	0.35	369	0.13	0.34	369	0.039
Pacific Islander	0.00	0.05	369	0.01	0.09	369	-0.076
Caucasian	0.77	0.42	369	0.78	0.42	369	-0.019
Male	0.48	0.50	369	0.50	0.50	369	-0.049

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A10. Mean Covariate Differences for Grade 7 Math – Cohort B

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Reading Scale	730.47	67.39	42436	725.55	75.69	542	0.069
2012 Math Scale	733.33	95.94	39806	720.66	111.47	506	**0.122**
2012 Reading Scale	729.28	96.99	39510	720.96	104.39	501	0.083
School Free-Lunch	0.55	0.19	41088	0.63	0.13	502	**-0.517**
School Achievement	731.83	23.92	41068	720.38	27.21	502	**0.448**
African American	0.11	0.32	43178	0.15	0.36	550	**-0.114**
American Indian	0.23	0.42	43178	0.21	0.40	550	0.048
Asian	0.02	0.15	43178	0.03	0.17	550	-0.037
Hispanic	0.13	0.34	43178	0.23	0.42	550	**-0.271**
Pacific Islander	0.00	0.07	43178	0.00	0.04	550	0.046
Caucasian	0.69	0.46	43178	0.69	0.46	550	0.007
Male	0.50	0.50	43165	0.50	0.50	549	0.004
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Reading Scale	727.78	66.09	455	731.12	73.73	455	-0.048
2012 Math Scale	719.35	113.09	455	724.81	105.35	455	-0.050
2012 Reading Scale	720.98	100.63	455	723.56	96.83	455	-0.026
School Free-Lunch	0.62	0.17	455	0.64	0.13	455	-0.094
School Achievement	721.78	25.97	455	720.23	26.76	455	0.059
African American	0.14	0.35	455	0.17	0.37	455	-0.079
American Indian	0.19	0.39	455	0.19	0.40	455	-0.017
Asian	0.04	0.20	455	0.03	0.17	455	0.048
Hispanic	0.33	0.47	455	0.26	0.44	455	**0.160**
Pacific Islander	0.00	0.00	455	0.00	0.05	455	-0.094
Caucasian	0.69	0.46	455	0.67	0.47	455	0.042
Male	0.48	0.50	455	0.49	0.50	455	-0.018

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A11. Mean Covariate Differences for Grade 8 Reading – Cohort B

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	731.06	83.17	41404	732.58	91.92	615	-0.017
2012 Math Scale	734.45	97.53	38615	730.06	98.97	577	0.045
2012 Reading Scale	737.27	91.09	38817	732.36	102.29	581	0.051
School Free-Lunch	0.52	0.20	39393	0.55	0.16	587	**-0.165**
School Achievement	736.76	24.34	39393	738.05	26.04	587	-0.051
African American	0.11	0.31	42280	0.17	0.38	630	**-0.178**
American Indian	0.22	0.42	42280	0.19	0.39	630	0.089
Asian	0.02	0.15	42280	0.03	0.18	630	-0.056
Hispanic	0.13	0.33	42277	0.16	0.37	630	**-0.104**
Pacific Islander	0.00	0.07	42280	0.01	0.11	630	-0.089
Caucasian	0.69	0.46	42280	0.66	0.47	630	0.077
Male	0.50	0.50	42253	0.47	0.50	629	0.058
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
2013 Math Scale	745.03	84.42	543	737.56	92.54	543	0.084
2012 Math Scale	740.94	99.16	543	731.78	99.41	543	0.092
2012 Reading Scale	739.54	100.90	543	734.72	102.88	543	0.047
School Free-Lunch	0.55	0.20	543	0.55	0.16	543	-0.012
School Achievement	740.80	27.80	543	738.86	26.29	543	0.072
African American	0.16	0.37	543	0.17	0.38	543	-0.025
American Indian	0.21	0.41	543	0.19	0.39	543	0.046
Asian	0.02	0.14	543	0.03	0.18	543	-0.081
Hispanic	0.14	0.35	543	0.16	0.37	543	-0.067
Pacific Islander	0.01	0.10	543	0.01	0.10	543	0.000
Caucasian	0.67	0.47	543	0.66	0.47	543	0.027
Male	0.49	0.50	543	0.47	0.50	543	0.048

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A12. Mean Covariate Differences for Grade 8 Math – Cohort B

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Reading Scale	750.75	82.06	40960	745.13	78.98	1062	0.070
2012 Math Scale	734.53	97.66	38272	725.10	97.69	1000	0.097
2012 Reading Scale	737.56	91.71	38044	732.04	80.73	994	0.064
School Free-Lunch	0.53	0.20	39641	0.60	0.20	1022	**-0.349**
School Achievement	736.32	24.58	39641	724.72	23.27	1022	**0.485**
African American	0.11	0.31	41647	0.16	0.37	1078	**-0.142**
American Indian	0.22	0.42	41647	0.20	0.40	1078	0.059
Asian	0.02	0.15	41647	0.01	0.12	1078	0.062
Hispanic	0.13	0.33	41647	0.19	0.39	1078	**-0.171**
Pacific Islander	0.00	0.07	41647	0.01	0.09	1078	-0.033
Caucasian	0.69	0.46	41647	0.67	0.47	1078	0.060
Male	0.50	0.50	41637	0.50	0.50	1076	-0.005
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
2013 Reading Scale	750.33	86.83	932	749.99	77.36	932	0.004
2012 Math Scale	724.17	107.83	932	725.62	95.75	932	-0.014
2012 Reading Scale	729.19	95.85	932	733.32	80.97	932	-0.047
School Free-Lunch	0.60	0.19	932	0.60	0.20	932	0.034
School Achievement	724.72	25.69	932	724.47	23.17	932	0.010
African American	0.17	0.38	932	0.17	0.37	932	0.011
American Indian	0.16	0.37	932	0.19	0.40	932	-0.084
Asian	0.02	0.14	932	0.02	0.13	932	0.032
Hispanic	0.20	0.40	932	0.20	0.40	932	0.011
Pacific Islander	0.00	0.06	932	0.01	0.08	932	-0.047
Caucasian	0.68	0.47	932	0.66	0.47	932	0.043
Male	0.47	0.50	932	0.49	0.50	932	-0.043

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A13. Mean Covariate Differences for 2013 Algebra – Cohort C

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
Predicted Disruption	0.07	0.01	30721	0.07	0.02	2179	**-0.218**
Algebra Scale	746.45	50.08	31021	739.28	50.20	2196	**0.143**
2012 Math Scale	745.28	91.82	31021	736.73	98.00	2196	0.090
2012 Reading Scale	761.30	87.63	30804	755.68	89.18	2181	0.064
School Free-Lunch	0.47	0.13	30941	0.46	0.14	2195	0.033
School Achievement	752.89	29.75	30941	750.51	30.65	2195	0.079
African American	0.10	0.30	31021	0.13	0.33	2196	-0.081
American Indian	0.20	0.40	31021	0.22	0.41	2196	-0.049
Asian	0.03	0.16	31021	0.02	0.15	2196	0.023
Hispanic	0.00	0.00	31021	0.00	0.00	2196	0.000
Pacific Islander	0.01	0.08	31021	0.00	0.06	2196	0.047
Caucasian	0.68	0.46	31021	0.64	0.48	2196	0.086
Male	0.49	0.50	31018	0.47	0.50	2195	0.048
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
Predicted Disruption	0.07	0.02	2179	0.07	0.02	2179	-0.005
Algebra Scale	744.67	49.41	2179	739.55	50.17	2179	0.103
2012 Math Scale	741.35	95.62	2179	737.50	96.85	2179	0.040
2012 Reading Scale	757.07	94.64	2179	755.75	89.15	2179	0.014
School Free-Lunch	0.45	0.14	2179	0.46	0.14	2179	-0.067
School Achievement	752.48	32.20	2179	750.52	30.73	2179	0.062
African American	0.13	0.34	2179	0.13	0.33	2179	0.012
American Indian	0.22	0.42	2179	0.22	0.41	2179	0.009
Asian	0.03	0.17	2179	0.02	0.15	2179	0.037
Hispanic	0.00	0.00	2179	0.00	0.00	2179	0.000
Pacific Islander	0.00	0.06	2179	0.00	0.06	2179	0.000
Caucasian	0.64	0.48	2179	0.64	0.48	2179	-0.006
Male	0.49	0.50	2179	0.47	0.50	2179	0.035

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

A Statistical Investigation of Oklahoma Computer Disruptions

Table A14. Mean Covariate Differences for 2013 Algebra (7th & 8th) – Cohort C

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
Predicted Disruption	0.07	0.03	8632	0.08	0.03	621	**-0.345**
Algebra Scale	775.49	47.84	9077	763.11	55.51	648	**0.240**
2012 Math Scale	810.83	65.01	8738	799.77	71.70	622	**0.162**
2013 Math Scale	802.69	69.26	9077	791.65	75.09	648	**0.153**
2013 Read Scale	806.86	67.91	9050	799.37	73.35	648	**0.106**
African American	0.09	0.29	9077	0.10	0.30	648	-0.023
American Indian	0.17	0.37	9077	0.19	0.40	648	-0.074
Asian	0.05	0.23	9077	0.04	0.21	648	0.042
Caucasian	0.75	0.43	9077	0.71	0.45	648	0.091
Hispanic	0.10	0.29	9077	0.07	0.26	648	0.082
Pacific Islander	0.00	0.07	9077	0.00	0.04	648	0.059
Free Lunch	0.33	0.47	9077	0.41	0.49	648	**-0.152**
Grade	7.89	0.32	9077	7.91	0.29	661	-0.084
School FRL Proportion	0.33	0.15	8991	0.35	0.16	647	-0.098
School Achievement	804.81	17.56	8991	799.35	22.83	647	**0.270**
Male	0.49	0.50	9076	0.46	0.50	648	0.057

	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
Predicted Disruption	0.08	0.03	621	0.08	0.03	621	0.001
Algebra Scale	771.51	48.81	621	763.50	55.45	621	**0.154**
2012 Math Scale	798.61	89.06	621	800.07	71.36	621	-0.018
2013 Math Scale	795.61	71.39	621	792.02	74.68	621	0.049
2013 Read Scale	804.61	70.84	621	800.23	73.21	621	0.061
African American	0.08	0.27	621	0.10	0.29	621	-0.057
American Indian	0.20	0.40	621	0.20	0.40	621	0.012
Asian	0.05	0.22	621	0.05	0.21	621	0.015
Caucasian	0.71	0.45	621	0.71	0.45	621	-0.007
Hispanic	0.07	0.25	621	0.07	0.26	621	-0.025
Pacific Islander	0.00	0.04	621	0.00	0.04	621	0.000
Free Lunch	0.41	0.49	621	0.40	0.49	621	0.007
Grade	7.92	0.27	621	7.91	0.28	621	0.023
School FRL Proportion	0.33	0.15	621	0.35	0.15	621	-0.099
School Achievement	801.66	20.94	621	799.72	22.11	621	0.090
Male	0.45	0.50	621	0.46	0.50	621	-0.019

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

Table A15. Mean Covariate Differences for 2013 English Group 1 – Cohort D

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
Predicted Disruption	0.05	0.02	7991	0.06	0.03	427	**-0.377**
2013 English	761.18	71.68	35702	756.61	69.93	3113	0.065
2013 History Scale	706.90	98.52	7995	693.56	96.85	431	**0.137**
School Free-Lunch	0.46	0.09	35700	0.46	0.07	3112	0.060
African American	0.11	0.31	35702	0.10	0.30	3113	0.027
American Indian	0.20	0.40	35702	0.17	0.38	3113	0.078
Asian	0.03	0.17	35702	0.02	0.15	3113	0.044
Hispanic	0.00	0.00	35702	0.00	0.00	3113	0.000
Pacific Islander	0.01	0.09	35702	0.01	0.08	3113	0.037
Caucasian	0.67	0.47	35702	0.54	0.50	3113	**0.280**
Male	0.49	0.50	35696	0.49	0.50	3108	-0.007
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	*N*	Mean	SD	*N*	Cohen's *d*
Predicted Disruption	0.06	0.03	427	0.06	0.03	427	-0.025
2013 English	746.11	73.47	427	748.36	74.80	427	-0.030
2013 History Scale	692.31	95.73	427	694.80	96.44	427	-0.026
School Free-Lunch	0.46	0.09	427	0.45	0.06	427	0.087
African American	0.13	0.33	427	0.11	0.31	427	0.058
American Indian	0.21	0.41	427	0.23	0.42	427	-0.045
Asian	0.03	0.17	427	0.02	0.14	427	0.076
Hispanic	0.00	0.00	427	0.00	0.00	427	0.000
Pacific Islander	0.02	0.14	427	0.01	0.12	427	0.054
Caucasian	0.56	0.50	427	0.57	0.50	427	-0.009
Male	0.48	0.50	427	0.49	0.50	427	-0.033

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.

A Statistical Investigation of Oklahoma Computer Disruptions

Table A16. Mean Covariate Differences for 2013 English Group 2 – Cohort D

	Before Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
Predicted Disruption	0.08	0.04	30130	0.12	0.10	2688	**-0.595**
2013 English Scale	761.18	71.68	35702	756.61	69.93	3113	0.065
2011 Reading Scale	704.12	171.71	30134	709.98	167.72	2692	-0.035
2011 Math Scale	721.04	108.29	30134	724.22	104.05	2692	-0.030
African American	0.11	0.31	35702	0.10	0.30	3113	0.027
American Indian	0.20	0.40	35702	0.17	0.38	3113	0.078
Asian	0.03	0.17	35702	0.02	0.15	3113	0.044
Caucasian	0.67	0.47	35702	0.54	0.50	3113	**0.280**
Hispanic	0.00	0.00	35702	0.00	0.00	3113	0.000
Pacific Islander	0.01	0.09	35702	0.01	0.08	3113	0.037
Free Lunch	0.47	0.50	35702	0.38	0.49	3113	**0.176**
Male	0.49	0.50	35696	0.49	0.50	3108	-0.007
	After Matching						
	Non-Disrupted			Disrupted			
	Mean	SD	N	Mean	SD	N	Cohen's d
Predicted Disruption	0.12	0.10	2688	0.12	0.10	2688	-0.001
2013 English Scale	765.24	67.92	2688	763.56	64.10	2688	0.025
2011 Reading Scale	708.85	165.13	2688	710.03	167.59	2688	-0.007
2011 Math Scale	726.26	97.85	2688	724.43	103.74	2688	0.018
African American	0.10	0.30	2688	0.09	0.29	2688	0.029
American Indian	0.18	0.39	2688	0.18	0.38	2688	0.004
Asian	0.02	0.14	2688	0.02	0.13	2688	0.014
Caucasian	0.53	0.50	2688	0.55	0.50	2688	-0.028
Hispanic	0.00	0.00	2688	0.00	0.00	2688	0.000
Pacific Islander	0.01	0.08	2688	0.00	0.07	2688	0.025
Free Lunch	0.39	0.49	2688	0.37	0.48	2688	0.044
Male	0.50	0.50	2688	0.49	0.50	2688	0.021

Note. Bolded values indicate Cohen's *d* greater than .10 or -.10.