



OKLAHOMA STATE DEPARTMENT OF EDUCATION

Janet Barresi, State Superintendent of Public Instruction

Oklahoma School Testing Program Oklahoma Core Curriculum Tests

Grades 3–8 Assessments 2013–2014 Technical Report

FINAL

Submitted to
Oklahoma State Department of Education
October 2014



CTB

Copyright

Developed and published under contract with Oklahoma State Department of Education by CTB/McGraw-Hill LLC, 20 Ryan Ranch Road, Monterey, California 93940-5703. Copyright © 2014 by Oklahoma State Department of Education. Portions copyright © CTB/McGraw-Hill LLC. All rights reserved. Only Oklahoma State educators and citizens may copy and/or download and print the document, located online at <http://ok.gov/sde/accountability-state-testing-results>. Any other use or reproduction of this document, in whole or in part, requires written approval of Oklahoma State Department of Education and the publisher.

Revision History

Version 1.0

Table of Contents

Copyright	i
Revision History	ii
List of Tables	v
List of Figures	viii
List of Appendices	xi
Acronyms and Abbreviations	xii
Introduction.....	1
Section 1—Overview	2
Section 1.1—Purpose	2
Section 1.2—Oklahoma Academic Standards	3
Section 2—Item and Test Development.....	5
Section 2.1—Aligning Test to Oklahoma Academic Standards	5
Section 2.2—Item Pool Development and Selection	8
Section 2.3—Configuration of the Spring 2014 Tests	10
Section 2.4—Operational and Field Test Items by Content Area.....	10
Section 3—Administration	12
Section 3.1—Packaging and Shipping.....	12
Section 3.2—Materials Return.....	12
Section 3.3—Materials Discrepancies Process	13
Section 4—Scoring	14
Section 4.1—Handscoring	14
Section 5—Sampling Plan & Field Test Design.....	18
Section 5.1—Sampling Plan	18
Section 5.2—Field Test Design	19
Section 5.3—Data Checking Activities	19
5.3.1 <i>Suppressed/Omitted/Invalidated cases</i>	19
5.3.2 <i>Duplicate cases</i>	19
5.3.3 <i>Non-public schools</i>	19
5.3.4 <i>Second-timers</i>	19
Section 6—Methods.....	20
Section 6.1—Classical Item Analyses	20
Section 6.2—Differential Item Functioning (DIF)	21
Section 6.3—Calibration & Item Fit	22
Section 6.4—Equating	25
Section 6.5—Writing Scoring.....	26
Section 7—Results.....	27
Section 7.1—Item Level Performance	27
Section 7.2—Standards Level Performance.....	29
Section 7.3—Test Level Performance	29
Section 7.4—Proficiency Level Performance.....	30
Section 8—Summary of Reliability & Validity.....	32
Section 8.1—Item Level Reliability.....	32

Section 8.2—Test Level Reliability	32
Section 8.3—Test Level Validity	34
Section 8.4—Performance Level Reliability	35
References	39
Tables	43
Figures.....	135
Appendices.....	207
Appendix A	207
Standards, Objectives/Skills, and Processes Assessed by Subject	207
Appendix B	224
Spring 2014 OK Grades 3 to 8 Writing Linking Study	224

List of Tables

Table 1. Testable Standards for OCCT Grades 3 to 8	3
Table 2. Testable Standards for OCCT Grades 3 to 8	5
Table 3. Criteria for Aligning the Test with Oklahoma Academic Standards Standards and Objectives	7
Table 4. Percentage of Items by Depth of Knowledge Levels for OCCT 3–8 Assessments.....	9
Table 5. Configuration of the OCCT 3–8 Tests for Spring 2014	11
Table 6. Inter-Rater Reliability Coefficients for the Operational Writing Prompts	17
Table 7. Equating Sample Size for Spring 2014 and Respective Percentage of the Population for Each Grade and Content	18
Table 8. Writing Analytic Traits and Scoring Weights	26
Table 9. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2014	43
Table 10. Summary of <i>P</i> -Values and Item-Test Correlations Statistics for Operational Test Forms, Spring 2012 to Spring 2014.....	49
Table 11. Summary of Range of <i>P</i> -Values and Item-Test Correlations Statistics for Operational and Field Test, Spring 2014	50
Table 12. Summary of Range of Omission Rates for Operational and Field Test by Item Type, Spring 2014	51
Table 13. Spring 2014 Summary of Operational and Field Test Items Flagged for Mantel- Haenszel Differential Item Functioning, by Item Type: Gender Male/Female	52
Table 14. Spring 2014 Summary of Operational and Field Test Items Flagged for Mantel- Haenszel Differential Item Functioning, by Item Type: Ethnicity White/Asian	53
Table 15. Spring 2014 Summary of Operational and Field Test Items Flagged for Mantel- Haenszel Differential Item Functioning, by Item Type: Ethnicity White/African American	54
Table 16. Spring 2014 Summary of Operational and Field Test Items Flagged for Mantel- Haenszel Differential Item Functioning, by Item Type: Ethnicity White/Hispanic	55
Table 17. Mathematics Grades 3–5 Standards Level Summary Data, Spring 2014.....	56
Table 18. Mathematics Grades 6–8 Standards Level Summary Data, Spring 2014.....	57
Table 19. Reading Grades 3–5 Standards Level Summary Data, Spring 2014	58
Table 20. Reading Grades 6–8 Standards Level Summary Data, Spring 2014	59
Table 21. Science Grades 5 & 8 Standards Level Summary Data, Spring 2014.....	60
Table 22. Social Studies Grade 5 & U.S. History Grade 8 Standards Level Summary Data, Spring 2014.....	61
Table 23. Spring 2013 Scale Score Statistics.....	62
Table 24. Spring 2014 Scale Score Statistics.....	63
Table 25. Spring 2014, Mathematics Grades 3 & 4 State and Subgroup Scale Score Descriptive Data.....	64
Table 26. Spring 2014, Mathematics Grades 5 & 6 State and Subgroup Scale Score Descriptive Data.....	65
Table 27. Spring 2014, Mathematics Grades 7 & 8 State and Subgroup Scale Score Descriptive Data.....	66

Table 28. Spring 2014, Reading Grades 3 & 4 State and Subgroup Scale Score Descriptive Data.....	67
Table 29. Spring 2014, Reading Grades 5 & 6 State and Subgroup Scale Score Descriptive Data.....	68
Table 30. Spring 2014, Reading Grades 7 & 8 State and Subgroup Scale Score Descriptive Data.....	69
Table 31. Spring 2014, Science Grades 5 & 8 State and Subgroup Scale Score Descriptive Data.....	70
Table 32. Spring 2014, Social Studies Grade 5 State and Subgroup Scale Score Descriptive Data.....	71
Table 33. Spring 2014, U.S. History Grade 8 State and Subgroup Scale Score Descriptive Data.....	72
Table 34. Spring 2014, Subgroup Scale Score Mean Differences, <i>t</i> -test: Accommodated/Non Accommodated.....	73
Table 35. Spring 2014, Subgroup Scale Score Mean Differences, <i>t</i> -test: ELL/Non ELL.....	74
Table 36. Spring 2014, Subgroup Scale Score Mean Differences, <i>t</i> -test: Female/Male	75
Table 37. Spring 2014, Mathematics Subgroup Scale Score Mean Differences, <i>t</i> -test: IEP/Non IEP.....	76
Table 38. Spring 2014, Mathematics Subgroup Scale Score Mean Differences, <i>t</i> -test: Low SES/High SES	77
Table 39. Spring 2014, Mathematics Subgroup Scale Score Mean Differences, <i>t</i> -test: Section 504/Non Section 504.....	78
Table 40. Spring 2014, Mathematics Subgroup Mean Differences, ANOVA: Ethnicity.....	79
Table 41. Spring 2014, Reading Subgroup Mean Differences, ANOVA: Ethnicity	80
Table 42. Spring 2014, Science Subgroup Mean Differences, ANOVA: Ethnicity.....	81
Table 43. Spring 2014, Social Studies Subgroup Mean Differences, ANOVA: Ethnicity	81
Table 44. Spring 2014, U.S. History Subgroup Mean Differences, ANOVA: Ethnicity	81
Table 45. Mathematics, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity	82
Table 46. Reading, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity	88
Table 47. Science, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity	94
Table 48. Social Studies, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity	96
Table 49. U.S. History, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity	97
Table 50. Spring 2014, Mean Scale Score and Standard Deviations for State and Each Proficiency Level.....	98
Table 51. Spring 2013 Proficiency Level Impact Data, (% rounded)	99
Table 52. Spring 2014 Proficiency Level Impact Data, (% rounded)	100
Table 53. Spring 2014, Mathematics State and Subgroup Proficiency Level Impact Data.....	101
Table 54. Spring 2014, Reading State and Subgroup Proficiency Level Impact Data.....	104
Table 55. Spring 2014, Science State and Subgroup Proficiency Level Impact Data.....	107
Table 56. Spring 2014, Social Studies State and Subgroup Proficiency Level Impact Data	108

Table 57. Spring 2014, U.S. History State and Subgroup Proficiency Level Impact Data	109
Table 58. Spring 2014 Operational Test Parameters Correlations	110
Table 59. Scale Score Statistics for Operational Test in Spring 2013 and Spring 2014.....	111
Table 60. Spring 2014, Proficiency Levels Cut Scores and Scale Bounds.....	112
Table 61. Spring 2014, Proportion of Students Within Each Performance Level	113
Table 62. Differences in Overall Pass Rate for Spring 2013 and Spring 2014	114
Table 63. Spring 2014, Summary of Range of P-Values and Item-Test Correlations Statistics for Operational and Field Test, by Item Type	115
Table 64. Spring 2013 and Spring 2014 Test Reliability Data	116
Table 65. Mathematics, Raw Score to Scale Score Conversions & Standard Error of Measurement.....	117
Table 66. Reading, Raw Score to Scale Score Conversions & Standard Error of Measurement.....	121
Table 67. Science, Raw Score to Scale Score Conversions & Standard Error of Measurement.....	125
Table 68. Social Studies, Raw Score to Scale Score Conversions & Standard Error of Measurement.....	127
Table 69. U.S. History, Raw Score to Scale Score Conversions & Standard Error of Measurement.....	128
Table 70. Spring 2014, Total Group Factor Analysis Results: Eigenvalues	129
Table 71. Spring 2014, Mathematics Subgroup Factor Analysis Results: Eigenvalues	130
Table 72. Spring 2014, Reading Subgroup Factor Analysis Results: Eigenvalues	131
Table 73. Spring 2014, Science Subgroup Factor Analysis Results: Eigenvalues	132
Table 74. Spring 2014, Social Studies Subgroup Factor Analysis Results: Eigenvalues	132
Table 75. Spring 2014, U.S. History Subgroup Factor Analysis Results: Eigenvalues	132
Table 76. Spring 2014, Proficiency Level Cut Scores and Standard Error of Measurement (SEM).....	133
Table 77. Estimates of Accuracy and Consistency of Performance Classification	133
Table 78. Accuracy & Consistency Estimates by Cut Score	134
Table 79. Accuracy & Consistency Estimates by Cut Score: False Positive and False Negative Rates.....	134

List of Figures

Figure 1. Spring 2014 Grade 3 Mathematics operational scale score histogram.....	135
Figure 2. Spring 2014 Grade 4 Mathematics operational scale score histogram.....	136
Figure 3. Spring 2014 Grade 5 Mathematics operational scale score histogram.....	137
Figure 4. Spring 2014 Grade 6 Mathematics operational scale score histogram.....	138
Figure 5. Spring 2014 Grade 7 Mathematics operational scale score histogram.....	139
Figure 6. Spring 2014 Grade 8 Mathematics operational scale score histogram.....	140
Figure 7. Spring 2014 Grade 3 Reading operational scale score histogram.....	141
Figure 8. Spring 2014 Grade 4 Reading operational scale score histogram.....	142
Figure 9. Spring 2014 Grade 5 Reading operational scale score histogram.....	143
Figure 10. Spring 2014 Grade 6 Reading operational scale score histogram.....	144
Figure 11. Spring 2014 Grade 7 Reading operational scale score histogram.....	145
Figure 12. Spring 2014 Grade 8 Reading operational scale score histogram.....	146
Figure 13. Spring 2014 Grade 5 Science operational scale score histogram.....	147
Figure 14. Spring 2014 Grade 8 Science operational scale score histogram.....	148
Figure 15. Spring 2014 Grade 5 Social Studies operational scale score histogram.....	149
Figure 16. Spring 2014 Grade 8 U.S. History operational scale score histogram.....	150
Figure 17. Spring 2014 Grade 3 Mathematics operational test characteristic curve and standard error of measurement curve.....	151
Figure 18. Spring 2014 Grade 4 Mathematics operational test characteristic curve and standard error of measurement curve.....	152
Figure 19. Spring 2014 Grade 5 Mathematics operational test characteristic curve and standard error of measurement curve.....	153
Figure 20. Spring 2014 Grade 6 Mathematics operational test characteristic curve and standard error of measurement curve.....	154
Figure 21. Spring 2014 Grade 7 Mathematics operational test characteristic curve and standard error of measurement curve.....	155
Figure 22. Spring 2014 Grade 8 Mathematics operational test characteristic curve and standard error of measurement curve.....	156
Figure 23. Spring 2014 Grade 3 Reading operational test characteristic curve and standard error of measurement curve.....	157
Figure 24. Spring 2014 Grade 4 Reading operational test characteristic curve and standard error of measurement curve.....	158
Figure 25. Spring 2014 Grade 5 Reading operational test characteristic curve and standard error of measurement curve.....	159
Figure 26. Spring 2014 Grade 6 Reading operational test characteristic curve and standard error of measurement curve.....	160
Figure 27. Spring 2014 Grade 7 Reading operational test characteristic curve and standard error of measurement curve.....	161
Figure 28. Spring 2014 Grade 8 Reading operational test characteristic curve and standard error of measurement curve.....	162

Figure 29. Spring 2014 Grade 5 Science operational test characteristic curve and standard error of measurement curve.....	163
Figure 30. Spring 2014 Grade 8 Science operational test characteristic curve and standard error of measurement curve.....	164
Figure 31. Spring 2014 Grade 5 Social Studies operational test characteristic curve and standard error of measurement curve.....	165
Figure 32. Spring 2014 Grade 8 U.S. History operational test characteristic curve and standard error of measurement curve.....	166
Figure 33. Spring 2014 Grade 3 Mathematics scree plot: Total	167
Figure 34. Spring 2014 Grade 3 Mathematics scree plot: Accommodated	167
Figure 35. Spring 2014 Grade 3 Mathematics scree plot: English Language Learner	168
Figure 36. Spring 2014 Grade 3 Mathematics scree plot: Free Lunch	168
Figure 37. Spring 2014 Grade 3 Mathematics scree plot: Individualized Education Program ..	169
Figure 38. Spring 2014 Grade 4 Mathematics scree plot: Total	169
Figure 39. Spring 2014 Grade 4 Mathematics scree plot: Accommodated	170
Figure 40. Spring 2014 Grade 4 Mathematics scree plot: English Language Learner	170
Figure 41. Spring 2014 Grade 4 Mathematics scree plot: Free Lunch	171
Figure 42. Spring 2014 Grade 4 Mathematics scree plot: Individualized Education Program ..	171
Figure 43. Spring 2014 Grade 5 Mathematics scree plot: Total	172
Figure 44. Spring 2014 Grade 5 Mathematics scree plot: Accommodated	172
Figure 45. Spring 2014 Grade 5 Mathematics scree plot: English Language Learner	173
Figure 46. Spring 2014 Grade 5 Mathematics scree plot: Free Lunch	173
Figure 47. Spring 2014 Grade 5 Mathematics scree plot: Individualized Education Program ..	174
Figure 48. Spring 2014 Grade 6 Mathematics scree plot: Total	174
Figure 49. Spring 2014 Grade 6 Mathematics scree plot: Accommodated	175
Figure 50. Spring 2014 Grade 6 Mathematics scree plot: English Language Learner	175
Figure 51. Spring 2014 Grade 6 Mathematics scree plot: Free Lunch	176
Figure 52. Spring 2014 Grade 6 Mathematics scree plot: Individualized Education Program ..	176
Figure 53. Spring 2014 Grade 7 Mathematics scree plot: Total	177
Figure 54. Spring 2014 Grade 7 Mathematics scree plot: Accommodated	177
Figure 55. Spring 2014 Grade 7 Mathematics scree plot: English Language Learner	178
Figure 56. Spring 2014 Grade 7 Mathematics scree plot: Free Lunch	178
Figure 57. Spring 2014 Grade 7 Mathematics scree plot: Individualized Education Program ..	179
Figure 58. Spring 2014 Grade 8 Mathematics scree plot: Total	179
Figure 59. Spring 2014 Grade 8 Mathematics scree plot: Accommodated	180
Figure 60. Spring 2014 Grade 8 Mathematics scree plot: English Language Learner	180
Figure 61. Spring 2014 Grade 8 Mathematics scree plot: Free Lunch	181
Figure 62. Spring 2014 Grade 8 Mathematics scree plot: Individualized Education Program ..	181
Figure 63. Spring 2014 Grade 3 Reading scree plot: Total	182
Figure 64. Spring 2014 Grade 3 Reading scree plot: Accommodated	182
Figure 65. Spring 2014 Grade 3 Reading scree plot: English Language Learner	183
Figure 66. Spring 2014 Grade 3 Reading scree plot: Free Lunch	183
Figure 67. Spring 2014 Grade 3 Reading scree plot: Individualized Education Program.....	184
Figure 68. Spring 2014 Grade 4 Reading scree plot: Total	184
Figure 69. Spring 2014 Grade 4 Reading scree plot: Accommodated	185

Figure 70. Spring 2014 Grade 4 Reading scree plot: English Language Learner	185
Figure 71. Spring 2014 Grade 4 Reading scree plot: Free Lunch	186
Figure 72. Spring 2014 Grade 4 Reading scree plot: Individualized Education Program.....	186
Figure 73. Spring 2014 Grade 5 Reading scree plot: Total	187
Figure 74. Spring 2014 Grade 5 Reading scree plot: Accommodated	187
Figure 75. Spring 2014 Grade 5 Reading scree plot: English Language Learner	188
Figure 76. Spring 2014 Grade 5 Reading scree plot: Free Lunch	188
Figure 77. Spring 2014 Grade 5 Reading scree plot: Individualized Education Program.....	189
Figure 78. Spring 2014 Grade 6 Reading scree plot: Total	189
Figure 79. Spring 2014 Grade 6 Reading scree plot: Accommodated	190
Figure 80. Spring 2014 Grade 6 Reading scree plot: English Language Learner	190
Figure 81. Spring 2014 Grade 6 Reading scree plot: Free Lunch	191
Figure 82. Spring 2014 Grade 6 Reading scree plot: Individualized Education Program.....	191
Figure 83. Spring 2014 Grade 7 Reading scree plot: Total	192
Figure 84. Spring 2014 Grade 7 Reading scree plot: Accommodated	192
Figure 85. Spring 2014 Grade 7 Reading scree plot: English Language Learner	193
Figure 86. Spring 2014 Grade 7 Reading scree plot: Free Lunch	193
Figure 87. Spring 2014 Grade 7 Reading scree plot: Individualized Education Program.....	194
Figure 88. Spring 2014 Grade 8 Reading scree plot: Total	194
Figure 89. Spring 2014 Grade 8 Reading scree plot: Accommodated	195
Figure 90. Spring 2014 Grade 8 Reading scree plot: English Language Learner	195
Figure 91. Spring 2014 Grade 8 Reading scree plot: Free Lunch	196
Figure 92. Spring 2014 Grade 8 Reading scree plot: Individualized Education Program.....	196
Figure 93. Spring 2014 Grade 5 Science scree plot: Total	197
Figure 94. Spring 2014 Grade 5 Science scree plot: Accommodated	197
Figure 95. Spring 2014 Grade 5 Science scree plot: English Language Learner	198
Figure 96. Spring 2014 Grade 5 Science scree plot: Free Lunch	198
Figure 97. Spring 2014 Grade 5 Science scree plot: Individualized Education Program.....	199
Figure 98. Spring 2014 Grade 8 Science scree plot: Total	199
Figure 99. Spring 2014 Grade 8 Science scree plot: Accommodated	200
Figure 100. Spring 2014 Grade 8 Science scree plot: English Language Learner	200
Figure 101. Spring 2014 Grade 8 Science scree plot: Free Lunch	201
Figure 102. Spring 2014 Grade 8 Science scree plot: Individualized Education Program.....	201
Figure 103. Spring 2014 Grade 5 Social Studies scree plot: Total	202
Figure 104. Spring 2014 Grade 5 Social Studies scree plot: Accommodated	202
Figure 105. Spring 2014 Grade 5 Social Studies scree plot: English Language Learner	203
Figure 106. Spring 2014 Grade 5 Social Studies scree plot: Free Lunch	203
Figure 107. Spring 2014 Grade 5 Social Studies scree plot: Individualized Education Program.....	204
Figure 108. Spring 2014 Grade 8 U.S. History scree plot: Total.....	204
Figure 109. Spring 2014 Grade 8 U.S. History scree plot: Accommodated.....	205
Figure 110. Spring 2014 Grade 8 U.S. History scree plot: English Language Learner.....	205
Figure 111. Spring 2014 Grade 8 U.S. History scree plot: Free Lunch.....	206
Figure 112. Spring 2014 Grade 8 U.S. History scree plot: Individualized Education Program.	206

List of Appendices

Table A1. OCCT Test Blueprint and Actual Item Counts: Grade 3 Mathematics	207
Table A2. OCCT Test Blueprint and Actual Item Counts: Grade 4 Mathematics	208
Table A3. OCCT Test Blueprint and Actual Item Counts: Grade 5 Mathematics	209
Table A4. OCCT Test Blueprint and Actual Item Counts: Grade 6 Mathematics	210
Table A5. OCCT Test Blueprint and Actual Item Counts: Grade 7 Mathematics	211
Table A6. OCCT Test Blueprint and Actual Item Counts: Grade 8 Mathematics	212
Table A7. OCCT Test Blueprint and Actual Item Counts: Grade 3 Reading.....	213
Table A8. OCCT Test Blueprint and Actual Item Counts: Grade 4 Reading.....	214
Table A9. OCCT Test Blueprint and Actual Item Counts: Grade 5 Reading.....	215
Table A10. OCCT Test Blueprint and Actual Item Counts: Grade 6 Reading.....	216
Table A11. OCCT Test Blueprint and Actual Item Counts: Grade 7 Reading.....	217
Table A12. OCCT Test Blueprint and Actual Item Counts: Grade 8 Reading.....	218
Table A13. OCCT Test Blueprint and Actual Item Counts: Grade 5 Science	219
Table A14. OCCT Test Blueprint and Actual Item Counts: Grade 8 Science	220
Table A15. OCCT Test Blueprint and Actual Item Counts: Grade 5 Social Studies	221
Table A16. OCCT Test Blueprint and Actual Item Counts: Grade 7 Social Studies (Geography).....	222
Table A17. OCCT Test Blueprint and Actual Item Counts: Grade 8 Social Studies (U.S. History)	223

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
C³ Oklahoma’s Core curriculum, the College, Career and Citizen Ready
CCSSO Council of Chief State School Officers
CE Critical Element
CFA Confirmatory Factor 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
FN False Negative
FP False Positive
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
OAAP Oklahoma Alternate Assessment Program
OAC Oklahoma Administrative Code
OAS Oklahoma Academic Standards
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, grades 3 through 8 (OCCT 3–8) test administrations and provides data evidences in supporting the validity and reliability of the tests.

For the OCCT 3–8, Reading and Mathematics tests are administered in grades 3–8; Science, Social Studies, and Writing in Grade 5; Geography in Grade 7; and Science, U.S. History, and Writing in Grade 8. 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 Oklahoma Alternate Assessment Program (OAAP), or portfolio assessment, which includes all of the grades 3–8 contents.

Within the state of Oklahoma, the development of the *Priority Academic Student Skills (PASS)* and most recently the Oklahoma Academic Standards (OAS) 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 are all important steps in the process of developing a valid assessment system (Barton, 2007). This document serves to capture a small portion of the enormous amount of time and effort devoted to one of the OSTP assessments, the OCCT for grades 3 through 8, in relation to the importance, reliability, and validity of the assessment as part of the Oklahoma assessment system.

The American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education's (NCME) *Standards for Educational and Psychological Testing* (1999) are the guides by which this report provides various evidences of validity. The McGraw-Hill Education CTB work process in the Oklahoma OCCT 3–8 for Spring 2014 has paid close attention to the *Standards for Educational and Psychological Testing*, and this report provides evidence that is appropriate to a statewide summative assessment. Because the results of this assessment are used as part of state and federal accountability systems, attention has also been paid to the federal guidance provided in the Standards and Assessments Peer Review Guidance (US DOE, 2004). Evidence within this report also relates to the Critical Elements (CE) as part of the guidance for Peer Review.

Section 1—Overview

The Oklahoma Core Curriculum Tests (OCCT) assessments are state-mandated criterion-referenced tests that measure student proficiency in specific content areas. Each test has the purpose of measuring the student's knowledge relative to the Oklahoma Academic Standards, Oklahoma's content standards. In Spring 2014, the OCCT assessments were administered to all eligible students in grades 3 through 8. The OCCT covered: Mathematics and Reading for grades 3 through 8; Science and Writing for grades 5 and 8; and Social Studies for grades 5, 7 (Geography), and 8 (U.S. History). Along with the operational tests (OP), other form variations were administered for the OCCT: equivalent forms (EQ), braille forms (BR), and large-print forms. Field test forms were administered for Science grades 5 and 8, and Social Studies grades 5, 7, and 8.

In the Fall of 2012, McGraw-Hill Education CTB was contracted by the Oklahoma State Department of Education (SDE) to develop, administer, and maintain the Oklahoma School Testing Program (OSTP) OCCT and Oklahoma Modified Alternate Assessment Program (OMAAP) for Achieving Classroom Excellence (ACE) End-of-Instruction (EOI) and grades 3–8. The purpose of this technical report is to provide objective information regarding technical aspects of the OCCT 3–8 assessments by specifying the technical details of the work accomplished from Summer 2013 through the end of Spring 2014 on these tests. This volume is intended to be one source of information to Oklahoma K–12 educational stakeholders (including testing coordinators, educators, parents, and other interested citizens) about the development, implementation, scoring, and technical attributes of the OCCT 3–8 assessments.

Other sources of information regarding the OSTP-ACE G3–8 tests include the administration manual *OSTP 2013–2014 Test Preparation Manual* found at: http://ok.gov/sde/sites/ok.gov.sde/files/documents/files/2705543-W_tpm_w13OK.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 teachers, students, and parent guides found at: <http://ok.gov/sde/assessment-administrator-resources-administrators>.

The Spring 2014 OCCT 3–8 field test items for the Science grades 5 and 8, and Social Studies grades 5, 7, and 8 assessments were developed by McGraw-Hill Education CTB in collaboration with the Oklahoma SDE. The operational assessments for Mathematics and Reading grades 3–8, Science grades 5 and 8, and Social Studies grades 5, 7, and 8 were developed by McGraw-Hill Education CTB in collaboration with the SDE, and were administered by the SDE. Note that there were other forms applied concurrently with the operational forms in each of the above administrations and for each of the contents: the BR form, the large-print form, and the EQ form.

Section 1.1—Purpose

This report includes only data and analyses for the operational forms and contents for the Spring 2014 administration. It begins with a description of the Oklahoma content standards, which are

described in **Section 1.2—Oklahoma Academic Standards**. All operational and field test items for OCCT 3–8 Spring 2014 were subjected to cycles of reviews by the SDE and Pearson or by McGraw-Hill Education CTB. A description of the item development process, along with a description of the alignment process and test development, is presented in complete detail in **Section 2—Item and Test Development**. A detailed description of the administration processes is found in **Section 3—Administration**, and a discussion of the operational population and the research samples utilized in the analysis is found in **Section 5—Sampling Plan & Field Test Design**.

The Spring 2014 OCCT 3–8 scores for Mathematics and Reading grades 3 through 8, and Science grades 5 and 8 were based on a post-equating design, except for Reading Grade 3 which was based on pre-equating. The Social Studies grades 5 and 8 items were analyzed for new scaling. The Grade 7 Social Studies were field test items only. A complete description of the operational and field test item analyses and the calibration/scaling and equating analysis is found in **Section 6—Methods** and **Section 7—Results**.

Section 1.2—Oklahoma Academic Standards

McGraw-Hill Education CTB developed the Spring 2014 Oklahoma OCCT 3–8 assessments to measure the Oklahoma Academic Standards shown in Table 1. The objectives associated with the content and/or process standards tested are provided in **Appendix A**.

Table 1. Testable Standards for OCCT Grades 3–8

Mathematics Grades 3–8	
Standard 1.	Algebraic Reasoning: Patterns and Relationships
Standard 2.	Number Sense and Operation
Standard 3.	Geometry
Standard 4.	Measurement
Standard 5.	Data Analysis
Reading Grades 4–8 (Grade 3 in parentheses)	
Standard 1. (Standard 2.)	Vocabulary
Standard 3. (Standard 4.)	Comprehension/Critical Literacy
Standard 4. (Standard 5.)	Literature
Standard 5. (Standard 6.)	Research and Information
Science Grades 5 & 8	
Process/Inquiry Standards and Objectives	
Process 1.	Observe and Measure
Process 2.	Classify
Process 3.	Experiment
Process 4.	Interpret and Communicate

Table 1. Testable Standards for OCCT Grades 3–8 (*continued*)

Grade 5 Content Standards	
Standard 1.	Properties of Matter and Energy
Standard 2.	Organisms and Environments
Standard 3.	Structures of the Earth and the Solar System
Grade 8 Content Standards	
Standard 1.	Properties and Chemical Changes in Matter
Standard 2.	Motion and Forces
Standard 3.	Diversity and Adaptations of Organisms
Standard 4.	Structures/Forces of the Earth/Solar System
Standard 5.	Earth’s History
Social Studies Grade 5	
Standard 1.	James Towne Settlement and Plimoth Plantation
Standard 2.	Colonial America
Standard 3.	American Revolution
Standard 4.	Early Federal Period
Social Studies Grade 7 (Geography)	
Standard 1.	Geographic Tools/Geography Skills
Standard 2.	Human and Physical Characteristics of Regions
Standard 3.	Patterns of the Earth
Standard 4.	Human Systems
Standard 5.	Human/Environment Interaction
Social Studies Grade 8 (U.S. History)	
Standard 1.	Causes and Events of the American Revolution
Standard 2.	Foundations and Founders of the American Nation
Standard 3.	Developing the American Government System
Standard 4.	The Transformation of the United States to the Mid-1800s
Standard 5.	Causes, Events, and Leadership in the Civil War

Section 2—Item and Test Development

In the Spring 2014 administration, there was one operational form for the tests administered for Mathematics and Reading grades 3–8, and there was one operational form with embedded field-test items for Science grades 5 and 8, Grade 5 Social Studies and Grade 8 U.S. History. For the Grade 7 Geography test, field test forms were administered instead. There were six field test forms for Science grades 5 and 8, and Social Studies grades 5 and 8. There were two field test forms for Grade 7 Geography. A braille form and an equivalent form were produced for Mathematics and Reading grades 3–8, Science grades 5 and 8, and Social Studies grades 5 and 8. Because it was a field test year, no Equivalent form was produced for Grade 7 Geography.

The braille form is usually a mirror of the operational form. The equivalent is designated as a breach form. A student could receive an equivalent form for various reasons, including becoming ill during the test administration or experiencing any kind of security breach. The State Department of Education Office of Accountability and Assessments determines eligibility for an Equivalent form on a case-by-case basis.

Section 2.1—Aligning Test to Oklahoma Academic 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, or test questions appropriate for the content being assessed (e.g., multiple-choice or open-ended).

A list of the assessable standards for each subject is provided in Table 2. For Mathematics and Reading, the same assessable standards appear in each grade level.

Table 2. Testable Standards for OCCT Grades 3–8

Mathematics Grades 3–8	
Standard 1.	Algebraic Reasoning: Patterns and Relationships
Standard 2.	Number Sense and Operation
Standard 3.	Geometry
Standard 4.	Measurement
Standard 5.	Data Analysis

Table 2. Testable Standards for OCCT Grades 3–8 (*continued*)

Reading Grades 4–8 (Grade 3 in parentheses)	
Standard 1. (Standard 2.)	Vocabulary
Standard 3. (Standard 4.)	Comprehension/Critical Literacy
Standard 4. (Standard 5.)	Literature
Standard 5. (Standard 6.)	Research and Information
Science Grades 5 & 8	
Process/Inquiry Standards and Objectives	
Process 1.	Observe and Measure
Process 2.	Classify
Process 3.	Experiment
Process 4.	Interpret and Communicate
Grade 5 Content Standards	
Standard 1.	Properties of Matter and Energy
Standard 2.	Organisms and Environments
Standard 3.	Structures of the Earth and the Solar System
Grade 8 Content Standards	
Standard 1.	Properties and Chemical Changes in Matter
Standard 2.	Motion and Forces
Standard 3.	Diversity and Adaptations of Organisms
Standard 4.	Structures/Forces of the Earth/Solar System
Standard 5.	Earth’s History
Social Studies Grade 5	
Standard 1.	James Towne Settlement and Plimoth Plantation
Standard 2.	Colonial America
Standard 3.	American Revolution
Standard 4.	Early Federal Period
Social Studies Grade 7 (Geography)	
Standard 1.	Geographic Tools/Geography Skills
Standard 2.	Human and Physical Characteristics of Regions
Standard 3.	Physical Systems of the Earth
Standard 4.	Human Systems
Standard 5.	Human/Environment Interaction
Social Studies Grade 8 (U.S. History)	
Standard 1.	Causes and Events of the American Revolution
Standard 2.	The Revolutionary Era
Standard 3.	Developing the American Government System
Standard 4.	The Transformation of the United States to the Mid-1800s
Standard 5.	Causes, Events, and Leadership in the Civil War

Blueprints

The test blueprint defines the proportion of the content to be covered on the test that best reflects the proportional importance and coverage of standards in the classroom.

In addition to the test Blueprints provided by the SDE (see <http://www.ok.gov/sde/test-support-teachers-and-administrators> for blueprints), Table 3 describes four criteria for test alignment with the Oklahoma Academic Standards and its objectives.

Table 3. Criteria for Aligning the Test with Oklahoma Academic Standards Standards and Objectives

1. Categorical Concurrence	The test is constructed so that there are at least six items measuring each OAS 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 OAS 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 OAS 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 OAS 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 McGraw-Hill Publishing CTB staff had the opportunity to review actual student performances on the newly-developed and field-tested multiple-choice (MC) items across the subjects based on the Spring 2014 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) or NCLB. The purpose of the review meeting was to ensure that psychometrically sound, fair, and aligned items are used in the construction of the OCCT 3–8 assessments and entered into the respective item banks. McGraw-Hill Education CTB provided technical and psychometric expertise to provide 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 McGraw-Hill Education CTB. The SDE administrators and content specialists attended the meetings facilitated by McGraw-Hill Education CTB. CTB content specialists and research scientists provided training to the SDE staff on how to interpret and review the field test data. Meeting materials included: a document explaining the flagging criteria about item quality (e.g., p -value, model-fit, etc.), a document containing flagged items and item images. McGraw-Hill Education CTB discussed with the SDE the analyses performed and the criteria for flagging items. Flagged items were then reviewed, and decisions were made as to whether to accept the item, keep 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) model-fit. 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 left in the bank for future use. While 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 OCCT 3–8 assessment items are as follows:

- p -value $<.25$ or $>.90$
- point-biserial correlation $<.20$
- 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**)

Section 2.2—Item Pool Development and Selection

To ensure content validity of the Oklahoma OCCT 3–8 tests, McGraw-Hill Education CTB content experts closely studied the OAS Standards and worked with Oklahoma content area specialist, teachers, and assessment experts to gather a pool of existing items that measure Oklahoma’s Assessment Frameworks (i.e., OAS) for each subject. Once the need for field test items was determined, based on the availability of items for future test construction, a pool of items that measured Oklahoma Academic Standards in each subject was developed. These items were developed under universal design guidelines set by the SDE and carefully reviewed and discussed by Content and Bias/Sensitivity Review Committees to evaluate not only content validity, but also plain language and the quality and appropriateness of the items. These committees were comprised of Oklahoma teachers and SDE staff. The committee’s recommendations were used to select and/or revise items from the item pool used to construct the field test portions of the Spring 2014 assessments.

The source of the operational items included a pool of previously field-tested or operationally-administered items ranging from the Spring 2006 through the Spring 2013 administrations for Mathematics grades 3–8, Reading grades 3–8, and Science grades 5 and 8. The items were

calibrated live using data from the operational administrations to estimate parameters for these items.

Item selection and form development for Spring 2014 were completed as a collaborative effort between staff at the SDE and McGraw-Hill Education CTB 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 item response theory statistics, described in **Section 6—Methods**, editors selected items with the best content-relevant and statistical characteristics.

The OCCT 3–8 operational tests for the Spring 2014 cycle were built by including previously field tested and operationally used items. Content experts also targeted the pre-defined percentage of items measuring various Depth of Knowledge (DOK) levels for assembling the tests. Table 4 provides the DOK level percentages for the Spring 2014 operational assessments.

Table 4. Percentage of Items by Depth of Knowledge Levels for OCCT 3–8 Assessments

Subject	Grade	DOK Level					
		1		2		3	
		Target	Actual	Target	Actual	Target	Actual
Mathematics	3	20–25	20	65–70	70	5–15	10
	4	20–25	22	65–70	70	5–15	8
	5	20–25	20	65–70	66	5–15	14
	6	10–15	14	65–70	72	15–25	14
	7	10–15	18	65–70	64	15–25	18
	8	10–15	12	65–70	64	15–25	24
Reading	3	20–25	12	65–70	64	5–15	24
	4	20–25	16	65–70	70	5–15	14
	5	20–25	14	65–70	70	5–15	16
	6	10–15	3	65–70	70	15–25	22
	7	10–15	10	65–70	66	15–25	24
	8	10–15	8	65–70	72	15–25	20
Science	5	20–25	16	65–70	66	5–15	18
	8	10–15	16	65–70	53	15–25	31
Social Studies	5	20–25	24	65–70	66	5–15	10
	8	10–15	12	65–70	68	15–25	20

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 to avoid inclusion of items that may be biased in any

manner against a group, which would lead to inequitable test results. As described earlier, all field test items were analyzed statistically for DIF using the field test data. At the data review meetings, a McGraw-Hill Education CTB research scientist explained the significance of DIF, 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 (for example, cultural barriers) or language in an item that might result in bias (i.e., 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 Academic Standards standard and objective, and the DOK intended. The items were also reviewed to ensure 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 biasing 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 (Holland and Thayer, 1988; Michaelides, 2008). The results for Spring 2014 are found in **Section 7—Results** of this report.

Section 2.3—Configuration of the Spring 2014 Tests

For Spring 2014, McGraw-Hill Education CTB Content Development selected items from the available item pools that had been previously field tested and approved by the SDE staff for usage on the operational assessments. The operational items on the Mathematics and Reading tests had appeared previously in the years 2008–2013. The operational items on the Science tests had appeared previously in the years 2007–2013. No operational items appeared on the Grade 7 Geography test. Field test items were selected from items that were approved by the SDE staff and Oklahoma teachers for Science grades 5 and 8, and Grade 5 Social Studies and Grade 8 U.S. History. Most of the field test items on the Science and Social Studies tests had never been field tested; however, some items had been previously field tested, but required revisions and additional field testing. McGraw-Hill Education CTB Research analyzed the selected items and provided feedback to Content Development regarding the best set of items to serve as the Spring 2014 operational form.

Section 2.4—Operational and Field Test Items by Content Area

Table 5 provides an overview of the number of operational and field test items that composed the Spring 2014 OCCT 3–8 assessments. The Spring 2014 test was composed of one core operationally scored form for each subject, except for Grade 7 Geography. Field test items were embedded in the operational test forms for Science and Social Studies to build the item bank for future use. No field test items were used in the Mathematics and Reading tests. The forms in the Spring 2014 assessments were randomly assigned within classrooms to obtain randomly equivalent samples of examinees for the field test items.

Table 5. Configuration of the OCCT 3–8 Tests for Spring 2014

Subject	Grade	Forms	Item Counts (Per Form)			Maximum Possible Points on OP Test Items (Per Form)	
			OP	FT	Test	OP	FT
			MC	MC		MC	MC
Mathematics	3	1	50	.	50	50	.
	4	1	50	.	50	50	.
	5	1	50	.	50	50	.
	6	1	50	.	50	50	.
	7	1	50	.	50	50	.
	8	1	50	.	50	50	.
Reading	3	1	50	.	50	50	.
	4	1	50	.	50	50	.
	5	1	50	.	50	50	.
	6	1	50	.	50	50	.
	7	1	50	.	50	50	.
	8	1	50	.	50	50	.
Science	5	1	45	10	55	45	10
	8	1	45	10	55	45	10
Social Studies	5	1	50	10	60	50	10
	7	1	50	10	60	.	10
	8	1	50	10	60	50	10

Note: OP = Operational; FT = Field Test; MC = Multiple Choice.

Section 3—Administration

To ensure a valid and reliable assessment, the OCCT 3–8 assessments are first constructed in alignment with the *Oklahoma C³* Standards (now called the Oklahoma Academic Standards) by the Oklahoma SDE in collaboration with McGraw-Hill Education CTB. 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 3–8 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 3–8 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 3–8 assessment materials, McGraw-Hill Education CTB’s Transportation Department maintains the quality and security of material distribution and return by hiring reputable carriers that possess the ability to trace shipments. McGraw-Hill Education CTB 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 Oklahoma WAVE system. Oklahoma educators also use this system to provide McGraw-Hill Education CTB 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 McGraw-Hill Education CTB 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. McGraw-Hill Education CTB utilizes double-column boxes to distribute and collect test materials, and makes additional cartons available for 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 McGraw-Hill Education CTB's Data Processing Facility.

Section 3.3—Materials Discrepancies Process

The scanning process allows McGraw-Hill Education CTB to capture MC 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 McGraw-Hill Education CTB Oklahoma Help Desk, and reported. The results are subsequently reported to the SDE.

A predetermined date is set by the SDE and McGraw-Hill Education CTB 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 McGraw-Hill Education CTB 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. McGraw-Hill Education CTB takes security seriously and makes every effort to recover missing material.

Section 4—Scoring

The Oklahoma Spring 2014 OCCT grades 3–8 test books included items that were machine scored (MC items), and extended-writing prompts (WPs) that were scored by trained human or handscorers (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 3–8 test. Writing is assessed at grades 5 and 8. Each writing response receives two types of scores. First, a series of analytic scores focus on specific writing traits. These traits receive scores of one to four. 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. Condition codes are used if the student’s writing response is unscorable. The results are reported with the MC results.

Scoring Rubrics

Analytic scoring rubrics were provided by the Oklahoma SDE. The rubrics focus on five specific writing skills: 1) Ideas and Development, 2) Organization, Unity, and Coherence, 3) Word Choice, 4) Sentences and Paragraphs, and 5) Grammar, Usage, and Mechanics. Each is rated from four (the highest score) to one (the lowest score). In addition, the skill areas of Ideas and Development, and Organization, Unity, and Coherence require that the composition be written in a mode that is appropriate for both audience and purpose. Three modes are emphasized: Opinion/Argument, Informative, and Narrative.

Anchor Papers

The 2014 Operational Writing prompts for grades 5 and 8 were new and required range finding. Prior to scoring, the prompts underwent extensive range-finding with two SDE representatives, on March 27–28, 2014, in Oklahoma City. Rubrics and Writing prompts were reviewed. Anchor candidates were discussed and final anchors selected for use during the training and scoring activities. Three anchor papers per score point were selected for each writing skill area. The range-finding discussions were helpful in defining the parameters of the scoring requirements in the analytic Writing rubrics and in providing insights and summary statements for training of raters.

Section 4.1—Handscoring

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. The operational prompts for grades 5 and 8 were scored concurrently by two separate groups of raters. Raters and team leaders were trained using the following steps:

- Provide a general introduction to OCCT 3–8
- 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 McGraw-Hill Education CTB 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 McGraw-Hill Education CTB 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, McGraw-Hill Education CTB'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

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

McGraw-Hill Education CTB 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 McGraw-Hill Education CTB 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 McGraw-Hill Education CTB'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—Summary of Reliability and Validity describes the outcomes of inter-rater reliability analysis. The inter-rater reliability coefficients for the operational writing prompts are presented in Table 6.

Table 6. Inter-Rater Reliability Coefficients for the Operational Writing Prompts

Grade	Trait	PEID Item ID	Score Points	% of Agreement			Checkset Average Agreement Percentages
				Perfect	Adjacent	Perfect + Adjacent	
5	A	01556414	1–4	0.53	0.36	0.89	0.71
	B	01556416	1–4	0.54	0.36	0.90	0.72
	C	01556417	1–4	0.53	0.36	0.89	0.72
	D	01556418	1–4	0.54	0.36	0.90	0.71
	E	01556419	1–4	0.53	0.37	0.90	0.71
8	A	01556420	1–4	0.50	0.39	0.89	0.71
	B	01556422	1–4	0.50	0.39	0.89	0.71
	C	01556423	1–4	0.51	0.38	0.89	0.72
	D	01556424	1–4	0.51	0.38	0.89	0.73
	E	01556425	1–4	0.50	0.38	0.88	0.72

Section 5—Sampling Plan & Field Test Design

Section 5.1—Sampling Plan

A sample representative of the population of Oklahoma students was used for Spring 2014 equating because final scale scores and performance levels should be reported within two weeks of the closed testing window. To meet this 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 Research conducted a data integrity check and compared the sample selection (expected) to the 2013 sample to assure that the sample was representative. Table 7 shows equating sample size and respective percentage of the population for Mathematics grades 3–8, Reading grades 4–8 and Science grades 5 and 8 of the Spring 2014 administration. For grades 3, 4, and 5, samples were used for equating; while for grades 6, 7, and 8, almost whole population datasets were used. Grades 3, 4, and 5 students took paper-pencil tests and grades 6, 7, and 8 students took online tests. CTB Research received paper-pencil tests results last due to necessary pickup and transit time.

Table 7. Equating Sample Size for Spring 2014 and Respective Percentage of the Population for Each Grade and Content

Content	Grade	Samples	Population	Percent
Mathematics	3	17254	48887	35%
	4	15687	48758	32%
	5	14722	48299	30%
	6	43599	47474	92%
	7	42887	46374	92%
	8	34019	37626	90%
Reading	4	15723	48656	32%
	5	14768	48206	31%
	6	43724	47475	92%
	7	44106	47402	93%
	8	43719	47330	92%
Science	5	20425	48291	42%
	8	16830	47451	35%

Note: Reading Grade 3 was pre-equated; Social Studies and U.S. History used whole population for scaling.

Table 9, in the **Tables** section, provides the proportion of students in the sample and within the Spring 2014 population that came from each of the subgroups: gender, ethnicity, special population (ELL, IEP, Section 504, and accommodated), and socio-economic status (SES Low and SES High). SES Low flag is for students who received free lunch. It is clear from these tables that the sample is also representative of the state's population, even across most of the subgroups, with the exception of American Indian/Alaskan, which is overrepresented in grades

3–5 Mathematics. The differences between the sample and the state tend to be less than +/- 5% with median difference of 0.07 (absolute value).

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 where 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. Ten multiple-choice field test items per form were placed in common positions across Science, U.S. History, and Social Studies forms. Geography form was built with field test items only.

Section 5.3—Data Checking Activities

During the field test data analysis, CTB Research conducted detailed data checking and applied the following data cleaning exclusionary rules.

5.3.1 Suppressed/Omitted/Invalidated cases

Eliminate suppressed/omitted/invalidated cases flagged in the WinScore files. Eliminate cases that have five or fewer valid attempts.

5.3.2 Duplicate cases

Check and eliminate any duplicate cases by checking student ID (if available), first and last name, middle initial, GIS_CD (GIS code normally contain the district and school ID), teacher name, school, birthday, gender, and response vectors.

5.3.3 Non-public schools

The non-public schools are 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

Exclude students taking the test for the second time.

Section 6—Methods

The Spring 2014 OCCT 3–8 program was based on the application of a post-equating method using anchor items and equating samples except for Reading Grade 3, based on pre-equating. 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. McGraw-Hill Education CTB 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 three-parameter logistic (3PL) item response theory (IRT) model for the MC items. The following section describes the methods used in the analyses of the operational test items.

Section 6.1—Classical Item Analyses

Item Level Analyses

Each 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 too difficult because fewer than 25% of the students are achieving the correct answer. P -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.

It is also important to track the rate at which students do not respond to, or omit, items. Omitted items receive a zero score. The rate of omission often provides some information about testing times, or speededness, particularly if there is a high rate of items omitted at the end of a test session. It also provides an indication of items that may simply be unclear or illogically presented. When more than 5% of students omit an item, the item is reviewed by both CTB Research and Content Development and shared with the SDE.

A summary comparison of the classical statistics between the Spring 2012, Spring 2013, and Spring 2014 OCCT 3–8 results is presented in Table 10. 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, except Mathematics Grade 8 which had a p -value difference of -0.09. 2014 mean p -values were lower than those of 2013 because the OMAAP test was not administered in 2014; therefore, OMAAP students took OCCT tests in 2014. Also, high performing Grade 8 students took the OCCT EOI Algebra I test for the Spring 2014 administration.

A detailed summary of the item level classical raw score statistics and omission rates for Spring 2014 and a comparison to Spring 2013 is provided in **Section 7—Results**.

Section 6.2—Differential Item Functioning (DIF)

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 or skills or statistical Type I errors. 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. For OCCT 3–8, DIF analyses are conducted across gender (males/females) and ethnicity—focal subgroups African American (not Hispanic), American Indian/Alaskan Native, Asian, Hispanic, and Multiracial versus the reference group White (not Hispanic).

The Mantel-Haenszel DIF statistic (Holland and Thayer, 1988; Michaelides, 2008) was used for the OCCT 3–8 operational tests. It matches students across the reference and focal groups based on their overall test performance, and provides a chi-square statistic 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

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 on the MC items.

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 3–8 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 is an item. An item will be more difficult if the difficulty parameter is higher. 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

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

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.

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). 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 at least 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 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

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

with

$$E_{ik} = \frac{1}{N_{ik}} \sum_{j \in \text{cell } k}^{N_{ik}} P_i(\hat{\theta}_j). \quad (3)$$

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

$$df = I(m_j - 1) - m_j, \quad (4)$$

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_{Ij} was transformed to Z_{Q_I}

where

$$Z_{Q_I} = (Q_{Ij} - df) / (2df)^{1/2}. \quad (5)$$

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 $Z_{Q_I} \text{Crit}$ used to flag items was calculated using the expression

$$Z_{Q_I} \text{Crit} = \left(\frac{N}{1500} \right) * 4, \quad (6)$$

where N is the calibration sample size.

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

Section 6.4—Equating

Test Scaling and Equating

Once all item-level analyses were conducted, Spring 2014 OCCT 3–8 Mathematics, Reading, and Science forms were 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 2014 operational forms by statistics and content. CTB PARDUX (version 1.66, 2011) software was applied to equating. Social Studies and U.S. History underwent a method of scaling. TCC plots for Mathematics, Reading, Science, Social Studies, and U.S. History are found in Figures 17–32.

Stability of Anchor Item

Stability of anchor items for equating procedure is important. The following method was applied to evaluate and refine the anchor item sets before 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 outlying anchor 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 item 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 before and after equating.
- 3) An outlier for a -parameter or b -parameter can be a candidate based on anchor item plot, which shows the relationships of anchor item parameters before and after equating (Kolen and Brennan, 2004).
- 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 Reading and Mathematics.
- 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, McGraw-Hill Education CTB recommended that 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 were still scored as part of the whole test. After an anchor item was removed from the anchor set based on the above criteria, the anchor file needed to be adjusted and a second version of the calibration and equating were produced. All outputs in the second version were to be evaluated following

the same guidelines as the original calibration runs. No anchor items were flagged and removed in the Spring 2014 administration data analyses.

Section 6.5—Writing Scoring

Writing prompts were administered in Spring 2014. Students in grades 5 and 8 responded to one operational writing prompt. The writing score is a weighted composite of five analytic scores that focus on specific domains of writing skills. These skills are listed in Table 8. Each student's response to a prompt is read by two independent raters; the raters' scores for each domain are averaged. The domain scores range from 1 (the lowest score) to 4 (the highest score).

Table 8. Writing Analytic Traits and Scoring Weights

Writing Analytic Traits	Weight
Ideas and Development (ID)	30%
Organization, Unity, and Coherence (OUC)	25%
Word Choice (WC)	15%
Sentences and Paragraphs (SP)	15%
Grammar, Usage, and Mechanics (GUM)	15%

The composite score (CS) is calculated as a weighted composite of the average of two independent ratings for each of the five analytic traits:

$$CS = 15(0.30ID + 0.25OUC + 0.15WC + 0.15SP + 0.15GUM)$$

2014 Writing prompts were selected and modified from 2013 Field test items. To place 2014 Writing prompts on 2013 scale, an equipercentile-linking method was applied. Concordance tables from this method were applied to produce final 2014 operational scores.

Detailed information can be found in **Appendix B**.

Section 7—Results

This section provides the data analysis results for the Spring 2014 OCCT 3–8. 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. There were no item suppressions in the operational OCCT 3–8 tests.

Section 7.1—Item Level Performance

A summary comparison of the classical item statistics between the Spring 2012, Spring 2013, and Spring 2014 OCCT 3–8 results is presented in Table 10. Typically, differences of less than about $|0.05|$ are desirable. As can be seen in Table 10, between Spring 2013 and Spring 2014, p -values had a slight decrease across grades and content areas, with the largest differences seen in Mathematics Grade 8 (-0.09). OMAAP tests were not administered in the 2014 administration. The decrease of mean p -values stems from the fact that most OMAAP students were low performance students. Many high performing Grade 8 students took OCCT EOI Algebra I in the 2014 administration, and this caused much lower mean p -values for Mathematics Grade 8. The mean item-test correlation showed the largest change in Reading Grade 7.

A summary of the range of p -values and item-test correlations for all operational and field test items for Spring 2014 is presented in Table 11. (Note that 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.) For analysis, the Writing trait was treated as an item. As shown in Table 11, the average p -values for the operational test items are from the mid 0.50s to mid 0.70s in Mathematics; from the high 0.60s to mid 0.70s in Reading; 0.70 and 0.60 in Science grade 5 and grade 8 respectively; 0.63 in Social Studies; 0.61 in U.S. History; and in the high 0.50s to mid 0.60s in Writing. The range of the p -values dips below 0.25 in Mathematics Grade 7 and 8, and Science Grade 8. Item-test correlations across most grades and content areas are within typical and acceptable ranges; except for Science Grade 5 where one item shows an item-test correlation lower than 0.15. For the field test items, the average p -values are in the high 0.40s and mid 0.50s for Science, and mid 0.40s for Social Studies and U.S. History. The item-test correlations for field test items are in the mid 0.20s to low 0.30s for Science, Social Studies, U.S. History, and Geography.

The item omission rates for operational and field test items are presented in Table 12. The operational items for Mathematics, Reading, Science, Social Studies, and U.S. History had omission rate across grades 4–8 less than 1.5% (well below the 5% criteria), indicating acceptable administration times for the number of items in each test session. The MC field test items for Science, Social Studies, U.S. History, and Geography show omission rates well below the 5% criteria for the MC items. For Writing grades 5 and 8, the omission rates were 0.00% because only a few students did not try Writing prompts.

DIF results are reported for Mathematics, Reading, Science, Social Studies, U.S. History, and Geography in Table 13 for gender and Tables 14–16 for ethnicity. The results indicate that the

majority of operational test items did not exhibit bias. For operational items on gender DIF, there were a total of 21 items (2.63%) flagged for moderate “B” DIF and 1 item (0.13%) flagged for severe “C” DIF. For operational items in the four ethnicity groups DIF analyses, there were between 0.88% to 7.00% of items flagged for moderate “B” DIF and between 0.13% and 1.50% of items flagged for severe “C” DIF. There is no table for ethnicity group American Indian/Alaskan Native included as there were no flagged items reported.

DIF results for the field test items in Mathematics, Reading, Science, Social Studies, U.S. History, and Geography show that for gender DIF, there were a total of 4 items (1.51%) flagged for moderate “B” DIF and no items flagged for severe “C” DIF. For field test items in the four ethnicity groups included in the DIF analyses, there were between 1.13% to 6.42% of items flagged for moderate “B” DIF, and between 0.38% and 3.40% of items flagged for severe “C” DIF. There is no table for ethnicity group American Indian/Alaskan Native included as there were no flagged items reported.

All of the items flagged were reviewed by CTB 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 flagged operational and/or field test items in future test forms.

Problematic Items

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 must be 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 2014 OCCT 3–8 operational and field test data analyses, items were reviewed for their classical statistics, and when those statistics were outside the range of acceptable 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 for a new standard or new approach that was expected to be difficult). For the Spring 2014 OCCT 3–8 operational test, there were no items suppressed. Out of the 284 field-tested items, the ones with less than desirable p -values and item-test correlations were reviewed by McGraw-Hill Education CTB content experts and Research, and 19 were considered to have less than desirable statistics and were suppressed from the item pool.

Section 7.2—Standards Level Performance

A review of the item difficulty across standards within each grade and content area is provided to illustrate at which standards items were more or less difficult for students. The summaries are presented in Tables 17–22. The tables provide the number of operational items, the reliability (coefficient alpha), and standard error of measurement (SEM) (formulas for which are found in **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.

The reliability at each standard, which is influenced by the number of items contributing to each standard, ranges from 0.49 to 0.84 in Mathematics, from 0.36 to 0.84 in Reading, and from 0.47 to 0.72 in Science. Ranges for Social Studies were from 0.48 to 0.73 and for U.S. History from 0.43 to 0.77. Across all content areas, the standard errors are no greater than 2.13 (Reading Grade 3 is the highest), and the maximum amount of IRT information is 0.22. IRT locations should be reviewed within each grade by standard, as should the p -values.

IRT locations and p -values can also be reviewed within each grade by standard in Tables 17–22 as well. The IRT scale locations provide an indication as to the average b -parameters or location values of a set of items contributing to each of the standards. The distinction from average p -values is that the IRT locations provide information about where the items are found along the scale score continuum, such that higher values indicate a lower probability of students with low estimated ability of answering 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 OCCT 3–8 applies a number-correct scoring method based on the 3PL IRT model, which is used to estimate scale scores corresponding to each raw score. In this method, all students who have the same raw score get the same scale score regardless of which items are correct.

Tables 23 and 24 provide the state-level scale score summary statistics across grades and content areas for Spring 2013 and Spring 2014, respectively. (Spring 2014 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**.) The Spring 2013 results were shown as a reference and for comparison purpose. Histograms and associated skewness and kurtosis of the data for Spring 2014 Mathematics, Reading, Science, Social Studies, and U.S. History are provided in Figures 1–16. 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 Tables 25–33. An independent sample *t*-test was conducted on the mean differences between accommodated and non-accommodated students, ELL and non-ELL students, gender groups, IEP and non-IEP, Section 504 and non-Section 504, and Low/High SES 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. Results of the *t*-tests and ANOVAs are found in Tables 34–39 and Tables 40–44, respectively.

As presented in Tables 34–39, results of the *t*-tests show that females outperform males in most grades and contents, except at Mathematics Grade 3. Mean differences are not statistically significant at Mathematics Grade 4, Science Grade 8 and Social Studies Grade 5. At the subgroup level, the results show that accommodated students tend to have the lowest performance of the “special population” subgroups. Overall results of the *t*-tests within each category indicate that accommodated, ELL, IEP, and Low SES students all score significantly lower than the rest of the population in all grades and content areas (mean differences ranging from 34 to 124 fewer scale score points), as expected. For the Section 504 group, the same is true in all grades and content areas, but the average differences compared with the rest of the state are less than 8 scale score points.

Statistically significant differences exist between the ethnicity groups in all content areas and grades as presented in the ANOVA results in Tables 40–44.

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 Tables 45–49. In comparing ethnicities across all grades and content areas, students identified as Asian and White (not Hispanic) tended to outperform the other ethnicities. As shown in Tables 45–49, Asian outperformed all other ethnicities in all content areas and grades. White (not Hispanic) outperformed all other ethnicities, except Asian, in all content areas and grades. Most pairs were significantly different, with the following exceptions which were **not** significant: American Indian/Alaskan Native is not significantly different from Multiracial in Mathematics and Reading grades 3, 7, and 8.

Section 7.4—Proficiency Level Performance

Table 50 shows the scale score means and standard deviations for the state and for students in each proficiency level. Tables 51 and 52 provide 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 and 2014, respectively. Table 51 provides Spring 2013 data as a reference. Tables 50–52 do not include the number of students considered *Undetermined* (invalid) in the denominator of calculation.

Impact data across proficiency levels are also provided for each gender, ethnicity, and special population subgroups in Tables 53–57, where comparative performance across subgroups mimics what was provided for the scale score descriptions.

Section 8—Summary of Reliability & Validity

This section summarizes some of the evidence in the earlier sections and provides additional evidence to support the degree to which the OCCT 3–8 tests are reliable and valid. For the OCCT 3–8, 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 *OSTEP 2013–2014 Test Preparation Manual* (http://ok.gov/sde/sites/ok.gov.sde/files/documents/files/2705543-W_tpm_w13OK.pdf) under General Guidance, which 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 hand scorers, as was described in **Section 4—Scoring**. Table 6 for grades 5 and 8 provide the relevant inter-rater reliability 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, a type of internal consistency measure, is one measure of the correlation between each item and the overall test as described in **Section 6—Methods**, results of which were described in **Section 7—Results**. The item-test correlations for each content area, grade, and item type are shown in Table 63. The operational item-test average correlations range from 0.39 to 0.44 (Mathematics); from 0.40 to 0.45 (Reading); 0.37 and 0.39 (Science); 0.38 (Social Studies); 0.41 (U.S. History); and 0.94 and 0.97 (Writing). One operational item in Science Grade 5 presented an item-test correlation of 0.15. That item was investigated by Content Development and found to be correctly scored. Any operational items with extremely low point biserial that may remain in the OCCT 3–8 item pool will be avoided in 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 some bias or systematic over- or under-representation of subgroup performance when compared to the total group performance. As discussed in **Section 7—Results** and is apparent in Tables 13–16 (last rows), the percentage of operational and field test items that exhibited DIF at the moderate and severe levels was 2.44% for gender and between 1.13% and 8.83% for the four ethnicity groups, not including American Indian/Alaskan Native which had no flagged items.

Section 8.2—Test Level Reliability

Total test reliability statistics (alpha and 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 (in this case measured by Cronbach’s alpha [α ; 1951]) may range from 0.00 to 1.00, where 1.00 refers to a perfectly reliable test. The OCCT 3–8 reliability data are based on Oklahoma-specific representative samples from each grade (the scaling sample), and the results for 2014 are typical of the results obtained for all previous OCCT 3–8 operational tests. The total test reliabilities of the operational forms were evaluated first by Cronbach’s α (1951) index of internal consistency. The specific calculation for Cronbach’s α is

$$\hat{\alpha} = \frac{k}{k-1} \left(1 - \frac{\sum \hat{\sigma}_i^2}{\hat{\sigma}_x^2} \right), \quad (7)$$

where k is the number of items on the test form, $\hat{\sigma}_i^2$ is the variance of item i , and $\hat{\sigma}_x^2$ is the total test variance and the summation is over all the items ($i = 1, \dots, k$) 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 64 shows the reliability coefficients for each scored operational test form for each content area and grade for both Spring 2013 and Spring 2014. Alpha reliability coefficients for Spring 2013 and Spring 2014 are quite similar. Reliability for Spring 2014 ranged between 0.85 (Science Grade 8) and 0.92 (Reading Grade 5). Such a range is indicative of the high reliability of Spring 2014 OCCT 3–8 operational tests. As is evident in Tables 25–33, for Spring 2014 state and subgroup data, the coefficients are quite high and similar to the state, even at the subgroup levels. The mean (and range) of the state-level reliability coefficients for each content area are as follows: Mathematics 0.91 (range 0.89–0.91), Reading 0.91 (range 0.89–0.92), Science 0.86 (range 0.85–0.87), Social Studies 0.88, and U.S. History 0.90. At the subgroup level, the lowest reliability (0.77) was found for the ELL students in Science Grade 8.

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:

$$SEM = SD_{TT}(\sqrt{1 - \hat{\alpha}}), \quad (8)$$

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:

$$CSEM = SD_{SS}(\sqrt{1 - \hat{\alpha}}), \quad (9)$$

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 Tables 25–33.

Scale score specific SEMs are given in Tables 65–69, which also provide the raw scores associated with each scale score.

Section 8.3—Test Level Validity

Exploratory Factor Analysis (EFA) and Confirmatory Factor Analyses (CFA) were conducted to evaluate the unidimensionality assumption of the OCCT 3–8 test scores for the total population and various subgroups such as accommodated, ELL, Section 504, Low SES, and IEP. In factor analyses, the “construct” is referred to as a factor. 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 Tables 70–75, KMO values for the total group ranged from 0.96 to 0.98, and for each subgroup from 0.91 to 0.97 (Accommodated), from 0.87 to 0.96 (ELL), from 0.95 to 0.98 (Free Lunch), and from 0.92 to 0.98 (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, which means there will be times the percentages can be greater than 100%. Tables 70–75 also show the total amount of variance that exists in each form, as well as the percent 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 81–90% (Mathematics), 92–103% (Reading), 103–105% (Science), 104% (Social Studies), and 100% (U.S. History). The range of variance by the first eigenvalue in each content area and subgroup is as follows:

- Accommodated: 82–87% (Mathematics), 94–99% (Reading), 96% and 101% (Science), 96% (Social Studies), and 95% (U.S. History).
- ELL: 79–85% (Mathematics), 90–95% (Reading), 86% and 95% (Science), 87% (Social Studies), and 88% (U.S. History).

- Free Lunch: 81–90% (Mathematics), 92–103% (Reading), 102% and 105% (Science), 104% (Social Studies), and 100% (U.S. History).
- IEP: 85–90% (Mathematics), 95–101% (Reading), 98% and 101% (Science), 99% (Social Studies), and 97% (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 Science is mostly higher than the other five content areas for the total population and each subgroup.

As a rule, “essential unidimensionality” is assumed when the ratio of the first eigenvalue to the second eigenvalue is at least three. The final columns of Tables 70–75 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 four; therefore, the OCCT 3–8 tests are demonstrating 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. Examinations of the scree plots (Figures 33–112) for all grades and content areas for the total population and each subgroup indicate 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 all open-ended items are scored twice).

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 Social Studies and U.S. History standard setting workshop and the Bookmark Standard Setting Procedure used to set the cut scores are given in the Oklahoma State Testing Program Standard Setting Technical Report for OSTP Grade 5 Social Studies, Grade 8 U.S. History, and End-of-Instruction (EOI) U.S. History (CTB/McGraw-Hill, 2014). It may be useful to note that 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 76 shows the Spring 2014 SEMs estimated for each of the cut scores for each content area and grade.

Comparison of these SEMs to the SEMs associated with other OCCT 3–8 scale scores for each test (shown in Tables 65–69) reveals that these values are almost always among the lowest, meaning that the OCCT 3–8 tests tend to measure most accurately near the cut score. 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 65–69; there are more scale scores possible at each raw score than can be shown in these tables.)

Not only is it important that the amount of measurement error around the cut score is minimal, but also important is the expected consistency with which students would be classified into performance levels if given the test over 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 as follows.

	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):

$$\text{kappa} = \frac{P - P_c}{1 - P_c}, \quad (10)$$

where P is defined as the sum of diagonal values of the contingency table (shaded above) 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:

$$P_c = (P_{1.} \times P_{.1}) + (P_{2.} \times P_{.2}) + (P_{3.} \times P_{.3}). \quad (11)$$

The Livingston and Lewis (1995) method based on the binomial error model and the four-parameter beta true score distribution was applied to OCCT 3–8. Tables 77 and 78 show the classification consistency and classification accuracy indices. Note that 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 77 shows the average consistency is 0.69 across all grades and content areas, and ranges from 0.60 (Science Grade 8) to 0.77 (Reading Grade 3). The average accuracy is 0.77 across all grades and content areas, and ranges from 0.70 (Science Grade 8) to 0.83 (Reading Grade 3). 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 $(\text{Consistency} - \text{Chance}^1) / (1 - \text{Chance})$. 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 0. This is true of the OCCT 3–8 data in Table 77. The average Kappa is 0.55 over all grades and content areas and ranges from 0.46 (Science Grade 8) to 0.60 (Reading Grade 3).

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 the true classification given the modeled form. Table 78 shows consistency and accuracy at the cut score level. The average consistency across grades and cut score level is 0.89, ranging from 0.83 (Science Grade 8, at the Proficient and Advanced proficiency levels) to 0.97 (Reading Grade 3, at the Advanced proficiency level). The average accuracy across grades and

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.

cut score level is 0.92, ranging from 0.88 (Science grades 5 and 8, at the Proficient and Advanced proficiency levels) to 0.98 (Reading Grade 3, at the Advanced proficiency level). Finally, Table 79 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.

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.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20 (1), 37–46.
- 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.
- CTB/McGraw-Hill. (2013). *Oklahoma School Testing Program: 2013–2014 Test Preparation Manual*. Monterey, CA: CTB/McGraw-Hill LLC.
- CTB/McGraw-Hill. (2014). *Oklahoma State Testing Program: Standard setting technical report for OSTP Grade 5 Social Studies, Grade 8 U.S. History, and End-of-Instruction U.S. History*. 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 Publishing Company.
- 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). *MULTILOG; 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 9. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2014

Content	Grade	Female			Male		
		Sample	State	Diff.	Sample	State	Diff.
Mathematics	3	49.05	48.99	0.06	50.95	51.01	-0.06
	4	48.63	48.64	-0.01	51.37	51.36	0.01
	5	49.37	48.95	0.43	50.63	51.05	-0.43
	6	49.83	49.18	0.65	50.17	50.82	-0.65
	7	49.36	48.89	0.47	50.64	51.11	-0.47
	8	48.87	48.33	0.53	51.13	51.67	-0.53
Reading	3	48.97	48.99	-0.02	51.03	51.01	0.02
	4	48.60	48.63	-0.03	51.40	51.37	0.03
	5	49.30	48.94	0.36	50.70	51.06	-0.36
	6	49.87	49.18	0.69	50.13	50.82	-0.69
	7	49.33	48.90	0.44	50.67	51.10	-0.44
	8	49.52	49.04	0.48	50.48	50.96	-0.48
Science	5	50.00	48.94	1.05	50.00	51.06	-1.05
	8	49.15	49.06	0.09	50.85	50.94	-0.09
Social Studies	5	48.89	48.97	-0.08	51.11	51.03	0.08
U.S. History	8	48.98	49.06	-0.08	51.02	50.94	0.08

Table 9. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2014 (*continued*)

Content	Grade	African American (Not Hispanic)			American Indian/Alaskan		
		Sample	State	Diff.	Sample	State	Diff.
Mathematics	3	6.69	8.98	-2.30	18.41	14.68	3.72
	4	6.77	9.13	-2.37	18.60	14.91	3.69
	5	6.73	9.14	-2.41	19.64	15.11	4.53
	6	8.93	9.19	-0.26	15.75	15.67	0.08
	7	8.98	9.10	-0.12	16.23	16.19	0.05
	8	10.03	10.15	-0.12	17.40	17.42	-0.02
Reading	3	8.75	9.00	-0.25	14.68	14.73	-0.05
	4	6.78	9.16	-2.38	18.69	14.94	3.74
	5	6.80	9.16	-2.36	19.75	15.15	4.60
	6	8.99	9.26	-0.26	15.72	15.65	0.07
	7	8.87	9.00	-0.12	15.94	15.97	-0.02
	8	9.31	9.51	-0.20	16.33	16.45	-0.12
Science	5	9.18	9.14	0.03	17.73	15.10	2.62
	8	7.67	9.49	-1.82	19.33	16.39	2.94
Social Studies	5	9.14	9.14	0.01	15.25	15.11	0.14
U.S. History	8	9.47	9.48	-0.01	16.46	16.39	0.07

Table 9. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2014 (*continued*)

Content	Grade	Hispanic			Asian		
		Sample	State	Diff.	Sample	State	Diff.
Mathematics	3	13.95	16.31	-2.36	1.32	1.86	-0.54
	4	14.04	16.02	-1.98	1.64	1.96	-0.32
	5	13.92	15.54	-1.62	1.32	1.83	-0.51
	6	14.81	14.80	0.01	1.91	1.83	0.09
	7	14.39	14.44	-0.05	1.81	1.75	0.06
	8	14.62	14.52	0.10	1.09	1.08	0.01
Reading	3	16.33	16.17	0.16	1.82	1.80	0.02
	4	13.93	15.88	-1.95	1.60	1.93	-0.33
	5	13.83	15.44	-1.60	1.31	1.78	-0.47
	6	14.79	14.69	0.10	1.91	1.81	0.10
	7	14.31	14.11	0.20	2.02	1.94	0.09
	8	13.66	13.53	0.14	1.88	1.80	0.08
Science	5	12.96	15.56	-2.60	1.56	1.83	-0.27
	8	10.42	13.68	-3.26	1.17	1.84	-0.67
Social Studies	5	15.58	15.53	0.05	1.82	1.84	-0.02
U.S. History	8	13.81	13.69	0.12	1.83	1.84	0.00

Table 9. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2014 (*continued*)

Content	Grade	Hawaiian/Pacific Islander			White (Not Hispanic)			Multiracial		
		Sample	State	Diff.	Sample	State	Diff.	Sample	State	Diff.
Mathematics	3	0.15	0.28	-0.13	52.54	50.69	1.84	6.95	7.19	-0.25
	4	0.11	0.26	-0.14	52.16	50.58	1.58	6.68	7.13	-0.46
	5	0.12	0.30	-0.18	51.74	51.30	0.44	6.52	6.78	-0.26
	6	0.25	0.28	-0.04	51.19	50.61	0.58	7.17	7.62	-0.45
	7	0.25	0.27	-0.02	51.56	51.09	0.47	6.78	7.16	-0.38
	8	0.21	0.23	-0.02	50.26	50.16	0.10	6.40	6.44	-0.04
Reading	3	0.24	0.27	-0.03	50.97	50.82	0.15	7.20	7.20	-0.01
	4	0.12	0.25	-0.13	52.24	50.70	1.54	6.64	7.13	-0.49
	5	0.12	0.29	-0.17	51.62	51.39	0.23	6.56	6.80	-0.24
	6	0.24	0.27	-0.03	51.16	50.69	0.48	7.18	7.64	-0.46
	7	0.26	0.28	-0.02	51.91	51.61	0.30	6.69	7.10	-0.42
	8	0.21	0.23	-0.02	52.37	52.17	0.20	6.25	6.32	-0.07
Science	5	0.14	0.30	-0.15	52.23	51.28	0.94	6.21	6.78	-0.57
	8	0.12	0.23	-0.11	55.48	52.16	3.33	5.80	6.21	-0.41
Social Studies	5	0.30	0.30	0.00	51.11	51.30	-0.19	6.79	6.79	0.00
U.S. History	8	0.23	0.23	0.00	51.96	52.16	-0.20	6.23	6.21	0.01

Table 9. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2014 (*continued*)

Content	Grade	ELL			IEP			Section 504		
		Sample	State	Diff.	Sample	State	Diff.	Sample	State	Diff.
Mathematics	3	9.24	11.49	-2.25	17.37	16.60	0.77	0.76	0.81	-0.05
	4	6.10	7.57	-1.46	18.52	17.72	0.80	0.71	0.95	-0.23
	5	5.50	6.16	-0.65	18.03	17.46	0.57	0.80	1.18	-0.37
	6	4.69	5.39	-0.70	13.56	16.94	-3.38	1.06	1.07	-0.01
	7	4.60	4.89	-0.29	13.87	16.36	-2.49	1.13	1.11	0.02
	8	5.31	5.81	-0.51	16.13	19.07	-2.94	1.22	1.21	0.01
Reading	3	11.23	11.03	0.21	16.52	16.68	-0.16	0.81	0.80	0.00
	4	5.71	7.14	-1.44	18.60	17.76	0.84	0.71	0.96	-0.24
	5	5.25	5.80	-0.55	18.06	17.48	0.57	0.79	1.17	-0.38
	6	4.54	5.08	-0.54	13.69	17.00	-3.31	1.06	1.08	-0.02
	7	4.44	4.52	-0.08	13.71	16.07	-2.36	1.12	1.12	0.00
	8	4.24	4.57	-0.33	13.30	15.67	-2.36	1.13	1.13	-0.01
Science	5	2.47	6.06	-3.59	15.73	17.41	-1.68	0.87	1.18	-0.30
	8	1.15	4.91	-3.76	14.65	15.59	-0.94	1.02	1.10	-0.08
Social Studies	5	6.14	6.09	0.05	17.41	17.37	0.04	1.15	1.18	-0.03
U.S. History	8	4.94	4.93	0.01	15.63	15.62	0.01	1.09	1.10	-0.01

Table 9. Subgroup Representativeness of Scaling Sample Compared to Total Population, Spring 2014 (*continued*)

Content	Grade	SES - Low			SES - High			Accommodated		
		Sample	State	Diff.	Sample	State	Diff.	Sample	State	Diff.
Mathematics	3	38.12	38.08	0.04	61.88	61.92	-0.04	13.34	15.13	-1.79
	4	38.98	38.66	0.33	61.02	61.34	-0.33	13.63	14.59	-0.96
	5	38.53	39.05	-0.52	61.47	60.95	0.52	13.65	14.75	-1.10
	6	40.57	39.74	0.82	59.43	60.26	-0.82	7.04	10.49	-3.45
	7	42.30	41.81	0.49	57.70	58.19	-0.49	7.67	10.15	-2.48
	8	38.19	37.40	0.79	61.81	62.60	-0.79	8.50	11.53	-3.03
Reading	3	38.45	37.88	0.57	61.55	62.12	-0.57	13.45	13.75	-0.29
	4	38.98	38.64	0.34	61.02	61.36	-0.34	13.38	13.74	-0.36
	5	38.47	39.04	-0.57	61.53	60.96	0.57	13.37	14.06	-0.69
	6	40.60	39.80	0.80	59.40	60.20	-0.80	6.15	9.45	-3.30
	7	42.94	42.45	0.50	57.06	57.55	-0.50	6.19	8.39	-2.19
	8	44.20	43.20	1.00	55.80	56.80	-1.00	5.44	7.70	-2.26
Science	5	38.45	39.02	-0.58	61.55	60.98	0.58	11.31	14.51	-3.20
	8	44.14	43.26	0.87	55.86	56.74	-0.87	8.94	10.33	-1.39
Social Studies	5	38.90	39.05	-0.15	61.10	60.95	0.15	13.69	14.48	-0.79
U.S. History	8	43.04	43.24	-0.20	56.96	56.76	0.20	9.46	10.14	-0.67

Table 10. Summary of *P*-Values and Item-Test Correlations Statistics for Operational Test Forms, Spring 2012 to Spring 2014

Content	Grade	Operational Mean <i>P</i> -Values*				Operational Mean Item-Test Correlations*			
		2012	2013	2014	Diff. S14-S13	2012	2013	2014	Diff. S14-S13
Mathematics	3	0.76	0.76	0.74	-0.02	0.41	0.44	0.44	0.00
	4	0.76	0.76	0.72	-0.04	0.41	0.42	0.44	0.02
	5	0.73	0.73	0.69	-0.04	0.39	0.40	0.43	0.03
	6	0.66	0.66	0.64	-0.02	0.41	0.41	0.43	0.02
	7	0.63	0.62	0.60	-0.02	0.40	0.39	0.41	0.02
	8	0.64	0.65	0.56	-0.09	0.41	0.42	0.39	-0.03
Reading	3	0.72	0.72	0.67	-0.05	0.41	0.42	0.43	0.01
	4	0.74	0.72	0.71	-0.01	0.39	0.40	0.44	0.04
	5	0.76	0.77	0.72	-0.05	0.40	0.42	0.45	0.03
	6	0.70	0.73	0.68	-0.05	0.41	0.40	0.42	0.02
	7	0.77	0.77	0.73	-0.04	0.38	0.40	0.45	0.05
	8	0.77	0.76	0.72	-0.04	0.36	0.39	0.40	0.01
Science	5	.	0.68	0.70	0.02	.	0.37	0.39	0.02
	8	.	0.57	0.60	0.03	.	0.36	0.37	0.01
Social Studies	5	.	.	0.63	.	.	.	0.38	.
U.S. History	8	.	.	0.61	.	.	.	0.41	.
Writing	5	.	.	0.58	.	.	.	0.94	.
	8	.	.	0.64	.	.	.	0.97	.

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

Table 11. Summary of Range of *P*-Values and Item-Test Correlations Statistics for Operational and Field Test, Spring 2014

Content	Grade	Mean <i>P</i> -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
Mathematics	3	0.38	0.74	0.94	.	.	.	0.25	0.44	0.61	.	.	.
	4	0.45	0.72	0.93	.	.	.	0.28	0.44	0.55	.	.	.
	5	0.36	0.69	0.95	.	.	.	0.25	0.43	0.60	.	.	.
	6	0.29	0.64	0.94	.	.	.	0.23	0.43	0.56	.	.	.
	7	0.24	0.60	0.97	.	.	.	0.18	0.41	0.61	.	.	.
	8	0.21	0.56	0.92	.	.	.	0.16	0.39	0.58	.	.	.
Reading	3	0.39	0.67	0.93	.	.	.	0.23	0.43	0.60	.	.	.
	4	0.40	0.71	0.89	.	.	.	0.21	0.44	0.57	.	.	.
	5	0.38	0.72	0.90	.	.	.	0.27	0.45	0.59	.	.	.
	6	0.37	0.68	0.88	.	.	.	0.25	0.42	0.57	.	.	.
	7	0.49	0.73	0.92	.	.	.	0.27	0.45	0.63	.	.	.
	8	0.29	0.72	0.92	.	.	.	0.26	0.40	0.52	.	.	.
Science	5	0.43	0.70	0.97	0.17	0.49	0.89	0.15	0.39	0.50	0.08	0.26	0.48
	8	0.24	0.60	0.92	0.24	0.53	0.89	0.09	0.37	0.53	0.09	0.29	0.50
Social Studies	5	0.42	0.63	0.85	0.20	0.44	0.77	0.09	0.38	0.52	0.02	0.27	0.48
Geography	7	.	.	.	0.12	0.45	0.89	.	.	.	0.12	0.31	0.46
U.S. History	8	0.27	0.61	0.86	0.10	0.45	0.85	0.21	0.41	0.57	0.00	0.27	0.47
Writing	5	0.58	0.58	0.58	.	.	.	0.91	0.94	0.95	.	.	.
	8	0.63	0.64	0.64	.	.	.	0.97	0.97	0.98	.	.	.

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

Table 12. Summary of Range of Omission Rates for Operational and Field Test by Item Type, Spring 2014

Content	Grade	Item Type	Omission Rates*					
			Operational Items			Field Test Items		
			Low	Mean	High	Low	Mean	High
Mathematics	3	MC	0.09%	0.33%	1.44%	.	.	.
	4	MC	0.05%	0.15%	0.59%	.	.	.
	5	MC	0.03%	0.11%	0.41%	.	.	.
	6	MC	0.03%	0.07%	0.11%	.	.	.
	7	MC	0.02%	0.05%	0.10%	.	.	.
	8	MC	0.02%	0.07%	0.11%	.	.	.
Reading	3	MC	0.02%	0.32%	0.90%	.	.	.
	4	MC	0.07%	0.17%	0.39%	.	.	.
	5	MC	0.02%	0.11%	0.19%	.	.	.
	6	MC	0.01%	0.06%	0.10%	.	.	.
	7	MC	0.01%	0.06%	0.11%	.	.	.
	8	MC	0.01%	0.05%	0.09%	.	.	.
Science	5	MC	0.04%	0.11%	0.22%	0.01%	0.11%	0.26%
	8	MC	0.03%	0.10%	0.16%	0.00%	0.11%	0.28%
Social Studies	5	MC	0.04%	0.11%	0.23%	0.03%	0.12%	0.26%
Geography	7	MC	.	.	.	0.00%	0.10%	0.30%
U.S. History	8	MC	0.03%	0.09%	0.16%	0.03%	0.08%	0.18%
Writing	5	CR	0.00%	0.00%	0.00%	.	.	.
	8	CR	0.00%	0.00%	0.00%	.	.	.

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

Table 13. Spring 2014 Summary of Operational and Field Test Items Flagged for Mantel-Haenszel Differential Item Functioning, by Item Type: Gender Male/Female

Content	Grade	Item Type	Operational Items		Field Test Items		Total DIF Flags B+C
			B	C	B	C	
Mathematics	3	MC	1	0	.	.	1
	4	MC	1	0	.	.	1
	5	MC	1	1	.	.	2
	6	MC
	7	MC	2	0	.	.	2
	8	MC	3	0	.	.	3
Reading	3	MC	1	0	.	.	1
	4	MC	2	0	.	.	2
	5	MC	1	0	.	.	1
	6	MC	1	0	.	.	1
	7	MC	4	0	.	.	4
	8	MC
Science	5	MC	1	0	0	0	1
	8	MC	2	0	3	0	5
Social Studies	5	MC	0	0	0	0	.
Geography	7	MC	.	.	0	0	0
U.S. History	8	MC	1	0	1	0	2
Total Items Flagged			21	1	4	0	26
Total Items Tested			800		265		1065
Percentage of Items Flagged			2.63%	0.13%	1.51%	0.00%	2.44%

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

Table 14. Spring 2014 Summary of Operational and Field Test Items Flagged for Mantel-Haenszel Differential Item Functioning, by Item Type: Ethnicity White/Asian

Content	Grade	Item Type	Operational Items		Field Test Items		Total DIF Flags B+C
			B	C	B	C	
Mathematics	3	MC	8	0	.	.	8
	4	MC	4	0	.	.	4
	5	MC	4	1	.	.	5
	6	MC	5	1	.	.	6
	7	MC	2	1	.	.	3
	8	MC	4	2	.	.	6
Reading	3	MC	5	0	.	.	5
	4	MC	2	0	.	.	2
	5	MC	2	0	.	.	2
	6	MC	4	1	.	.	5
	7	MC	4	3	.	.	7
	8	MC	7	2	.	.	9
Science	5	MC	2	0	2	2	6
	8	MC	2	1	6	3	12
Social Studies	5	MC	1	0	5	0	6
Geography	7	MC	.	.	0	3	3
U.S. History	8	MC	0	0	4	1	5
Total Items Flagged			56	12	17	9	94
Total Items Tested			800		265		1065
Percentage of Items Flagged			7.00%	1.50%	6.42%	3.40%	8.83%

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

Table 15. Spring 2014 Summary of Operational and Field Test Items Flagged for Mantel-Haenszel Differential Item Functioning, by Item Type: Ethnicity White/African American

Content	Grade	Item Type	Operational Items		Field Test Items		Total DIF Flags B+C
			B	C	B	C	
Mathematics	3	MC	2	0	.	.	2
	4	MC	3	0	.	.	3
	5	MC	3	0	.	.	3
	6	MC	1	0	.	.	1
	7	MC	1	0	.	.	1
	8	MC	1	1	.	.	2
Reading	3	MC	0	0	.	.	0
	4	MC	0	0	.	.	0
	5	MC	0	0	.	.	0
	6	MC	1	0	.	.	1
	7	MC	1	0	.	.	1
	8	MC	1	0	.	.	1
Science	5	MC	0	0	1	1	2
	8	MC	1	0	1	0	2
Social Studies	5	MC	0	0	0	0	0
Geography	7	MC	.	.	7	1	8
U.S. History	8	MC	0	0	0	1	1
Total Items Flagged			15	1	9	3	28
Total Items Tested			800		265		1065
Percentage of Items Flagged			1.88%	0.13%	3.40%	1.13%	2.63%

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

Table 16. Spring 2014 Summary of Operational and Field Test Items Flagged for Mantel-Haenszel Differential Item Functioning, by Item Type: Ethnicity White/Hispanic

Content	Grade	Item Type	Operational Items		Field Test Items		Total DIF Flags B+C
			B	C	B	C	
Mathematics	3	MC	0	0	.	.	0
	4	MC	0	0	.	.	0
	5	MC	1	0	.	.	1
	6	MC	0	0	.	.	0
	7	MC	0	0	.	.	0
	8	MC	0	0	.	.	0
Reading	3	MC	1	0	.	.	1
	4	MC	1	0	.	.	1
	5	MC	0	0	.	.	0
	6	MC	1	0	.	.	1
	7	MC	2	1	.	.	3
	8	MC	1	0	.	.	1
Science	5	MC	0	0	0	1	1
	8	MC	0	0	1	0	1
Social Studies	5	MC	0	0	1	0	1
Geography	7	MC	.	.	1	0	1
U.S. History	8	MC	0	0	0	0	0
Total Items Flagged			7	1	3	1	12
Total Items Tested			800		265		1065
Percentage of Items Flagged			0.88%	0.13%	1.13%	0.38%	1.13%

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

Table 17. Mathematics Grades 3–5 Standards Level Summary Data, Spring 2014

Grade/ Obj.	Standard Reference	No. of Items	Average	Average	Objective	Average P-value							Alpha	SEM
			Difficulty (IRT Loc)	IRT Information	% Correct State Mean	State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass			
3.1	Algebraic Reasoning: Patterns and Relationships	7	630.86	0.04	81.60	0.82	0.51	0.73	0.87	0.96	0.90	0.54	0.92	
3.2	Number Sense & Operation	20	672.05	0.07	76.26	0.77	0.39	0.63	0.83	0.96	0.88	0.84	1.63	
3.3	Geometry	7	694.00	0.03	68.26	0.69	0.43	0.58	0.71	0.86	0.77	0.49	1.04	
3.4	Measurement	9	725.00	0.07	62.88	0.63	0.28	0.44	0.67	0.91	0.76	0.71	1.24	
3.5	Data Analysis	7	635.71	0.08	81.29	0.81	0.42	0.70	0.89	0.98	0.92	0.70	0.87	
4.1	Algebraic Reasoning: Patterns and Relationships	7	628.71	0.05	82.48	0.83	0.50	0.75	0.90	0.98	0.92	0.63	0.89	
4.2	Number Sense & Operation	18	698.94	0.07	71.03	0.71	0.40	0.58	0.77	0.94	0.82	0.79	1.65	
4.3	Geometry	9	697.22	0.04	70.51	0.71	0.40	0.58	0.76	0.93	0.81	0.65	1.19	
4.4	Measurement	9	695.56	0.06	68.21	0.68	0.33	0.53	0.76	0.92	0.81	0.69	1.22	
4.5	Data Analysis	7	682.57	0.05	72.90	0.73	0.37	0.61	0.80	0.94	0.85	0.64	1.03	
5.1	Algebraic Reasoning: Patterns and Relationships	13	707.15	0.05	69.85	0.70	0.41	0.58	0.73	0.90	0.80	0.70	1.44	
5.2	Number Sense & Operation	16	694.56	0.07	70.89	0.71	0.36	0.56	0.76	0.93	0.83	0.81	1.54	
5.3	Geometry	7	726.14	0.04	64.81	0.65	0.36	0.52	0.68	0.86	0.75	0.56	1.12	
5.4	Measurement	7	738.00	0.06	62.80	0.63	0.28	0.46	0.67	0.87	0.76	0.63	1.10	
5.5	Data Analysis	7	671.29	0.07	75.95	0.76	0.42	0.65	0.81	0.94	0.87	0.64	0.96	

Note: Obj. = Objective; P.L. = Performance Level.

Table 18. Mathematics Grades 6–8 Standards Level Summary Data, Spring 2014

Grade/ Obj.	Standard Reference	No. of Items	Average	Average	Objective	Average P-value						Alpha	SEM
			Difficulty (IRT Loc)	IRT Information	% Correct State Mean	State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass		
6.1	Algebraic Reasoning: Patterns and Relationships	13	691.46	0.12	68.71	0.69	0.36	0.56	0.76	0.93	0.81	0.78	1.39
6.2	Number Sense & Operation	15	702.93	0.09	67.31	0.67	0.39	0.53	0.73	0.93	0.79	0.77	1.56
6.3	Geometry	8	736.38	0.05	59.79	0.60	0.35	0.48	0.63	0.86	0.69	0.58	1.20
6.4	Measurement	7	759.14	0.19	52.65	0.53	0.26	0.36	0.55	0.86	0.64	0.63	1.12
6.5	Data Analysis	7	714.86	0.08	62.73	0.63	0.36	0.51	0.67	0.89	0.73	0.57	1.10
7.1	Algebraic Reasoning: Patterns and Relationships	15	732.13	0.06	62.10	0.62	0.38	0.53	0.68	0.85	0.73	0.70	1.61
7.2	Number Sense & Operation	11	720.82	0.06	63.77	0.64	0.39	0.54	0.69	0.89	0.75	0.65	1.38
7.3	Geometry	7	734.71	0.05	60.72	0.61	0.35	0.51	0.66	0.86	0.72	0.55	1.13
7.4	Measurement	9	768.22	0.22	49.82	0.50	0.24	0.33	0.53	0.86	0.62	0.72	1.25
7.5	Data Analysis	8	705.13	0.10	64.04	0.64	0.38	0.52	0.70	0.90	0.76	0.65	1.10
8.1	Algebraic Reasoning: Patterns and Relationships	16	747.31	0.09	51.98	0.52	0.27	0.42	0.60	0.79	0.66	0.71	1.76
8.2	Number Sense & Operation	11	739.09	0.09	50.75	0.51	0.25	0.40	0.59	0.81	0.66	0.69	1.44
8.3	Geometry	9	675.22	0.09	65.47	0.65	0.41	0.58	0.73	0.90	0.78	0.61	1.20
8.4	Measurement	7	699.14	0.09	61.45	0.62	0.35	0.53	0.70	0.88	0.76	0.54	1.14
8.5	Data Analysis	7	718.29	0.08	57.35	0.57	0.29	0.52	0.67	0.81	0.71	0.56	1.09

Note: Obj. = Objective; P.L. = Performance Level.

Table 19. Reading Grades 3–5 Standards Level Summary Data, Spring 2014

Grade/ Obj.	Standard Reference	No. of Items	Average	Average	Objective	Average P-value							
			Difficulty (IRT Loc)	IRT Information	% Correct State Mean	State	P.L. 1	P.L. 2	P.L. 3	P.L. 4	Pass	Alpha	SEM
3.2	Vocabulary	12	.	.	71.90	0.72	0.37	0.59	0.83	0.98	0.83	0.75	1.37
3.4	Comprehension/Critical Literacy	26	.	.	65.73	0.66	0.33	0.51	0.76	0.96	0.77	0.84	2.13
3.5	Literature	7	.	.	67.21	0.68	0.31	0.51	0.79	0.99	0.80	0.69	1.06
3.6	Research and Information	5	.	.	.	0.62	0.33	0.51	0.71	0.96	0.72	0.46	0.97
4.1	Vocabulary	11	684.00	0.08	72.25	0.72	0.41	0.65	0.83	0.97	0.84	0.71	1.31
4.3	Comprehension/Critical Literacy	24	674.08	0.06	72.68	0.73	0.42	0.64	0.83	0.98	0.84	0.84	1.95
4.4	Literature	9	719.56	0.06	65.60	0.66	0.35	0.53	0.76	0.97	0.78	0.65	1.28
4.5	Research and Information	6	695.83	0.07	68.10	0.68	0.39	0.58	0.79	0.97	0.80	0.56	0.98
5.1	Vocabulary	11	689.82	0.06	70.54	0.71	0.38	0.60	0.80	0.95	0.82	0.70	1.30
5.3	Comprehension/Critical Literacy	19	674.89	0.08	73.79	0.74	0.38	0.63	0.85	0.97	0.86	0.83	1.66
5.4	Literature	13	670.15	0.06	73.36	0.73	0.39	0.64	0.83	0.96	0.85	0.74	1.41
5.5	Research and Information	7	695.86	0.04	68.18	0.68	0.40	0.58	0.76	0.93	0.79	0.51	1.10

Note: Obj. = Objective; P.L. = Performance Level.

Table 20. Reading Grades 6–8 Standards Level Summary Data, Spring 2014

Grade/ Obj.	Standard Reference	No. of Items	Average	Average	Objective	Average P-value					Pass	Alpha	SEM
			Difficulty (IRT Loc)	IRT Information	% Correct State Mean	State	P.L. 1	P.L. 2	P.L. 3	P.L. 4			
6.1	Vocabulary	8	706.38	0.07	66.14	0.66	0.33	0.55	0.74	0.92	0.78	0.62	1.16
6.3	Comprehension/Critical Literacy	21	681.57	0.06	70.80	0.71	0.36	0.60	0.81	0.94	0.83	0.82	1.86
6.4	Literature	14	714.14	0.05	67.67	0.68	0.38	0.56	0.75	0.93	0.79	0.69	1.58
6.5	Research and Information	7	741.29	0.07	59.73	0.60	0.28	0.44	0.69	0.90	0.72	0.58	1.15
7.1	Vocabulary	11	695.82	0.11	72.76	0.73	0.41	0.62	0.79	0.94	0.82	0.70	1.27
7.3	Comprehension/Critical Literacy	20	705.85	0.11	70.79	0.71	0.36	0.58	0.78	0.94	0.82	0.82	1.79
7.4	Literature	12	679.25	0.08	76.45	0.76	0.45	0.67	0.83	0.95	0.86	0.70	1.31
7.5	Research and Information	7	687.86	0.13	76.46	0.77	0.41	0.64	0.84	0.98	0.88	0.65	0.98
8.1	Vocabulary	4	739.50	0.04	.	0.66	0.34	0.52	0.72	0.89	0.75	0.36	0.86
8.3	Comprehension/Critical Literacy	22	681.68	0.07	74.89	0.75	0.43	0.64	0.81	0.94	0.84	0.79	1.77
8.4	Literature	14	708.93	0.04	70.00	0.70	0.38	0.58	0.76	0.92	0.79	0.68	1.56
8.5	Research and Information	10	719.60	0.04	68.54	0.69	0.38	0.56	0.74	0.92	0.77	0.59	1.32

Note: Obj. = Objective; P.L. = Performance Level.

Table 21. Science Grades 5 & 8 Standards Level Summary Data, Spring 2014

Grade/ Obj.	Standard Reference	No. of Items	Average	Average	Objective	Average P-value					Pass	Alpha	SEM
			Difficulty (IRT Loc)	IRT Information	% Correct State Mean	State	P.L. 1	P.L. 2	P.L. 3	P.L. 4			
5.P1	Observe and Measure	9	646.67	0.08	75.14	0.75	0.49	0.72	0.85	0.95	0.88	0.61	1.15
5.P2	Classify	9	697.56	0.04	63.67	0.64	0.41	0.59	0.73	0.86	0.76	0.47	1.33
5.P3	Experiment	13	667.69	0.07	69.83	0.70	0.42	0.65	0.81	0.93	0.84	0.69	1.48
5.P4	Interpret and Communicate	14	647.64	0.07	70.07	0.70	0.45	0.65	0.80	0.93	0.84	0.69	1.51
5.S1	Properties of Matter and Energy	18	672.28	0.06	68.06	0.68	0.42	0.63	0.79	0.91	0.82	0.72	1.78
5.S2	Organisms and Environments	10	665.90	0.07	70.20	0.70	0.44	0.66	0.81	0.92	0.84	0.62	1.28
5.S3	Structures of the Earth and the Solar System	13	662.85	0.07	67.48	0.68	0.43	0.62	0.78	0.91	0.81	0.65	1.51
8.P1	Observe and Measure	11	692.73	0.18	62.54	0.63	0.37	0.56	0.73	0.87	0.77	0.63	1.38
8.P2	Classify	7	674.86	0.09	71.76	0.72	0.45	0.68	0.82	0.92	0.85	0.50	1.08
8.P3	Experiment	14	724.00	0.11	52.43	0.52	0.29	0.44	0.62	0.81	0.67	0.64	1.69
8.P4	Interpret and Communicate	13	708.92	0.18	60.41	0.60	0.38	0.55	0.68	0.84	0.73	0.58	1.56
8.S1	Properties and Chemical Changes in Matter	8	703.88	0.15	60.13	0.60	0.33	0.53	0.70	0.86	0.75	0.54	1.23
8.S2	Motion and Forces	8	720.00	0.14	53.72	0.54	0.28	0.45	0.63	0.83	0.69	0.55	1.27
8.S3	Diversity and Adaptations of Organisms	7	658.43	0.14	76.61	0.77	0.52	0.73	0.87	0.95	0.89	0.54	0.99
8.S4	Structures/Forces of the Earth/Solar System	11	738.55	0.17	48.95	0.49	0.28	0.41	0.57	0.76	0.63	0.52	1.53
8.S5	Earth's History	8	694.63	0.13	64.88	0.65	0.41	0.60	0.73	0.87	0.77	0.50	1.18

Note: Obj. = Objective; P.L. = Performance Level.

Table 22. Social Studies Grade 5 & U.S. History Grade 8 Standards Level Summary Data, Spring 2014

Grade/ Obj.	Standard Reference	No. of Items	Average	Average	Objective	Average P-value					Alpha	SEM	
			Difficulty (IRT Loc)	IRT Information	% Correct State Mean	State	P.L. 1	P.L. 2	P.L. 3	P.L. 4			Pass
5.1	James Towne Settlement and Plimoth Plantation	8	695.88	0.06	67.84	0.68	0.37	0.51	0.63	0.81	0.74	0.48	1.22
5.2	Colonial America	10	690.00	0.08	67.91	0.68	0.33	0.48	0.63	0.83	0.76	0.61	1.35
5.3	American Revolution	18	721.33	0.09	63.92	0.64	0.30	0.44	0.59	0.79	0.72	0.73	1.85
5.4	Early Federal Period	14	742.50	0.09	55.92	0.56	0.26	0.35	0.48	0.72	0.63	0.68	1.72
8.1	Causes and Events of the American Revolution	8	705.88	0.15	62.56	0.62	0.30	0.45	0.63	0.83	0.75	0.63	1.19
8.2	The Revolutionary Era	6	707.50	0.05	62.70	0.63	0.33	0.50	0.65	0.80	0.73	0.43	1.10
8.3	Developing the American Government System	10	719.30	0.13	55.61	0.56	0.30	0.42	0.53	0.74	0.66	0.60	1.38
8.4	The Transformation of the United States to the Mid-1800s	16	683.63	0.10	65.62	0.66	0.32	0.51	0.67	0.85	0.78	0.77	1.67
8.5	Causes, Events, and Leadership in the Civil War	10	719.90	0.08	57.02	0.57	0.30	0.42	0.55	0.77	0.68	0.63	1.38

Note: Obj. = Objective; P.L. = Performance Level.

Table 23. Spring 2013 Scale Score Statistics

Content	Grade	N		Scale Score Percentile									
		Count	Mean	SD	LOSS	N Min.	25th	50th	75th	N Max.	HOSS	Alpha	SEM
Mathematics	3	46316	739.00	88.33	400	151	691	740	792	864	990	0.91	2.62
	4	45383	745.43	90.14	400	133	693	744	798	858	990	0.90	2.69
	5	44295	740.71	86.98	400	137	690	741	788	430	990	0.89	2.81
	6	43221	737.10	78.84	400	129	687	737	787	241	990	0.90	2.94
	7	43146	732.30	80.70	400	239	688	730	782	126	990	0.89	3.02
	8	41377	732.09	83.25	400	143	685	735	786	309	990	0.90	3.00
Reading	3	45683	741.22	86.35	400	221	690	748	799	167	990	0.90	2.81
	4	44704	729.59	77.54	400	130	683	733	771	106	990	0.89	2.85
	5	43798	735.55	84.47	400	187	683	739	788	468	990	0.90	2.64
	6	42971	731.18	77.53	400	49	685	730	782	115	990	0.89	2.78
	7	43368	729.88	67.56	400	91	688	728	766	239	990	0.88	2.66
	8	42341	750.16	82.06	400	74	701	753	803	178	990	0.88	2.72
Science	5	44805	695.10	72.00	400	221	656	700	742	96	990	0.86	2.84
	8	44209	694.21	57.11	400	252	665	700	733	5	990	0.85	2.98

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

Table 24. Spring 2014 Scale Score Statistics

Content	Grade	N		Scale Score Percentile							HOSS	Alpha	SEM
		Count	Mean	SD	LOSS	N Min.	25th	50th	75th	N Max.			
Mathematics	3	48887	736.52	96.29	400	131	680	738	794	576	990	0.91	2.65
	4	48758	732.36	96.17	400	175	674	736	797	585	990	0.91	2.76
	5	48299	734.05	96.88	400	220	674	737	793	469	990	0.91	2.82
	6	47474	726.73	82.97	400	180	680	729	784	340	990	0.91	2.90
	7	46374	726.69	85.80	400	254	674	729	780	104	990	0.90	2.95
	8	37626	698.41	81.03	400	348	651	703	752	21	990	0.89	3.04
Reading	3	48752	733.10	91.18	400	304	682	737	789	96	990	0.91	2.94
	4	48656	720.53	87.68	400	267	674	721	776	328	990	0.91	2.85
	5	48206	723.42	94.26	400	400	671	724	776	256	990	0.92	2.77
	6	47475	725.32	86.99	400	176	678	725	776	129	990	0.90	2.94
	7	47402	730.89	76.62	400	207	692	732	775	466	990	0.92	2.74
	8	47330	740.62	88.73	400	169	689	750	795	483	990	0.89	2.85
Science	5	48291	695.83	74.78	400	164	649	702	739	88	990	0.87	2.75
	8	47451	694.88	59.45	400	222	666	697	734	18	990	0.85	2.90
Social Studies	5	48234	704.36	73.70	400	421	665	708	745	56	990	0.88	3.12
U.S. History	8	47432	683.80	79.74	400	743	639	689	735	46	990	0.90	3.05
Writing	5	47604	35.67	8.51	15	3665	31	35	42	224	60	0.97	0.62
	8	47128	36.47	9.63	15	3913	32	38	45	629	60	0.99	0.44

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

Table 25. Spring 2014, Mathematics Grades 3 & 4 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
3	50	Whole State	48887	736.52	96.29	400	990	0.91	2.65
		Female	23936	735.39	94.96	400	990	0.91	2.66
		Male	24923	737.67	97.55	400	990	0.92	2.64
		Asian	908	783.75	100.04	400	990	0.92	2.33
		African American	4391	685.35	99.30	400	990	0.92	2.89
		Hispanic	7974	706.30	93.45	400	990	0.91	2.82
		American Indian	7178	736.72	91.39	400	990	0.90	2.66
		White	24783	753.59	92.06	400	990	0.90	2.55
		Multiracial	3516	737.71	94.72	400	990	0.91	2.65
		ELL	5618	686.83	91.84	400	990	0.91	2.91
		IEP	8115	680.36	102.72	400	990	0.92	2.92
		Section 504	395	728.90	90.23	400	990	0.90	2.70
		Low SES	30270	716.06	94.32	400	990	0.91	2.77
Accommodated	7398	663.78	92.33	400	990	0.90	3.01		
4	50	Whole State	48758	732.36	96.17	400	990	0.91	2.76
		Female	23702	732.63	94.24	400	990	0.91	2.76
		Male	25023	732.15	97.99	400	990	0.91	2.76
		Asian	956	783.58	98.55	423	990	0.91	2.44
		African American	4453	678.36	97.60	400	990	0.91	3.00
		Hispanic	7810	708.31	92.41	400	990	0.91	2.89
		American Indian	7271	727.45	90.91	400	990	0.90	2.80
		White	24664	749.15	92.90	400	990	0.91	2.67
		Multiracial	3478	734.10	94.60	400	990	0.91	2.76
		ELL	3689	668.13	91.32	400	990	0.90	3.06
		IEP	8639	666.33	98.43	400	990	0.91	3.06
		Section 504	462	728.52	84.45	465	990	0.89	2.81
		Low SES	29910	709.62	92.53	400	990	0.91	2.89
Accommodated	7114	649.91	89.62	400	990	0.89	3.13		

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 26. Spring 2014, Mathematics Grades 5 & 6 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
5	50	Whole State	48299	734.05	96.88	400	990	0.91	2.82
		Female	23633	738.84	94.46	400	990	0.91	2.81
		Male	24651	729.50	98.91	400	990	0.91	2.82
		Asian	884	799.66	98.11	400	990	0.91	2.46
		African American	4414	690.86	99.74	400	990	0.91	3.01
		Hispanic	7505	713.86	94.81	400	990	0.91	2.92
		American Indian	7299	726.16	91.45	400	990	0.90	2.86
		White	24778	748.19	94.40	400	990	0.91	2.74
		Multiracial	3275	732.56	94.74	400	990	0.91	2.83
		ELL	2974	661.91	93.60	400	990	0.89	3.11
		IEP	8434	657.30	96.86	400	990	0.89	3.10
		Section 504	568	730.40	86.02	414	990	0.89	2.87
		Low SES Accommodated	29437 7122	712.39 646.01	93.44 90.50	400 400	990 990	0.90 0.87	2.93 3.14
6	50	Whole State	47474	726.73	82.97	400	990	0.91	2.90
		Female	23346	727.73	79.76	400	990	0.91	2.90
		Male	24123	725.80	85.93	400	990	0.91	2.90
		Asian	867	784.27	86.94	498	990	0.92	2.57
		African American	4365	685.13	82.99	400	990	0.89	3.04
		Hispanic	7025	708.65	79.96	400	990	0.90	2.97
		American Indian	7437	719.15	79.19	400	990	0.90	2.94
		White	24027	740.53	80.40	400	990	0.91	2.84
		Multiracial	3619	723.73	81.42	400	990	0.91	2.91
		ELL	2558	660.70	82.78	400	990	0.87	3.10
		IEP	8040	649.31	86.09	400	990	0.87	3.12
		Section 504	508	727.78	76.78	400	990	0.90	2.94
		Low SES Accommodated	28606 4979	707.82 639.75	80.23 81.13	400 400	990 990	0.90 0.85	2.99 3.14

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 27. Spring 2014, Mathematics Grades 7 & 8 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
7	50	Whole State	46374	726.69	85.80	400	990	0.90	2.95
		Female	22669	730.11	81.51	400	990	0.89	2.94
		Male	23700	723.44	89.59	400	990	0.91	2.95
		Asian	810	776.53	88.79	400	990	0.91	2.72
		African American	4221	685.83	86.97	400	990	0.88	3.04
		Hispanic	6698	706.76	82.39	400	990	0.88	3.02
		American Indian	7506	720.12	82.60	400	990	0.89	2.98
		White	23693	741.21	82.60	400	990	0.90	2.90
		Multiracial	3320	718.88	88.97	400	990	0.90	2.97
		ELL	2269	663.04	83.93	400	966	0.85	3.13
		IEP	7589	647.27	87.03	400	990	0.85	3.12
		Section 504	517	730.47	78.02	453	966	0.89	2.95
		Low SES	26985	705.81	83.71	400	990	0.89	3.02
		Accommodated	4708	639.09	83.37	400	966	0.82	3.13
8	50	Whole State	37626	698.41	81.03	400	990	0.89	3.04
		Female	18184	702.41	77.54	400	990	0.88	3.02
		Male	19437	694.69	84.00	400	990	0.89	3.04
		Asian	406	728.68	87.18	400	945	0.91	2.89
		African American	3819	664.44	81.32	400	902	0.86	3.11
		Hispanic	5463	684.72	79.38	400	945	0.88	3.08
		American Indian	6555	698.89	78.97	400	990	0.88	3.04
		White	18873	708.17	79.64	400	990	0.89	3.00
		Multiracial	2423	700.36	79.63	400	990	0.89	3.04
		ELL	2187	649.80	84.74	400	945	0.86	3.14
		IEP	7176	632.23	88.50	400	990	0.86	3.15
		Section 504	454	700.45	74.74	400	990	0.87	3.04
		Low SES	23555	685.34	81.65	400	990	0.88	3.07
		Accommodated	4339	628.25	86.03	400	990	0.84	3.16

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 28. Spring 2014, Reading Grades 3 & 4 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
3	50	Whole State	48752	733.10	91.18	400	990	0.91	2.94
		Female	23872	743.26	87.97	400	990	0.91	2.89
		Male	24857	723.38	93.11	400	990	0.91	2.98
		Asian	876	760.83	93.39	400	990	0.92	2.77
		African American	4390	691.10	93.43	400	990	0.91	3.11
		Hispanic	7882	700.99	88.73	400	990	0.90	3.09
		American Indian	7181	731.40	87.88	400	990	0.90	2.96
		White	24778	750.10	87.01	400	990	0.90	2.85
		Multiracial	3512	735.71	90.99	400	990	0.91	2.92
		ELL	5375	675.82	85.19	400	990	0.88	3.19
		IEP	8132	656.51	104.68	400	990	0.91	3.15
		Section 504	391	729.88	76.56	423	915	0.88	3.00
		Low SES Accommodated	30286	712.12	89.95	400	990	0.90	3.04
		6702	637.50	92.83	400	975	0.88	3.22	
4	50	Whole State	48656	720.53	87.68	400	990	0.91	2.85
		Female	23649	729.34	84.53	400	990	0.90	2.80
		Male	24977	712.22	89.79	400	990	0.92	2.89
		Asian	940	748.65	90.42	400	990	0.91	2.68
		African American	4456	682.23	87.33	400	990	0.91	3.05
		Hispanic	7726	694.10	83.19	400	990	0.90	3.01
		American Indian	7271	715.53	85.29	400	990	0.91	2.88
		White	24671	736.01	85.81	400	990	0.91	2.75
		Multiracial	3469	722.72	84.44	400	990	0.91	2.84
		ELL	3476	647.66	79.43	400	990	0.88	3.20
		IEP	8641	642.41	94.56	400	990	0.91	3.15
		Section 504	466	722.40	79.98	400	990	0.89	2.87
		Low SES Accommodated	29854	699.73	85.06	400	990	0.91	2.97
		6684	622.76	84.23	400	990	0.88	3.23	

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 29. Spring 2014, Reading Grades 5 & 6 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
5	50	Whole State	48206	723.42	94.26	400	990	0.92	2.77
		Female	23588	732.76	90.36	400	990	0.91	2.72
		Male	24614	714.49	97.00	400	990	0.92	2.81
		Asian	858	761.42	97.96	400	990	0.91	2.54
		African American	4415	682.81	95.80	400	990	0.92	2.97
		Hispanic	7441	697.08	91.15	400	990	0.91	2.92
		American Indian	7303	717.11	90.65	400	990	0.91	2.81
		White	24772	739.00	91.74	400	990	0.91	2.68
		Multiracial	3278	725.66	92.65	400	990	0.91	2.76
		ELL	2794	633.94	84.95	400	887	0.89	3.17
		IEP	8428	631.97	98.02	400	990	0.91	3.13
		Section 504	565	726.32	82.55	400	990	0.88	2.80
		Low SES	29388	700.99	91.44	400	990	0.91	2.90
Accommodated	6778	616.45	88.98	400	990	0.89	3.19		
6	50	Whole State	47475	725.32	86.99	400	990	0.90	2.94
		Female	23347	734.77	83.41	400	990	0.90	2.90
		Male	24123	716.20	89.36	400	990	0.91	2.97
		Asian	860	768.55	88.63	400	990	0.90	2.67
		African American	4394	687.00	85.35	400	990	0.90	3.10
		Hispanic	6974	701.26	82.33	400	990	0.89	3.06
		American Indian	7430	717.73	83.50	400	990	0.90	2.99
		White	24063	740.37	85.58	400	990	0.90	2.86
		Multiracial	3625	724.37	84.78	400	990	0.90	2.95
		ELL	2411	644.08	83.37	400	966	0.88	3.21
		IEP	8069	635.83	88.80	400	966	0.89	3.21
		Section 504	512	729.93	77.60	404	966	0.88	2.95
		Low SES	28582	704.31	83.98	400	990	0.90	3.05
Accommodated	4488	623.14	84.37	400	966	0.87	3.23		

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 30. Spring 2014, Reading Grades 7 & 8 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
7	50	Whole State	47402	730.89	76.62	400	990	0.92	2.74
		Female	23175	740.45	73.70	400	990	0.91	2.67
		Male	24222	721.76	78.22	400	990	0.92	2.80
		Asian	918	767.21	83.50	400	990	0.92	2.46
		African American	4266	698.98	73.60	400	990	0.91	2.97
		Hispanic	6690	709.47	73.66	400	990	0.91	2.91
		American Indian	7568	724.37	74.72	400	990	0.91	2.79
		White	24462	743.60	74.76	400	990	0.91	2.64
		Multiracial	3367	727.25	76.26	400	990	0.92	2.77
		ELL	2143	658.85	70.74	400	990	0.89	3.18
		IEP	7618	652.94	78.93	400	990	0.91	3.15
		Section 504	531	733.65	63.03	400	990	0.89	2.75
		Low SES	27282	712.18	73.85	400	990	0.91	2.89
Accommodated	3975	640.54	76.28	400	895	0.89	3.19		
8	50	Whole State	47330	740.62	88.73	400	990	0.89	2.85
		Female	23210	753.00	85.19	400	990	0.88	2.78
		Male	24119	728.72	90.43	400	990	0.89	2.91
		Asian	852	776.17	90.06	400	990	0.88	2.63
		African American	4500	698.09	89.41	400	990	0.89	3.04
		Hispanic	6402	715.55	85.70	400	990	0.88	2.98
		American Indian	7784	736.89	84.88	400	990	0.88	2.88
		White	24694	754.77	86.54	400	990	0.88	2.77
		Multiracial	2991	741.55	86.45	400	990	0.88	2.85
		ELL	2161	655.17	81.74	400	990	0.86	3.22
		IEP	7415	647.59	92.68	400	990	0.88	3.20
		Section 504	537	742.61	82.80	400	990	0.86	2.86
		Low SES	26882	717.44	87.39	400	990	0.88	2.97
Accommodated	3645	635.55	87.83	400	990	0.87	3.23		

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 31. Spring 2014, Science Grades 5 & 8 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
5	45	Whole State	48291	695.83	74.78	400	990	0.87	2.75
		Female	23632	696.78	70.83	400	990	0.86	2.75
		Male	24653	694.93	78.36	400	990	0.88	2.75
		Asian	885	722.96	81.77	400	990	0.88	2.58
		African American	4416	654.00	77.15	400	990	0.87	2.92
		Hispanic	7513	672.82	71.41	400	990	0.86	2.88
		American Indian	7294	693.62	70.98	400	990	0.86	2.78
		White	24765	709.77	71.95	400	990	0.86	2.68
		Multiracial	3275	698.54	70.75	400	990	0.86	2.74
		ELL	2925	631.11	71.04	400	918	0.84	3.02
		IEP	8408	642.07	79.24	400	990	0.87	2.96
		Section 504	568	695.80	65.08	475	990	0.84	2.78
		Low SES	29446	678.76	72.90	400	990	0.87	2.84
		Accommodated	7008	631.27	74.39	400	918	0.85	3.00
8	45	Whole State	47451	694.88	59.45	400	990	0.85	2.90
		Female	23276	695.37	54.81	400	990	0.84	2.90
		Male	24165	694.43	63.57	400	990	0.87	2.89
		Asian	871	718.76	62.95	400	990	0.87	2.75
		African American	4504	661.86	66.75	400	990	0.84	3.01
		Hispanic	6490	678.17	59.63	400	990	0.84	2.97
		American Indian	7778	691.95	56.59	400	868	0.84	2.91
		White	24750	705.31	55.44	400	990	0.85	2.85
		Multiracial	2948	695.55	57.04	400	990	0.85	2.90
		ELL	2330	642.80	61.80	400	816	0.77	3.05
		IEP	7398	644.94	67.11	400	868	0.81	3.03
		Section 504	524	697.24	58.05	400	868	0.85	2.89
		Low SES	26922	680.31	60.14	400	990	0.84	2.96
		Accommodated	4902	638.20	67.38	400	816	0.79	3.05

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 32. Spring 2014, Social Studies Grade 5 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
5	50	Whole State	48234	704.36	73.70	400	990	0.88	3.12
		Female	23616	703.88	69.61	400	990	0.87	3.14
		Male	24610	704.85	77.42	400	990	0.89	3.10
		Asian	886	734.95	79.88	400	990	0.90	2.91
		African American	4407	667.54	78.22	400	990	0.86	3.25
		Hispanic	7490	683.48	71.88	400	990	0.86	3.22
		American Indian	7287	700.46	69.19	400	990	0.86	3.16
		White	24746	716.99	70.73	400	990	0.87	3.06
		Multiracial	3275	707.52	71.82	400	990	0.87	3.12
		ELL	2937	642.86	75.46	400	951	0.82	3.30
		IEP	8379	655.27	79.67	400	990	0.86	3.27
		Section 504	568	708.89	73.00	400	990	0.86	3.12
		Low SES	29400	687.13	71.80	400	990	0.86	3.21
Accommodated	6984	644.88	76.08	400	951	0.83	3.30		

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 33. Spring 2014, U.S. History Grade 8 State and Subgroup Scale Score Descriptive Data

Grade	No. of Items	Subgroup	Sample Size	Scale Score		Min Scale Score Obtained	Max Scale Score Obtained	Coefficient	
				Mean	SD			Alpha	SEM
8	50	Whole State	47432	683.80	79.74	400	990	0.90	3.05
		Female	23265	680.26	75.10	400	990	0.88	3.08
		Male	24156	687.23	83.82	400	990	0.90	3.01
		Asian	871	725.30	81.17	400	990	0.91	2.77
		African American	4497	645.86	85.03	400	990	0.88	3.17
		Hispanic	6494	662.72	80.54	400	990	0.88	3.14
		American Indian	7774	679.34	73.92	400	990	0.88	3.08
		White	24740	696.23	76.53	400	990	0.89	2.99
		Multiracial	2947	683.72	78.56	400	990	0.89	3.05
		ELL	2339	615.71	81.32	400	834	0.83	3.26
		IEP	7408	621.14	86.20	400	990	0.86	3.23
		Section 504	521	683.37	81.63	400	922	0.90	3.04
		Low SES	26922	663.51	78.75	400	990	0.88	3.14
Accommodated	4808	611.36	83.92	400	990	0.84	3.25		

Note: SEM = Standard Error of Measurement; ELL = English Language Learner; IEP = Individualized Education Program.

Table 34. Spring 2014, Subgroup Scale Score Mean Differences, *t*-test: Accommodated/Non Accommodated

Content	Grade	<i>t</i>	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Mathematics	3	-73.70	10131.59	<.0001	-85.70	91.26
	4	-83.92	9724.92	<.0001	-96.54	89.93
	5	-89.05	9689.15	<.0001	-103.27	89.70
	6	-80.38	6076.33	<.0001	-97.17	77.44
	7	-76.35	5737.77	<.0001	-97.50	80.59
	8	-57.87	5250.95	<.0001	-79.31	76.97
Reading	3	-92.28	8410.36	<.0001	-110.83	82.81
	4	-103.25	8585.32	<.0001	-113.33	78.52
	5	-107.77	8806.96	<.0001	-124.48	83.74
	6	-85.67	5365.31	<.0001	-112.85	80.48
	7	-78.45	4628.98	<.0001	-98.62	71.57
	8	-75.49	4205.39	<.0001	-113.84	83.38
Science	5	-79.38	9176.14	<.0001	-75.53	69.89
	8	-63.31	5673.29	<.0001	-63.21	56.25
Social Studies	5	-71.66	8992.94	<.0001	-69.55	69.52
U.S. History	8	-63.79	5706.49	<.0001	-80.61	75.94

Note: SEM at or closest above the cut scores.

Table 35. Spring 2014, Subgroup Scale Score Mean Differences, *t*-test: ELL/Non ELL

Content	Grade	<i>t</i>	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Mathematics	3	-42.93	7266.57	<.0001	-56.13	94.61
	4	-44.31	4362.41	<.0001	-69.49	94.40
	5	-43.34	3389.30	<.0001	-76.87	95.10
	6	-41.51	2845.76	<.0001	-69.79	81.46
	7	-37.03	2511.03	<.0001	-66.93	84.58
	8	-27.73	2431.55	<.0001	-51.61	80.12
Reading	3	-51.97	6923.99	<.0001	-64.38	88.92
	4	-55.80	4123.61	<.0001	-78.47	85.32
	5	-57.08	3209.62	<.0001	-94.99	91.61
	6	-49.07	2685.20	<.0001	-85.59	84.93
	7	-48.11	2376.98	<.0001	-75.46	74.99
	8	-49.59	2399.99	<.0001	-89.54	86.74
Science	5	-50.75	3335.52	<.0001	-68.90	72.95
	8	-41.83	2545.94	<.0001	-54.77	58.26
Social Studies	5	-45.72	3289.97	<.0001	-65.50	72.02
U.S. History	8	-41.61	2566.47	<.0001	-71.62	78.22

Note: SEM at or closest above the cut scores.

Table 36. Spring 2014, Subgroup Scale Score Mean Differences, *t*-test: Female/Male

Content	Grade	<i>t</i>	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Mathematics	3	-2.62	48848.02	0.0089	-2.28	96.29
	4	0.54	48711.81	0.5878	0.47	96.18
	5	10.62	48281.29	<.0001	9.34	96.76
	6	2.53	47384.41	0.0113	1.93	82.96
	7	8.40	46251.64	<.0001	6.68	85.74
	8	9.27	37612.26	<.0001	7.72	80.94
Reading	3	24.23	48714.16	<.0001	19.88	90.63
	4	21.66	48622.40	<.0001	17.12	87.27
	5	21.41	48161.41	<.0001	18.27	93.81
	6	23.41	47405.86	<.0001	18.57	86.48
	7	26.79	47383.91	<.0001	18.70	76.05
	8	30.07	47305.44	<.0001	24.28	87.90
Science	5	2.72	48117.73	0.0065	1.85	74.77
	8	1.73	46870.31	0.0828	0.94	59.44
Social Studies	5	-1.44	48021.64	0.1491	-0.97	73.69
U.S. History	8	-9.55	47174.53	<.0001	-6.97	79.66

Note: SEM at or closest above the cut scores.

Table 37. Spring 2014, Mathematics Subgroup Scale Score Mean Differences, *t*-test: IEP/Non IEP

Content	Grade	<i>t</i>	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Mathematics	3	-54.92	10789.02	<.0001	-67.33	92.98
	4	-69.82	11908.05	<.0001	-80.25	91.16
	5	-81.25	11617.79	<.0001	-92.99	90.22
	6	-90.69	10511.20	<.0001	-93.21	75.24
	7	-88.60	10007.70	<.0001	-94.96	78.28
	8	-72.99	9441.46	<.0001	-81.78	74.38
Reading	3	-74.94	10108.16	<.0001	-91.92	84.50
	4	-87.43	11189.76	<.0001	-94.98	79.81
	5	-97.00	11002.73	<.0001	-110.83	84.34
	6	-102.00	10504.89	<.0001	-107.82	76.98
	7	-96.37	9789.57	<.0001	-92.88	68.60
	8	-96.58	9375.39	<.0001	-110.31	79.16
Science	5	-70.00	11216.79	<.0001	-65.10	70.59
	8	-71.80	9177.17	<.0001	-59.16	55.44
Social Studies	5	-63.58	11087.61	<.0001	-59.42	70.18
U.S. History	8	-69.69	9460.57	<.0001	-74.26	75.04

Note: SEM at or closest above the cut scores.

Table 38. Spring 2014, Mathematics Subgroup Scale Score Mean Differences, *t*-test: Low SES/High SES

Content	Grade	<i>t</i>	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Mathematics	3	-62.92	40816.10	<.0001	-53.71	92.69
	4	-69.24	40670.80	<.0001	-58.84	91.81
	5	-64.08	40533.09	<.0001	-55.46	93.03
	6	-63.96	40910.77	<.0001	-47.58	79.64
	7	-65.02	42835.32	<.0001	-49.94	82.19
	8	-42.27	31570.46	<.0001	-34.94	79.24
Reading	3	-69.55	41674.33	<.0001	-55.38	87.13
	4	-69.77	41252.68	<.0001	-53.82	83.67
	5	-69.03	41317.37	<.0001	-57.47	90.00
	6	-68.18	41226.70	<.0001	-52.80	83.06
	7	-64.72	43652.86	<.0001	-44.09	73.45
	8	-69.01	45545.09	<.0001	-53.65	84.66
Science	5	-66.09	41462.30	<.0001	-43.75	71.67
	8	-64.83	46534.54	<.0001	-33.68	57.06
Social Studies	5	-67.81	41554.72	<.0001	-44.13	70.49
U.S. History	8	-67.09	45679.45	<.0001	-46.93	76.27

Note: SEM at or closest above the cut scores.

Table 39. Spring 2014, Mathematics Subgroup Scale Score Mean Differences, *t*-test: Section 504/Non Section 504

Content	Grade	<i>t</i>	Degrees of Freedom	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Mathematics	3	-1.58	48885.00	0.1146	-7.68	96.29
	4	-0.86	48756.00	0.3878	-3.88	96.17
	5	-0.90	48297.00	0.3662	-3.70	96.88
	6	0.29	47472.00	0.7751	1.06	82.97
	7	1.01	46372.00	0.3143	3.82	85.80
	8	0.54	37624.00	0.5902	2.06	81.03
Reading	3	-0.70	48750.00	0.4837	-3.24	91.18
	4	0.46	48654.00	0.6434	1.89	87.68
	5	0.73	48204.00	0.4633	2.93	94.26
	6	1.20	47473.00	0.2286	4.65	86.98
	7	0.83	47400.00	0.4052	2.78	76.62
	8	0.52	47328.00	0.6016	2.01	88.73
Science	5	-0.01	48289.00	0.9924	-0.03	74.78
	8	0.91	47449.00	0.3613	2.38	59.45
Social Studies	5	1.47	48232.00	0.1414	4.57	73.70
U.S. History	8	-0.12	47430.00	0.9008	-0.44	79.74

Note: SEM at or closest above the cut scores.

Table 40. Spring 2014, Mathematics Subgroup Mean Differences, ANOVA: Ethnicity

Grade	Categories	Sum of Squares	df	Mean Square	<i>F</i>	Sig.
3	Between Groups	28033920.10	5	5606784.00	645.33	<.0001
	Within Groups	423498398.30	48744	8688.20	.	.
	Total	451532318.40	48749	.	.	.
4	Between Groups	27148455.20	5	5429691.00	625.05	<.0001
	Within Groups	422405315.40	48626	8686.80	.	.
	Total	449553770.60	48631	.	.	.
5	Between Groups	20509100.20	5	4101820.00	458.14	<.0001
	Within Groups	431084377.00	48149	8953.10	.	.
	Total	451593477.20	48154	.	.	.
6	Between Groups	17754129.50	5	3550825.90	546.67	<.0001
	Within Groups	307451183.90	47334	6495.40	.	.
	Total	325205313.40	47339	.	.	.
7	Between Groups	17241515.00	5	3448303.00	493.88	<.0001
	Within Groups	322867423.60	46242	6982.10	.	.
	Total	340108938.60	46247	.	.	.
8	Between Groups	7610353.70	5	1522070.70	239.37	<.0001
	Within Groups	238659327.20	37533	6358.70	.	.
	Total	246269680.90	37538	.	.	.

Note: df = Degrees of Freedom

Table 41. Spring 2014, Reading Subgroup Mean Differences, ANOVA: Ethnicity

Grade	Categories	Sum of Squares	df	Mean Square	<i>F</i>	Sig.
3	Between Groups	23751530.60	5	4750306.10	607.54	<.0001
	Within Groups	380100689.50	48613	7818.90	.	.
	Total	403852220.10	48618	.	.	.
4	Between Groups	18783831.80	5	3756766.40	514.46	<.0001
	Within Groups	354362944.60	48527	7302.40	.	.
	Total	373146776.30	48532	.	.	.
5	Between Groups	20003046.90	5	4000609.40	472.25	<.0001
	Within Groups	407144125.20	48061	8471.40	.	.
	Total	427147172.10	48066	.	.	.
6	Between Groups	17979910.60	5	3595982.10	500.55	<.0001
	Within Groups	340093606.30	47340	7184.10	.	.
	Total	358073516.90	47345	.	.	.
7	Between Groups	12944580.60	5	2588916.10	462.90	<.0001
	Within Groups	264344206.50	47265	5592.80	.	.
	Total	277288787.20	47270	.	.	.
8	Between Groups	18291322.20	5	3658264.40	489.02	<.0001
	Within Groups	353217696.60	47217	7480.70	.	.
	Total	371509018.80	47222	.	.	.

Note: df = Degrees of Freedom

Table 42. Spring 2014, Science Subgroup Mean Differences, ANOVA: Ethnicity

Grade	Categories	Sum of Squares	df	Mean Square	<i>F</i>	Sig.
5	Between Groups	17228749.90	5	3445750.00	658.73	<.0001
	Within Groups	251826731.50	48142	5230.90	.	.
	Total	269055481.30	48147	.	.	.
8	Between Groups	9980614.10	5	1996122.80	601.31	<.0001
	Within Groups	157135128.10	47335	3319.60	.	.
	Total	167115742.20	47340	.	.	.

Note: df = Degrees of Freedom

Table 43. Spring 2014, Social Studies Subgroup Mean Differences, ANOVA: Ethnicity

Grade	Categories	Sum of Squares	df	Mean Square	<i>F</i>	Sig.
5	Between Groups	14162045.20	5	2832409.00	551.68	<.0001
	Within Groups	246874093.60	48085	5134.10	.	.
	Total	261036138.80	48090	.	.	.

Note: df = Degrees of Freedom

Table 44. Spring 2014, U.S. History Subgroup Mean Differences, ANOVA: Ethnicity

Grade	Categories	Sum of Squares	df	Mean Square	<i>F</i>	Sig.
8	Between Groups	14832837.60	5	2966567.50	490.97	<.0001
	Within Groups	285904413.80	47317	6042.30	.	.
	Total	300737251.40	47322	.	.	.

Note: df = Degrees of Freedom

Table 45. Mathematics, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
3	SS	American Indian/Alaskan	Asian	47.03(*)	2.53	38.72	55.33
			African Amer. (Not Hispanic)	-51.38(*)	2.53	-55.89	-46.86
			Hispanic	-30.42(*)	2.53	-34.26	-26.59
			Multiracial	0.98	2.53	-3.87	5.83
			White (Not Hispanic)	16.87(*)	2.53	13.71	20.03
		Asian	Amer. Indian/Alaskan	-47.03(*)	2.30	-54.59	-39.47
			African Amer. (Not Hispanic)	-98.4(*)	2.30	-106.23	-90.58
			Hispanic	-77.45(*)	2.30	-84.97	-69.93
			Multiracial	-46.05(*)	2.30	-54.03	-38.06
			White (Not Hispanic)	-30.16(*)	2.30	-37.41	-22.91
		African American (Not Hispanic)	Amer. Indian/Alaskan	51.38(*)	2.50	46.92	55.83
			Asian	98.4(*)	2.50	89.92	106.88
			Hispanic	20.95(*)	2.50	16.58	25.32
			Multiracial	52.36(*)	2.50	47.09	57.62
			White (Not Hispanic)	68.24(*)	2.50	64.44	72.05
		Hispanic	Amer. Indian/Alaskan	30.42(*)	2.53	26.58	34.27
			Asian	77.45(*)	2.53	69.18	85.73
			African Amer. (Not Hispanic)	-20.95(*)	2.53	-25.39	-16.51
			Multiracial	31.41(*)	2.53	26.62	36.19
			White (Not Hispanic)	47.29(*)	2.53	44.25	50.33
		Multiracial	Amer. Indian/Alaskan	-0.98	2.48	-5.73	3.77
			Asian	46.05(*)	2.48	37.46	54.64
			African Amer. (Not Hispanic)	-52.36(*)	2.48	-57.58	-47.14
			Hispanic	-31.41(*)	2.48	-36.08	-26.73
			White (Not Hispanic)	15.89(*)	2.48	11.73	20.04
		White (Not Hispanic)	Amer. Indian/Alaskan	-16.87(*)	2.56	-20.07	-13.66
			Asian	30.16(*)	2.56	22.08	38.24
			African Amer. (Not Hispanic)	-68.24(*)	2.56	-72.16	-64.33
Hispanic	-47.29(*)		2.56	-50.37	-44.22		
Multiracial	-15.89(*)		2.56	-20.19	-11.58		

(*)Significant differences

Table 45. Mathematics, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity
(continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
4	SS	American Indian/Alaskan	Asian	56.13(*)	2.53	48.02	64.24
			African Amer. (Not Hispanic)	-49.08(*)	2.53	-53.57	-44.60
			Hispanic	-19.14(*)	2.53	-22.98	-15.29
			Multiracial	6.65(*)	2.53	1.79	11.51
			White (Not Hispanic)	21.71(*)	2.53	18.56	24.85
		Asian	Amer. Indian/Alaskan	-56.13(*)	2.31	-63.53	-48.73
			African Amer. (Not Hispanic)	-105.21(*)	2.31	-112.88	-97.54
			Hispanic	-75.27(*)	2.31	-82.64	-67.90
			Multiracial	-49.48(*)	2.31	-57.34	-41.62
			White (Not Hispanic)	-34.43(*)	2.31	-41.52	-27.33
		African American (Not Hispanic)	Amer. Indian/Alaskan	49.08(*)	2.50	44.66	53.51
			Asian	105.21(*)	2.50	96.92	113.51
			Hispanic	29.95(*)	2.50	25.58	34.31
			Multiracial	55.73(*)	2.50	50.47	61.00
			White (Not Hispanic)	70.79(*)	2.50	67.00	74.58
		Hispanic	Amer. Indian/Alaskan	19.14(*)	2.53	15.29	22.99
			Asian	75.27(*)	2.53	67.18	83.36
			African Amer. (Not Hispanic)	-29.95(*)	2.53	-34.38	-25.51
			Multiracial	25.79(*)	2.53	20.97	30.60
			White (Not Hispanic)	40.84(*)	2.53	37.78	43.91
		Multiracial	Amer. Indian/Alaskan	-6.65(*)	2.47	-11.41	-1.90
			Asian	49.48(*)	2.47	41.06	57.90
			African Amer. (Not Hispanic)	-55.73(*)	2.47	-60.95	-50.52
			Hispanic	-25.79(*)	2.47	-30.49	-21.09
			White (Not Hispanic)	15.05(*)	2.47	10.88	19.23
		White (Not Hispanic)	Amer. Indian/Alaskan	-21.71(*)	2.56	-24.89	-18.52
			Asian	34.43(*)	2.56	26.55	42.30
			African Amer. (Not Hispanic)	-70.79(*)	2.56	-74.68	-66.90
Hispanic	-40.84(*)		2.56	-43.95	-37.74		
Multiracial	-15.05(*)		2.56	-19.38	-10.73		

(*)Significant differences

Table 45. Mathematics, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity (continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
5	SS	American Indian/Alaskan	Asian	73.5(*)	2.53	64.97	82.03
			African Amer. (Not Hispanic)	-35.3(*)	2.53	-39.87	-30.73
			Hispanic	-12.3(*)	2.53	-16.23	-8.36
			Multiracial	6.4(*)	2.53	1.36	11.44
			White (Not Hispanic)	22.02(*)	2.53	18.83	25.21
		Asian	Amer. Indian/Alaskan	-73.5(*)	2.30	-81.26	-65.74
			African Amer. (Not Hispanic)	-108.8(*)	2.30	-116.83	-100.77
			Hispanic	-85.8(*)	2.30	-93.54	-78.05
			Multiracial	-67.1(*)	2.30	-75.36	-58.84
			White (Not Hispanic)	-51.48(*)	2.30	-58.93	-44.02
		African American (Not Hispanic)	Amer. Indian/Alaskan	35.3(*)	2.50	30.79	39.81
			Asian	108.8(*)	2.50	100.09	117.51
			Hispanic	23(*)	2.50	18.52	27.49
			Multiracial	41.7(*)	2.50	36.25	47.15
			White (Not Hispanic)	57.32(*)	2.50	53.46	61.19
		Hispanic	Amer. Indian/Alaskan	12.3(*)	2.53	8.36	16.24
			Asian	85.8(*)	2.53	77.27	94.32
			African Amer. (Not Hispanic)	-23(*)	2.53	-27.55	-18.46
			Multiracial	18.69(*)	2.53	13.68	23.71
			White (Not Hispanic)	34.32(*)	2.53	31.16	37.48
		Multiracial	Amer. Indian/Alaskan	-6.4(*)	2.47	-11.32	-1.48
			Asian	67.1(*)	2.47	58.24	75.96
			African Amer. (Not Hispanic)	-41.7(*)	2.47	-47.09	-36.31
			Hispanic	-18.69(*)	2.47	-23.59	-13.80
			White (Not Hispanic)	15.63(*)	2.47	11.28	19.97
		White (Not Hispanic)	Amer. Indian/Alaskan	-22.02(*)	2.56	-25.26	-18.79
			Asian	51.48(*)	2.56	43.17	59.78
			African Amer. (Not Hispanic)	-57.32(*)	2.56	-61.29	-53.36
Hispanic	-34.32(*)		2.56	-37.52	-31.12		
Multiracial	-15.63(*)		2.56	-20.14	-11.12		

(*)Significant differences

Table 45. Mathematics, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity (continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
6	SS	American Indian/Alaskan	Asian	65.12(*)	2.53	57.79	72.44
			African Amer. (Not Hispanic)	-34.02(*)	2.53	-37.91	-30.13
			Hispanic	-10.51(*)	2.53	-13.90	-7.11
			Multiracial	4.57(*)	2.53	0.44	8.71
			White (Not Hispanic)	21.37(*)	2.53	18.67	24.08
		Asian	Amer. Indian/Alaskan	-65.12(*)	2.30	-71.77	-58.47
			African Amer. (Not Hispanic)	-99.14(*)	2.30	-106.03	-92.25
			Hispanic	-75.62(*)	2.30	-82.29	-68.96
			Multiracial	-60.54(*)	2.30	-67.55	-53.54
			White (Not Hispanic)	-43.74(*)	2.30	-50.15	-37.34
		African American (Not Hispanic)	Amer. Indian/Alaskan	34.02(*)	2.50	30.19	37.86
			Asian	99.14(*)	2.50	91.66	106.62
			Hispanic	23.52(*)	2.50	19.64	27.39
			Multiracial	38.6(*)	2.50	34.07	43.12
			White (Not Hispanic)	55.4(*)	2.50	52.09	58.71
		Hispanic	Amer. Indian/Alaskan	10.51(*)	2.53	7.11	13.90
			Asian	75.62(*)	2.53	68.29	82.96
			African Amer. (Not Hispanic)	-23.52(*)	2.53	-27.45	-19.59
			Multiracial	15.08(*)	2.53	10.91	19.25
			White (Not Hispanic)	31.88(*)	2.53	29.12	34.65
		Multiracial	Amer. Indian/Alaskan	-4.57(*)	2.48	-8.62	-0.52
			Asian	60.54(*)	2.48	52.99	68.10
			African Amer. (Not Hispanic)	-38.6(*)	2.48	-43.09	-34.10
			Hispanic	-15.08(*)	2.48	-19.17	-10.99
			White (Not Hispanic)	16.8(*)	2.48	13.24	20.36
		White (Not Hispanic)	Amer. Indian/Alaskan	-21.37(*)	2.56	-24.12	-18.63
			Asian	43.74(*)	2.56	36.60	50.89
			African Amer. (Not Hispanic)	-55.4(*)	2.56	-58.80	-52.00
Hispanic	-31.88(*)		2.56	-34.68	-29.08		
Multiracial	-16.8(*)		2.56	-20.49	-13.12		

(*)Significant differences

Table 45. Mathematics, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity (continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
7	SS	American Indian/Alaskan	Asian	56.41(*)	2.53	48.58	64.24
			African Amer. (Not Hispanic)	-34.29(*)	2.53	-38.36	-30.21
			Hispanic	-13.36(*)	2.53	-16.92	-9.80
			Multiracial	-1.24	2.53	-5.65	3.18
			White (Not Hispanic)	21.09(*)	2.53	18.29	23.90
		Asian	Amer. Indian/Alaskan	-56.41(*)	2.30	-63.50	-49.32
			African Amer. (Not Hispanic)	-90.7(*)	2.30	-98.05	-83.34
			Hispanic	-69.77(*)	2.30	-76.90	-62.63
			Multiracial	-57.65(*)	2.30	-65.16	-50.13
			White (Not Hispanic)	-35.32(*)	2.30	-42.17	-28.46
		African American (Not Hispanic)	Amer. Indian/Alaskan	34.29(*)	2.50	30.27	38.30
			Asian	90.7(*)	2.50	82.69	98.70
			Hispanic	20.93(*)	2.50	16.83	25.03
			Multiracial	33.05(*)	2.50	28.21	37.89
			White (Not Hispanic)	55.38(*)	2.50	51.89	58.86
		Hispanic	Amer. Indian/Alaskan	13.36(*)	2.53	9.81	16.91
			Asian	69.77(*)	2.53	61.91	77.63
			African Amer. (Not Hispanic)	-20.93(*)	2.53	-25.08	-16.77
			Multiracial	12.12(*)	2.53	7.64	16.61
			White (Not Hispanic)	34.45(*)	2.53	31.53	37.38
		Multiracial	Amer. Indian/Alaskan	1.24	2.48	-3.07	5.55
			Asian	57.65(*)	2.48	49.54	65.76
			African Amer. (Not Hispanic)	-33.05(*)	2.48	-37.85	-28.25
			Hispanic	-12.12(*)	2.48	-16.51	-7.73
			White (Not Hispanic)	22.33(*)	2.48	18.50	26.17
		White (Not Hispanic)	Amer. Indian/Alaskan	-21.09(*)	2.56	-23.93	-18.25
			Asian	35.32(*)	2.56	27.66	42.97
			African Amer. (Not Hispanic)	-55.38(*)	2.56	-58.96	-51.80
Hispanic	-34.45(*)		2.56	-37.42	-31.49		
Multiracial	-22.33(*)		2.56	-26.30	-18.36		

(*)Significant differences

Table 45. Mathematics, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity (continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
8	SS	American Indian/Alaskan	Asian	29.8(*)	2.54	19.44	40.15
			African Amer. (Not Hispanic)	-34.45(*)	2.54	-38.57	-30.32
			Hispanic	-14.17(*)	2.54	-17.88	-10.46
			Multiracial	1.47	2.54	-3.34	6.29
			White (Not Hispanic)	9.28(*)	2.54	6.38	12.18
		Asian	Amer. Indian/Alaskan	-29.8(*)	2.24	-38.93	-20.66
			African Amer. (Not Hispanic)	-64.24(*)	2.24	-73.57	-54.92
			Hispanic	-43.96(*)	2.24	-53.15	-34.77
			Multiracial	-28.32(*)	2.24	-37.90	-18.74
			White (Not Hispanic)	-20.52(*)	2.24	-29.48	-11.55
		African American (Not Hispanic)	Amer. Indian/Alaskan	34.45(*)	2.51	30.38	38.51
			Asian	64.24(*)	2.51	53.81	74.68
			Hispanic	20.28(*)	2.51	16.06	24.49
			Multiracial	35.92(*)	2.51	30.73	41.11
			White (Not Hispanic)	43.73(*)	2.51	40.18	47.27
		Hispanic	Amer. Indian/Alaskan	14.17(*)	2.53	10.47	17.86
			Asian	43.96(*)	2.53	33.59	54.34
			African Amer. (Not Hispanic)	-20.28(*)	2.53	-24.53	-16.02
			Multiracial	15.64(*)	2.53	10.72	20.57
			White (Not Hispanic)	23.45(*)	2.53	20.35	26.55
		Multiracial	Amer. Indian/Alaskan	-1.47	2.47	-6.15	3.21
			Asian	28.32(*)	2.47	17.76	38.88
			African Amer. (Not Hispanic)	-35.92(*)	2.47	-41.03	-30.81
			Hispanic	-15.64(*)	2.47	-20.45	-10.84
			White (Not Hispanic)	7.81(*)	2.47	3.56	12.05
		White (Not Hispanic)	Amer. Indian/Alaskan	-9.28(*)	2.56	-12.21	-6.35
			Asian	20.52(*)	2.56	10.26	30.77
			African Amer. (Not Hispanic)	-43.73(*)	2.56	-47.35	-40.10
Hispanic	-23.45(*)		2.56	-26.59	-20.31		
Multiracial	-7.81(*)		2.56	-12.22	-3.39		

(*)Significant differences

Table 46. Reading, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
3	SS	American Indian/Alaskan	Asian	29.43(*)	2.53	21.42	37.44
			African Amer. (Not Hispanic)	-40.3(*)	2.53	-44.59	-36.02
			Hispanic	-30.41(*)	2.53	-34.06	-26.77
			Multiracial	4.31	2.53	-0.30	8.91
			White (Not Hispanic)	18.7(*)	2.53	15.70	21.70
		Asian	Amer. Indian/Alaskan	-29.43(*)	2.30	-36.70	-22.16
			African Amer. (Not Hispanic)	-69.73(*)	2.30	-77.25	-62.21
			Hispanic	-59.84(*)	2.30	-67.08	-52.61
			Multiracial	-25.12(*)	2.30	-32.80	-17.45
			White (Not Hispanic)	-10.73(*)	2.30	-17.72	-3.74
		African American (Not Hispanic)	Amer. Indian/Alaskan	40.3(*)	2.50	36.08	44.53
			Asian	69.73(*)	2.50	61.57	77.90
			Hispanic	9.89(*)	2.50	5.73	14.05
			Multiracial	44.61(*)	2.50	39.61	49.61
			White (Not Hispanic)	59(*)	2.50	55.39	62.62
		Hispanic	Amer. Indian/Alaskan	30.41(*)	2.53	26.76	34.07
			Asian	59.84(*)	2.53	51.86	67.83
			African Amer. (Not Hispanic)	-9.89(*)	2.53	-14.11	-5.67
			Multiracial	34.72(*)	2.53	30.17	39.27
			White (Not Hispanic)	49.11(*)	2.53	46.21	52.01
		Multiracial	Amer. Indian/Alaskan	-4.31	2.48	-8.82	0.20
			Asian	25.12(*)	2.48	16.85	33.39
			African Amer. (Not Hispanic)	-44.61(*)	2.48	-49.57	-39.65
			Hispanic	-34.72(*)	2.48	-39.16	-30.28
			White (Not Hispanic)	14.39(*)	2.48	10.44	18.34
		White (Not Hispanic)	Amer. Indian/Alaskan	-18.7(*)	2.56	-21.74	-15.66
			Asian	10.73(*)	2.56	2.94	18.53
			African Amer. (Not Hispanic)	-59(*)	2.56	-62.72	-55.29
			Hispanic	-49.11(*)	2.56	-52.05	-46.18
			Multiracial	-14.39(*)	2.56	-18.48	-10.30

(*)Significant differences

Table 46. Reading, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity
(continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
4	SS	American Indian/Alaskan	Asian	33.11(*)	2.53	25.62	40.61
			African Amer. (Not Hispanic)	-33.31(*)	2.53	-37.42	-29.19
			Hispanic	-21.43(*)	2.53	-24.96	-17.90
			Multiracial	7.19(*)	2.53	2.73	11.65
			White (Not Hispanic)	20.48(*)	2.53	17.59	23.36
		Asian	Amer. Indian/Alaskan	-33.11(*)	2.31	-39.95	-26.28
			African Amer. (Not Hispanic)	-66.42(*)	2.31	-73.49	-59.34
			Hispanic	-54.54(*)	2.31	-61.35	-47.73
			Multiracial	-25.92(*)	2.31	-33.17	-18.68
			White (Not Hispanic)	-12.64(*)	2.31	-19.19	-6.09
		African American (Not Hispanic)	Amer. Indian/Alaskan	33.31(*)	2.50	29.25	37.36
			Asian	66.42(*)	2.50	58.76	74.08
			Hispanic	11.88(*)	2.50	7.86	15.89
			Multiracial	40.49(*)	2.50	35.66	45.32
			White (Not Hispanic)	53.78(*)	2.50	50.31	57.25
		Hispanic	Amer. Indian/Alaskan	21.43(*)	2.53	17.89	24.97
			Asian	54.54(*)	2.53	47.06	62.02
			African Amer. (Not Hispanic)	-11.88(*)	2.53	-15.95	-7.80
			Multiracial	28.62(*)	2.53	24.19	33.04
			White (Not Hispanic)	41.9(*)	2.53	39.08	44.73
		Multiracial	Amer. Indian/Alaskan	-7.19(*)	2.47	-11.55	-2.83
			Asian	25.92(*)	2.47	18.15	33.70
			African Amer. (Not Hispanic)	-40.49(*)	2.47	-45.28	-35.71
			Hispanic	-28.62(*)	2.47	-32.94	-24.30
			White (Not Hispanic)	13.29(*)	2.47	9.45	17.12
		White (Not Hispanic)	Amer. Indian/Alaskan	-20.48(*)	2.56	-23.40	-17.55
			Asian	12.64(*)	2.56	5.36	19.92
			African Amer. (Not Hispanic)	-53.78(*)	2.56	-57.35	-50.21
Hispanic	-41.9(*)		2.56	-44.76	-39.05		
Multiracial	-13.29(*)		2.56	-17.26	-9.31		

(*)Significant differences

Table 46. Reading, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity
(continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
5	SS	American Indian/Alaskan	Asian	44.31(*)	2.53	35.91	52.72
			African Amer. (Not Hispanic)	-34.3(*)	2.53	-38.74	-29.86
			Hispanic	-20.02(*)	2.53	-23.86	-16.19
			Multiracial	8.55(*)	2.53	3.65	13.45
			White (Not Hispanic)	21.89(*)	2.53	18.79	25.00
		Asian	Amer. Indian/Alaskan	-44.31(*)	2.30	-51.95	-36.68
			African Amer. (Not Hispanic)	-78.62(*)	2.30	-86.51	-70.72
			Hispanic	-64.34(*)	2.30	-71.97	-56.71
			Multiracial	-35.77(*)	2.30	-43.88	-27.65
			White (Not Hispanic)	-22.42(*)	2.30	-29.77	-15.07
		African American (Not Hispanic)	Amer. Indian/Alaskan	34.3(*)	2.50	29.92	38.68
			Asian	78.62(*)	2.50	70.04	87.19
			Hispanic	14.28(*)	2.50	9.91	18.64
			Multiracial	42.85(*)	2.50	37.55	48.15
			White (Not Hispanic)	56.19(*)	2.50	52.44	59.95
		Hispanic	Amer. Indian/Alaskan	20.02(*)	2.53	16.19	23.86
			Asian	64.34(*)	2.53	55.93	72.74
			African Amer. (Not Hispanic)	-14.28(*)	2.53	-18.70	-9.85
			Multiracial	28.57(*)	2.53	23.69	33.46
			White (Not Hispanic)	41.92(*)	2.53	38.84	45.00
		Multiracial	Amer. Indian/Alaskan	-8.55(*)	2.47	-13.33	-3.77
			Asian	35.77(*)	2.47	27.04	44.49
			African Amer. (Not Hispanic)	-42.85(*)	2.47	-48.09	-37.61
			Hispanic	-28.57(*)	2.47	-33.34	-23.81
			White (Not Hispanic)	13.34(*)	2.47	9.12	17.57
		White (Not Hispanic)	Amer. Indian/Alaskan	-21.89(*)	2.56	-25.04	-18.75
			Asian	22.42(*)	2.56	14.23	30.62
			African Amer. (Not Hispanic)	-56.19(*)	2.56	-60.05	-52.34
			Hispanic	-41.92(*)	2.56	-45.04	-38.80
			Multiracial	-13.34(*)	2.56	-17.73	-8.96

(*)Significant differences

Table 46. Reading, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity
(continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
6	SS	American Indian/Alaskan	Asian	50.82(*)	2.53	43.09	58.55
			African Amer. (Not Hispanic)	-30.73(*)	2.53	-34.82	-26.65
			Hispanic	-16.47(*)	2.53	-20.05	-12.89
			Multiracial	6.64(*)	2.53	2.29	10.99
			White (Not Hispanic)	22.64(*)	2.53	19.79	25.49
		Asian	Amer. Indian/Alaskan	-50.82(*)	2.30	-57.84	-43.81
			African Amer. (Not Hispanic)	-81.56(*)	2.30	-88.82	-74.29
			Hispanic	-67.29(*)	2.30	-74.33	-60.25
			Multiracial	-44.18(*)	2.30	-51.57	-36.79
			White (Not Hispanic)	-28.18(*)	2.30	-34.94	-21.42
		African American (Not Hispanic)	Amer. Indian/Alaskan	30.73(*)	2.50	26.70	34.76
			Asian	81.56(*)	2.50	73.66	89.45
			Hispanic	14.26(*)	2.50	10.19	18.34
			Multiracial	37.38(*)	2.50	32.63	42.12
			White (Not Hispanic)	53.37(*)	2.50	49.90	56.85
		Hispanic	Amer. Indian/Alaskan	16.47(*)	2.53	12.89	20.04
			Asian	67.29(*)	2.53	59.54	75.04
			African Amer. (Not Hispanic)	-14.26(*)	2.53	-18.39	-10.14
			Multiracial	23.11(*)	2.53	18.72	27.50
			White (Not Hispanic)	39.11(*)	2.53	36.19	42.03
		Multiracial	Amer. Indian/Alaskan	-6.64(*)	2.48	-10.90	-2.38
			Asian	44.18(*)	2.48	36.21	52.15
			African Amer. (Not Hispanic)	-37.38(*)	2.48	-42.09	-32.66
			Hispanic	-23.11(*)	2.48	-27.42	-18.81
			White (Not Hispanic)	16(*)	2.48	12.25	19.75
		White (Not Hispanic)	Amer. Indian/Alaskan	-22.64(*)	2.56	-25.53	-19.76
			Asian	28.18(*)	2.56	20.64	35.72
			African Amer. (Not Hispanic)	-53.37(*)	2.56	-56.94	-49.81
			Hispanic	-39.11(*)	2.56	-42.07	-36.15
			Multiracial	-16(*)	2.56	-19.87	-12.13

(*)Significant differences

Table 46. Reading, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity
(continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
7	SS	American Indian/Alaskan	Asian	42.84(*)	2.53	36.22	49.46
			African Amer. (Not Hispanic)	-25.39(*)	2.53	-29.02	-21.76
			Hispanic	-14.9(*)	2.53	-18.08	-11.72
			Multiracial	2.88	2.53	-1.05	6.81
			White (Not Hispanic)	19.23(*)	2.53	16.74	21.73
		Asian	Amer. Indian/Alaskan	-42.84(*)	2.31	-48.88	-36.81
			African Amer. (Not Hispanic)	-68.23(*)	2.31	-74.52	-61.95
			Hispanic	-57.74(*)	2.31	-63.82	-51.66
			Multiracial	-39.96(*)	2.31	-46.39	-33.53
			White (Not Hispanic)	-23.61(*)	2.31	-29.41	-17.80
		African American (Not Hispanic)	Amer. Indian/Alaskan	25.39(*)	2.50	21.82	28.97
			Asian	68.23(*)	2.50	61.44	75.03
			Hispanic	10.49(*)	2.50	6.84	14.15
			Multiracial	28.27(*)	2.50	23.97	32.58
			White (Not Hispanic)	44.63(*)	2.50	41.53	47.72
		Hispanic	Amer. Indian/Alaskan	14.9(*)	2.53	11.73	18.07
			Asian	57.74(*)	2.53	51.09	64.39
			African Amer. (Not Hispanic)	-10.49(*)	2.53	-14.20	-6.79
			Multiracial	17.78(*)	2.53	13.79	21.77
			White (Not Hispanic)	34.13(*)	2.53	31.52	36.74
		Multiracial	Amer. Indian/Alaskan	-2.88	2.48	-6.71	0.95
			Asian	39.96(*)	2.48	33.07	46.85
			African Amer. (Not Hispanic)	-28.27(*)	2.48	-32.54	-24.01
			Hispanic	-17.78(*)	2.48	-21.69	-13.87
			White (Not Hispanic)	16.35(*)	2.48	12.95	19.76
		White (Not Hispanic)	Amer. Indian/Alaskan	-19.23(*)	2.56	-21.76	-16.71
			Asian	23.61(*)	2.56	17.16	30.06
			African Amer. (Not Hispanic)	-44.63(*)	2.56	-47.81	-41.44
Hispanic	-34.13(*)		2.56	-36.78	-31.49		
Multiracial	-16.35(*)		2.56	-19.88	-12.83		

(*)Significant differences

Table 46. Reading, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity
(continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
8	SS	American Indian/Alaskan	Asian	39.28(*)	2.54	31.36	47.20
			African Amer. (Not Hispanic)	-38.79(*)	2.54	-42.90	-34.68
			Hispanic	-21.34(*)	2.54	-25.04	-17.64
			Multiracial	4.66	2.54	-0.06	9.38
			White (Not Hispanic)	17.88(*)	2.54	15.03	20.73
		Asian	Amer. Indian/Alaskan	-39.28(*)	2.30	-46.47	-32.09
			African Amer. (Not Hispanic)	-78.07(*)	2.30	-85.51	-70.63
			Hispanic	-60.62(*)	2.30	-67.88	-53.36
			Multiracial	-34.62(*)	2.30	-42.35	-26.88
			White (Not Hispanic)	-21.4(*)	2.30	-28.34	-14.46
		African American (Not Hispanic)	Amer. Indian/Alaskan	38.79(*)	2.50	34.74	42.84
			Asian	78.07(*)	2.50	69.99	86.16
			Hispanic	17.45(*)	2.50	13.24	21.66
			Multiracial	43.45(*)	2.50	38.35	48.56
			White (Not Hispanic)	56.67(*)	2.50	53.16	60.18
		Hispanic	Amer. Indian/Alaskan	21.34(*)	2.53	17.65	25.03
			Asian	60.62(*)	2.53	52.65	68.59
			African Amer. (Not Hispanic)	-17.45(*)	2.53	-21.70	-13.20
			Multiracial	26(*)	2.53	21.16	30.84
			White (Not Hispanic)	39.22(*)	2.53	36.16	42.28
		Multiracial	Amer. Indian/Alaskan	-4.66(*)	2.47	-9.25	-0.07
			Asian	34.62(*)	2.47	26.34	42.90
			African Amer. (Not Hispanic)	-43.45(*)	2.47	-48.48	-38.42
			Hispanic	-26(*)	2.47	-30.73	-21.28
			White (Not Hispanic)	13.22(*)	2.47	9.09	17.35
		White (Not Hispanic)	Amer. Indian/Alaskan	-17.88(*)	2.56	-20.76	-15.00
			Asian	21.4(*)	2.56	13.67	29.13
			African Amer. (Not Hispanic)	-56.67(*)	2.56	-60.27	-53.08
Hispanic	-39.22(*)		2.56	-42.33	-36.11		
Multiracial	-13.22(*)		2.56	-17.51	-8.92		

(*)Significant differences

Table 47. Science, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
5	SS	American Indian/Alaskan	Asian	29.33(*)	2.53	22.82	35.85
			African Amer. (Not Hispanic)	-39.62(*)	2.53	-43.11	-36.13
			Hispanic	-20.81(*)	2.53	-23.81	-17.80
			Multiracial	4.91(*)	2.53	1.06	8.76
			White (Not Hispanic)	16.15(*)	2.53	13.71	18.59
		Asian	Amer. Indian/Alaskan	-29.33(*)	2.30	-35.26	-23.41
			African Amer. (Not Hispanic)	-68.95(*)	2.30	-75.09	-62.82
			Hispanic	-50.14(*)	2.30	-56.06	-44.22
			Multiracial	-24.42(*)	2.30	-30.73	-18.11
			White (Not Hispanic)	-13.18(*)	2.30	-18.88	-7.49
		African American (Not Hispanic)	Amer. Indian/Alaskan	39.62(*)	2.50	36.17	43.06
			Asian	68.95(*)	2.50	62.30	75.61
			Hispanic	18.81(*)	2.50	15.39	22.24
			Multiracial	44.53(*)	2.50	40.36	48.70
			White (Not Hispanic)	55.77(*)	2.50	52.82	58.72
		Hispanic	Amer. Indian/Alaskan	20.81(*)	2.53	17.79	23.82
			Asian	50.14(*)	2.53	43.63	56.65
			African Amer. (Not Hispanic)	-18.81(*)	2.53	-22.29	-15.34
			Multiracial	25.72(*)	2.53	21.88	29.55
			White (Not Hispanic)	36.95(*)	2.53	34.54	39.37
		Multiracial	Amer. Indian/Alaskan	-4.91(*)	2.47	-8.67	-1.15
			Asian	24.42(*)	2.47	17.65	31.19
			African Amer. (Not Hispanic)	-44.53(*)	2.47	-48.65	-40.41
			Hispanic	-25.72(*)	2.47	-29.46	-21.97
			White (Not Hispanic)	11.24(*)	2.47	7.92	14.56
		White (Not Hispanic)	Amer. Indian/Alaskan	-16.15(*)	2.56	-18.62	-13.68
			Asian	13.18(*)	2.56	6.84	19.53
			African Amer. (Not Hispanic)	-55.77(*)	2.56	-58.80	-52.74
Hispanic	-36.95(*)		2.56	-39.40	-34.51		
Multiracial	-11.24(*)		2.56	-14.69	-7.79		

(*)Significant differences

Table 47. Science, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity
(continued)

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
8	SS	American Indian/Alaskan	Asian	26.82(*)	2.54	21.60	32.04
			African Amer. (Not Hispanic)	-30.09(*)	2.54	-32.82	-27.35
			Hispanic	-13.78(*)	2.54	-16.24	-11.32
			Multiracial	3.6(*)	2.54	0.44	6.76
			White (Not Hispanic)	13.36(*)	2.54	11.46	15.26
		Asian	Amer. Indian/Alaskan	-26.82(*)	2.31	-31.56	-22.07
			African Amer. (Not Hispanic)	-56.9(*)	2.31	-61.82	-51.99
			Hispanic	-40.6(*)	2.31	-45.39	-35.80
			Multiracial	-23.21(*)	2.31	-28.34	-18.09
			White (Not Hispanic)	-13.45(*)	2.31	-18.03	-8.87
		African American (Not Hispanic)	Amer. Indian/Alaskan	30.09(*)	2.50	27.39	32.79
			Asian	56.9(*)	2.50	51.57	62.24
			Hispanic	16.31(*)	2.50	13.51	19.10
			Multiracial	33.69(*)	2.50	30.28	37.11
			White (Not Hispanic)	43.45(*)	2.50	41.11	45.78
		Hispanic	Amer. Indian/Alaskan	13.78(*)	2.53	11.33	16.23
			Asian	40.6(*)	2.53	35.34	45.85
			African Amer. (Not Hispanic)	-16.31(*)	2.53	-19.13	-13.49
			Multiracial	17.38(*)	2.53	14.15	20.62
			White (Not Hispanic)	27.14(*)	2.53	25.11	29.17
		Multiracial	Amer. Indian/Alaskan	-3.6(*)	2.46	-6.67	-0.53
			Asian	23.21(*)	2.46	17.74	28.69
			African Amer. (Not Hispanic)	-33.69(*)	2.46	-37.05	-30.33
			Hispanic	-17.38(*)	2.46	-20.54	-14.23
			White (Not Hispanic)	9.76(*)	2.46	6.99	12.52
		White (Not Hispanic)	Amer. Indian/Alaskan	-13.36(*)	2.56	-15.28	-11.44
			Asian	13.45(*)	2.56	8.36	18.55
			African Amer. (Not Hispanic)	-43.45(*)	2.56	-45.84	-41.06
Hispanic	-27.14(*)		2.56	-29.20	-25.08		
Multiracial	-9.76(*)		2.56	-12.64	-6.88		

(*)Significant differences

Table 48. Social Studies, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
5	SS	American Indian/Alaskan	Asian	34.49(*)	2.53	28.04	40.94
			African Amer. (Not Hispanic)	-32.93(*)	2.53	-36.39	-29.47
			Hispanic	-16.98(*)	2.53	-19.97	-14.00
			Multiracial	7.06(*)	2.53	3.24	10.87
			White (Not Hispanic)	16.53(*)	2.53	14.11	18.95
		Asian	Amer. Indian/Alaskan	-34.49(*)	2.30	-40.36	-28.62
			African Amer. (Not Hispanic)	-67.42(*)	2.30	-73.49	-61.34
			Hispanic	-51.47(*)	2.30	-57.33	-45.61
			Multiracial	-27.44(*)	2.30	-33.68	-21.19
			White (Not Hispanic)	-17.96(*)	2.30	-23.60	-12.32
		African American (Not Hispanic)	Amer. Indian/Alaskan	32.93(*)	2.50	29.51	36.34
			Asian	67.42(*)	2.50	60.83	74.01
			Hispanic	15.95(*)	2.50	12.55	19.34
			Multiracial	39.98(*)	2.50	35.85	44.11
			White (Not Hispanic)	49.46(*)	2.50	46.53	52.38
		Hispanic	Amer. Indian/Alaskan	16.98(*)	2.53	13.99	19.97
			Asian	51.47(*)	2.53	45.02	57.92
			African Amer. (Not Hispanic)	-15.95(*)	2.53	-19.39	-12.50
			Multiracial	24.04(*)	2.53	20.24	27.84
			White (Not Hispanic)	33.51(*)	2.53	31.12	35.90
		Multiracial	Amer. Indian/Alaskan	-7.06(*)	2.47	-10.78	-3.33
			Asian	27.44(*)	2.47	20.73	34.14
			African Amer. (Not Hispanic)	-39.98(*)	2.47	-44.07	-35.90
			Hispanic	-24.04(*)	2.47	-27.75	-20.33
			White (Not Hispanic)	9.47(*)	2.47	6.18	12.77
		White (Not Hispanic)	Amer. Indian/Alaskan	-16.53(*)	2.56	-18.98	-14.08
			Asian	17.96(*)	2.56	11.68	24.24
			African Amer. (Not Hispanic)	-49.46(*)	2.56	-52.46	-46.45
Hispanic	-33.51(*)		2.56	-35.93	-31.09		
Multiracial	-9.47(*)		2.56	-12.89	-6.06		

(*)Significant differences

Table 49. U.S. History, Pair-Wise Dunnett’s C Post-Hoc Comparison of Spring 2014 Scale Score Mean Differences by Ethnicity

Grade	Dependent Variable	(I) Ethnicity	(J) Ethnicity	Mean Difference (J-I)	Dunnett's C	95% Confidence Interval	
						Lower Bound	Upper Bound
8	SS	American Indian/Alaskan	Asian	45.96(*)	2.54	38.92	53.01
			African Amer. (Not Hispanic)	-33.47(*)	2.54	-37.17	-29.78
			Hispanic	-16.62(*)	2.54	-19.93	-13.31
			Multiracial	4.38(*)	2.54	0.12	8.65
			White (Not Hispanic)	16.89(*)	2.54	14.33	19.45
		Asian	Amer. Indian/Alaskan	-45.96(*)	2.31	-52.37	-39.56
			African Amer. (Not Hispanic)	-79.43(*)	2.31	-86.07	-72.80
			Hispanic	-62.58(*)	2.31	-69.05	-56.11
			Multiracial	-41.58(*)	2.31	-48.49	-34.67
			White (Not Hispanic)	-29.07(*)	2.31	-35.25	-22.89
		African American (Not Hispanic)	Amer. Indian/Alaskan	33.47(*)	2.50	29.83	37.11
			Asian	79.43(*)	2.50	72.23	86.63
			Hispanic	16.85(*)	2.50	13.08	20.62
			Multiracial	37.86(*)	2.50	33.25	42.46
			White (Not Hispanic)	50.36(*)	2.50	47.21	53.51
		Hispanic	Amer. Indian/Alaskan	16.62(*)	2.53	13.32	19.92
			Asian	62.58(*)	2.53	55.50	69.67
			African Amer. (Not Hispanic)	-16.85(*)	2.53	-20.66	-13.04
			Multiracial	21(*)	2.53	16.64	25.37
			White (Not Hispanic)	33.51(*)	2.53	30.77	36.25
		Multiracial	Amer. Indian/Alaskan	-4.38(*)	2.46	-8.53	-0.24
			Asian	41.58(*)	2.46	34.19	48.97
			African Amer. (Not Hispanic)	-37.86(*)	2.46	-42.39	-33.32
			Hispanic	-21(*)	2.46	-25.26	-16.75
			White (Not Hispanic)	12.51(*)	2.46	8.77	16.24
		White (Not Hispanic)	Amer. Indian/Alaskan	-16.89(*)	2.56	-19.48	-14.30
			Asian	29.07(*)	2.56	22.20	35.95
			African Amer. (Not Hispanic)	-50.36(*)	2.56	-53.59	-47.13
			Hispanic	-33.51(*)	2.56	-36.29	-30.73
			Multiracial	-12.51(*)	2.56	-16.39	-8.62

(*)Significant differences

Table 50. Spring 2014, Mean Scale Score and Standard Deviations for State and Each Proficiency Level

Content	Grade	N Count	Total		Pass		Unsatisfactory		Limited Knowledge		Proficient		Advanced	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Mathematics	3	48887	736.52	96.29	787.56	64.77	573.1	54.55	669.24	19.14	749.1	28.48	859.5	50.88
	4	48758	732.36	96.17	785.82	63.51	580.4	54.91	672.38	17.27	750.2	28.87	860.7	49.83
	5	48299	734.05	96.88	787.68	62.90	577.6	56.74	672.08	18.47	744.6	24.11	846.5	50.6
	6	47474	726.73	82.97	771.97	52.59	605.7	52.99	680.08	10.6	746.1	27.06	838.2	43.19
	7	46374	726.69	85.80	775.87	53.11	610.7	56.67	687.23	9.31	749	27.07	843.7	40.74
	8	37626	698.41	81.03	756.93	43.19	579	57.59	671.64	17.37	733	19.95	808.3	33.81
Reading	3	48752	733.10	91.18	779.75	55.16	584.5	62.04	677.33	13.18	774.3	47.1	938.6	30.91
	4	48656	720.53	87.68	769.61	56.22	595.7	56.62	678.56	10.99	758.1	39.89	905.8	40.64
	5	48206	723.42	94.26	776.62	59.46	576.9	59.18	671.08	15.68	758	37.28	887.7	43.91
	6	47475	725.32	86.99	774.38	55.58	588.5	56.73	677.41	15.15	754	32.02	867.6	44.26
	7	47402	730.89	76.62	767.33	52.89	617.3	53.46	686.11	8.46	743.9	25.47	842	48.27
	8	47330	740.62	88.73	782.95	57.94	592.6	59.41	680.5	12.87	762.3	35.57	879.6	42.49
Science	5	48291	695.83	74.78	751.15	44.40	594.1	48.95	673.93	14.71	730.3	19.82	810.5	41.48
	8	47451	694.88	59.45	738.89	29.74	610.2	51.06	679.93	12.44	723.6	14.3	776.4	24.01
Social Studies	5	48234	704.36	73.70	733.55	49.31	554.1	62.74	641.43	13.7	688.1	13.9	760.2	42.67
U.S. History	8	47432	683.80	79.74	727.87	46.40	551.3	63.18	638.88	14.01	686.8	14.12	757.2	38.52

Note: Undetermined (invalid) students not included; Pass = Proficient + Advanced.

Table 51. Spring 2013 Proficiency Level Impact Data, (% rounded)

Content	Grade	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
Mathematics	3	46316	69.63	9.58	20.79	47.86	21.77
	4	45383	73.05	10.39	16.56	51.01	22.03
	5	44295	70.17	10.75	19.08	45.60	24.56
	6	43222	71.83	15.28	12.89	49.57	22.25
	7	43146	68.89	21.27	9.84	51.75	17.14
	8	41377	66.71	12.84	20.46	38.05	28.66
Reading	3	45683	71.67	11.86	16.47	67.63	4.04
	4	44704	68.42	14.11	17.47	62.63	5.80
	5	43798	68.67	11.42	19.91	60.33	8.35
	6	42971	66.08	12.90	21.01	55.77	10.31
	7	43368	71.68	15.16	13.16	61.24	10.44
	8	42341	77.10	10.68	12.22	63.13	13.97
Science	5	44805	51.00	20.99	28.01	35.15	15.85
	8	44209	52.43	19.81	27.77	38.39	14.03

Note: Undetermined (invalid) students not included.

Table 52. Spring 2014 Proficiency Level Impact Data, (% rounded)

Content	Grade	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
Mathematics	3	48887	67.02	12.51	20.47	43.68	23.34
	4	48758	65.28	15.29	19.42	44.22	21.06
	5	48299	65.65	14.74	19.61	37.88	27.77
	6	47474	66.35	19.25	14.39	47.69	18.66
	7	46374	64.13	22.71	13.15	45.92	18.22
	8	37626	52.94	19.83	27.23	36.09	16.84
Reading	3	48752	69.22	16.28	14.50	66.92	2.30
	4	48656	64.53	20.26	15.21	59.51	5.02
	5	48206	64.67	16.89	18.44	55.38	9.29
	6	47475	64.21	16.13	19.66	52.70	11.52
	7	47402	69.78	17.26	12.96	53.12	16.65
	8	47330	71.04	14.40	14.55	58.53	12.52
Science	5	48291	51.28	22.16	26.56	37.96	13.32
	8	47451	49.77	20.65	29.59	35.39	14.38
Social Studies	5	48234	76.56	8.69	14.75	28.31	48.25
U.S. History	8	47432	65.32	15.08	19.60	27.25	38.07

Note: Undetermined (invalid) students not included.

Table 53. Spring 2014, Mathematics State and Subgroup Proficiency Level Impact Data

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
3	Whole State	48887	67.02	12.51	20.47	43.68	23.34
	Female	23936	66.28	12.55	21.17	43.64	22.64
	Male	24923	67.76	12.46	19.78	43.71	24.05
	Asian	908	82.38	7.38	10.24	39.21	43.17
	African American	4391	45.52	27.51	26.96	34.55	10.98
	Hispanic	7974	54.26	19.59	26.15	40.69	13.57
	American Indian	7178	67.61	11.10	21.29	45.61	22.00
	White	24783	74.18	8.21	17.61	45.82	28.36
	Multiracial	137	48.18	25.55	26.28	35.04	13.14
	ELL	5618	45.55	24.94	29.51	36.31	9.24
	IEP	8115	43.19	29.62	27.18	32.36	10.83
	Section 504	395	64.05	10.63	25.32	43.54	20.51
	Low SES	30270	58.80	16.69	24.51	42.35	16.45
	Accommodated	7398	35.24	33.75	31.01	29.43	5.81
4	Whole State	48758	65.28	15.29	19.42	44.22	21.06
	Female	23702	65.30	14.90	19.80	44.69	20.61
	Male	25023	65.28	15.66	19.06	43.77	21.51
	Asian	956	80.75	5.96	13.28	39.54	41.21
	African American	4453	42.08	32.47	25.44	33.53	8.56
	Hispanic	7810	55.36	21.04	23.60	41.86	13.51
	American Indian	7271	63.91	14.92	21.17	46.25	17.66
	White	24664	72.44	10.84	16.72	46.52	25.92
	Multiracial	126	52.38	26.19	21.43	41.27	11.11
	ELL	3689	36.76	35.43	27.81	31.04	5.72
	IEP	8639	36.67	37.10	26.23	29.33	7.34
	Section 504	462	64.50	14.29	21.21	46.75	17.75
	Low SES	29910	56.13	20.48	23.39	42.48	13.65
	Accommodated	7114	28.63	42.34	29.03	25.20	3.43

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 53. Spring 2014, Mathematics State and Subgroup Proficiency Level Impact Data (*continued*)

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
5	Whole State	48299	65.65	14.74	19.61	37.88	27.77
	Female	23633	67.96	13.28	18.76	38.76	29.20
	Male	24651	63.45	16.13	20.42	37.03	26.42
	Asian	884	86.20	5.77	8.03	32.69	53.51
	African American	4414	47.21	28.23	24.56	31.58	15.63
	Hispanic	7505	57.18	19.59	23.24	36.50	20.68
	American Indian	7299	62.72	14.96	22.32	38.95	23.77
	White	24778	71.76	11.06	17.18	39.26	32.49
	Multiracial	144	58.33	23.61	18.06	35.42	22.92
	ELL	2974	34.53	38.94	26.53	27.03	7.50
	IEP	8434	32.80	39.93	27.27	25.03	7.77
	Section 504	568	63.20	12.68	24.12	40.14	23.06
	Low SES	29437	56.96	19.54	23.50	37.39	19.57
Accommodated	7122	27.21	44.07	28.71	22.65	4.56	
6	Whole State	47474	66.35	19.25	14.39	47.69	18.66
	Female	23346	66.83	18.14	15.03	48.76	18.08
	Male	24123	65.90	20.33	13.78	46.67	19.23
	Asian	867	84.43	7.84	7.73	38.29	46.14
	African American	4365	45.25	36.36	18.40	37.62	7.63
	Hispanic	7025	57.71	24.70	17.59	45.98	11.73
	American Indian	7437	62.96	20.79	16.26	48.06	14.90
	White	24027	73.35	14.29	12.37	50.22	23.12
	Multiracial	134	54.48	35.07	10.45	44.78	9.70
	ELL	2558	32.49	48.36	19.16	28.50	3.99
	IEP	8040	27.28	55.37	17.35	23.53	3.74
	Section 504	508	66.34	17.91	15.75	50.59	15.75
	Low SES	28606	57.48	25.04	17.48	45.85	11.63
Accommodated	4979	22.37	60.01	17.61	20.16	2.21	

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 53. Spring 2014, Mathematics State and Subgroup Proficiency Level Impact Data (*continued*)

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
7	Whole State	46374	64.13	22.71	13.15	45.92	18.22
	Female	22669	66.09	20.46	13.45	48.00	18.09
	Male	23700	62.27	24.86	12.87	43.92	18.35
	Asian	810	84.20	10.25	5.56	45.43	38.77
	African American	4221	45.06	39.21	15.73	38.26	6.80
	Hispanic	6698	54.85	29.37	15.78	43.82	11.03
	American Indian	7506	61.96	23.58	14.46	47.20	14.76
	White	23693	70.73	17.45	11.82	47.80	22.93
	Multiracial	126	52.38	32.54	15.08	39.68	12.70
	ELL	2269	32.70	50.55	16.75	29.26	3.44
	IEP	7589	25.14	59.86	15.00	21.94	3.20
	Section 504	517	65.96	21.66	12.38	47.58	18.38
	Low SES	26985	54.53	29.83	15.64	43.49	11.04
Accommodated	4708	21.03	64.40	14.57	19.05	1.98	
8	Whole State	37626	52.94	19.83	27.23	36.09	16.84
	Female	18184	55.42	17.36	27.22	38.53	16.89
	Male	19437	50.62	22.14	27.24	33.81	16.81
	Asian	406	70.44	13.05	16.50	39.66	30.79
	African American	3819	36.14	33.05	30.82	29.22	6.91
	Hispanic	5463	45.96	24.02	30.02	34.38	11.59
	American Indian	6555	53.29	19.30	27.41	36.74	16.55
	White	18873	57.96	16.46	25.59	37.77	20.18
	Multiracial	87	56.32	21.84	21.84	36.78	19.54
	ELL	2187	29.31	40.70	30.00	24.92	4.39
	IEP	7176	21.85	51.28	26.87	17.06	4.79
	Section 504	454	52.64	17.62	29.74	37.00	15.64
	Low SES	23555	45.90	24.70	29.40	33.19	12.71
Accommodated	4339	19.34	53.01	27.66	15.53	3.80	

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 54. Spring 2014, Reading State and Subgroup Proficiency Level Impact Data

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
3	Whole State	48752	69.22	16.28	14.50	66.92	2.30
	Female	23872	73.21	13.25	13.53	70.41	2.80
	Male	24857	65.40	19.18	15.42	63.58	1.82
	Asian	876	77.28	11.87	10.84	72.26	5.02
	African American	4390	50.50	29.89	19.61	49.77	0.73
	Hispanic	7882	54.63	26.11	19.26	53.87	0.76
	American Indian	7181	68.58	15.29	16.13	66.55	2.03
	White	24778	77.00	11.22	11.77	73.97	3.03
	Multiracial	133	53.38	34.59	12.03	53.38	2.51
	ELL	5375	42.27	34.90	22.83	42.07	0.20
	IEP	8132	36.02	46.53	17.45	35.26	0.76
	Section 504	391	70.59	12.79	16.62	69.05	1.53
	Low SES	30286	60.63	21.73	17.64	59.55	1.08
Accommodated	6702	26.57	53.25	20.17	26.45	0.12	
4	Whole State	48656	64.53	20.26	15.21	59.51	5.02
	Female	23649	68.33	16.91	14.76	62.50	5.83
	Male	24977	60.96	23.43	15.61	56.70	4.26
	Asian	940	74.79	12.77	12.45	64.79	10.00
	African American	4456	45.60	35.91	18.49	43.49	2.11
	Hispanic	7726	50.91	30.21	18.88	48.80	2.11
	American Indian	7271	62.89	20.62	16.49	58.81	4.08
	White	24671	72.03	14.67	13.30	65.43	6.59
	Multiracial	123	47.15	39.02	13.82	45.53	1.63
	ELL	3476	27.07	52.24	20.68	26.70	0.37
	IEP	8641	28.55	55.95	15.50	27.36	1.19
	Section 504	466	63.95	18.03	18.03	60.30	3.65
	Low SES	29854	54.96	27.00	18.04	52.43	2.53
Accommodated	6684	18.69	65.77	15.54	18.36	0.33	

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 54. Spring 2014, Reading State and Subgroup Proficiency Level Impact Data (*continued*)

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
5	Whole State	48206	64.67	16.89	18.44	55.38	9.29
	Female	23588	68.09	13.88	18.03	57.47	10.62
	Male	24614	61.40	19.77	18.83	53.39	8.02
	Asian	858	77.74	11.07	11.19	58.04	19.70
	African American	4415	46.39	29.94	23.67	42.31	4.08
	Hispanic	7441	52.37	24.24	23.38	47.49	4.88
	American Indian	7303	63.08	17.43	19.49	55.53	7.56
	White	24772	71.61	12.53	15.86	59.92	11.69
	Multiracial	139	47.48	25.90	26.62	43.88	3.60
	ELL	2794	21.62	52.47	25.91	21.01	0.61
	IEP	8428	25.13	53.62	21.25	23.46	1.67
	Section 504	565	64.60	12.39	23.01	56.64	7.96
	Low SES	29388	55.10	22.74	22.16	49.93	5.18
Accommodated	6778	17.39	60.34	22.26	16.73	0.66	
6	Whole State	47475	64.21	16.13	19.66	52.70	11.52
	Female	23347	68.42	12.76	18.82	55.54	12.88
	Male	24123	60.15	19.39	20.47	49.95	10.20
	Asian	860	80.81	7.44	11.74	55.47	25.35
	African American	4394	45.02	28.81	26.17	40.17	4.85
	Hispanic	6974	52.64	22.05	25.31	46.86	5.78
	American Indian	7430	61.16	16.90	21.94	52.36	8.80
	White	24063	71.50	12.06	16.44	56.61	14.89
	Multiracial	129	51.16	24.81	24.03	42.64	8.53
	ELL	2411	23.39	49.94	26.67	22.23	1.16
	IEP	8069	21.64	54.50	23.86	19.82	1.82
	Section 504	512	65.43	12.11	22.46	53.32	12.11
	Low SES	28582	54.41	21.51	24.08	48.00	6.41
Accommodated	4488	15.93	60.76	23.31	15.02	0.91	

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 54. Spring 2014, Reading State and Subgroup Proficiency Level Impact Data (*continued*)

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
7	Whole State	47402	69.78	17.26	12.96	53.12	16.65
	Female	23175	74.13	13.57	12.30	54.62	19.51
	Male	24222	65.63	20.80	13.58	51.70	13.93
	Asian	918	81.48	9.15	9.37	46.41	35.08
	African American	4266	51.83	29.65	18.52	44.73	7.10
	Hispanic	6690	57.95	24.48	17.56	48.19	9.76
	American Indian	7568	67.64	18.75	13.61	54.20	13.44
	White	24462	76.62	12.70	10.68	55.91	20.71
	Multiracial	131	58.02	25.19	16.79	45.80	12.21
	ELL	2143	26.78	52.68	20.53	24.87	1.91
	IEP	7618	26.66	57.34	16.00	23.97	2.69
	Section 504	531	71.75	12.99	15.25	57.82	13.94
	Low SES	27282	60.31	23.33	16.37	50.23	10.07
Accommodated	3975	20.30	64.35	15.35	19.07	1.23	
8	Whole State	47330	71.04	14.40	14.55	58.53	12.52
	Female	23210	76.06	10.97	12.98	60.78	15.27
	Male	24119	66.22	17.71	16.07	56.35	9.87
	Asian	852	82.98	7.75	9.27	59.39	23.59
	African American	4500	52.07	27.36	20.58	47.16	4.91
	Hispanic	6402	60.18	20.59	19.23	53.64	6.54
	American Indian	7784	70.39	14.80	14.81	59.84	10.55
	White	24694	77.04	10.63	12.32	61.30	15.74
	Multiracial	107	71.96	19.63	8.41	63.55	8.41
	ELL	2161	29.29	46.32	24.39	28.46	0.83
	IEP	7415	28.12	51.05	20.84	26.39	1.73
	Section 504	537	71.32	12.66	16.01	59.59	11.73
	Low SES	26882	61.51	20.26	18.24	54.52	6.99
Accommodated	3645	22.30	56.76	20.93	21.62	0.69	

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 55. Spring 2014, Science State and Subgroup Proficiency Level Impact Data

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
5	Whole State	48291	51.28	22.16	26.56	37.96	13.32
	Female	23632	51.32	20.42	28.26	39.08	12.24
	Male	24653	51.26	23.83	24.91	36.90	14.36
	Asian	885	65.08	13.79	21.13	40.11	24.97
	African American	4416	28.58	42.03	29.39	24.34	4.23
	Hispanic	7513	37.02	31.97	31.01	30.16	6.85
	American Indian	7294	49.34	21.98	28.68	38.02	11.32
	White	24765	59.57	16.03	24.39	42.48	17.09
	Multiracial	143	38.46	36.36	25.17	32.87	5.59
	ELL	2925	15.56	56.68	27.76	13.68	1.88
	IEP	8408	23.14	50.25	26.61	19.28	3.87
	Section 504	568	49.82	20.07	30.11	37.85	11.97
	Low SES	29446	41.04	29.09	29.88	32.85	8.18
Accommodated	7008	16.74	56.15	27.11	14.63	2.11	
8	Whole State	47451	49.77	20.65	29.59	35.39	14.38
	Female	23276	49.32	18.95	31.73	36.65	12.67
	Male	24165	50.20	22.28	27.52	34.17	16.03
	Asian	871	68.20	12.40	19.40	39.04	29.16
	African American	4504	27.44	39.79	32.77	22.38	5.06
	Hispanic	6490	37.21	29.51	33.28	29.71	7.50
	American Indian	7778	47.03	20.74	32.23	35.42	11.61
	White	24750	57.31	15.04	27.65	39.13	18.19
	Multiracial	110	50.91	23.64	25.45	37.27	13.64
	ELL	2330	14.46	54.33	31.20	13.26	1.20
	IEP	7398	16.94	52.00	31.06	13.91	3.03
	Section 504	524	48.28	20.42	31.30	33.02	15.27
	Low SES	26922	39.08	27.80	33.12	30.83	8.25
Accommodated	4902	13.24	56.14	30.62	11.57	1.67	

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 56. Spring 2014, Social Studies State and Subgroup Proficiency Level Impact Data

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
5	Whole State	48234	76.56	8.69	14.75	28.31	48.25
	Female	23616	77.05	7.74	15.21	30.16	46.89
	Male	24610	76.10	9.59	14.30	26.53	49.57
	Asian	886	86.12	5.64	8.24	20.65	65.46
	African American	4407	57.68	18.63	23.69	29.77	27.91
	Hispanic	7490	67.14	12.92	19.93	31.75	35.39
	American Indian	7287	76.57	8.51	14.92	31.27	45.30
	White	24746	82.26	5.95	11.79	26.39	55.88
	Multiracial	143	64.34	11.89	23.78	28.67	35.66
	ELL	2937	42.49	26.83	30.68	28.09	14.40
	IEP	8379	49.68	23.38	26.94	27.45	22.23
	Section 504	568	80.81	6.34	12.85	30.63	50.18
	Low SES	29400	69.17	11.89	18.94	31.54	37.64
	Accommodated	6984	43.76	26.37	29.87	27.79	15.97

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 57. Spring 2014, U.S. History State and Subgroup Proficiency Level Impact Data

Grade	Subgroup	N Count	Pass	Limited			
				Unsatisfactory	Knowledge	Proficient	Advanced
8	Whole State	47432	65.32	15.08	19.60	27.25	38.07
	Female	23265	63.51	14.59	21.90	29.04	34.47
	Male	24156	67.06	15.55	17.40	25.52	41.53
	Asian	871	82.20	7.81	9.99	19.52	62.69
	African American	4497	46.01	29.11	24.88	24.64	21.37
	Hispanic	6494	54.34	21.44	24.22	27.19	27.15
	American Indian	7774	63.91	14.37	21.73	30.63	33.28
	White	24740	71.50	11.35	17.16	26.85	44.65
	Multiracial	109	68.81	18.35	12.84	34.86	33.94
	ELL	2339	28.64	41.56	29.80	19.67	8.98
	IEP	7408	31.49	40.52	27.98	19.10	12.39
	Section 504	521	64.88	14.97	20.15	27.45	37.43
	Low SES	26922	55.08	20.83	24.10	27.75	27.33
	Accommodated	4808	26.71	44.63	28.66	17.74	8.96

Note: ELL = English Language Learner; IEP = Individualized Education Program; SES = Socio-economic Status.

Table 58. Spring 2014 Operational Test Parameters Correlations

Content	Grade	<i>a</i> -parameter					<i>b</i> -parameter					<i>c</i> -parameter				
		Corr	RMSD	Mean Diff	SD Ratio	Rdiff	Corr	RMSD	Mean Diff	SD Ratio	Rdiff	Corr	RMSD	Mean Diff	SD Ratio	Rdiff
Mathematics	3	0.92	0.001	0.000	1.045	-0.10	0.96	19.711	-1.165	0.989	-0.18	0.54	0.075	-0.004	0.647	-0.77
	4	0.93	0.001	0.000	0.863	-0.53	0.97	17.518	1.486	0.984	-0.19	0.65	0.051	0.011	0.761	-0.66
	5	0.85	0.001	0.000	1.092	-0.12	0.97	20.039	-4.736	1.092	0.25	0.38	0.090	-0.016	0.633	-0.79
	6	0.95	0.001	0.000	0.999	-0.17	0.96	14.414	1.690	0.959	-0.28	0.51	0.073	0.001	1.042	-0.47
	7	0.88	0.002	0.000	1.346	0.28	0.99	11.601	-1.253	0.922	-0.56	0.60	0.067	0.003	0.919	-0.52
	8	0.94	0.002	0.001	1.228	0.34	0.96	12.056	2.220	1.009	-0.10	0.60	0.074	0.007	0.881	-0.56
Reading	4	0.90	0.001	0.000	0.917	-0.40	0.95	15.725	0.341	0.920	-0.40	0.51	0.073	0.003	0.789	-0.66
	5	0.96	0.001	-0.001	0.902	-0.46	0.93	24.898	0.512	0.863	-0.53	0.25	0.076	0.004	0.646	-0.80
	6	0.94	0.001	0.000	0.982	-0.23	0.98	12.979	5.342	0.936	-0.42	0.86	0.046	0.015	0.942	-0.37
	7	0.94	0.002	0.000	0.888	-0.48	0.92	18.044	2.559	1.085	-0.01	0.50	0.076	0.011	0.502	-0.86
	8	0.96	0.001	0.000	1.144	0.31	0.91	21.266	2.784	1.067	-0.06	0.34	0.089	0.004	0.628	-0.80
Science	5	0.93	0.001	-0.001	0.855	-0.55	0.96	14.257	-5.270	0.942	-0.34	0.51	0.042	-0.028	1.345	-0.26
	8	0.96	0.002	0.000	0.935	-0.37	0.97	11.275	-0.840	1.020	-0.05	0.56	0.059	-0.005	0.900	-0.55

Table 59. Scale Score Statistics for Operational Test in Spring 2013 and Spring 2014

Content	Grade	Spring 2013					Spring 2014				
		N-Count	Alpha	Mean	SD	SEM	N-Count	Alpha	Mean	SD	SEM
Mathematics	3	46316	0.91	739.00	88.33	2.62	48887	0.91	736.52	96.29	2.65
	4	45383	0.90	745.43	90.14	2.69	48758	0.91	732.36	96.17	2.76
	5	44295	0.89	740.71	86.98	2.81	48299	0.91	734.05	96.88	2.82
	6	43221	0.90	737.09	78.86	2.94	47474	0.91	726.73	82.97	2.90
	7	43146	0.89	732.30	80.70	3.02	46374	0.90	726.69	85.80	2.95
	8	41377	0.90	732.09	83.25	3.00	37626	0.89	698.41	81.03	3.04
Reading	3	45683	0.90	741.22	86.35	2.81	48752	0.91	733.10	91.18	2.94
	4	44704	0.89	729.59	77.54	2.85	48656	0.91	720.53	87.68	2.85
	5	43798	0.90	735.55	84.47	2.64	48206	0.92	723.42	94.26	2.77
	6	42971	0.89	731.18	77.53	2.78	47475	0.90	725.32	86.99	2.94
	7	43368	0.88	729.88	67.56	2.66	47402	0.92	730.89	76.62	2.74
	8	42341	0.88	750.16	82.06	2.72	47330	0.89	740.62	88.73	2.85
Science	5	44805	0.86	695.10	72.00	2.84	48291	0.87	695.83	74.78	2.75
	8	44209	0.85	694.21	57.11	2.98	47451	0.85	694.88	59.45	2.90
Social Studies	5	48234	0.88	704.36	73.70	3.12
U.S. History	8	47432	0.90	683.80	79.74	3.05

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

Table 60. Spring 2014, Proficiency Levels Cut Scores and Scale Bounds

Content	Grade	Cut 1		Cut 2		Cut 3	
		LOSS	Perf. Level 2	Perf. Level 3	Perf. Level 4	HOSS	
Mathematics	3	400	633	700	798	990	
	4	400	639	700	805	990	
	5	400	638	700	791	990	
	6	400	664	700	795	990	
	7	400	674	700	800	990	
	8	400	642	700	774	990	
Reading	3	400	649	700	891	990	
	4	400	658	700	845	990	
	5	400	641	700	830	990	
	6	400	647	700	828	990	
	7	400	668	700	802	990	
	8	400	655	700	833	990	
Science	5	400	648	700	765	990	
	8	400	658	700	751	990	
Social Studies	5	400	615	660	711	990	
Geography	7	400	595	700	847	990	
U.S. History	8	400	612	662	715	990	
Writing	5	15	23	36	48	60	
	8	15	25	36	50	60	

Note: LOSS = Lowest Obtainable Scale Score; HOSS = Highest Obtainable Scale Score.

Table 61. Spring 2014, Proportion of Students Within Each Performance Level

Content	Grade	N-Count	Unsatisfactory	Limited Knowledge	Proficient	Advanced
Mathematics	3	48887	12.51	20.47	43.68	23.34
	4	48758	15.29	19.42	44.22	21.06
	5	48299	14.74	19.61	37.88	27.77
	6	47474	19.25	14.39	47.69	18.66
	7	46374	22.71	13.15	45.92	18.22
	8	37626	19.83	27.23	36.09	16.84
Reading	3	48752	16.28	14.50	66.92	2.30
	4	48656	20.26	15.21	59.51	5.02
	5	48206	16.89	18.44	55.38	9.29
	6	47475	16.13	19.66	52.70	11.52
	7	47402	17.26	12.96	53.12	16.65
	8	47330	14.40	14.55	58.53	12.52
Science	5	48291	22.16	26.56	37.96	13.32
	8	47451	20.65	29.59	35.39	14.38
Social Studies	5	48234	8.69	14.75	28.31	48.25
U.S. History	8	47432	15.08	19.60	27.25	38.07

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

Table 62. Differences in Overall Pass Rate for Spring 2013 and Spring 2014

Content	Grade	Pass Rates (% , rounded)		Differences in Pass Rates
		2013	2014	2014-2013
Mathematics	3	69.63	67.02	-2.61
	4	73.05	65.28	-7.77
	5	70.17	65.65	-4.52
	6	71.83	66.35	-5.48
	7	68.89	64.13	-4.76
	8	66.71	52.94	-13.77
Reading	3	71.67	69.22	-2.45
	4	68.42	64.53	-3.89
	5	68.67	64.67	-4.00
	6	66.08	64.21	-1.87
	7	71.68	69.78	-1.90
	8	77.10	71.04	-6.06
Science	5	51.00	51.28	0.28
	8	52.43	49.77	-2.66
Social Studies	5	.	76.56	.
U.S. History	8	.	65.32	.

Table 63. Spring 2014, Summary of Range of *P*-Values and Item-Test Correlations Statistics for Operational and Field Test, by Item Type

Content	Grade	Item Type	Mean <i>P</i> -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
Mathematics	3	MC	0.38	0.74	0.94	.	.	.	0.25	0.44	0.61	.	.	.
	4	MC	0.45	0.72	0.93	.	.	.	0.28	0.44	0.55	.	.	.
	5	MC	0.36	0.69	0.95	.	.	.	0.25	0.43	0.60	.	.	.
	6	MC	0.29	0.64	0.94	.	.	.	0.23	0.43	0.56	.	.	.
	7	MC	0.24	0.60	0.97	.	.	.	0.18	0.41	0.61	.	.	.
	8	MC	0.21	0.56	0.92	.	.	.	0.16	0.39	0.58	.	.	.
Reading	3	MC	0.39	0.67	0.93	.	.	.	0.23	0.43	0.60	.	.	.
	4	MC	0.40	0.71	0.89	.	.	.	0.21	0.44	0.57	.	.	.
	5	MC	0.38	0.72	0.90	.	.	.	0.27	0.45	0.59	.	.	.
	6	MC	0.37	0.68	0.88	.	.	.	0.25	0.42	0.57	.	.	.
	7	MC	0.49	0.73	0.92	.	.	.	0.27	0.45	0.63	.	.	.
	8	MC	0.29	0.72	0.92	.	.	.	0.26	0.40	0.52	.	.	.
Science	5	MC	0.43	0.70	0.97	0.17	0.49	0.89	0.15	0.39	0.50	0.08	0.26	0.48
	8	MC	0.24	0.60	0.92	0.24	0.53	0.89	0.09	0.37	0.53	0.09	0.29	0.50
Social Studies	5	MC	0.42	0.63	0.85	0.20	0.44	0.77	0.09	0.38	0.52	0.02	0.27	0.48
Geography	7	MC	.	.	.	0.12	0.45	0.89	.	.	.	0.12	0.31	0.46
U.S. History	8	MC	0.27	0.61	0.86	0.10	0.45	0.85	0.21	0.41	0.57	0.00	0.27	0.47
Writing	5	CR	0.58	0.58	0.58	.	.	.	0.91	0.94	0.95	.	.	.
	8	CR	0.63	0.64	0.64	.	.	.	0.97	0.97	0.98	.	.	.

Table 64. Spring 2013 and Spring 2014 Test Reliability Data

Content	Grade	Coefficient Alpha	
		2013	2014
Mathematics	3	0.91	0.91
	4	0.90	0.91
	5	0.89	0.91
	6	0.90	0.91
	7	0.89	0.90
	8	0.90	0.89
Reading	3	0.90	0.91
	4	0.89	0.91
	5	0.90	0.92
	6	0.89	0.90
	7	0.88	0.92
	8	0.88	0.89
Science	5	0.86	0.87
	8	0.85	0.85
Social Studies	5	.	0.88
U.S. History	8	.	0.90

Table 65. Mathematics, Raw Score to Scale Score Conversions & Standard Error of Measurement

Raw Score	Grade 3		Grade 4		Grade 5	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
0	400	147	400	163	400	179
1	400	147	400	163	400	179
2	400	147	400	163	400	179
3	400	147	400	163	400	179
4	400	147	400	163	400	179
5	400	147	400	163	400	179
6	400	147	400	163	400	179
7	400	147	400	163	400	179
8	400	147	400	163	400	179
9	400	147	400	163	400	179
10	401	146	400	163	400	179
11	446	100	423	140	414	164
12	476	73	465	97	466	112
13	497	58	493	71	498	80
14	515	50	514	58	522	63
15	530	44	531	49	540	53
16	543	39	545	43	556	46
17	555	36	558	39	570	42
18	566	34	570	36	582	38
19	576	32	581	34	594	35
20	586	30	591	32	604	33
21	595	29	600	31	614	32
22	603	28	609	29	624	30
23	611	27	618	28	633	29
24	620	26	627	28	641	28
25	627	25	635	27	650	27

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

Table 65. Mathematics, Raw Score to Scale Score Conversions & Standard Error of Measurement (*continued*)

Raw Score	Grade 3		Grade 4		Grade 5	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
26	635	25	643	26	658	26
27	643	24	651	26	666	26
28	650	24	658	25	674	25
29	658	24	666	25	682	25
30	665	23	674	24	690	25
31	672	23	681	24	698	24
32	680	23	689	24	706	24
33	688	23	696	23	713	24
34	695	23	704	23	721	24
35	703	23	712	23	729	24
36	711	24	719	23	737	24
37	720	24	727	23	746	24
38	729	24	736	24	754	24
39	738	25	744	24	763	24
40	747	25	753	24	772	25
41	758	26	763	25	782	25
42	769	27	773	26	793	26
43	780	28	784	27	804	28
44	794	30	797	29	817	29
45	809	32	811	31	832	32
46	827	36	828	35	849	35
47	850	42	850	40	871	40
48	882	53	880	50	900	48
49	940	82	933	75	949	69
50	990	115	990	116	990	95

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

Table 65. Mathematics, Raw Score to Scale Score Conversions & Standard Error of Measurement (*continued*)

Raw Score	Grade 6		Grade 7		Grade 8	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
0	400	201	400	204	400	203
1	400	201	400	204	400	203
2	400	201	400	204	400	203
3	400	201	400	204	400	203
4	400	201	400	204	400	203
5	400	201	400	204	400	203
6	400	201	400	204	400	203
7	400	201	400	204	400	203
8	400	201	400	204	400	203
9	400	201	400	204	400	203
10	446	155	453	151	458	145
11	498	104	502	102	504	100
12	528	73	531	73	533	73
13	550	58	553	59	555	58
14	568	49	571	51	573	50
15	583	43	587	45	588	43
16	596	39	601	41	602	39
17	608	35	614	38	613	35
18	619	33	625	35	624	32
19	629	31	636	33	634	30
20	638	29	647	31	643	28
21	647	28	656	30	651	27
22	656	27	666	28	659	25
23	664	26	674	27	667	24
24	672	25	683	26	675	24
25	680	24	691	25	682	23

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

Table 65. Mathematics, Raw Score to Scale Score Conversions & Standard Error of Measurement (*continued*)

Raw Score	Grade 6		Grade 7		Grade 8	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
26	687	23	699	24	689	23
27	694	23	707	24	696	22
28	702	22	714	23	703	22
29	709	22	721	22	710	22
30	716	21	729	22	717	21
31	722	21	736	22	724	21
32	729	20	743	22	731	21
33	736	20	750	21	738	21
34	743	20	758	21	745	21
35	750	19	765	21	752	21
36	756	19	773	21	760	21
37	763	19	780	21	767	21
38	770	19	788	21	775	21
39	777	19	796	21	783	21
40	784	19	804	21	791	21
41	791	19	812	21	800	22
42	799	19	821	22	809	22
43	807	19	831	23	819	23
44	816	20	842	25	830	25
45	826	22	855	28	843	26
46	838	24	871	32	858	29
47	853	29	891	37	876	34
48	874	37	919	45	902	41
49	913	56	966	65	945	60
50	990	121	990	79	990	88

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

Table 66. Reading, Raw Score to Scale Score Conversions & Standard Error of Measurement

Raw Score	Grade 3		Grade 4		Grade 5	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
0	400	204	400	175	400	173
1	400	204	400	175	400	173
2	400	204	400	175	400	173
3	400	204	400	175	400	173
4	400	204	400	175	400	173
5	400	204	400	175	400	173
6	400	204	400	175	400	173
7	400	204	400	175	400	173
8	400	204	400	175	400	173
9	400	204	400	175	400	173
10	423	180	454	121	400	173
11	488	116	491	84	459	114
12	522	82	515	62	493	80
13	546	62	533	50	516	60
14	564	51	548	42	533	49
15	579	44	561	37	547	41
16	592	39	571	34	559	36
17	603	35	581	31	570	33
18	613	32	590	29	579	30
19	623	30	599	27	588	28
20	631	28	607	26	596	26
21	639	27	614	25	604	25
22	647	26	621	24	611	24
23	655	25	628	23	618	23
24	662	24	635	23	625	23
25	669	23	642	22	631	22

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

Table 66. Reading, Raw Score to Scale Score Conversions & Standard Error of Measurement
 (continued)

Raw Score	Grade 3		Grade 4		Grade 5	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
26	675	23	648	22	638	22
27	682	22	655	22	644	22
28	689	22	661	21	651	22
29	695	22	667	21	657	22
30	702	22	674	21	664	22
31	709	22	680	21	671	22
32	716	22	687	21	678	22
33	723	22	693	21	685	22
34	730	23	700	21	692	23
35	737	23	707	22	700	23
36	745	23	714	22	708	24
37	753	24	721	22	716	24
38	761	24	729	23	724	25
39	770	25	737	23	733	25
40	779	25	746	24	743	26
41	789	26	755	25	753	27
42	800	27	765	26	764	28
43	812	29	776	27	776	30
44	825	31	789	29	790	32
45	840	33	803	32	806	35
46	859	37	820	35	825	40
47	882	44	842	41	851	48
48	915	55	873	51	887	61
49	975	86	925	74	953	97
50	990	96	990	116	990	123

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

Table 66. Reading, Raw Score to Scale Score Conversions & Standard Error of Measurement
(continued)

Raw Score	Grade 6		Grade 7		Grade 8	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
0	400	185	400	205	400	175
1	400	185	400	205	400	175
2	400	185	400	205	400	175
3	400	185	400	205	400	175
4	400	185	400	205	400	175
5	400	185	400	205	400	175
6	400	185	400	205	400	175
7	400	185	400	205	400	175
8	400	185	400	205	400	175
9	404	180	400	205	400	175
10	470	115	481	124	406	169
11	505	80	523	82	462	113
12	528	61	547	58	494	80
13	547	50	564	45	518	62
14	562	43	577	37	536	51
15	574	38	588	33	551	44
16	586	34	598	29	564	39
17	596	32	606	26	575	35
18	605	30	614	25	585	33
19	614	28	621	23	595	31
20	622	27	628	22	604	29
21	630	26	634	21	612	28
22	637	25	640	20	620	27
23	644	24	646	19	628	27
24	651	23	651	18	636	26
25	658	23	657	18	644	26

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

Table 66. Reading, Raw Score to Scale Score Conversions & Standard Error of Measurement
(continued)

Raw Score	Grade 6		Grade 7		Grade 8	
	Scale Score	SEM	Scale Score	SEM	Scale Score	SEM
26	665	22	662	17	651	26
27	671	22	667	17	659	25
28	678	22	672	17	666	25
29	685	22	677	17	674	25
30	691	22	682	17	682	25
31	698	22	687	16	689	25
32	704	22	692	16	697	25
33	711	22	697	17	705	26
34	718	22	703	17	714	26
35	725	22	708	17	722	26
36	733	23	714	17	731	27
37	741	23	720	18	740	27
38	749	24	726	18	750	28
39	757	25	732	19	760	29
40	766	25	739	19	771	30
41	776	27	747	20	782	31
42	787	28	755	21	795	32
43	799	30	765	23	809	34
44	812	32	775	24	825	36
45	828	35	787	27	843	39
46	847	39	802	30	864	43
47	872	46	821	35	891	50
48	906	57	848	44	928	61
49	966	84	895	68	990	89
50	990	99	990	148	990	89

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

Table 67. Science, Raw Score to Scale Score Conversions & Standard Error of Measurement

Raw Score	Grade 5		Grade 8	
	Scale Score	SEM	Scale Score	SEM
0	400	160	400	220
1	400	160	400	220
2	400	160	400	220
3	400	160	400	220
4	400	160	400	220
5	400	160	400	220
6	400	160	400	220
7	400	160	400	220
8	400	160	400	220
9	400	160	421	199
10	436	124	514	106
11	475	85	548	72
12	501	64	569	52
13	521	52	585	42
14	538	44	598	36
15	552	39	609	32
16	564	35	619	29
17	575	32	628	27
18	585	30	636	25
19	594	28	644	24
20	603	27	652	23
21	611	26	659	22
22	619	25	666	21
23	627	24	673	20

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

Table 67. Science, Raw Score to Scale Score Conversions & Standard Error of Measurement
(continued)

Raw Score	Grade 5		Grade 8	
	Scale Score	SEM	Scale Score	SEM
24	635	23	679	19
25	642	23	685	19
26	649	23	691	18
27	657	22	697	18
28	664	22	703	18
29	671	22	709	17
30	679	22	715	17
31	686	22	721	17
32	694	23	728	17
33	702	23	734	17
34	710	23	741	18
35	719	24	747	18
36	729	25	755	18
37	739	26	763	19
38	750	27	771	20
39	763	29	780	20
40	777	33	791	21
41	796	37	802	22
42	819	45	816	25
43	854	58	834	31
44	918	91	868	50
45	990	142	990	172

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

Table 68. Social Studies, Raw Score to Scale Score Conversions & Standard Error of Measurement

Raw Score	Grade 5		Raw Score	Grade 5	
	Scale Score	SEM		Scale Score	SEM
0	400	219	26	672	22
1	400	219	27	678	21
2	400	219	28	684	20
3	400	219	29	690	20
4	400	219	30	696	20
5	400	219	31	702	19
6	400	219	32	708	19
7	400	219	33	713	19
8	400	219	34	719	19
9	400	219	35	725	19
10	400	219	36	732	19
11	400	219	37	738	20
12	473	146	38	745	20
13	518	101	39	752	21
14	546	73	40	759	21
15	567	57	41	768	22
16	583	47	42	777	23
17	596	41	43	787	25
18	608	36	44	798	27
19	618	33	45	812	30
20	628	30	46	828	34
21	636	28	47	850	42
22	644	26	48	883	56
23	652	25	49	951	104
24	659	24	50	990	143
25	665	23			

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

Table 69. U.S. History, Raw Score to Scale Score Conversions & Standard Error of Measurement

Raw Score	Grade 8		Raw Score	Grade 8	
	Scale Score	SEM		Scale Score	SEM
0	400	210	26	658	21
1	400	210	27	665	20
2	400	210	28	671	20
3	400	210	29	677	20
4	400	210	30	683	20
5	400	210	31	689	19
6	400	210	32	695	19
7	400	210	33	702	19
8	400	210	34	708	19
9	400	210	35	715	20
10	400	210	36	721	20
11	400	210	37	728	20
12	465	145	38	735	20
13	514	96	39	743	20
14	541	69	40	751	21
15	560	53	41	759	21
16	575	44	42	768	22
17	587	38	43	777	23
18	598	33	44	788	24
19	608	30	45	800	26
20	616	28	46	815	30
21	624	26	47	834	36
22	632	24	48	863	49
23	639	23	49	922	88
24	646	22	50	990	146
25	652	22			

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

Table 70. Spring 2014, Total Group Factor Analysis Results: Eigenvalues

Content	Grade	KMO Statistic	Initial Eigenvalue		Ratio 1st to 2nd Eigenvalue
			Total	% Variance	
Mathematics	3	0.96	12.54	81%	7.47
	4	0.97	11.90	87%	9.76
	5	0.97	11.78	90%	8.78
	6	0.97	11.31	90%	6.74
	7	0.96	10.37	88%	6.58
	8	0.96	9.49	90%	8.94
Reading	3	0.97	11.72	92%	8.95
	4	0.98	11.58	101%	15.50
	5	0.98	13.04	96%	11.36
	6	0.98	10.57	103%	20.19
	7	0.98	12.44	102%	19.39
	8	0.98	9.10	102%	13.08
Science	5	0.97	7.54	105%	14.89
	8	0.96	6.60	103%	8.54
Social Studies	5	0.97	8.01	104%	20.50
U.S. History	8	0.98	9.94	100%	11.30

Note: KMO = Kaiser’s Measure of Sampling Adequacy

Table 71. Spring 2014, Mathematics Subgroup Factor Analysis Results: Eigenvalues

Grade	Subgroup	KMO Statistic	Initial Eigenvalue		Ratio 1st to 2nd Eigenvalue
			Total	% Variance	
3	Total Accommodated	0.95	10.83	82%	8.50
	ELL	0.95	11.46	79%	6.98
	Free Lunch	0.96	12.13	81%	7.37
	IEP	0.96	13.36	85%	10.57
4	Total Accommodated	0.94	9.18	82%	7.16
	ELL	0.95	10.39	81%	7.85
	Free Lunch	0.96	11.13	86%	9.03
	IEP	0.96	11.22	85%	8.96
5	Total Accommodated	0.95	8.12	86%	6.56
	ELL	0.95	9.55	85%	8.31
	Free Lunch	0.97	10.81	90%	8.18
	IEP	0.96	9.74	87%	7.31
6	Total Accommodated	0.93	6.65	87%	5.76
	ELL	0.94	8.20	83%	5.59
	Free Lunch	0.97	9.92	89%	6.32
	IEP	0.95	8.12	90%	6.40
7	Total Accommodated	0.91	5.84	83%	4.97
	ELL	0.92	7.04	81%	5.73
	Free Lunch	0.96	9.04	88%	6.15
	IEP	0.94	6.78	87%	5.46
8	Total Accommodated	0.93	6.98	86%	8.08
	ELL	0.92	8.01	82%	7.89
	Free Lunch	0.96	9.06	90%	8.63
	IEP	0.94	7.56	89%	8.22

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

Table 72. Spring 2014, Reading Subgroup Factor Analysis Results: Eigenvalues

Grade	Subgroup	KMO Statistic	Initial Eigenvalue		Ratio 1st to 2nd Eigenvalue
			Total	% Variance	
3	Total Accommodated	0.96	8.25	94%	11.32
	ELL	0.96	8.84	90%	8.76
	Free Lunch	0.97	10.92	92%	9.06
	IEP	0.97	11.68	95%	14.66
4	Total Accommodated	0.96	8.60	99%	16.86
	ELL	0.96	8.67	95%	14.08
	Free Lunch	0.98	11.04	100%	16.00
	IEP	0.98	11.59	99%	17.61
5	Total Accommodated	0.96	9.13	96%	12.73
	ELL	0.96	9.40	92%	12.68
	Free Lunch	0.98	12.35	97%	11.91
	IEP	0.97	11.65	96%	13.59
6	Total Accommodated	0.95	7.15	98%	14.68
	ELL	0.95	7.99	92%	13.93
	Free Lunch	0.98	9.74	103%	19.78
	IEP	0.97	8.55	101%	17.14
7	Total Accommodated	0.97	9.31	98%	16.36
	ELL	0.96	9.16	92%	16.00
	Free Lunch	0.98	11.83	102%	20.29
	IEP	0.98	10.75	100%	18.33
8	Total Accommodated	0.95	7.47	97%	14.15
	ELL	0.94	7.09	90%	12.82
	Free Lunch	0.97	8.90	103%	14.28
	IEP	0.97	8.48	101%	15.31

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

Table 73. Spring 2014, Science Subgroup Factor Analysis Results: Eigenvalues

Grade	Subgroup	KMO Statistic	Initial Eigenvalue		Ratio 1st to 2nd Eigenvalue
			Total	% Variance	
5	Total Accommodated	0.95	6.20	101%	12.27
	ELL	0.93	5.76	95%	11.46
	Free Lunch	0.97	7.13	105%	14.77
	IEP	0.96	7.35	101%	12.94
8	Total Accommodated	0.91	4.54	96%	7.37
	ELL	0.87	4.24	86%	5.69
	Free Lunch	0.95	5.94	102%	8.10
	IEP	0.92	5.00	98%	6.91

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

Table 74. Spring 2014, Social Studies Subgroup Factor Analysis Results: Eigenvalues

Grade	Subgroup	KMO Statistic	Initial Eigenvalue		Ratio 1st to 2nd Eigenvalue
			Total	% Variance	
5	Total Accommodated	0.93	5.35	96%	9.13
	ELL	0.90	5.00	87%	7.10
	Free Lunch	0.96	7.01	104%	16.54
	IEP	0.95	6.74	99%	12.17

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

Table 75. Spring 2014, U.S. History Subgroup Factor Analysis Results: Eigenvalues

Grade	Subgroup	KMO Statistic	Initial Eigenvalue		Ratio 1st to 2nd Eigenvalue
			Total	% Variance	
8	Total Accommodated	0.93	6.10	95%	12.01
	ELL	0.91	6.06	88%	10.43
	Free Lunch	0.97	8.68	100%	11.86
	IEP	0.95	7.12	97%	11.88

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

Table 76. Spring 2014, Proficiency Level Cut Scores and Standard Error of Measurement (SEM)

Content	Grade	Cut 1		Cut 2		Cut 3	
		Cut Score	SEM at Cut	Cut Score	SEM at Cut	Cut Score	SEM at Cut
Mathematics	3	633	25	700	23	798	32
	4	639	26	700	23	805	31
	5	638	28	700	24	791	26
	6	664	26	700	22	795	19
	7	674	27	700	24	800	21
	8	642	28	700	22	774	21
Reading	3	649	25	700	22	891	55
	4	658	21	700	21	845	51
	5	641	22	700	23	830	48
	6	647	23	700	22	828	35
	7	668	17	700	17	802	30
	8	655	25	700	26	833	39
Science	5	648	23	700	23	765	33
	8	658	22	700	18	751	18
Social Studies	5	615	33	660	23	711	19
U.S. History	8	612	28	662	20	715	20

Note: SEM at or closest above the cut scores.

Table 77. Estimates of Accuracy and Consistency of Performance Classification

Content	Grade	Accuracy	Consistency	False Positives	False Negatives	Kappa
Mathematics	3	0.78	0.70	0.11	0.11	0.58
	4	0.78	0.69	0.11	0.11	0.58
	5	0.77	0.68	0.11	0.12	0.57
	6	0.77	0.70	0.11	0.11	0.57
	7	0.77	0.69	0.12	0.12	0.56
	8	0.74	0.64	0.13	0.13	0.52
Reading	3	0.83	0.77	0.08	0.09	0.60
	4	0.80	0.74	0.10	0.10	0.59
	5	0.80	0.72	0.10	0.10	0.59
	6	0.79	0.71	0.10	0.11	0.57
	7	0.79	0.72	0.10	0.11	0.59
	8	0.78	0.70	0.11	0.11	0.54
Science	5	0.72	0.62	0.14	0.14	0.49
	8	0.70	0.60	0.15	0.15	0.46
Social Studies	5	0.74	0.66	0.12	0.14	0.50
U.S. History	8	0.74	0.65	0.13	0.14	0.52

Table 78. Accuracy & Consistency Estimates by Cut Score

Content	Grade	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
Mathematics	3	0.95	0.91	0.92	0.92	0.88	0.89
	4	0.94	0.91	0.93	0.91	0.88	0.89
	5	0.94	0.91	0.92	0.91	0.88	0.88
	6	0.93	0.91	0.93	0.90	0.87	0.90
	7	0.92	0.90	0.93	0.88	0.87	0.90
	8	0.91	0.89	0.93	0.88	0.85	0.90
Reading	3	0.94	0.91	0.98	0.91	0.88	0.97
	4	0.93	0.91	0.96	0.90	0.87	0.95
	5	0.94	0.91	0.95	0.91	0.88	0.93
	6	0.93	0.91	0.95	0.90	0.87	0.93
	7	0.93	0.92	0.94	0.91	0.88	0.91
	8	0.93	0.90	0.93	0.90	0.87	0.91
Science	5	0.90	0.88	0.92	0.87	0.84	0.89
	8	0.90	0.88	0.92	0.86	0.83	0.89
Social Studies	5	0.94	0.90	0.89	0.92	0.87	0.84
U.S. History	8	0.93	0.90	0.90	0.90	0.86	0.87

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

Table 79. Accuracy & Consistency Estimates by Cut Score: False Positive and False Negative Rates

Content	Grade	U/L+P+A		U+L/P+A		U+L+P/A	
		False	False	False	False	False	False
		Positive	Negative	Positive	Negative	Positive	Negative
Mathematics	3	0.02	0.03	0.04	0.05	0.05	0.04
	4	0.03	0.04	0.04	0.05	0.04	0.03
	5	0.02	0.04	0.04	0.05	0.05	0.04
	6	0.03	0.04	0.04	0.05	0.04	0.03
	7	0.04	0.05	0.04	0.05	0.04	0.03
	8	0.04	0.05	0.05	0.05	0.04	0.03
Reading	3	0.03	0.04	0.04	0.05	0.01	0.01
	4	0.03	0.04	0.04	0.05	0.02	0.01
	5	0.03	0.04	0.04	0.05	0.03	0.02
	6	0.03	0.04	0.04	0.05	0.03	0.02
	7	0.03	0.04	0.04	0.05	0.04	0.03
	8	0.03	0.04	0.04	0.05	0.04	0.03
Science	5	0.04	0.05	0.06	0.06	0.05	0.03
	8	0.04	0.06	0.06	0.06	0.05	0.03
Social Studies	5	0.02	0.04	0.04	0.05	0.06	0.06
U.S. History	8	0.03	0.04	0.05	0.05	0.05	0.04

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

Figures

Figure 1. Spring 2014 Grade 3 Mathematics operational scale score histogram

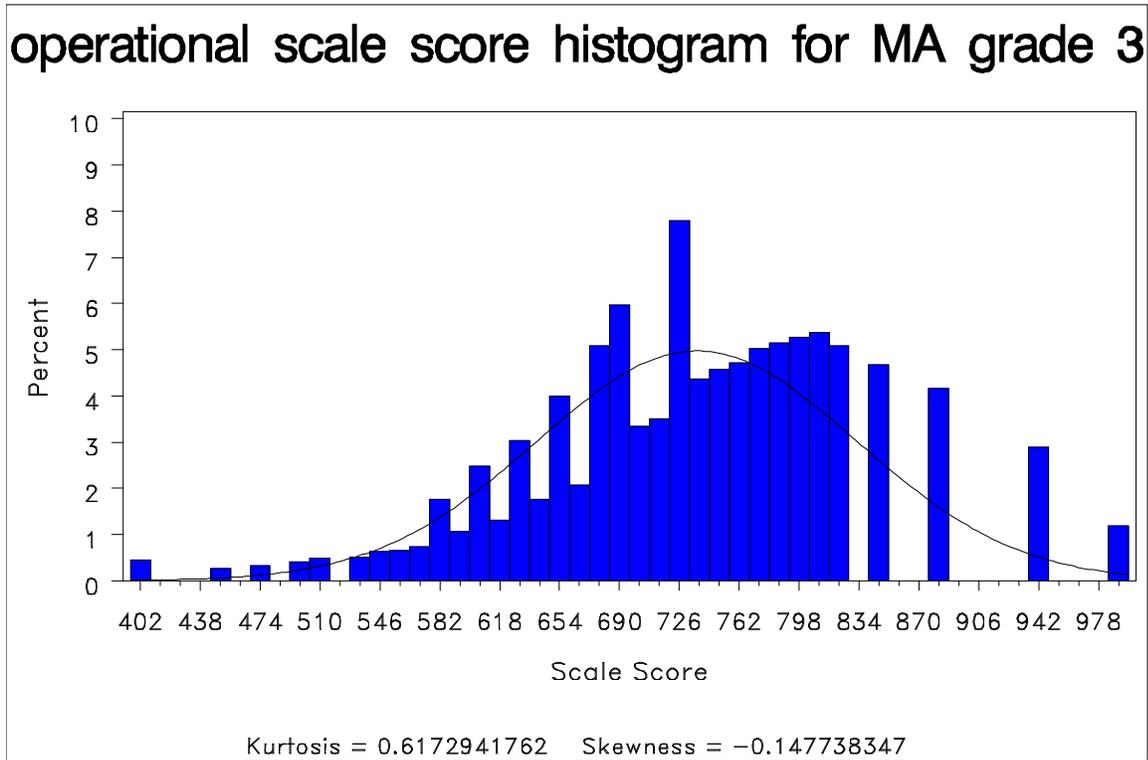


Figure 2. Spring 2014 Grade 4 Mathematics operational scale score histogram

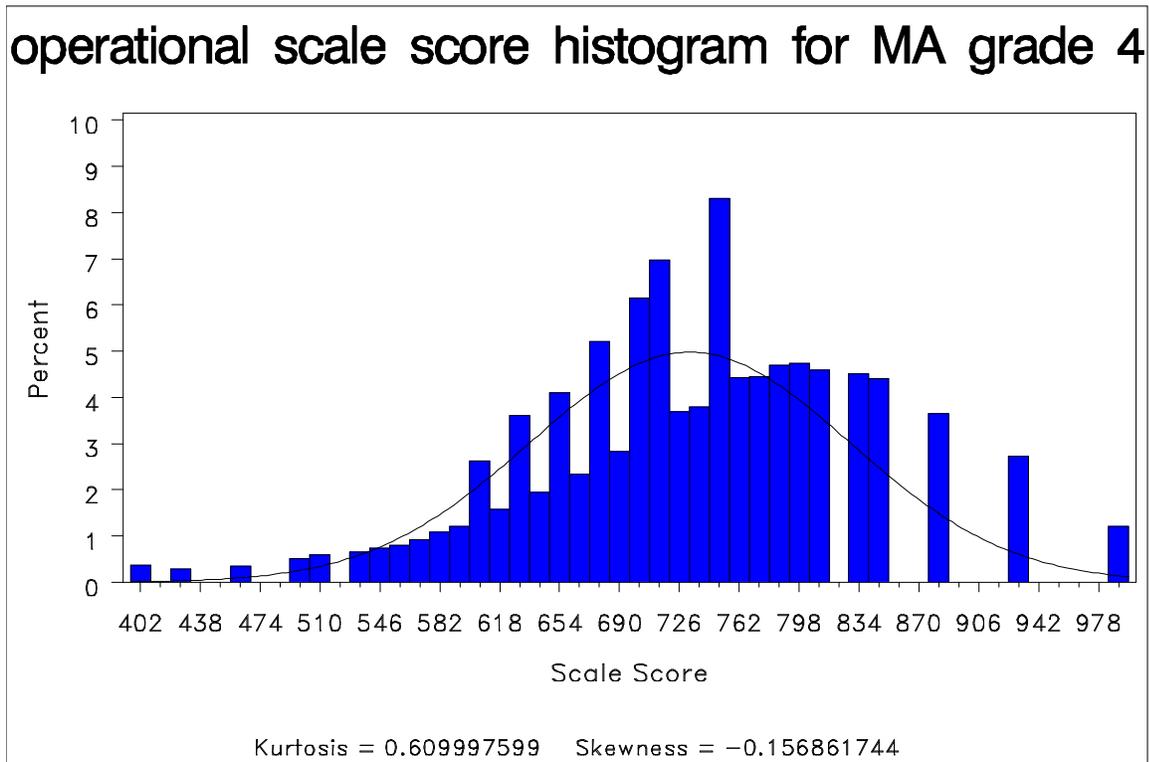


Figure 3. Spring 2014 Grade 5 Mathematics operational scale score histogram

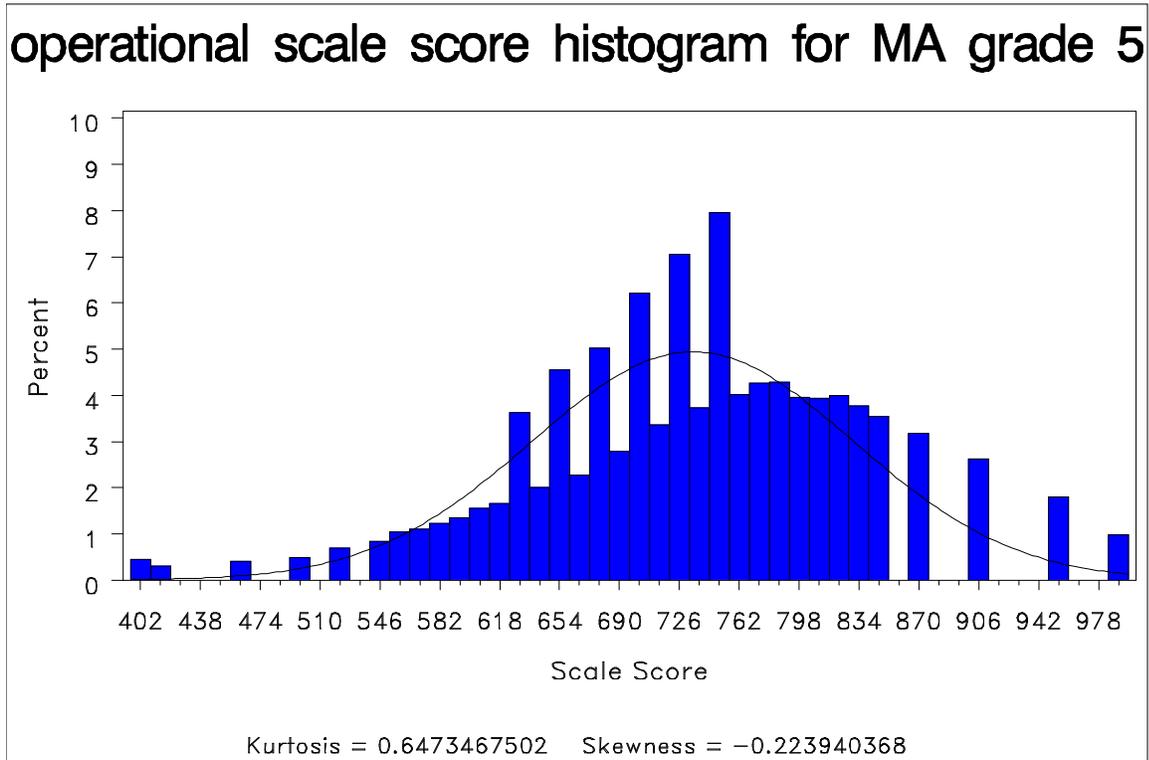


Figure 4. Spring 2014 Grade 6 Mathematics operational scale score histogram

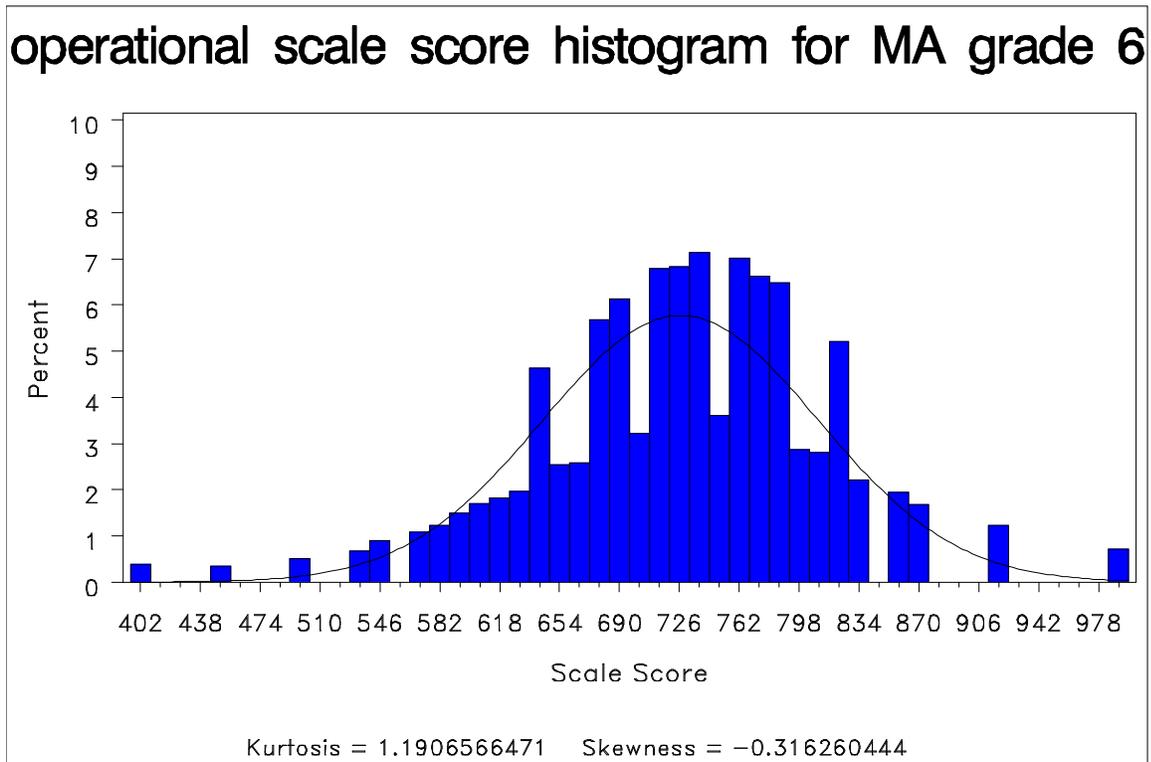


Figure 5. Spring 2014 Grade 7 Mathematics operational scale score histogram

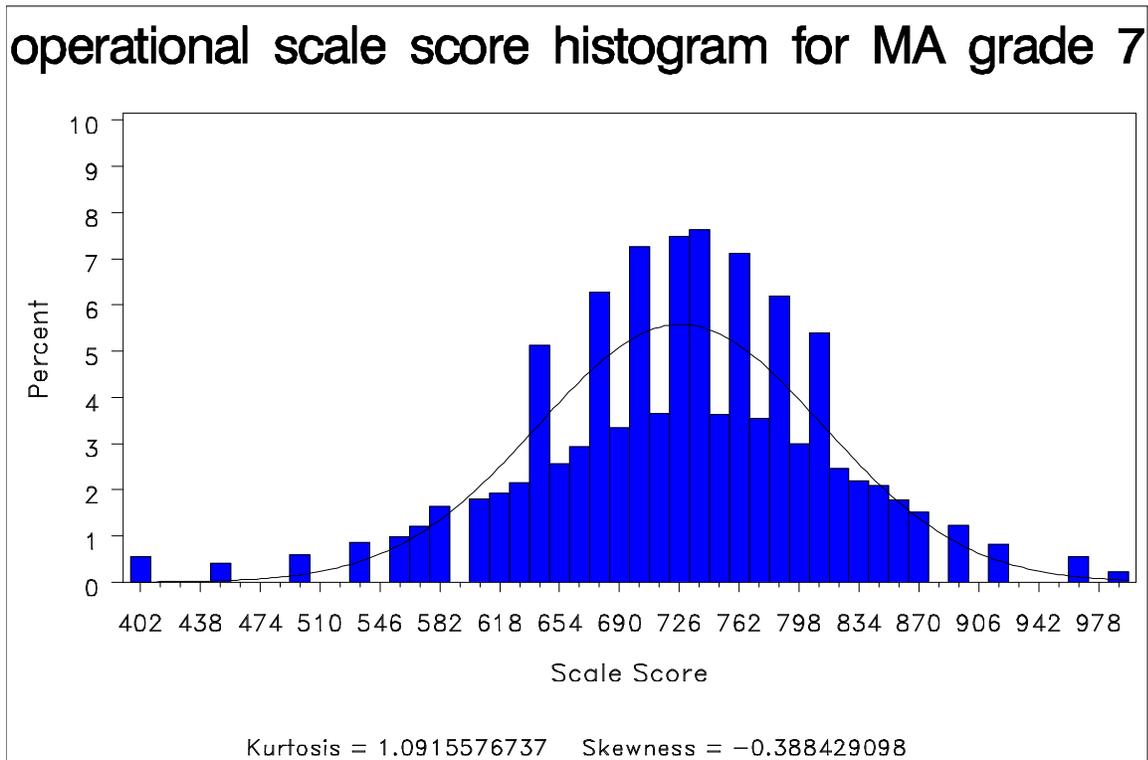


Figure 6. Spring 2014 Grade 8 Mathematics operational scale score histogram

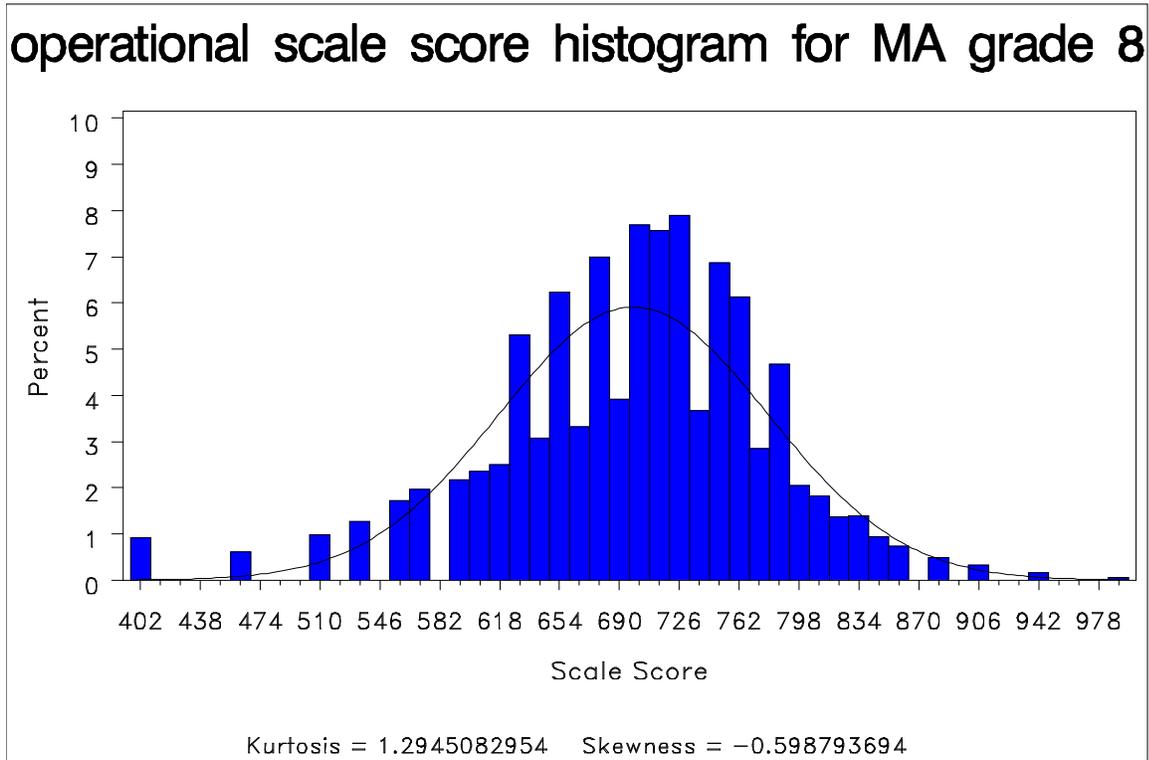


Figure 7. Spring 2014 Grade 3 Reading operational scale score histogram

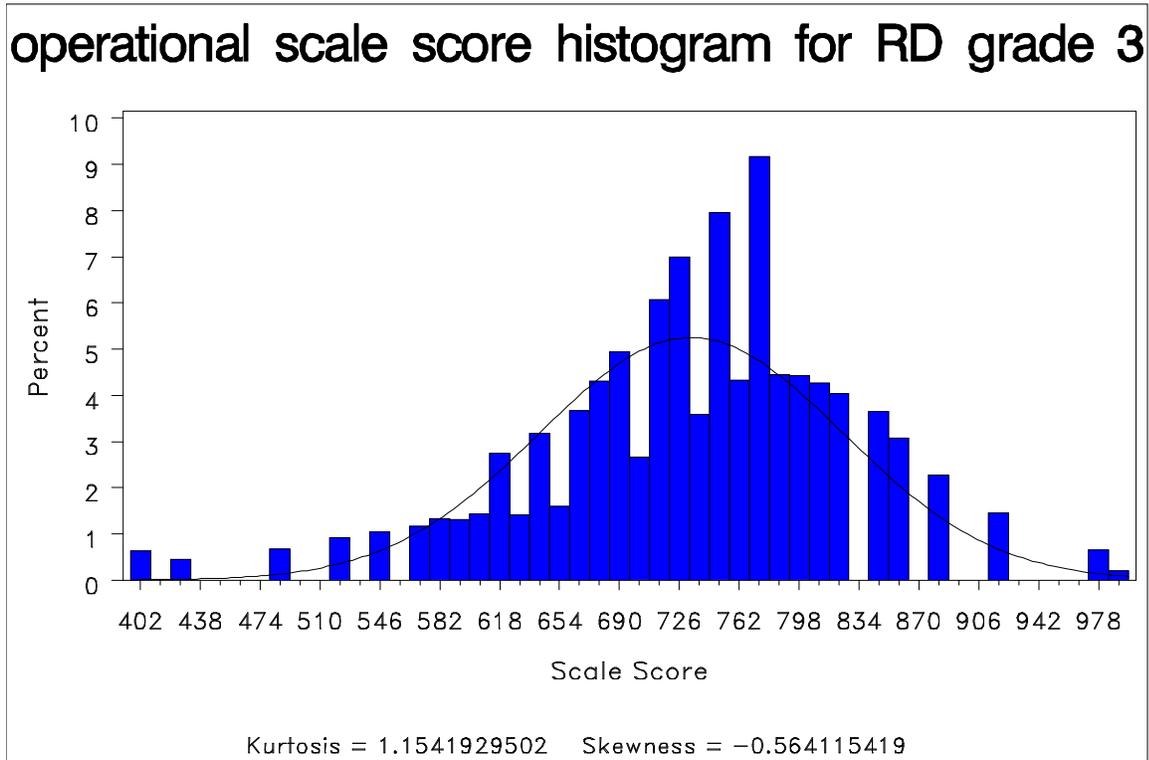


Figure 8. Spring 2014 Grade 4 Reading operational scale score histogram

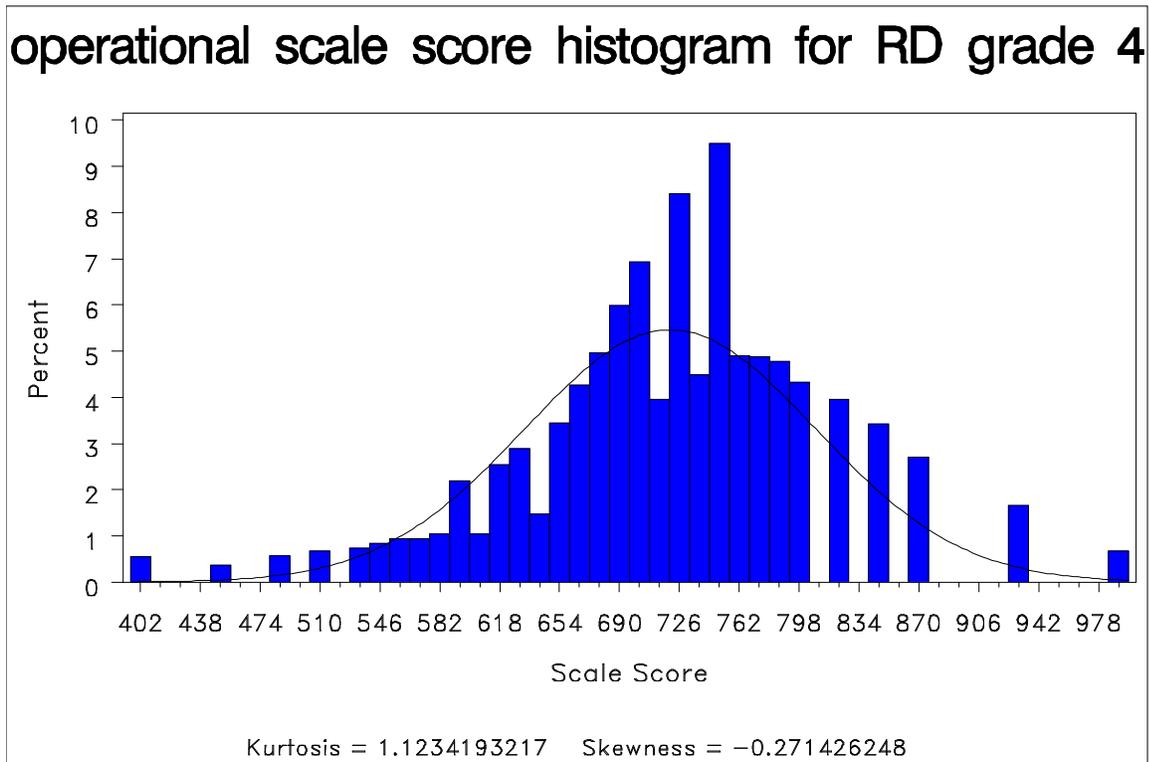


Figure 9. Spring 2014 Grade 5 Reading operational scale score histogram

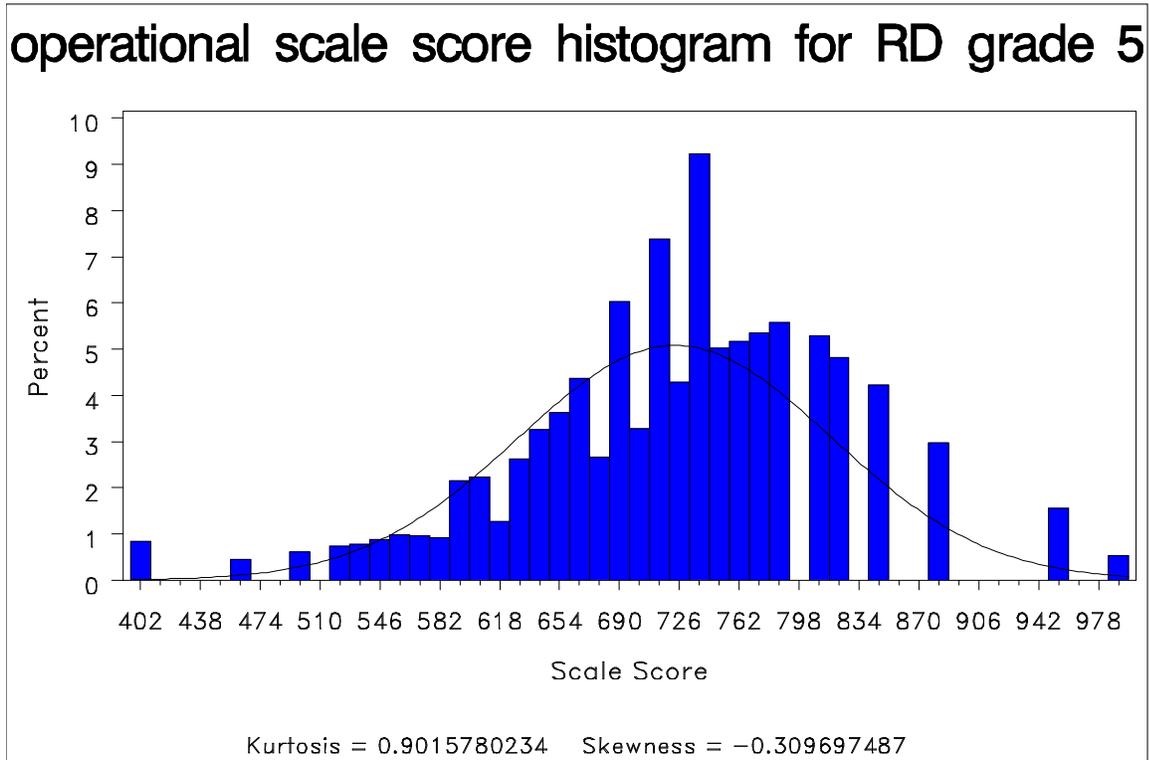


Figure 10. Spring 2014 Grade 6 Reading operational scale score histogram

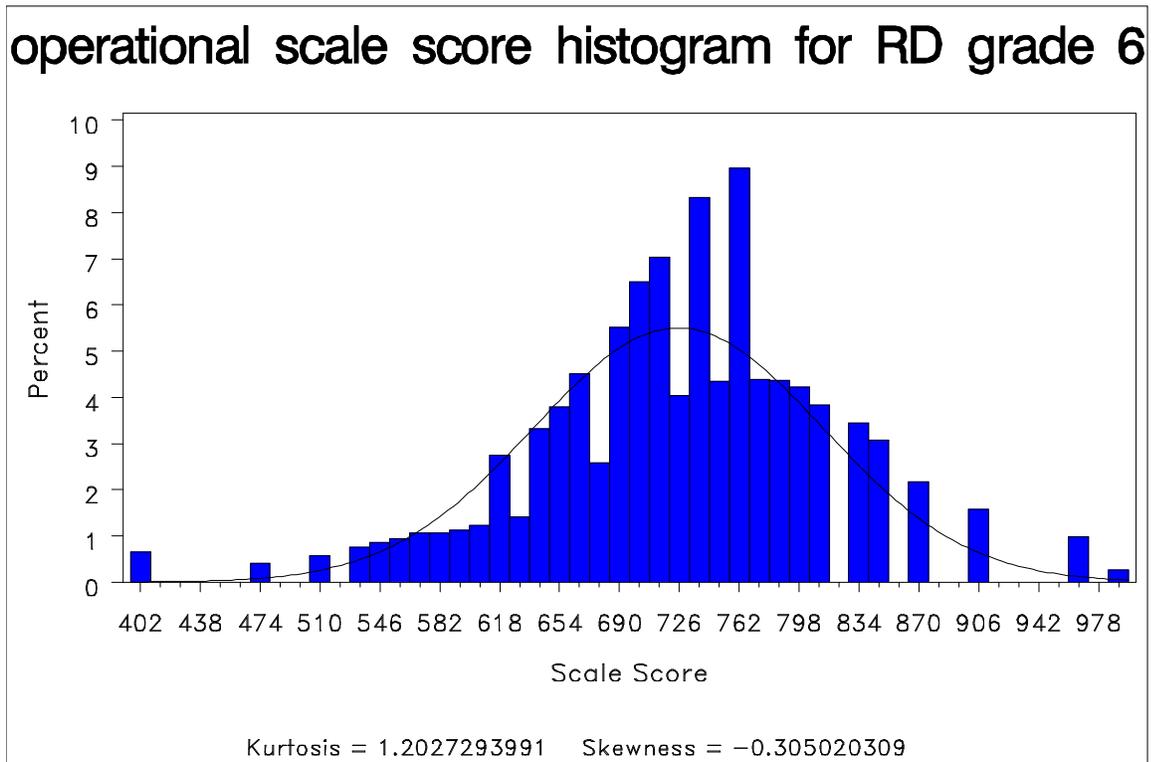


Figure 11. Spring 2014 Grade 7 Reading operational scale score histogram

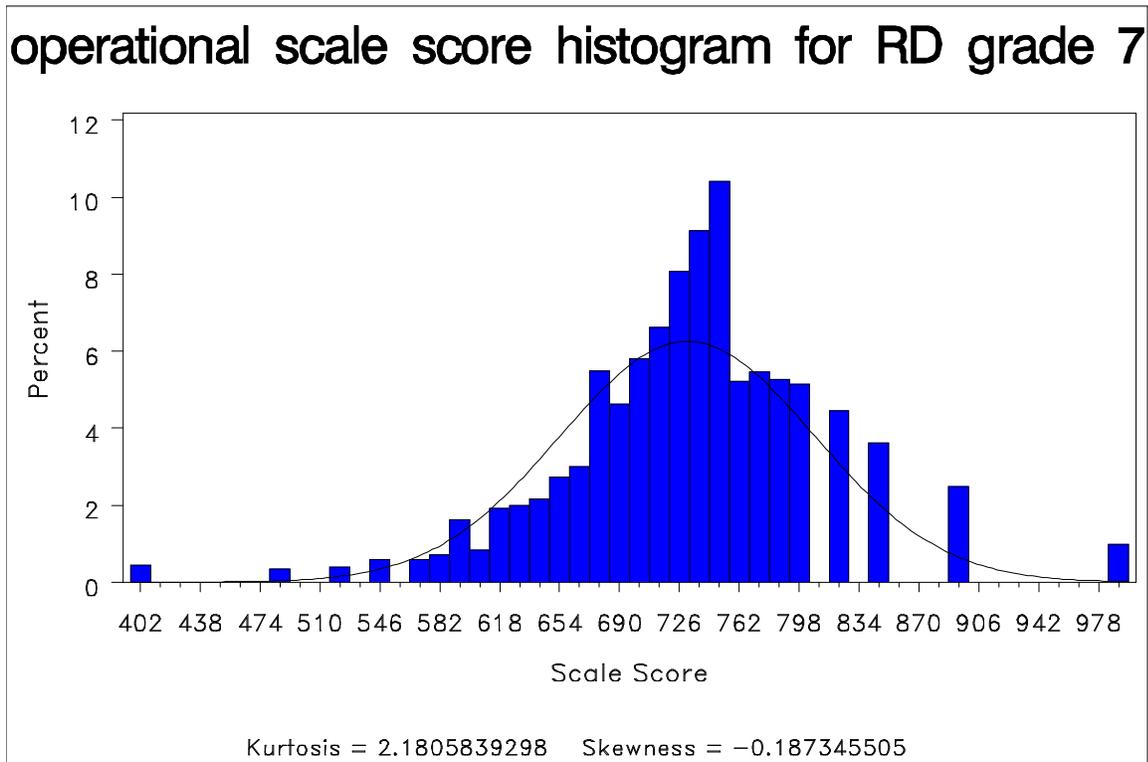


Figure 12. Spring 2014 Grade 8 Reading operational scale score histogram

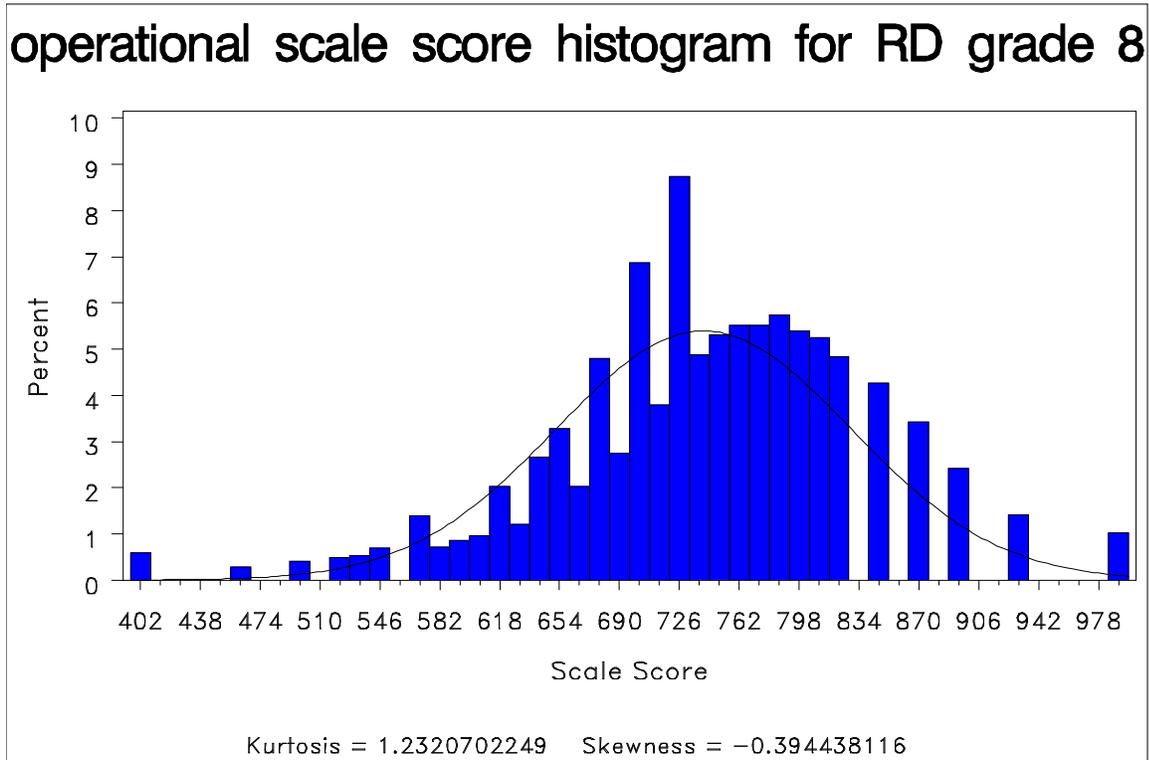


Figure 13. Spring 2014 Grade 5 Science operational scale score histogram

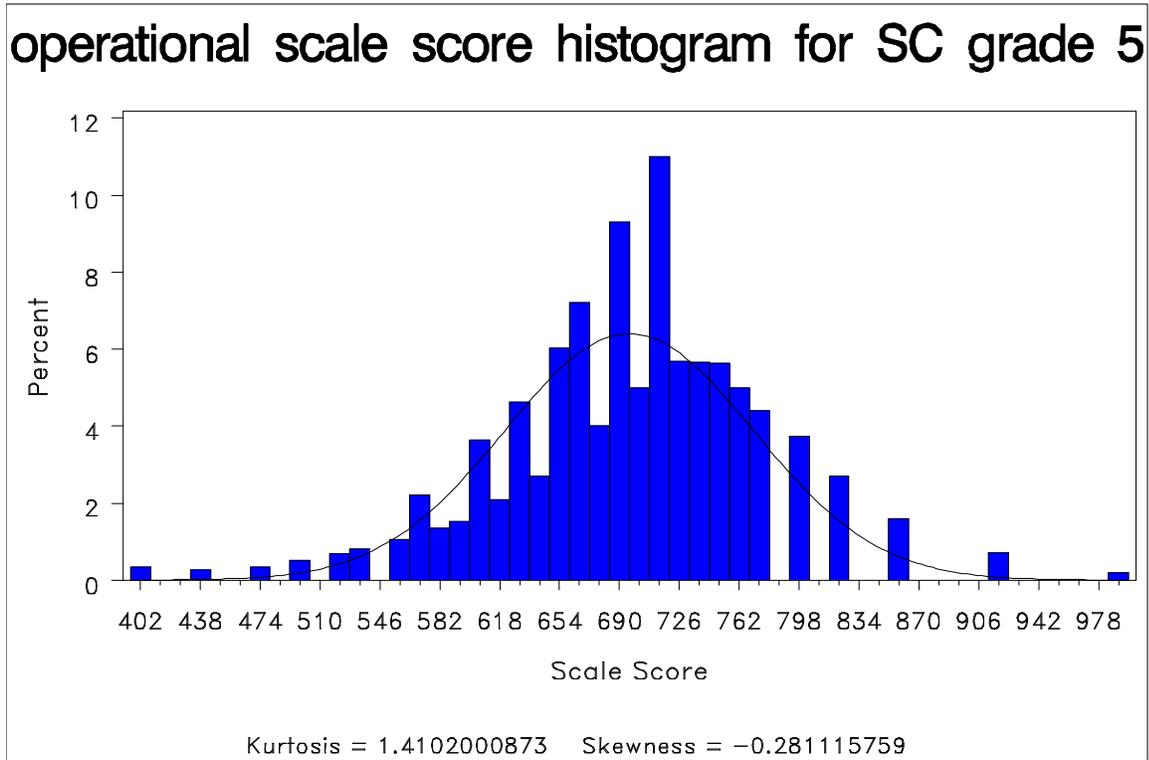


Figure 14. Spring 2014 Grade 8 Science operational scale score histogram

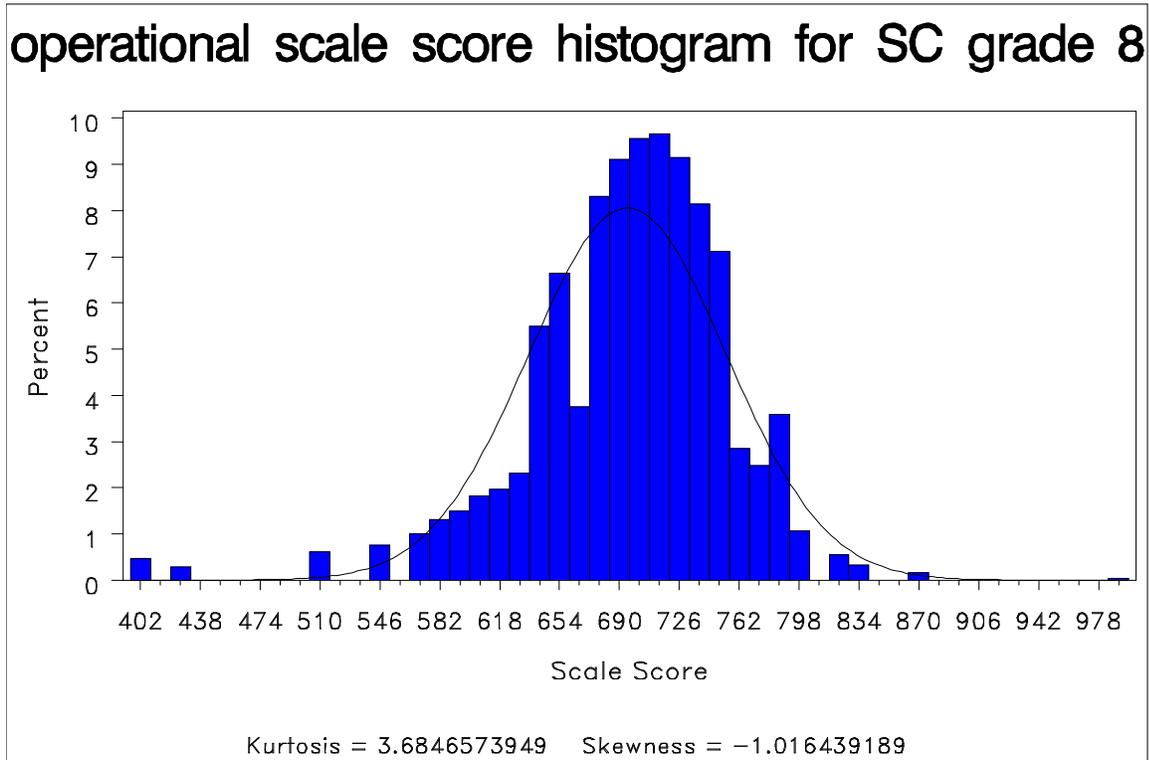


Figure 15. Spring 2014 Grade 5 Social Studies operational scale score histogram

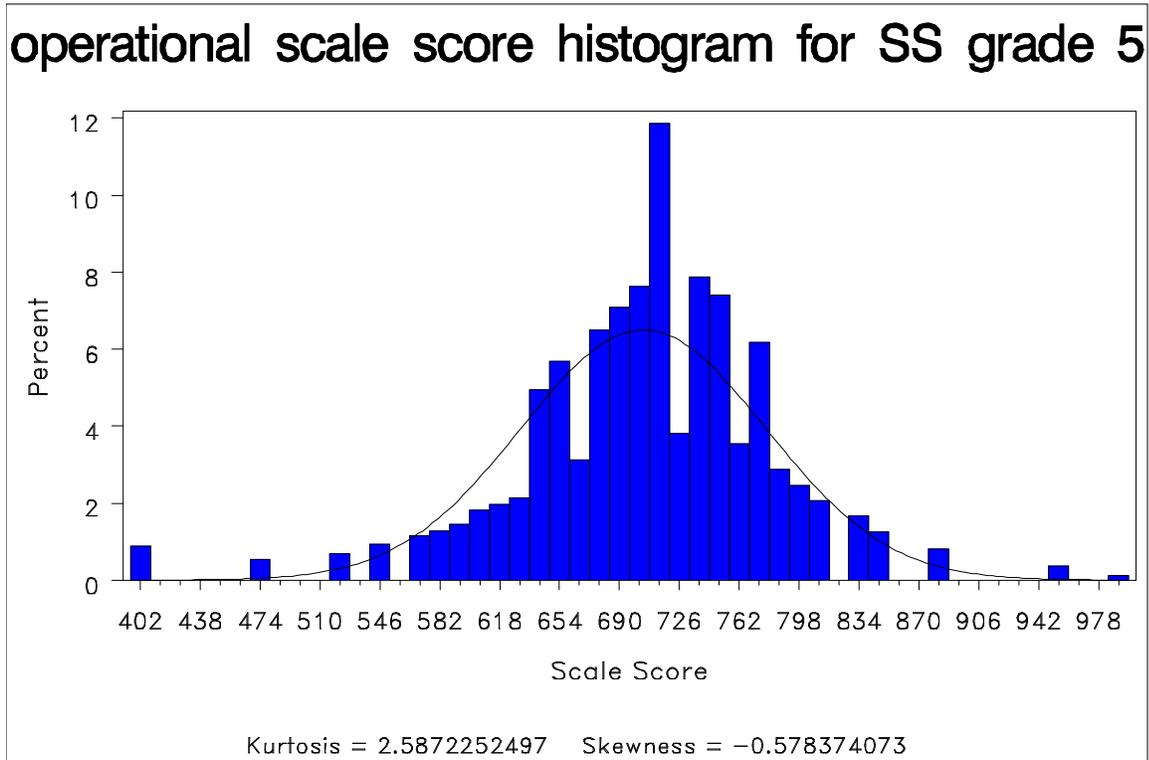


Figure 16. Spring 2014 Grade 8 U.S. History operational scale score histogram

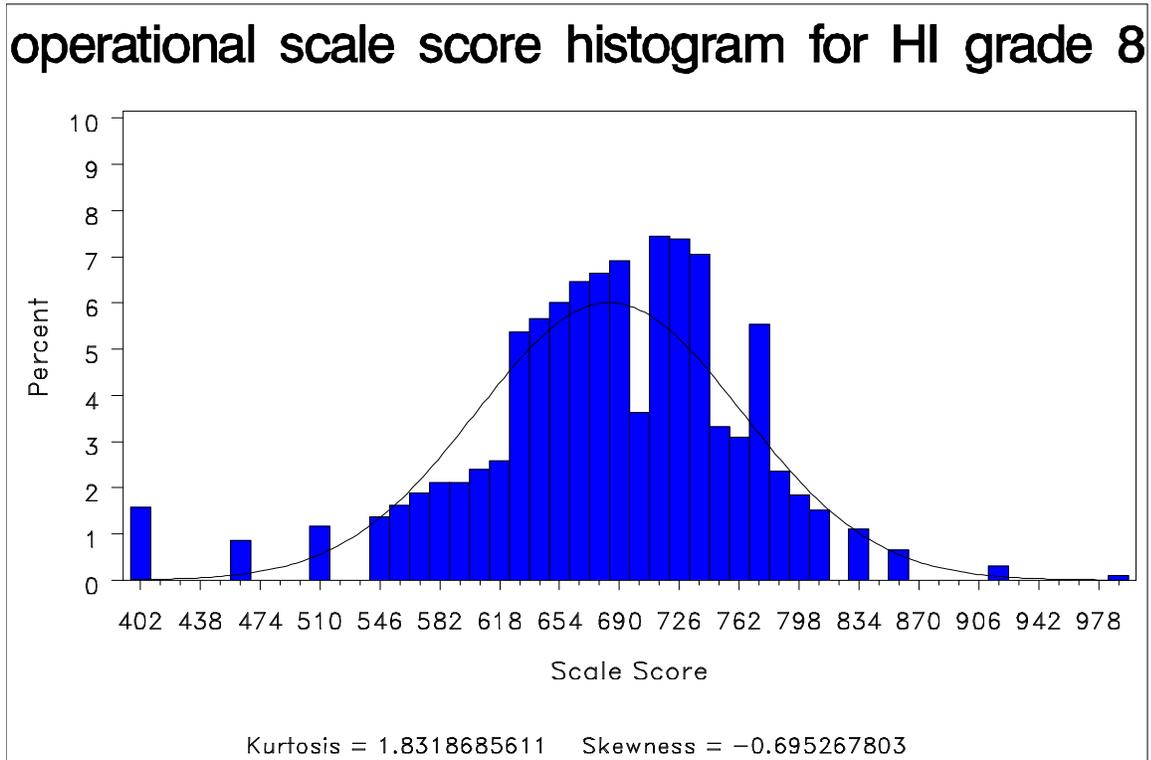


Figure 17. Spring 2014 Grade 3 Mathematics operational test characteristic curve and standard error of measurement curve

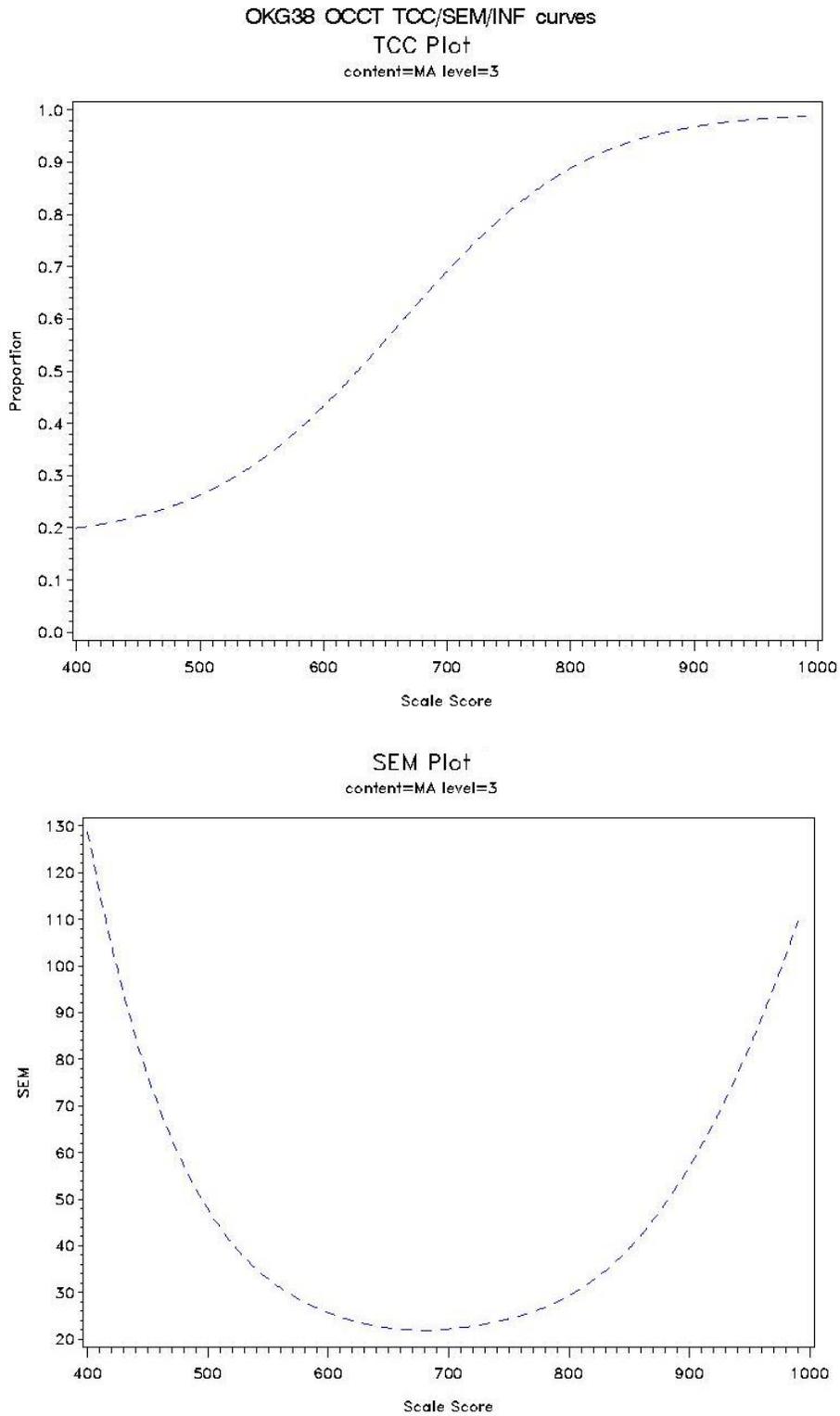


Figure 18. Spring 2014 Grade 4 Mathematics operational test characteristic curve and standard error of measurement curve

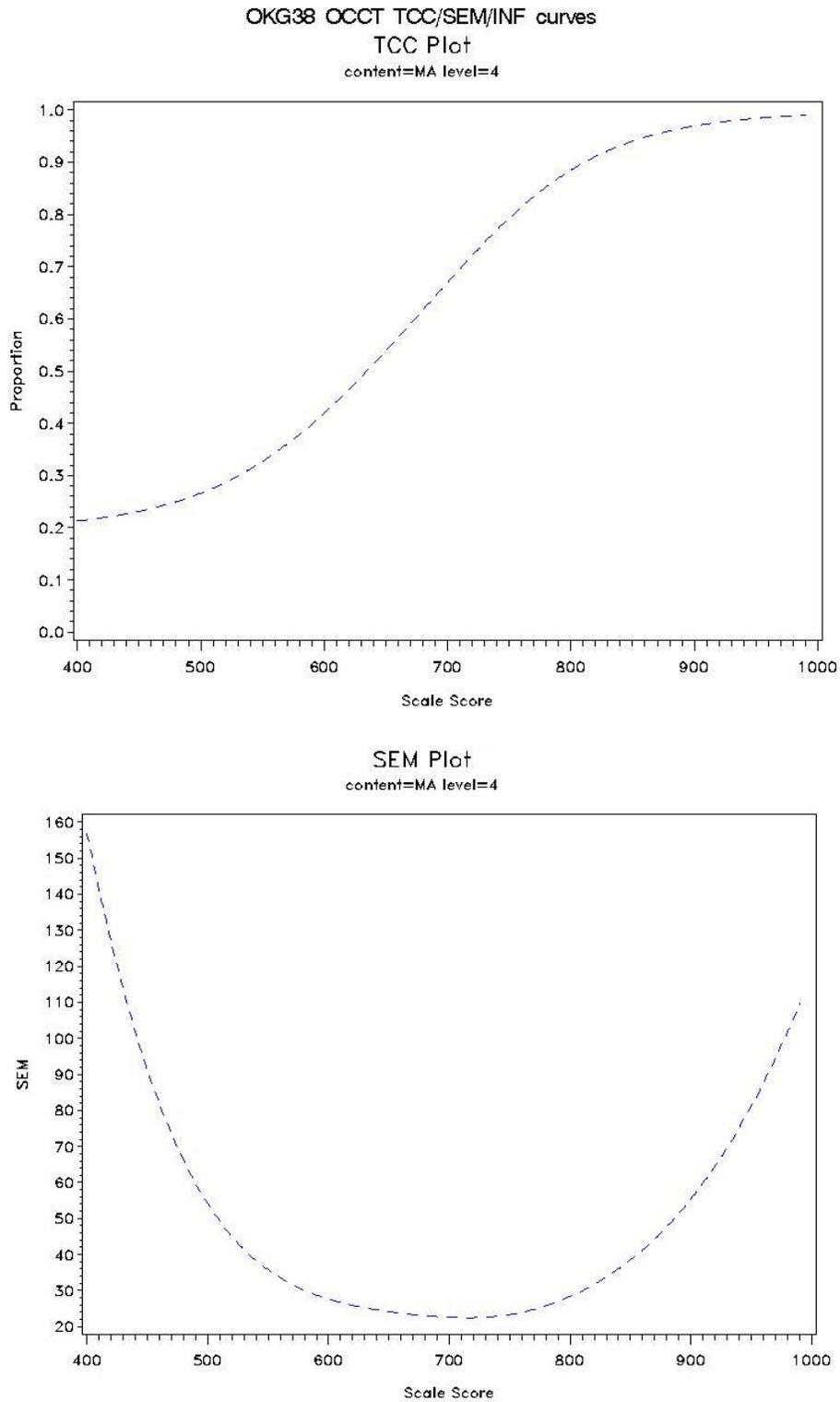


Figure 19. Spring 2014 Grade 5 Mathematics operational test characteristic curve and standard error of measurement curve

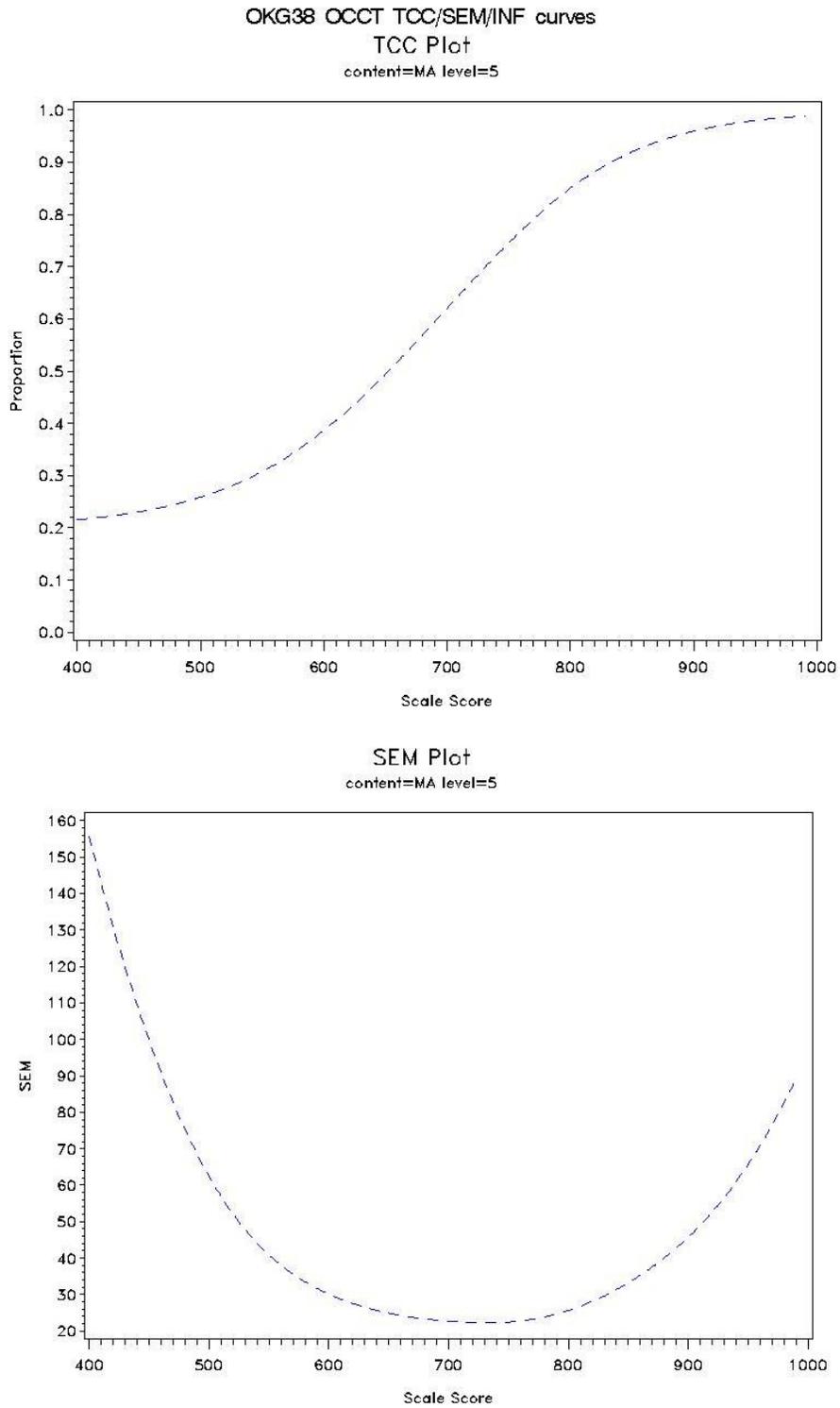


Figure 20. Spring 2014 Grade 6 Mathematics operational test characteristic curve and standard error of measurement curve

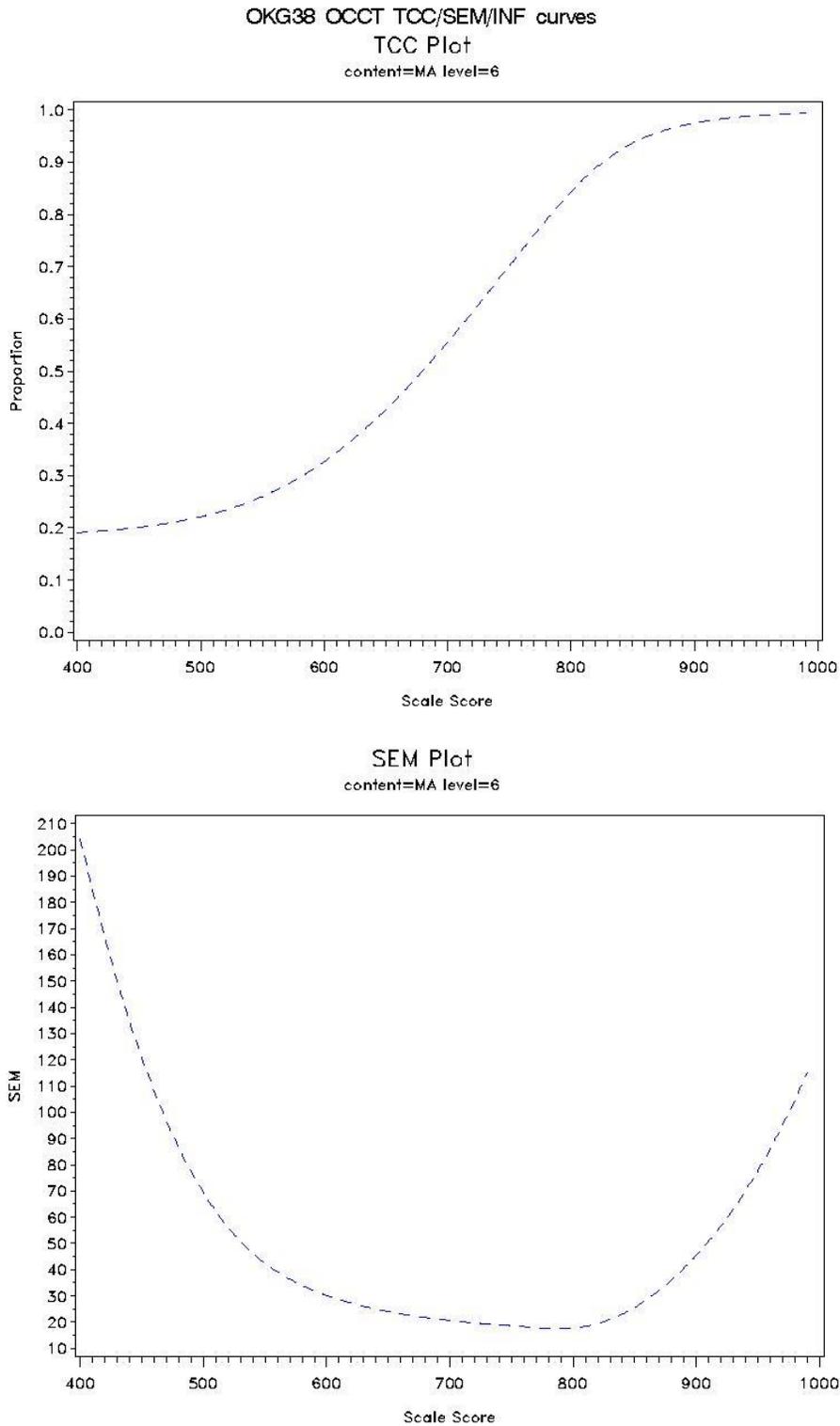


Figure 21. Spring 2014 Grade 7 Mathematics operational test characteristic curve and standard error of measurement curve

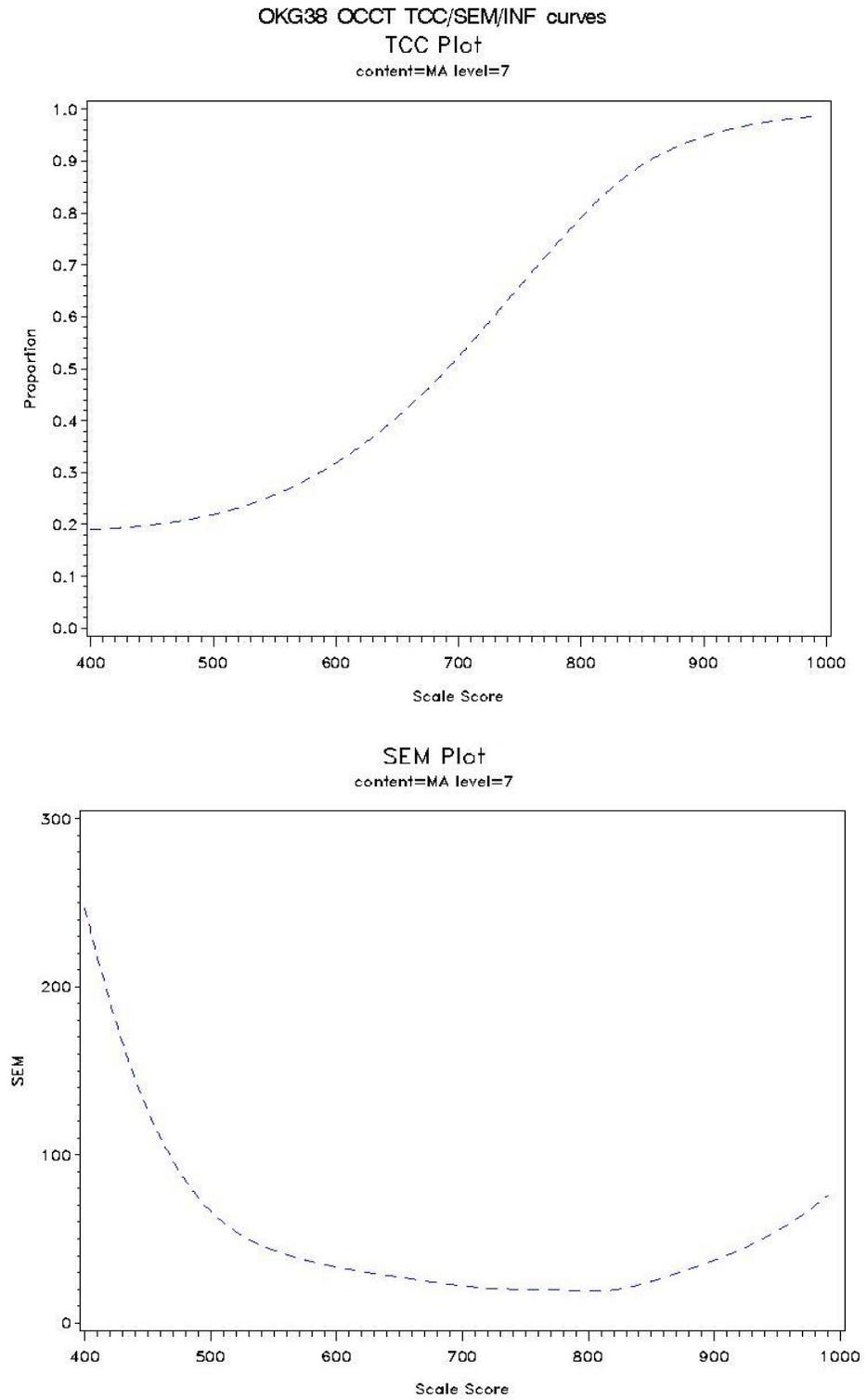


Figure 22. Spring 2014 Grade 8 Mathematics operational test characteristic curve and standard error of measurement curve

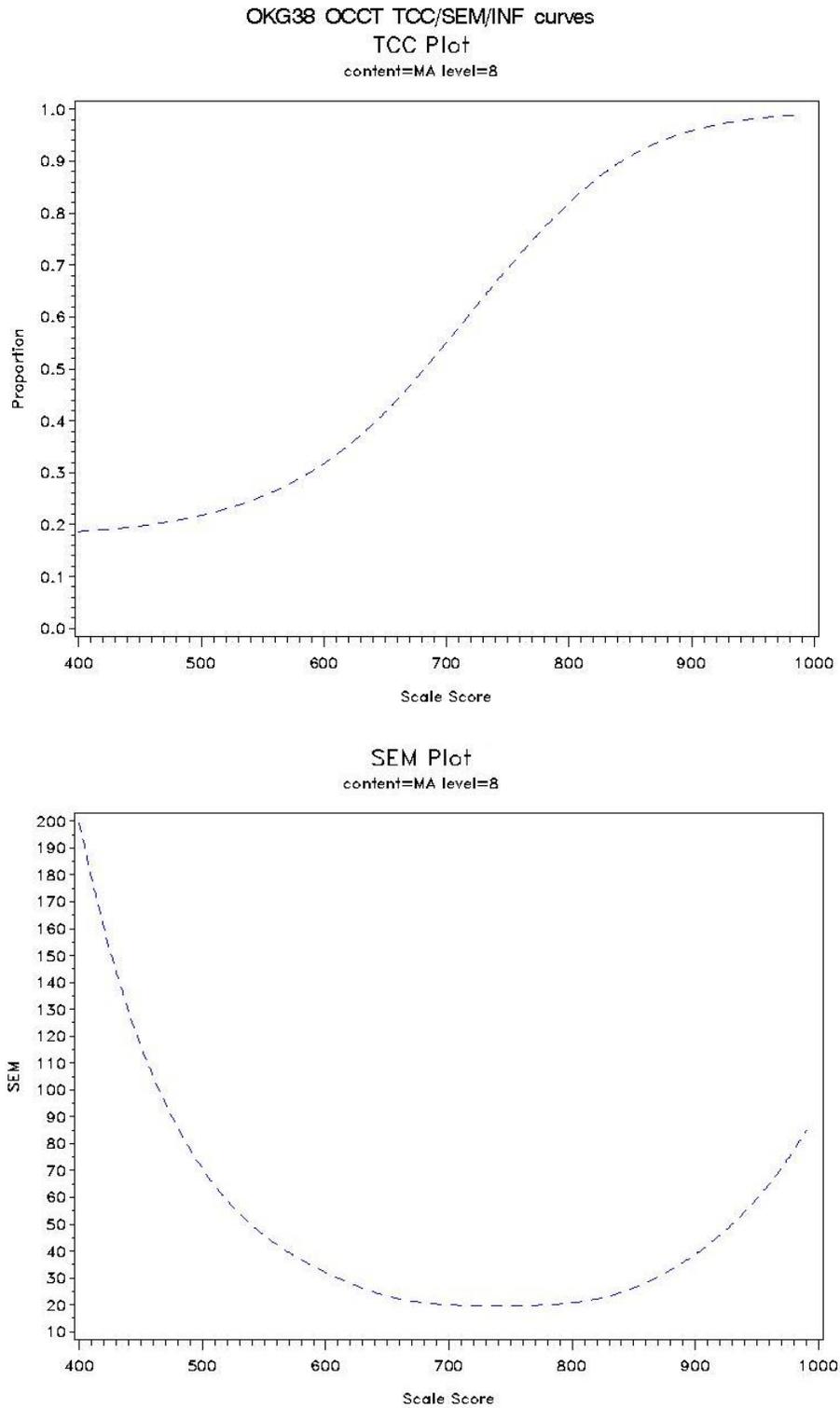


Figure 23. Spring 2014 Grade 3 Reading operational test characteristic curve and standard error of measurement curve

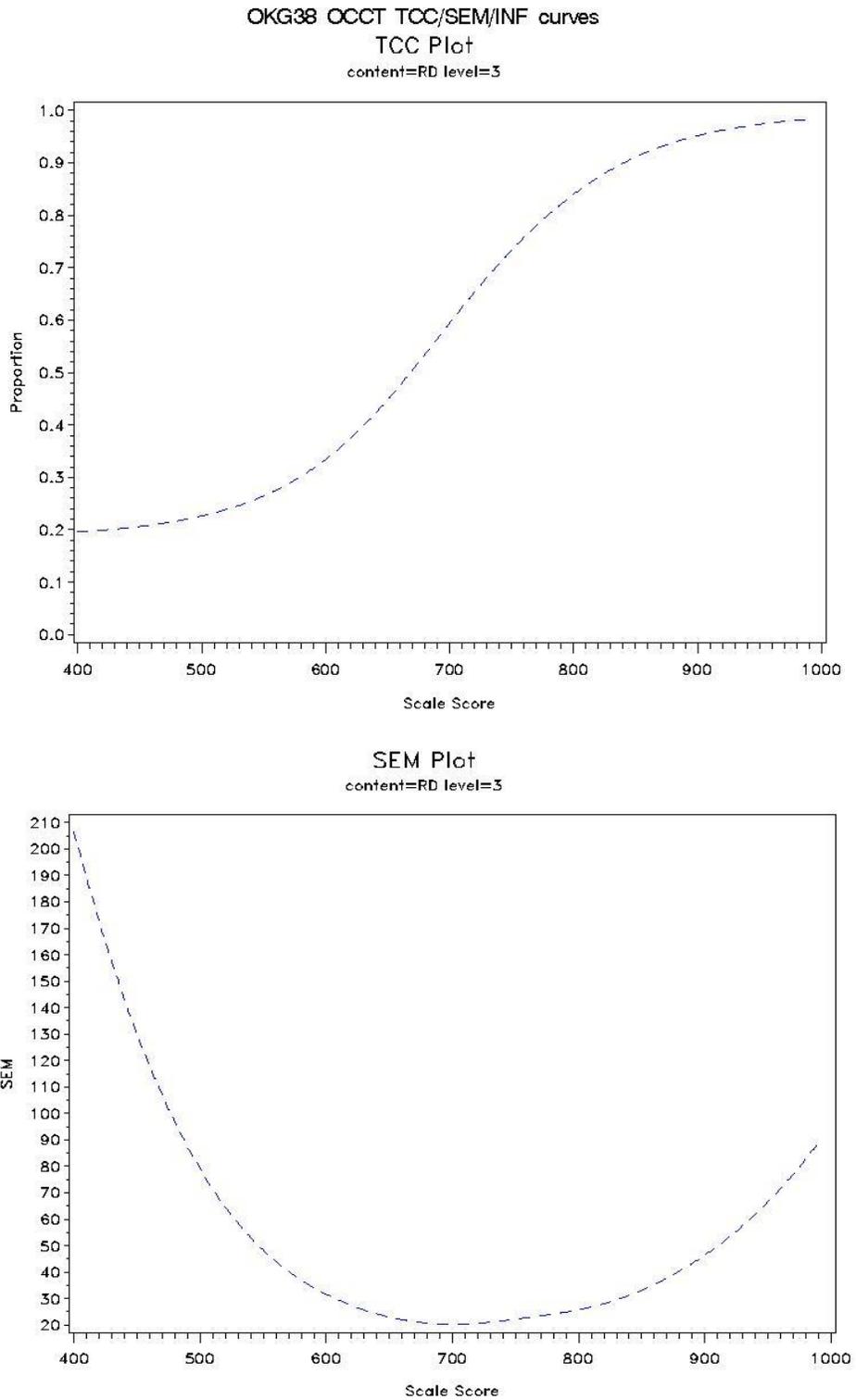


Figure 24. Spring 2014 Grade 4 Reading operational test characteristic curve and standard error of measurement curve

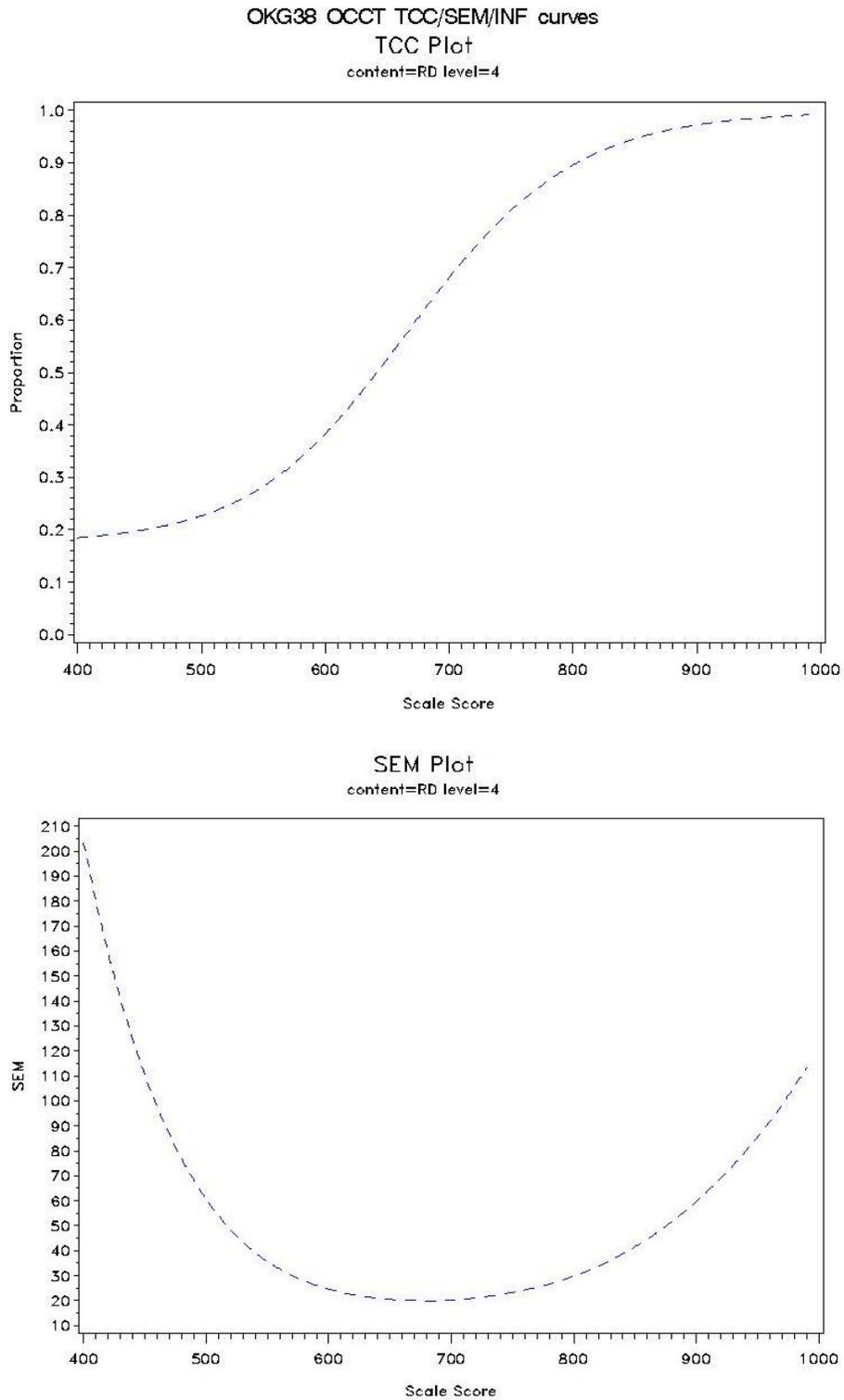


Figure 25. Spring 2014 Grade 5 Reading operational test characteristic curve and standard error of measurement curve

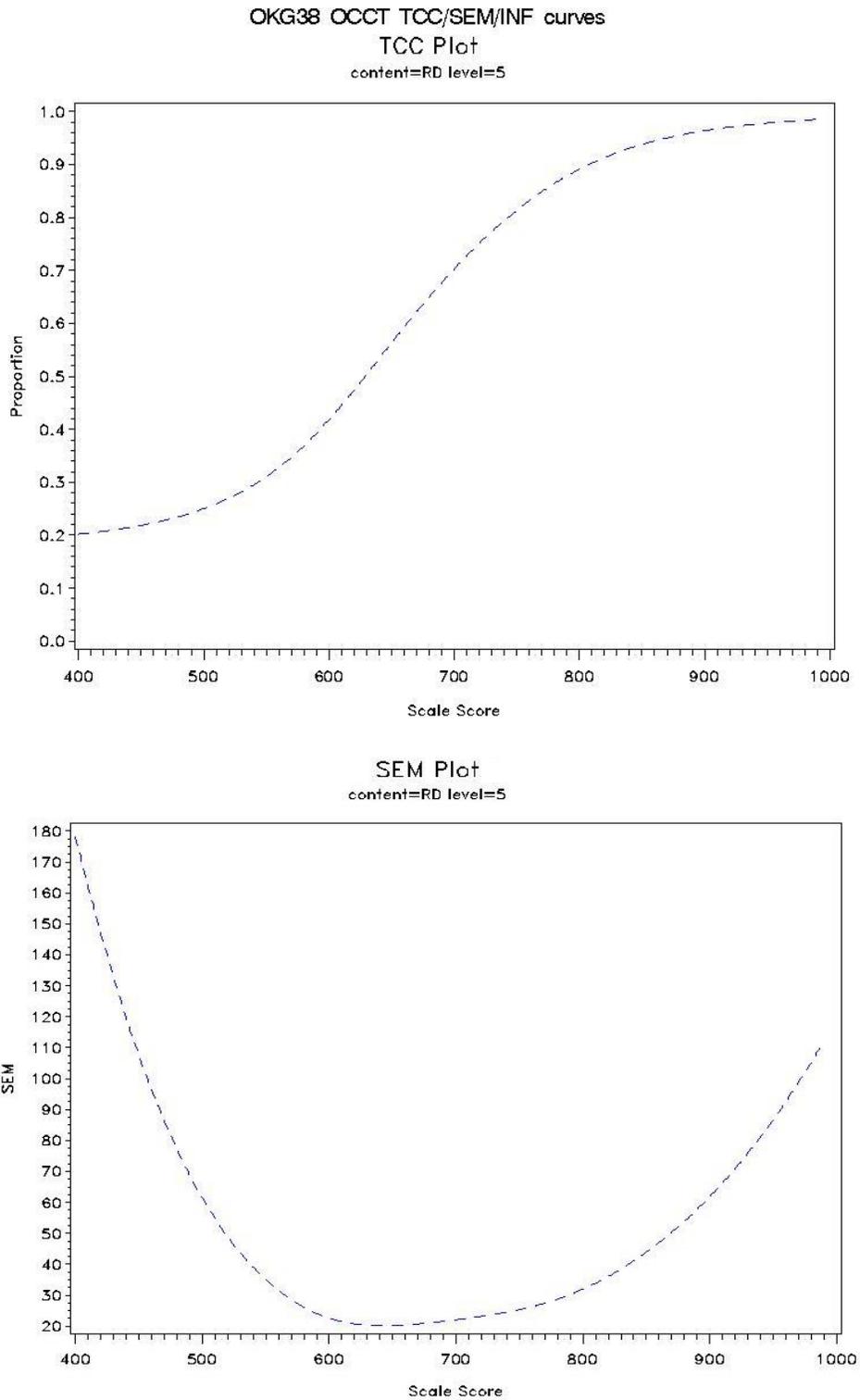


Figure 26. Spring 2014 Grade 6 Reading operational test characteristic curve and standard error of measurement curve

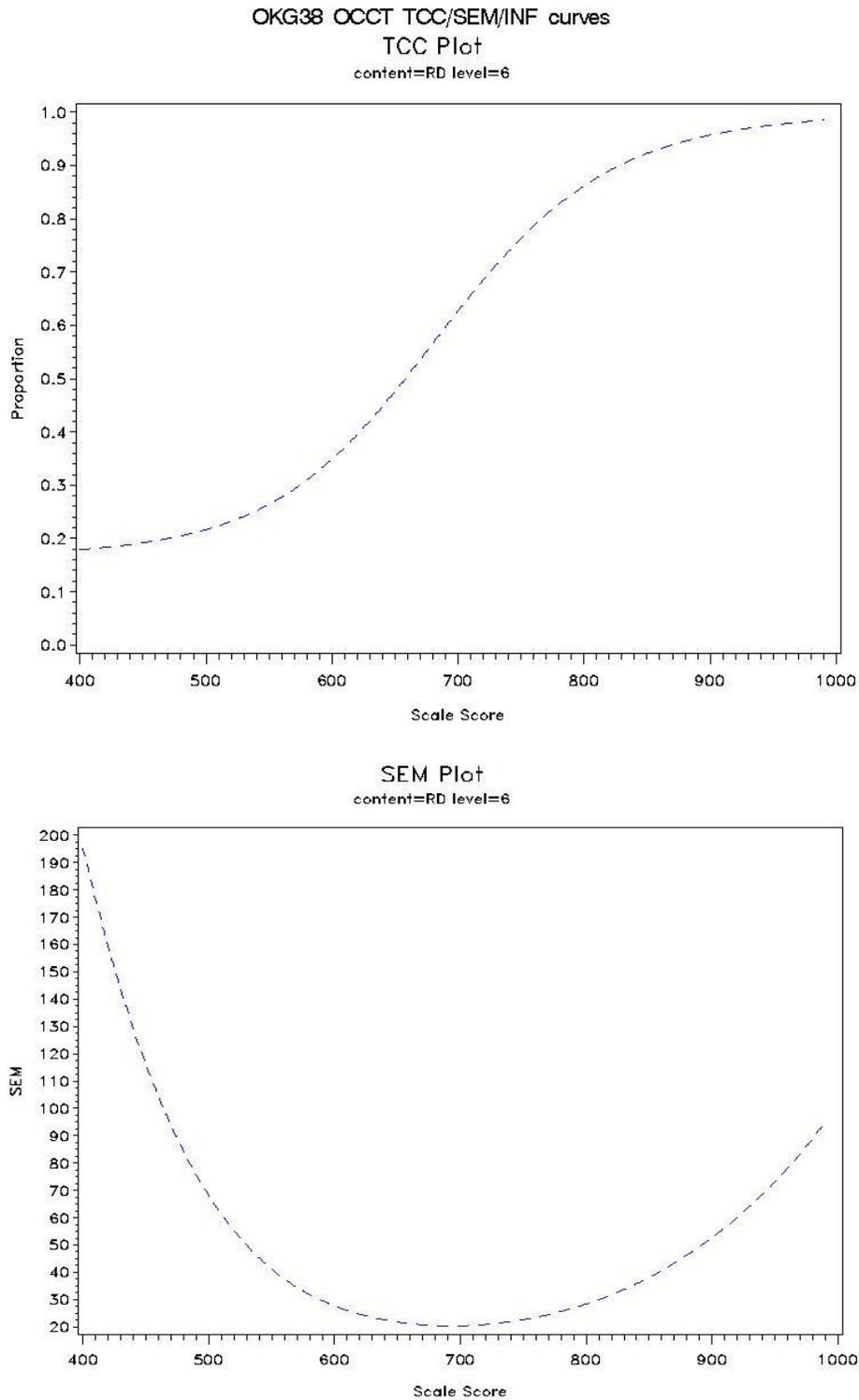


Figure 27. Spring 2014 Grade 7 Reading operational test characteristic curve and standard error of measurement curve

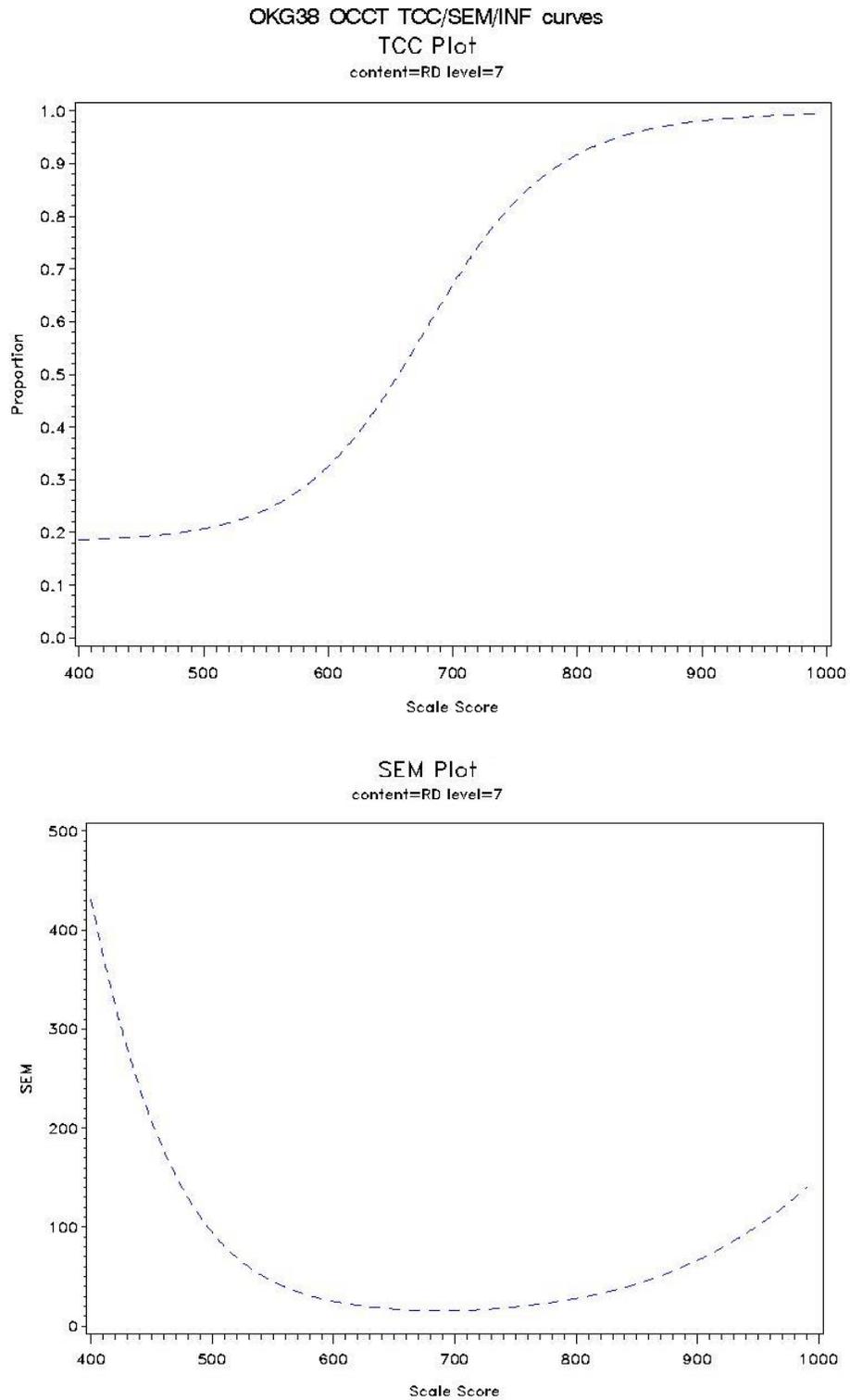


Figure 28. Spring 2014 Grade 8 Reading operational test characteristic curve and standard error of measurement curve

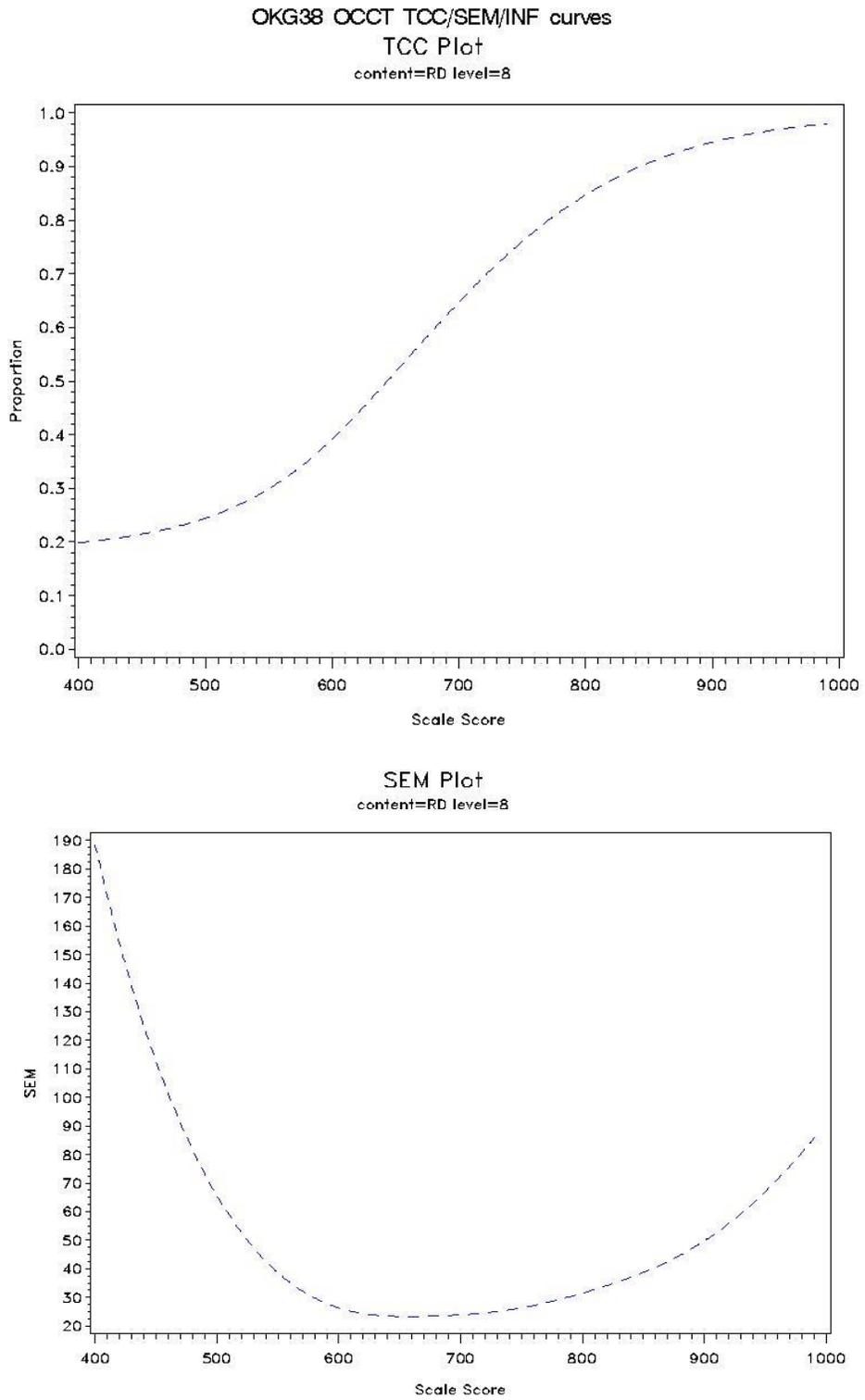


Figure 29. Spring 2014 Grade 5 Science operational test characteristic curve and standard error of measurement curve

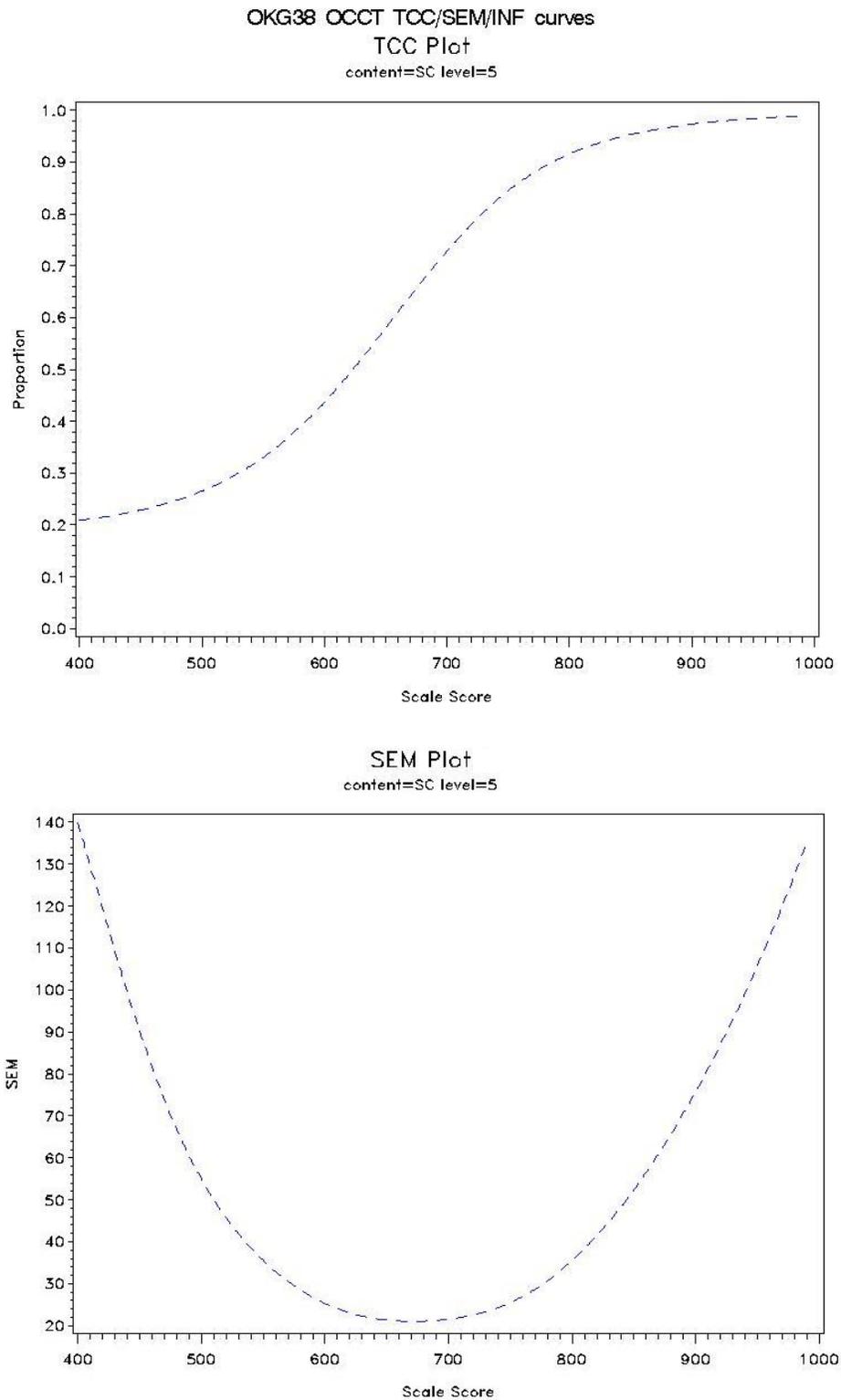


Figure 30. Spring 2014 Grade 8 Science operational test characteristic curve and standard error of measurement curve

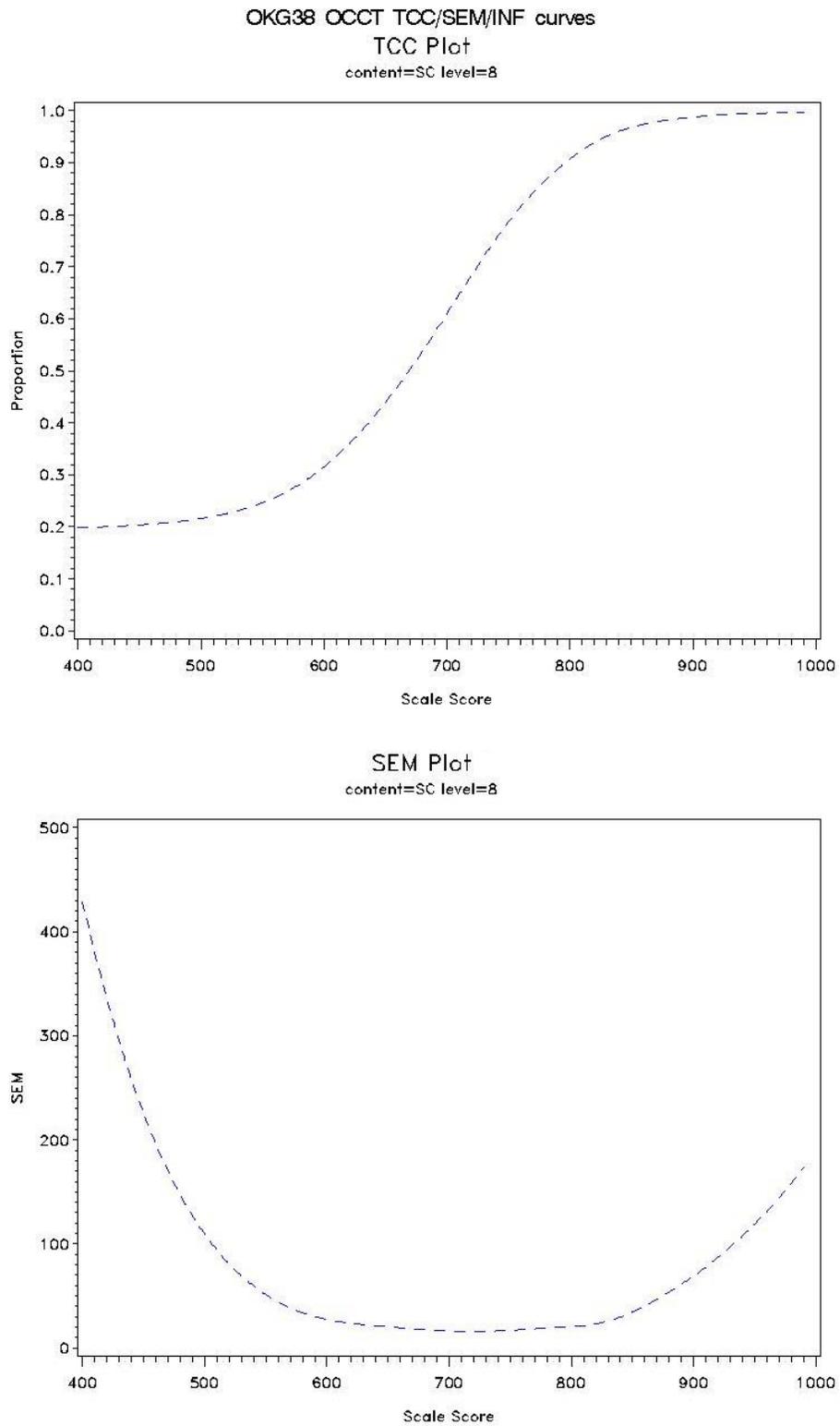


Figure 31. Spring 2014 Grade 5 Social Studies operational test characteristic curve and standard error of measurement curve

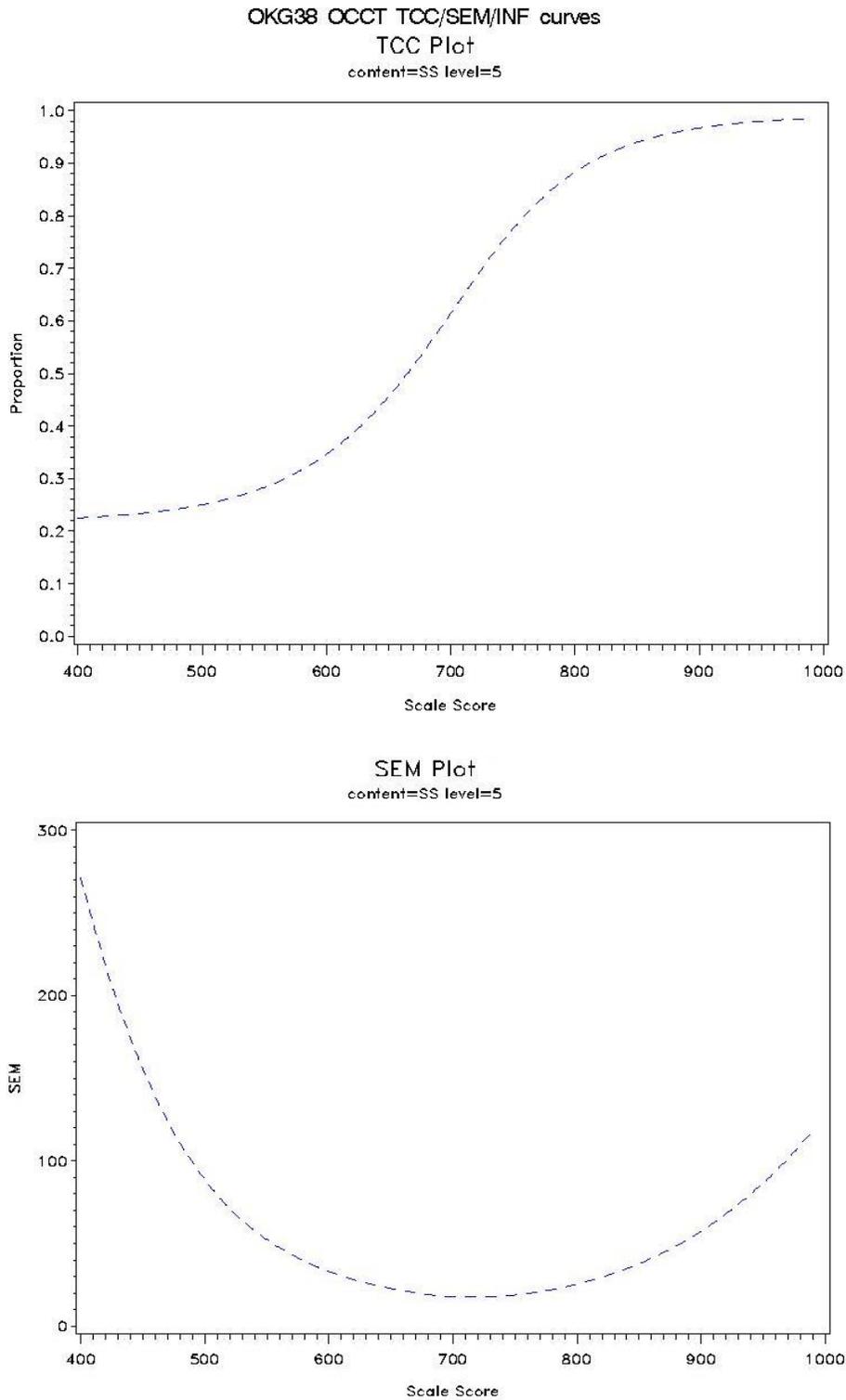


Figure 32. Spring 2014 Grade 8 U.S. History operational test characteristic curve and standard error of measurement curve

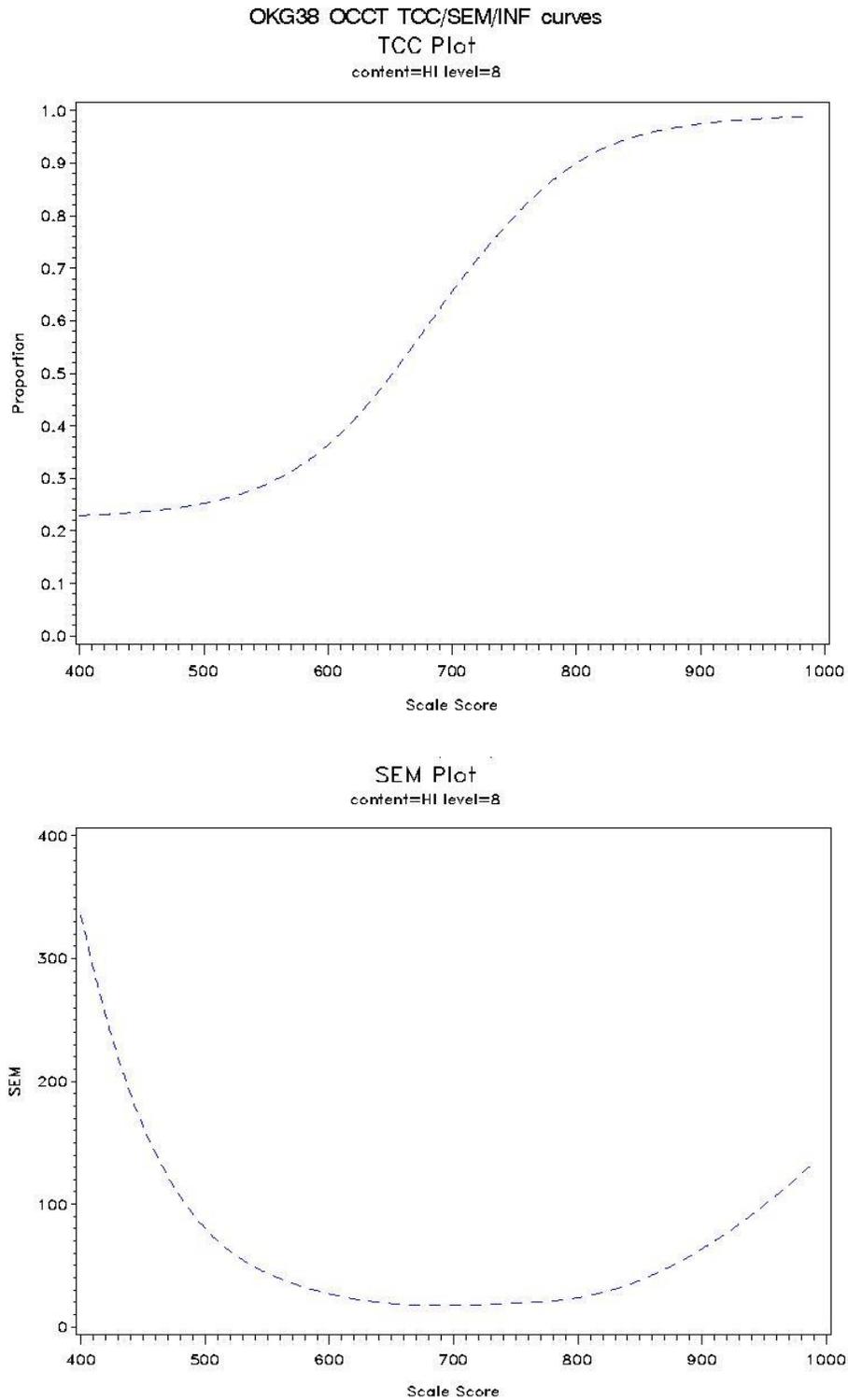


Figure 33. Spring 2014 Grade 3 Mathematics scree plot: Total

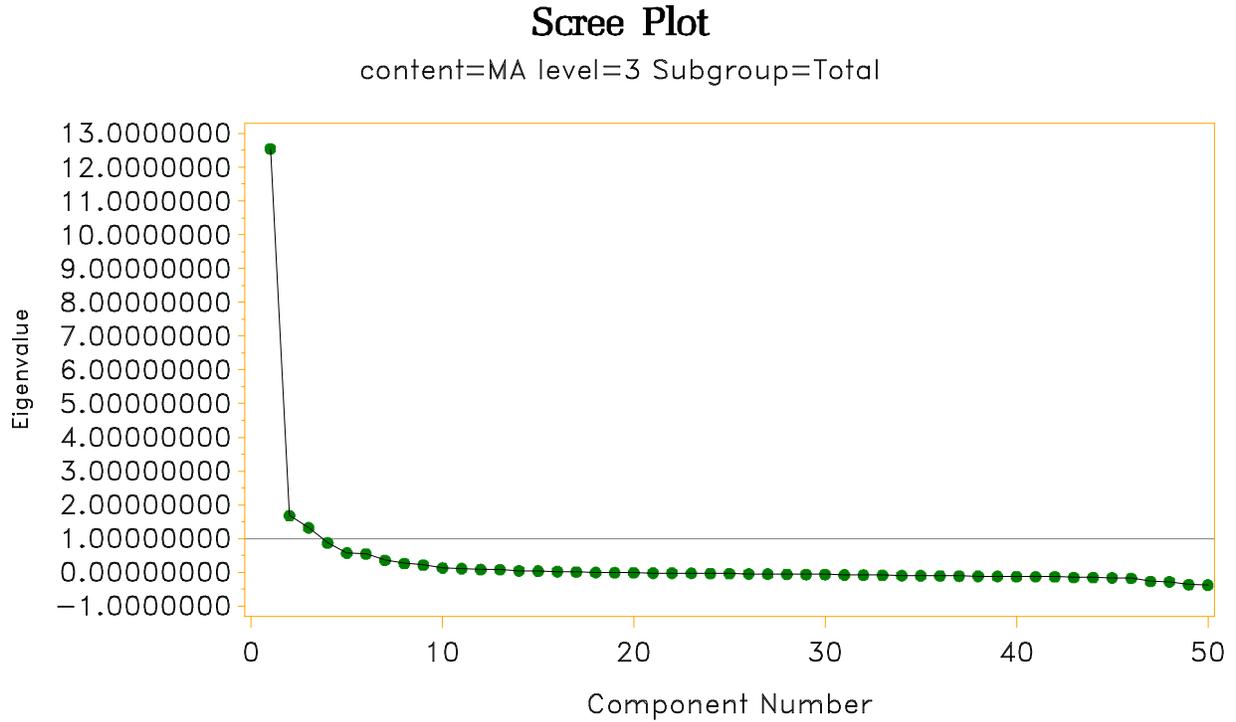


Figure 34. Spring 2014 Grade 3 Mathematics scree plot: Accommodated

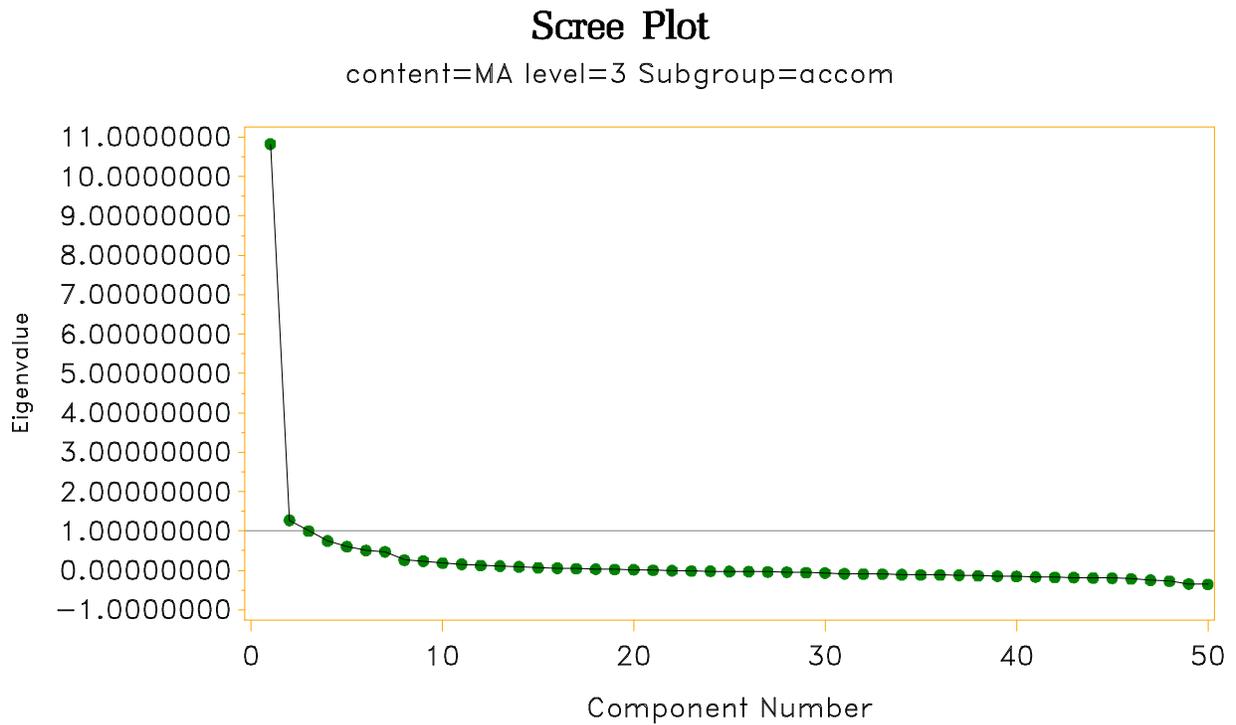


Figure 35. Spring 2014 Grade 3 Mathematics scree plot: English Language Learner

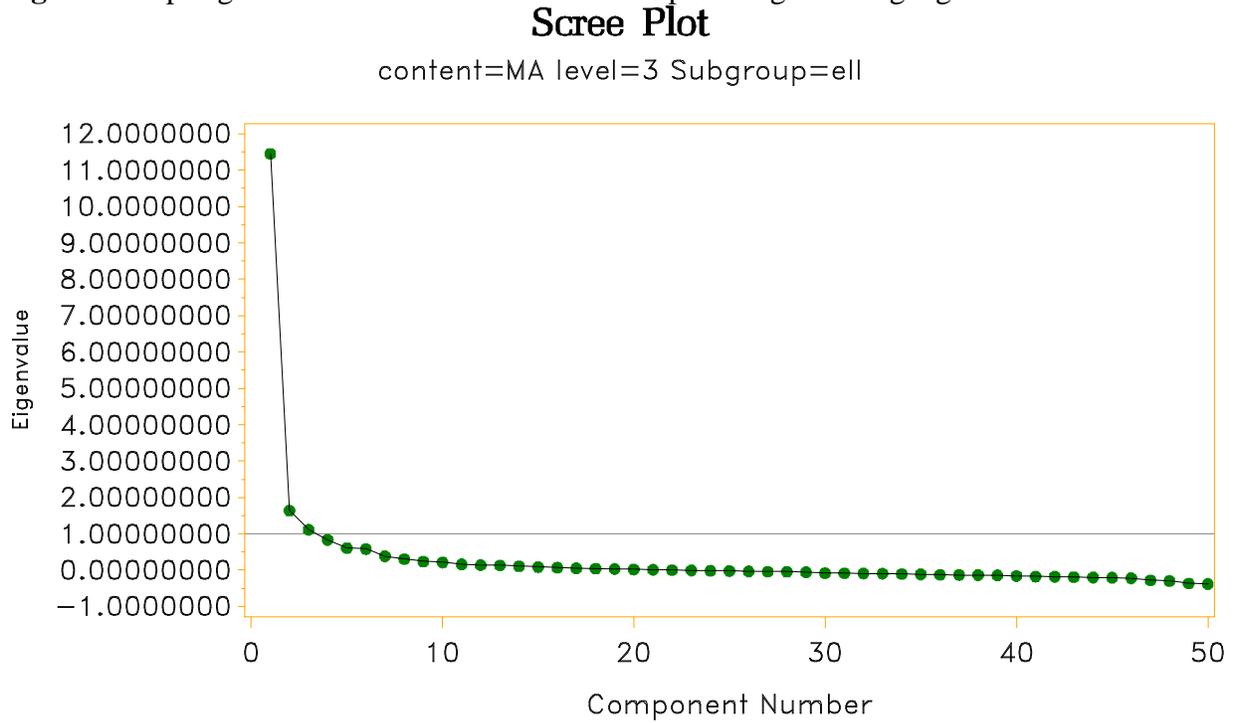


Figure 36. Spring 2014 Grade 3 Mathematics scree plot: Free Lunch

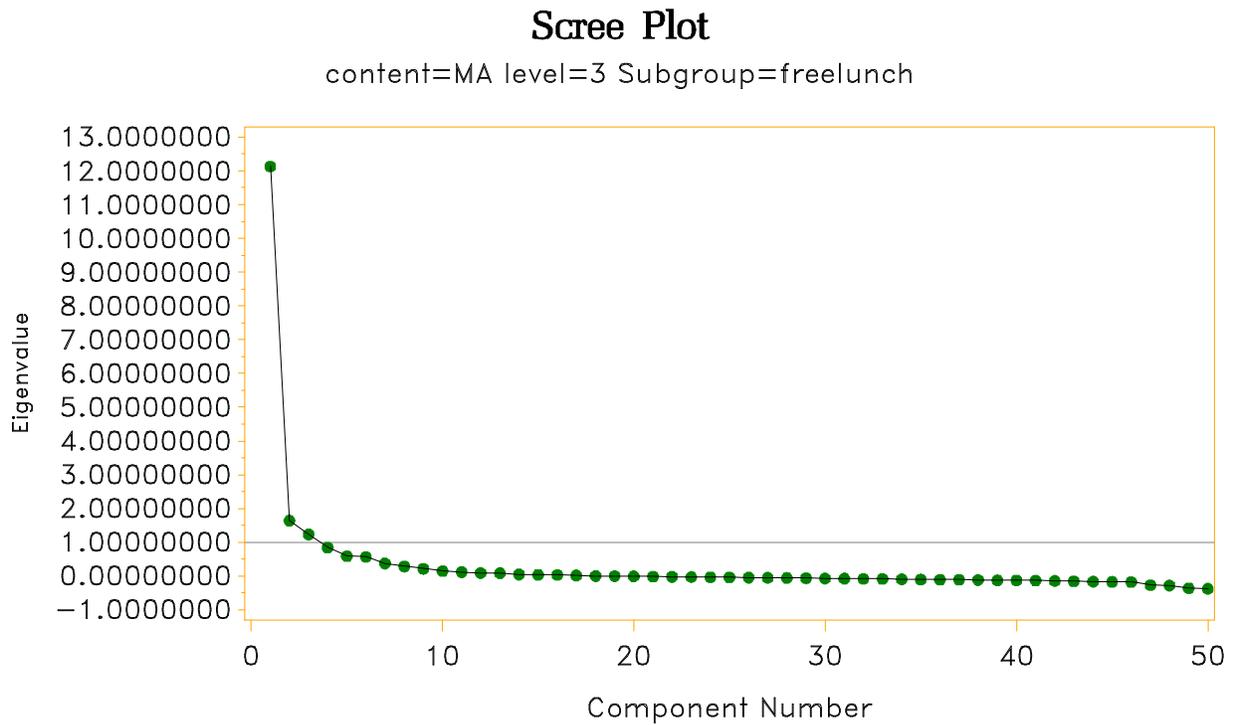


Figure 37. Spring 2014 Grade 3 Mathematics scree plot: Individualized Education Program

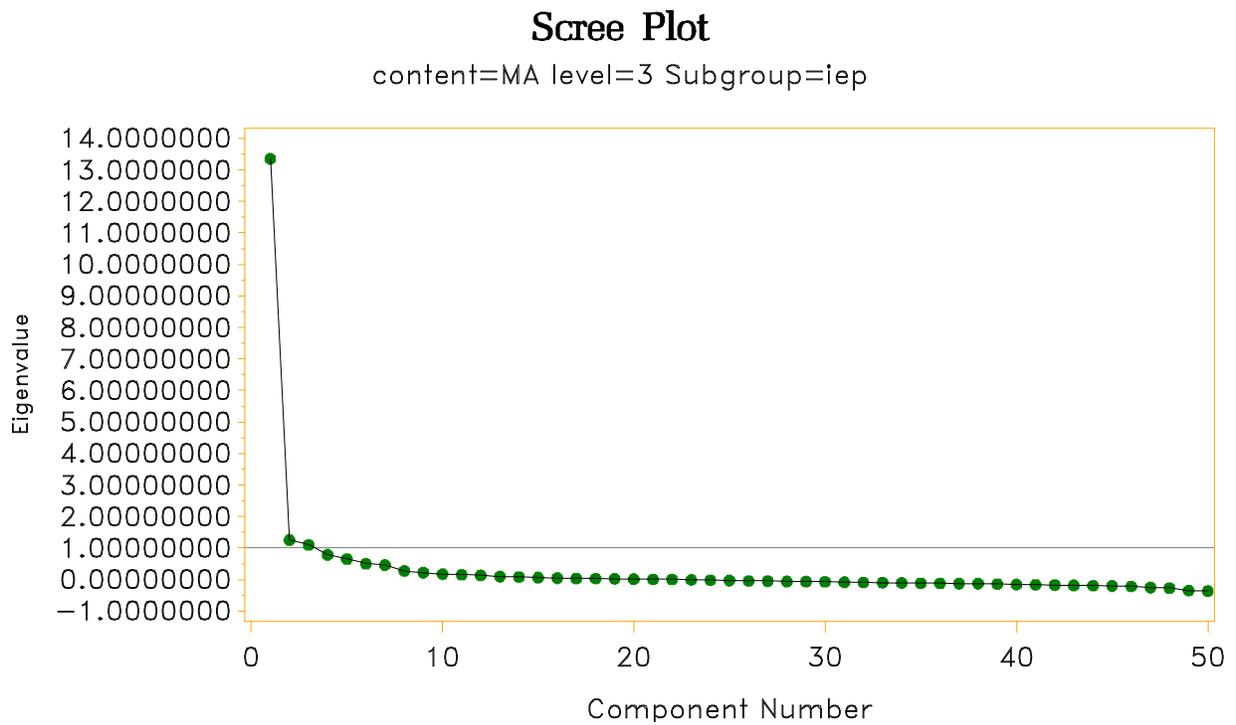


Figure 38. Spring 2014 Grade 4 Mathematics scree plot: Total

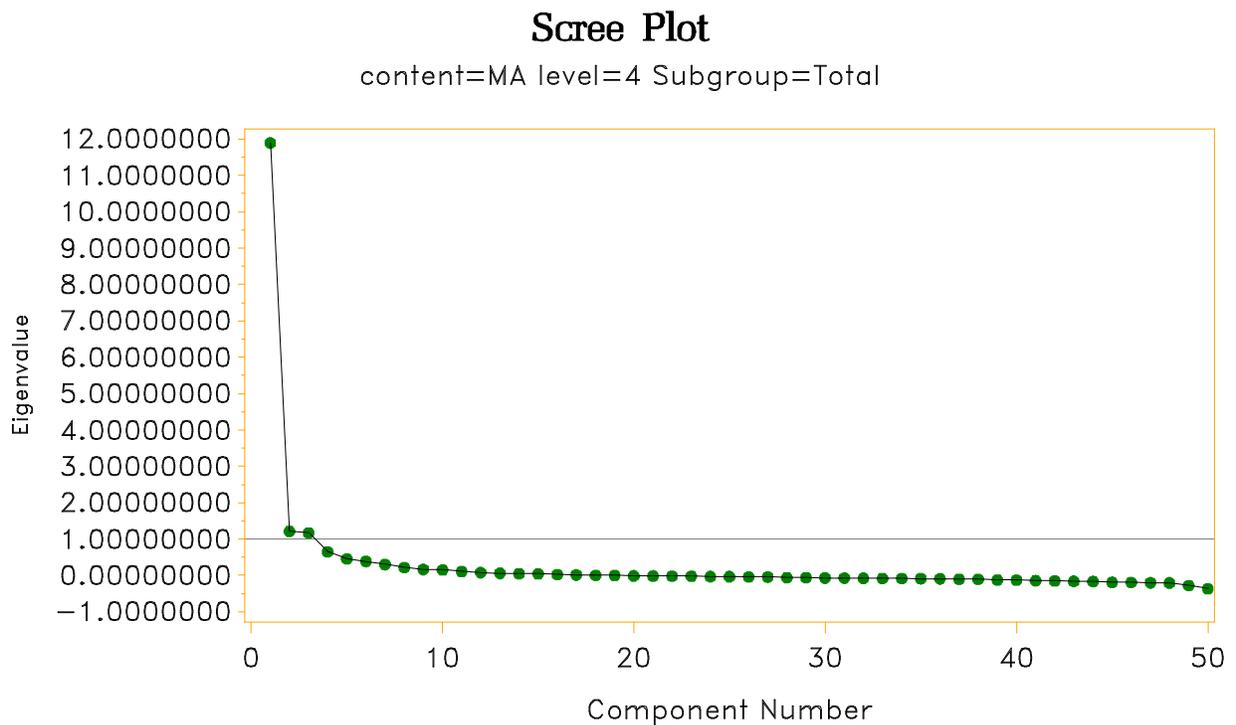


Figure 39. Spring 2014 Grade 4 Mathematics scree plot: Accommodated

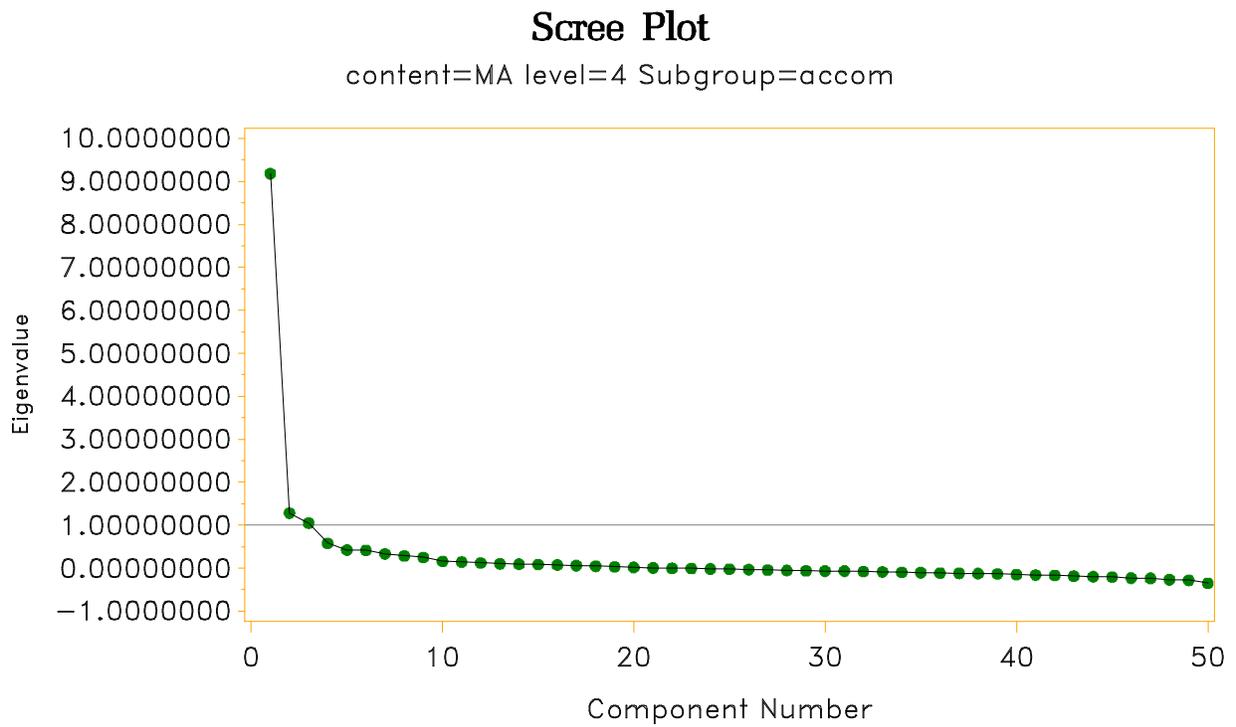


Figure 40. Spring 2014 Grade 4 Mathematics scree plot: English Language Learner

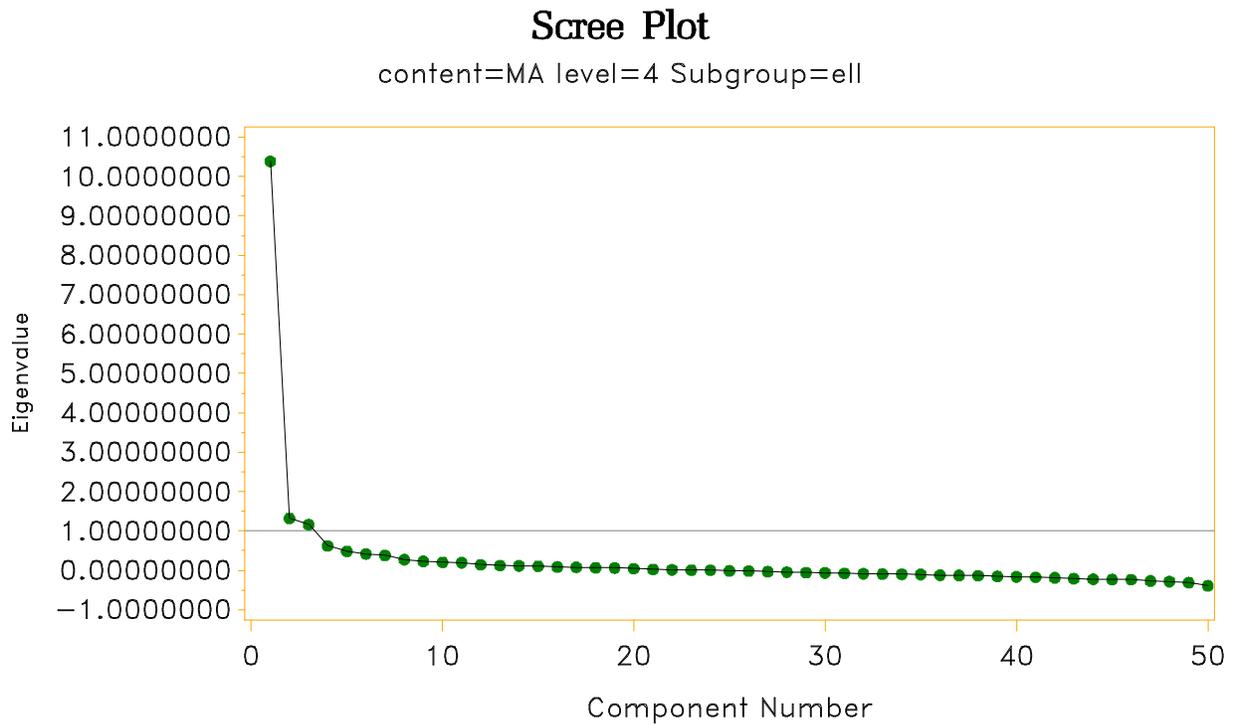


Figure 41. Spring 2014 Grade 4 Mathematics scree plot: Free Lunch

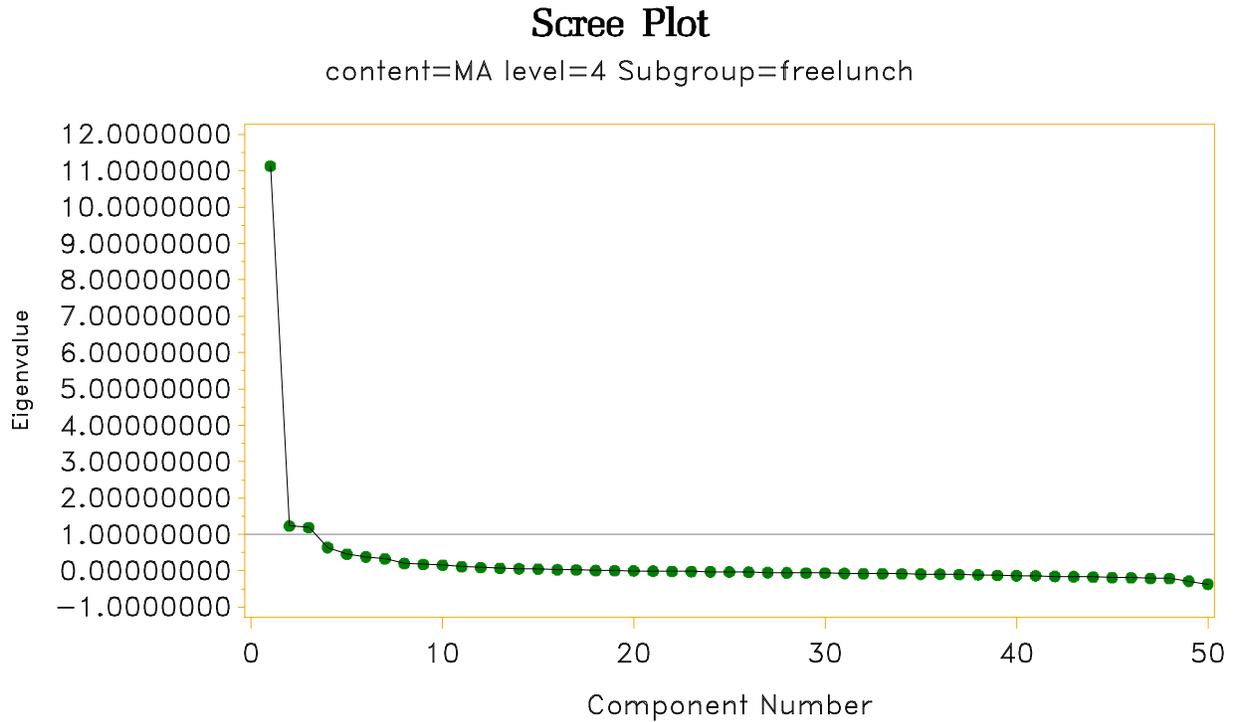


Figure 42. Spring 2014 Grade 4 Mathematics scree plot: Individualized Education Program

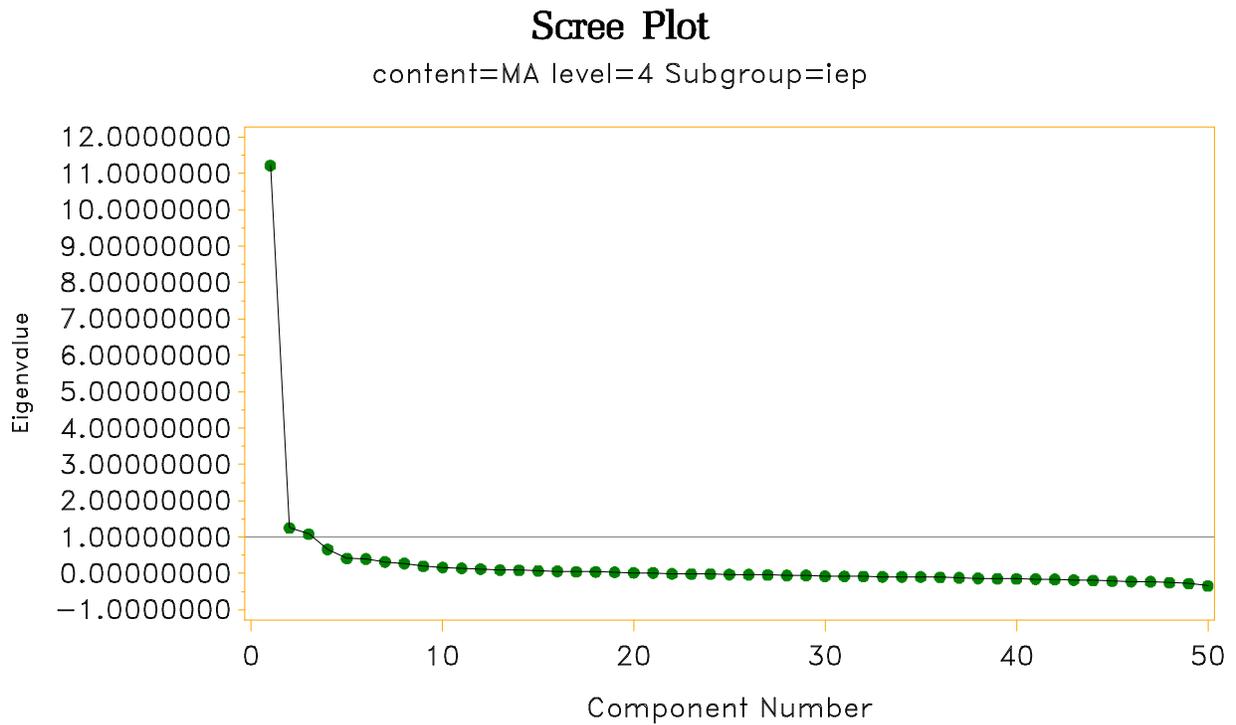


Figure 43. Spring 2014 Grade 5 Mathematics scree plot: Total

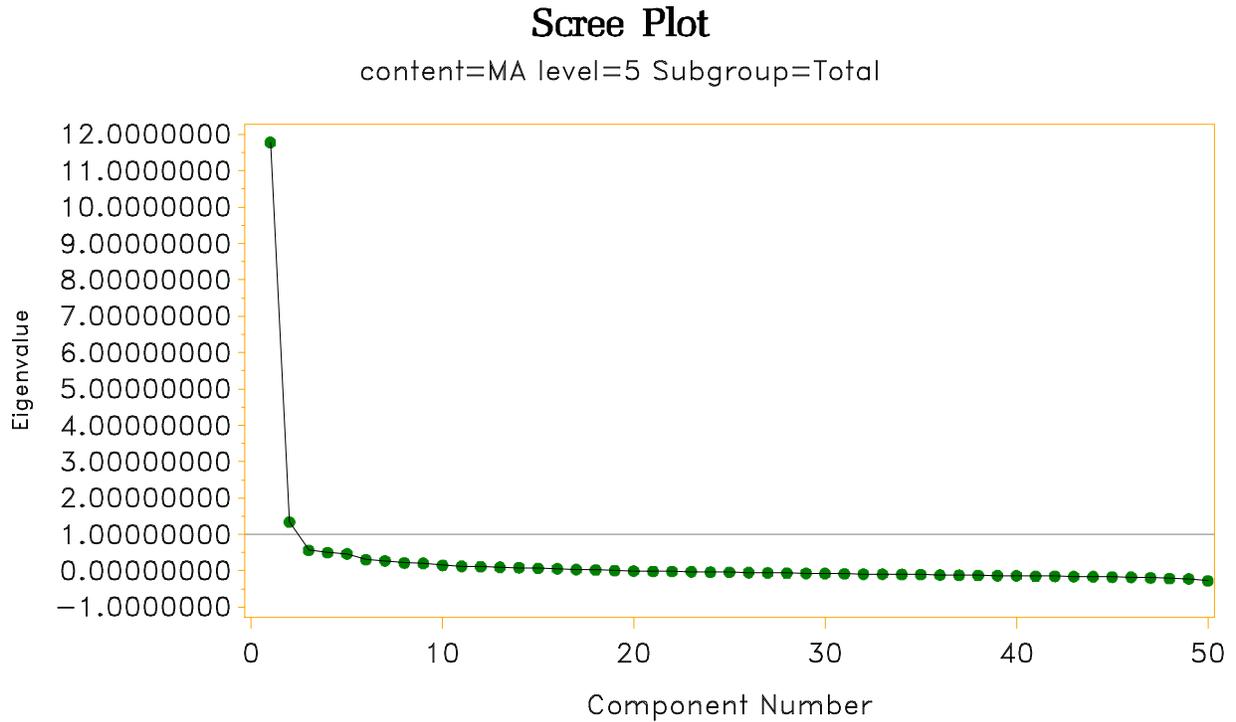


Figure 44. Spring 2014 Grade 5 Mathematics scree plot: Accommodated

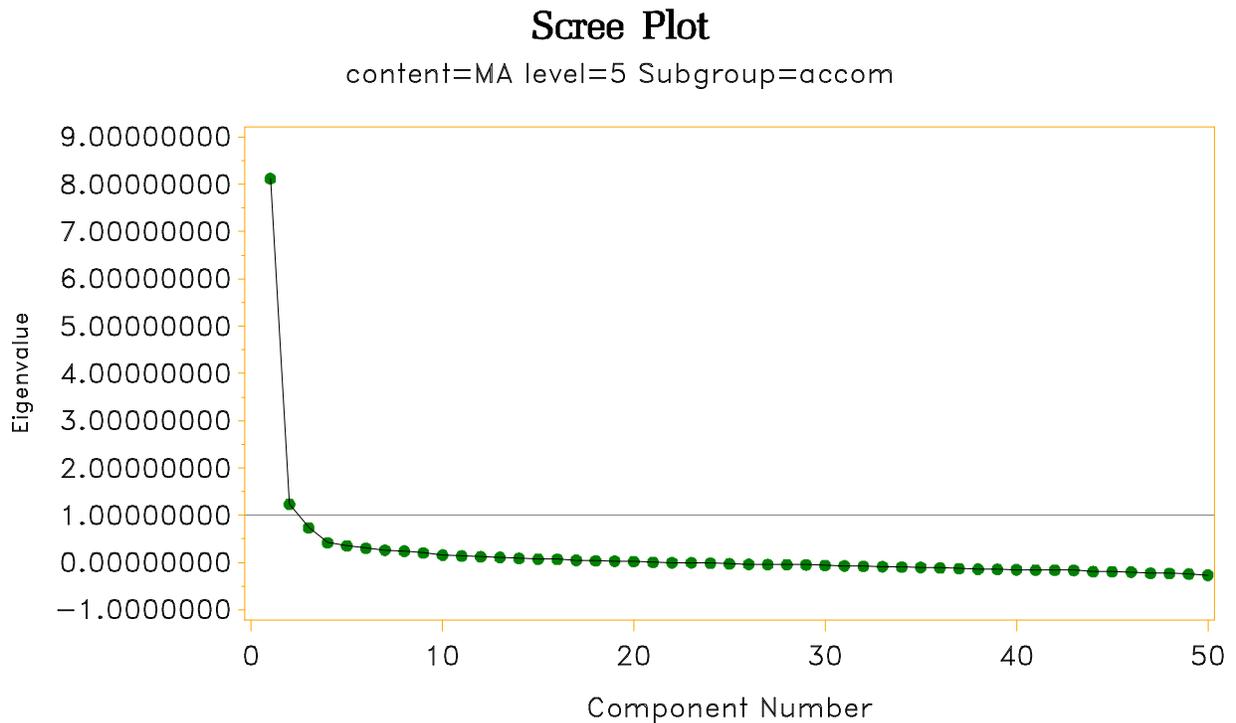


Figure 45. Spring 2014 Grade 5 Mathematics scree plot: English Language Learner

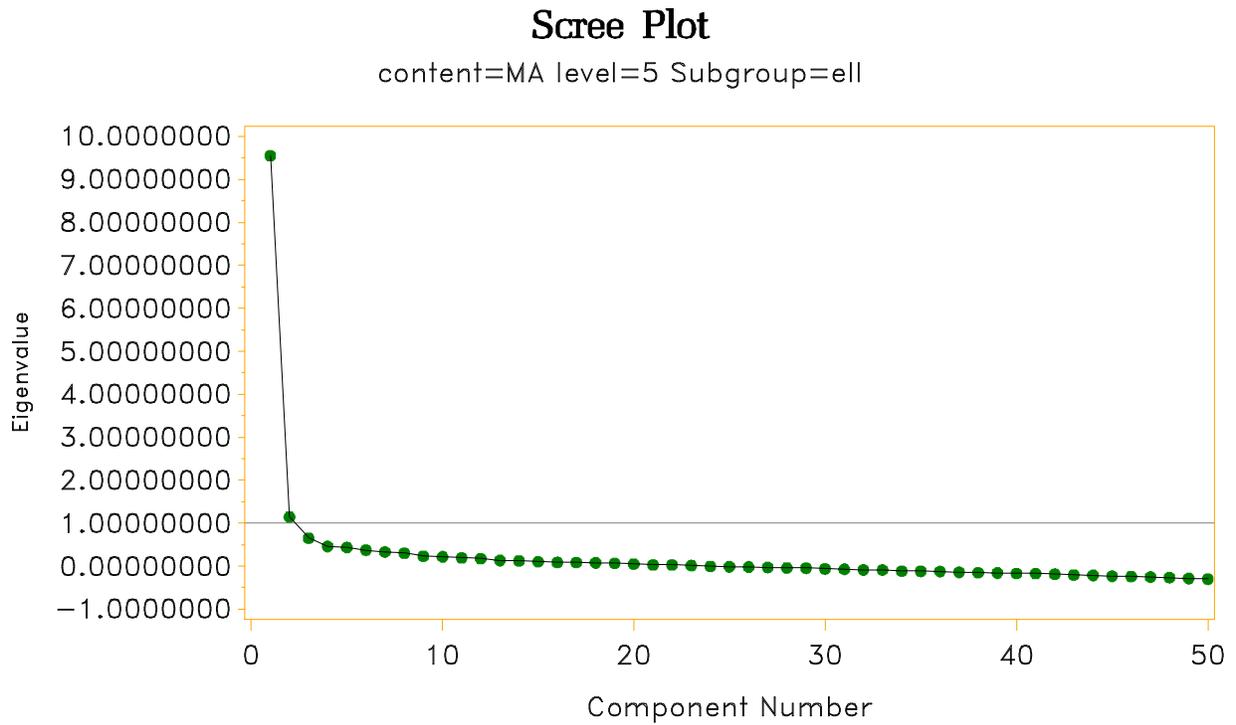


Figure 46. Spring 2014 Grade 5 Mathematics scree plot: Free Lunch

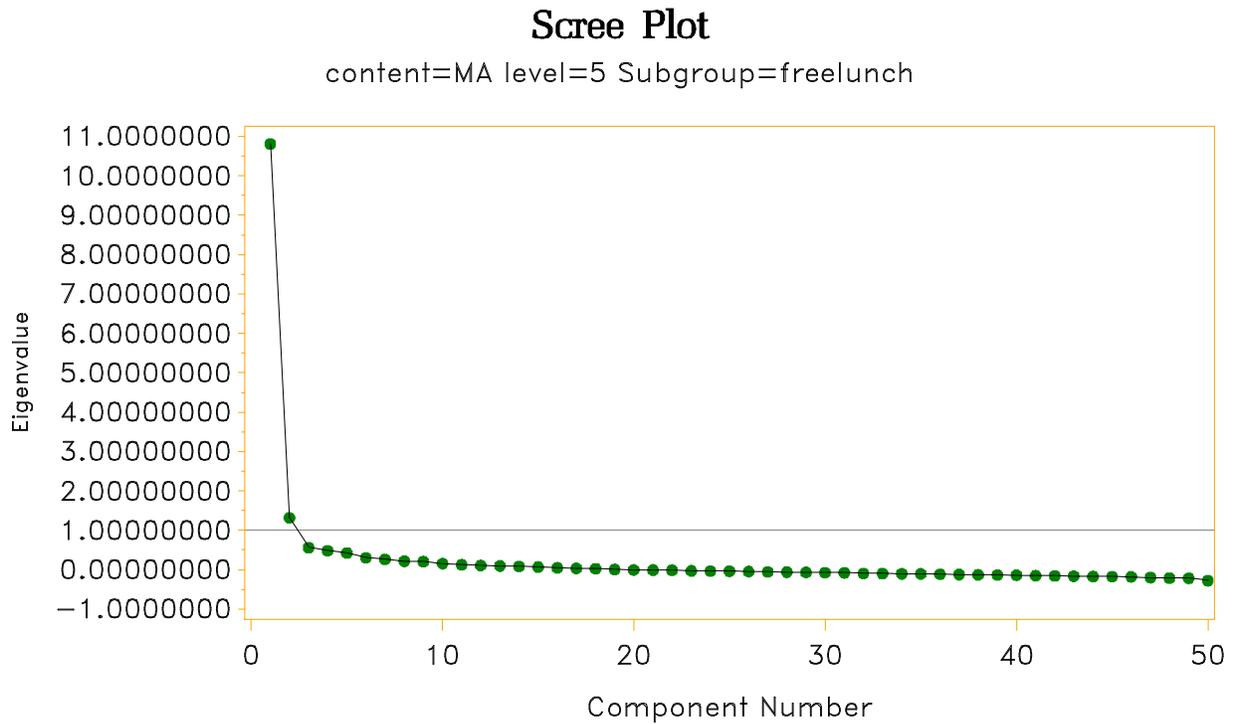


Figure 47. Spring 2014 Grade 5 Mathematics scree plot: Individualized Education Program

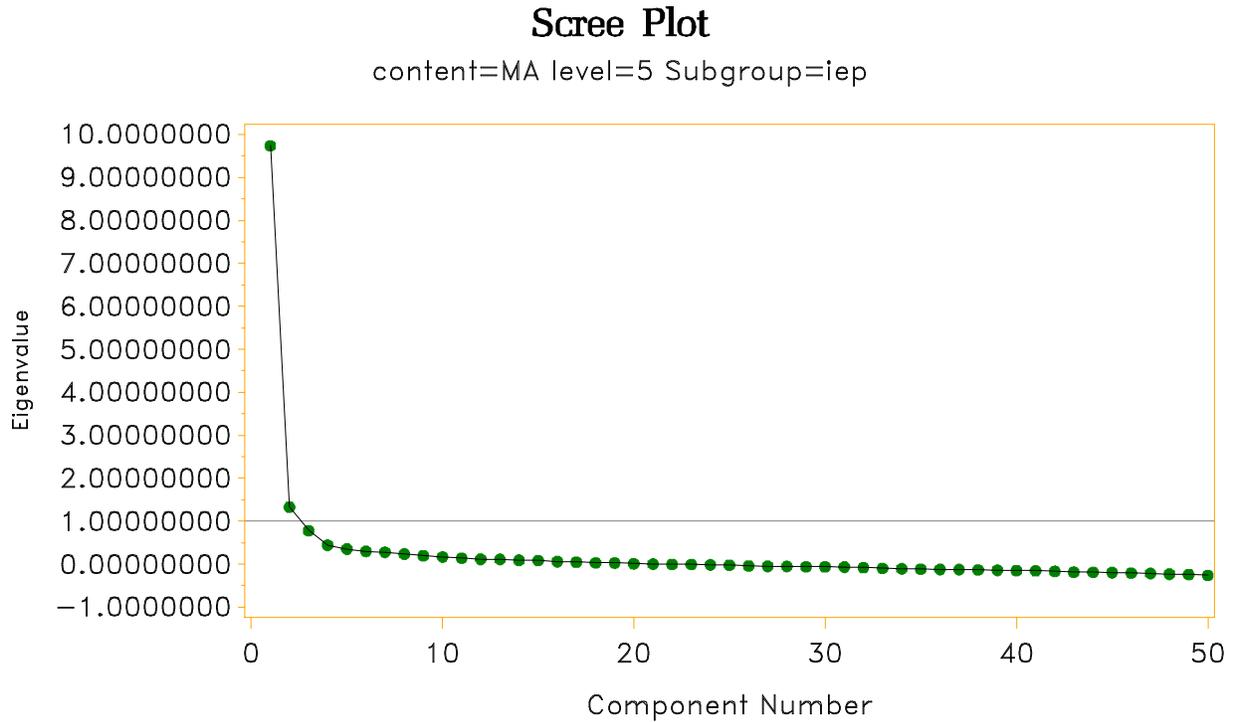


Figure 48. Spring 2014 Grade 6 Mathematics scree plot: Total

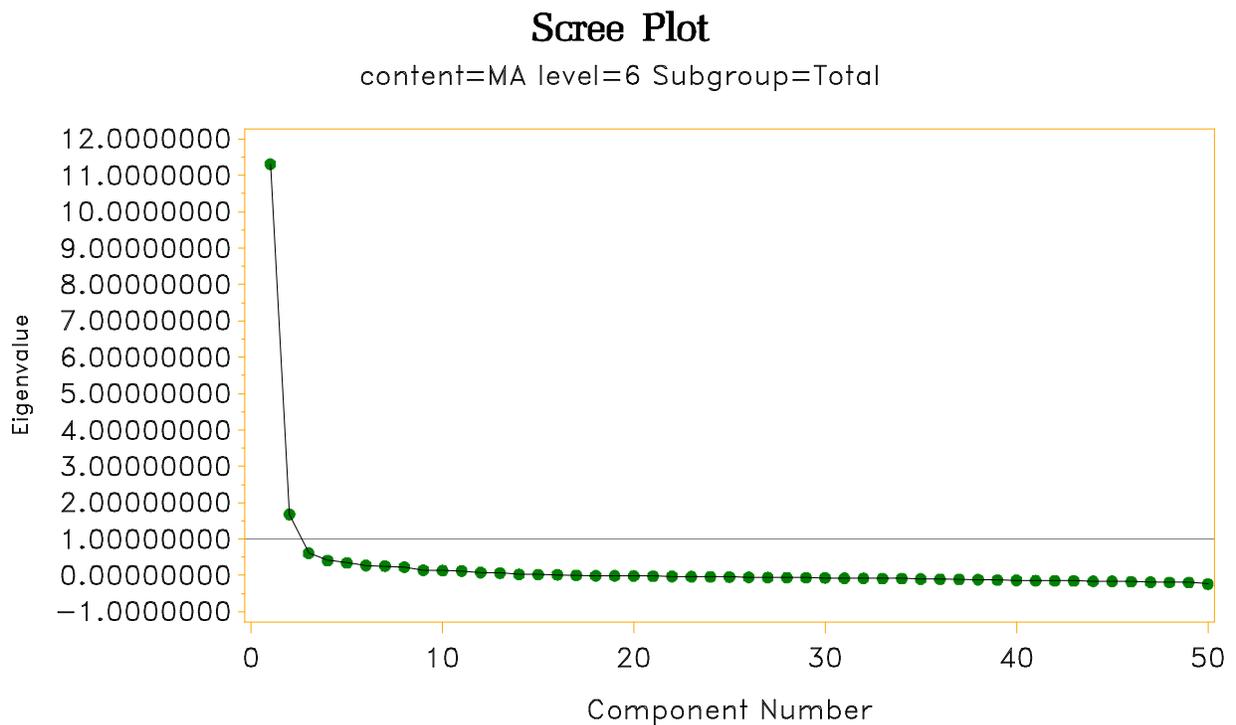


Figure 49. Spring 2014 Grade 6 Mathematics scree plot: Accommodated

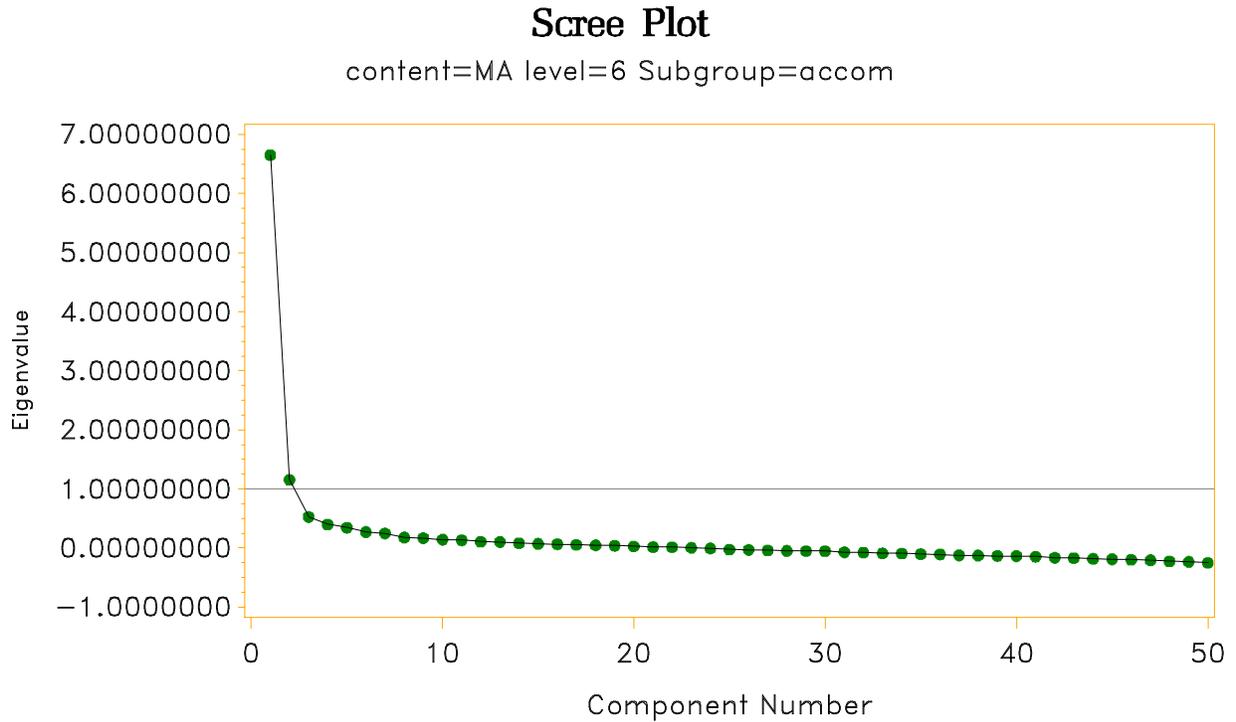


Figure 50. Spring 2014 Grade 6 Mathematics scree plot: English Language Learner

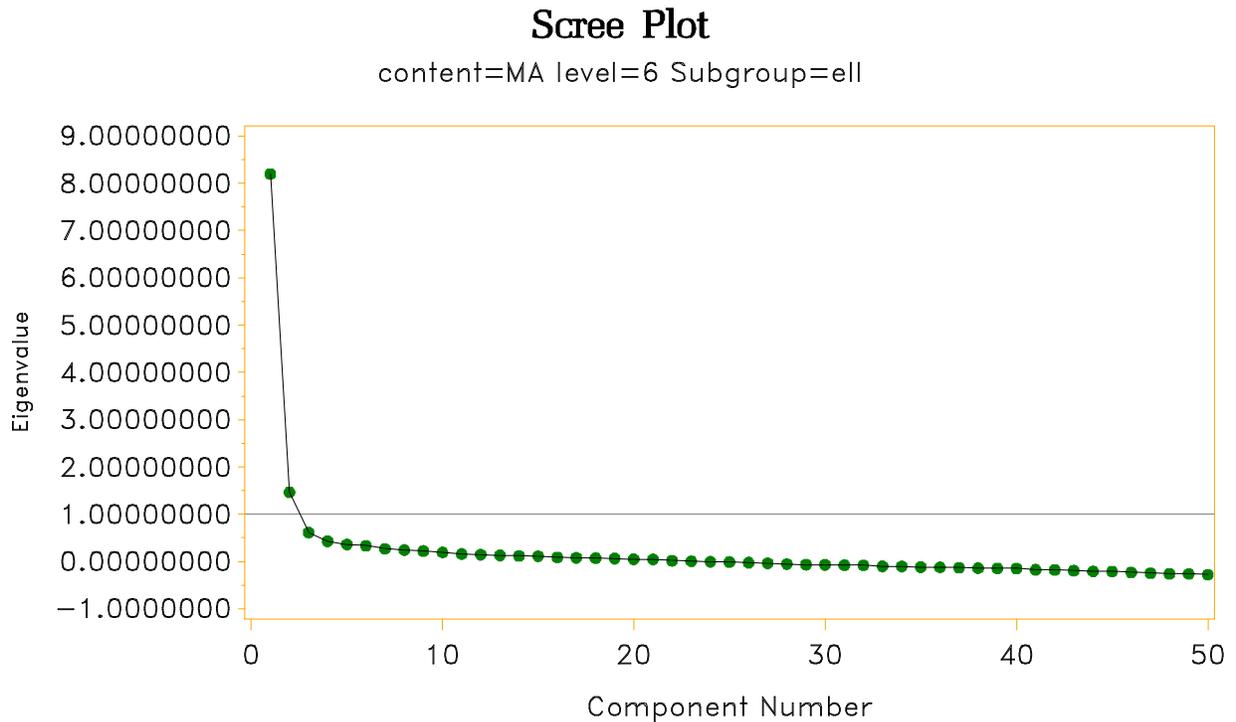


Figure 51. Spring 2014 Grade 6 Mathematics scree plot: Free Lunch

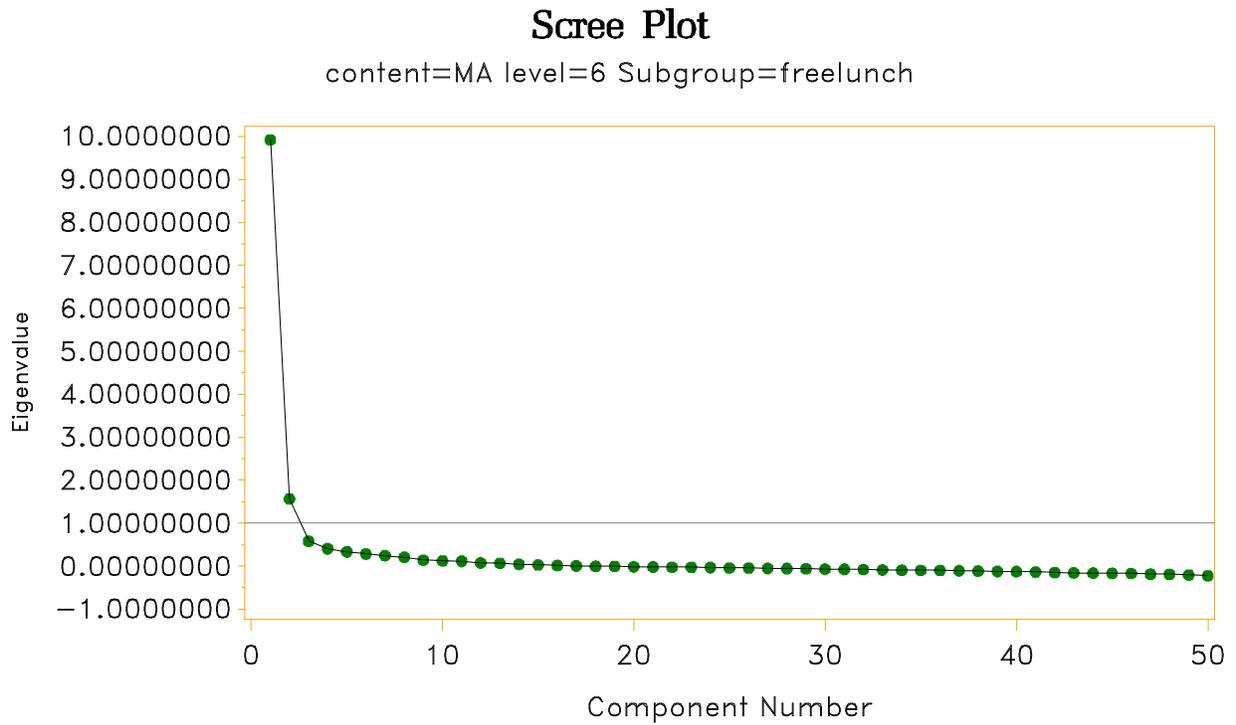


Figure 52. Spring 2014 Grade 6 Mathematics scree plot: Individualized Education Program

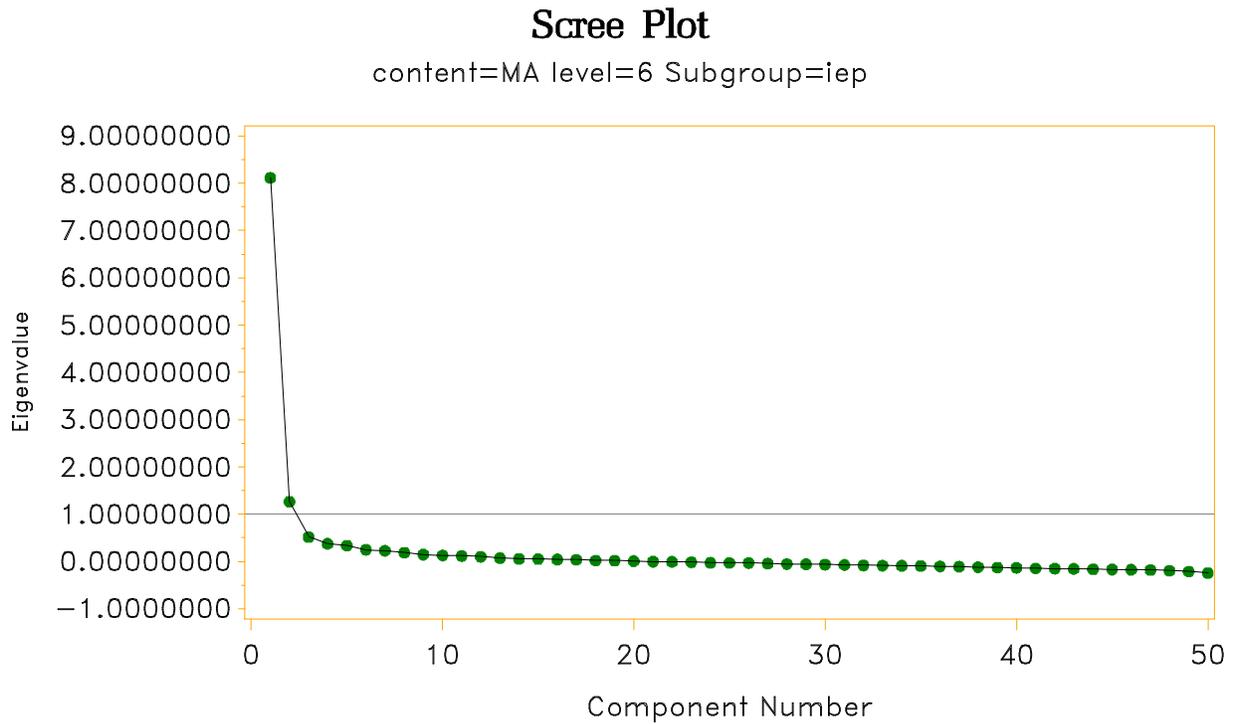


Figure 53. Spring 2014 Grade 7 Mathematics scree plot: Total

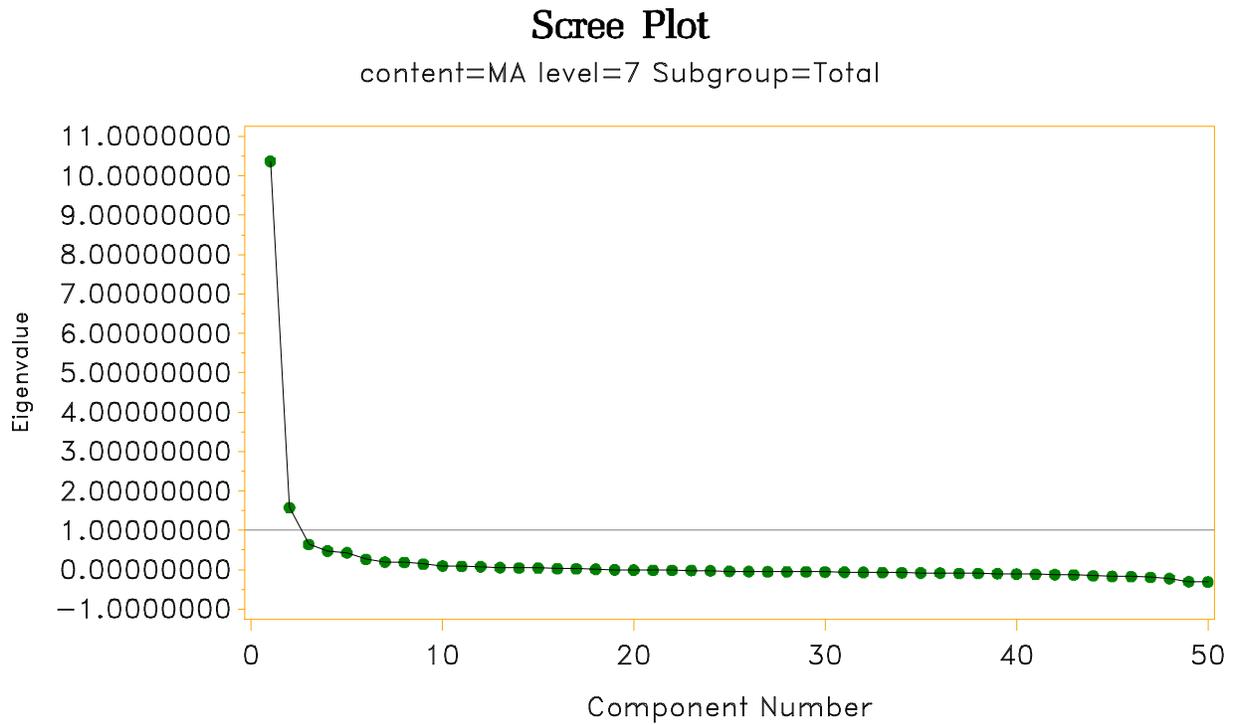


Figure 54. Spring 2014 Grade 7 Mathematics scree plot: Accommodated

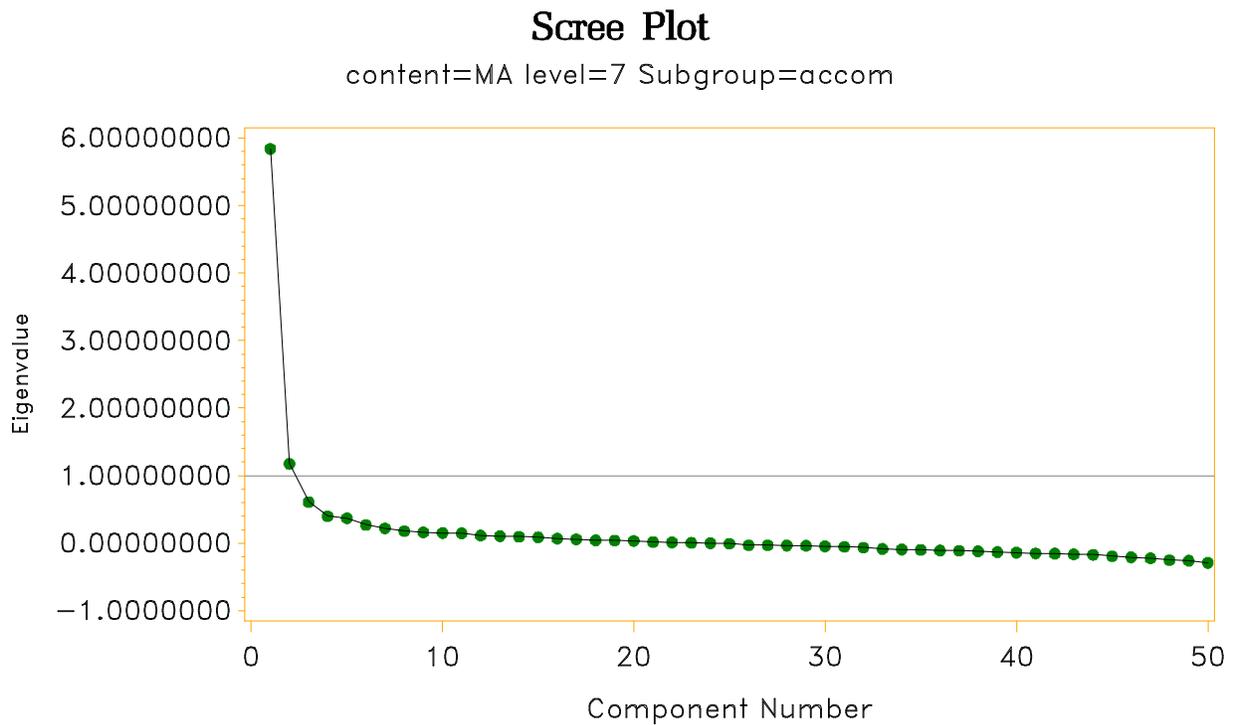


Figure 55. Spring 2014 Grade 7 Mathematics scree plot: English Language Learner

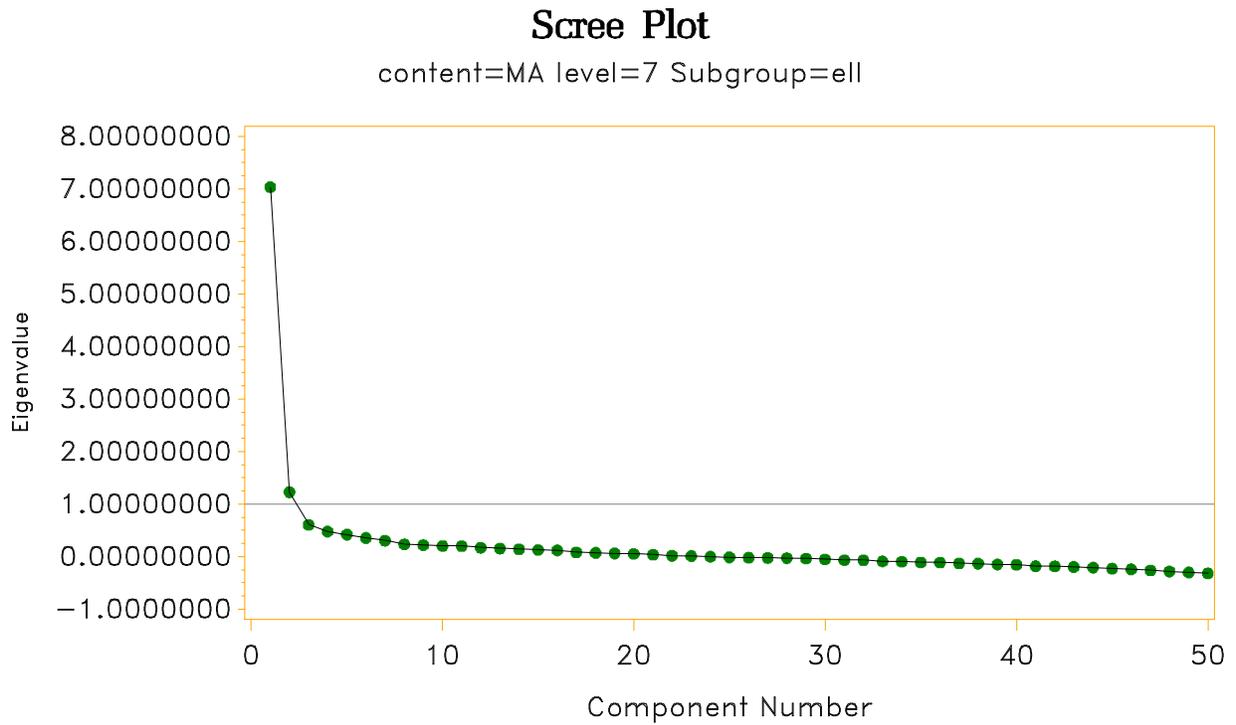


Figure 56. Spring 2014 Grade 7 Mathematics scree plot: Free Lunch

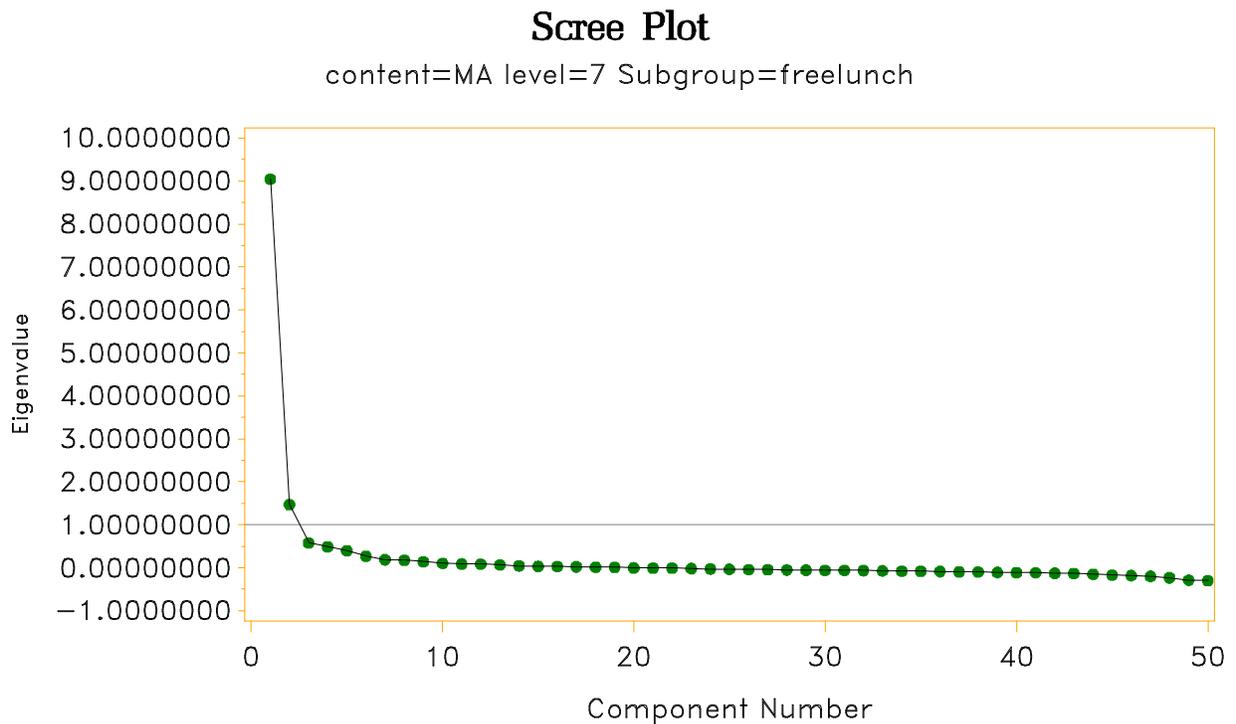


Figure 57. Spring 2014 Grade 7 Mathematics scree plot: Individualized Education Program

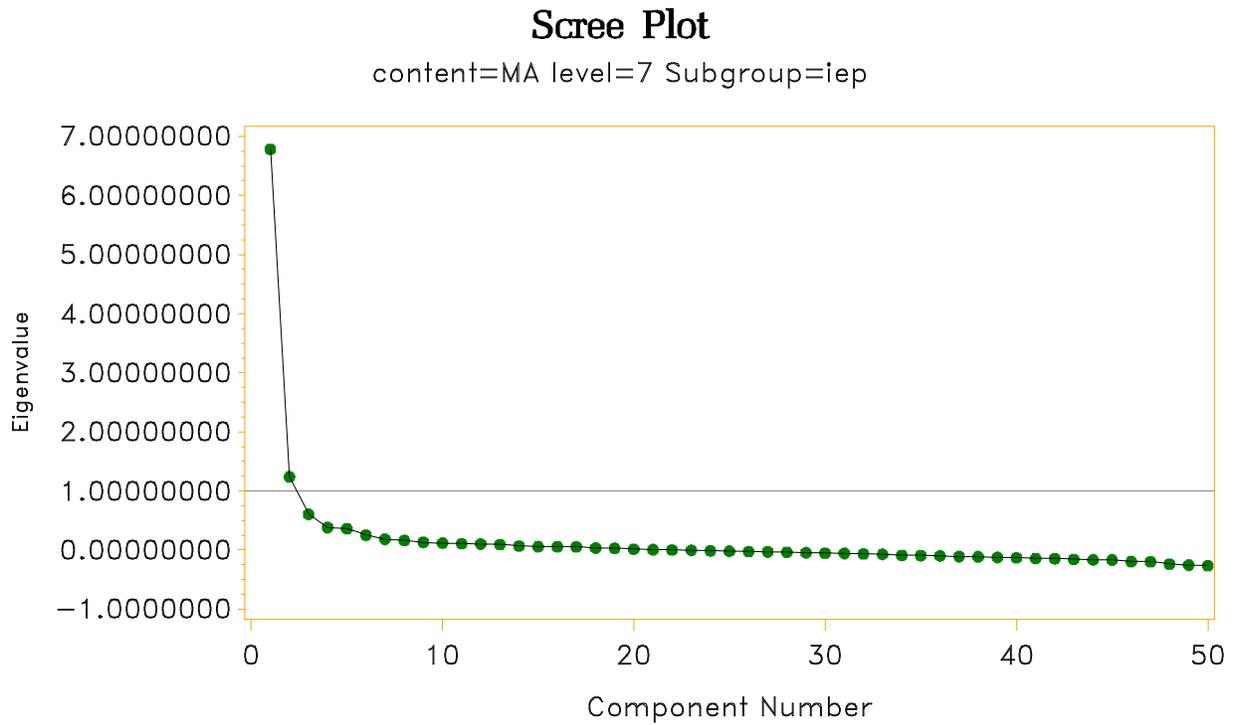


Figure 58. Spring 2014 Grade 8 Mathematics scree plot: Total

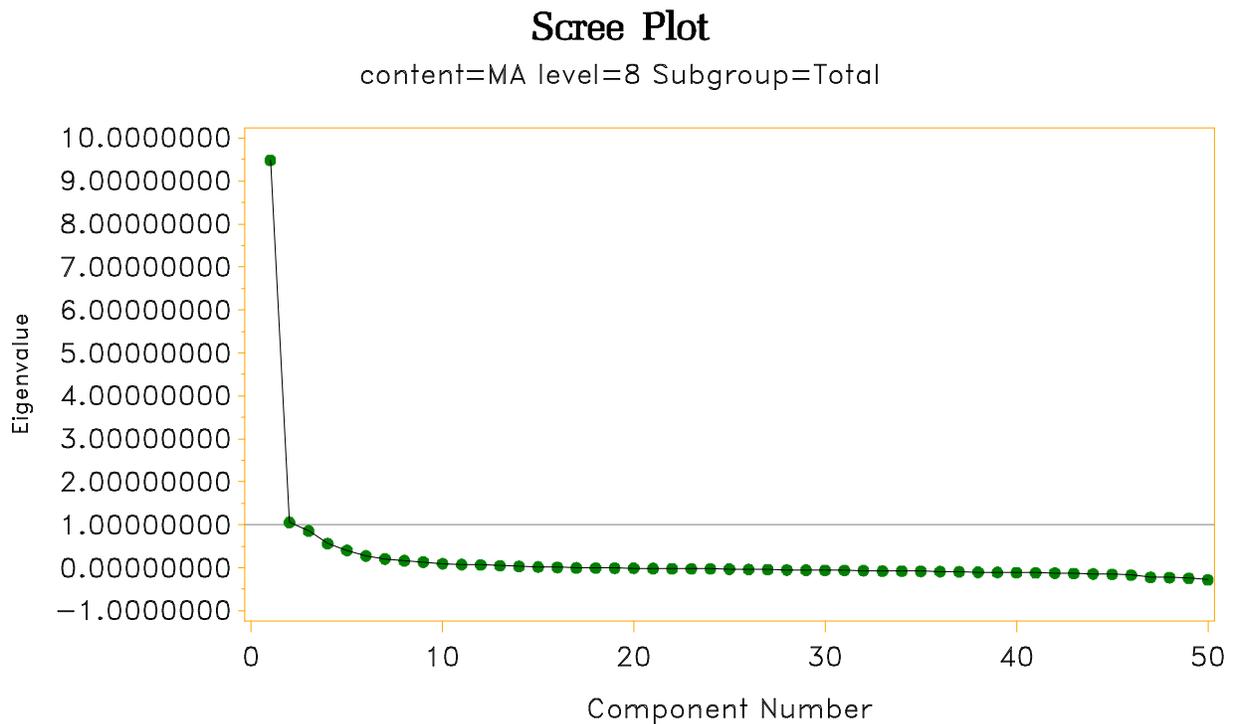


Figure 59. Spring 2014 Grade 8 Mathematics scree plot: Accommodated

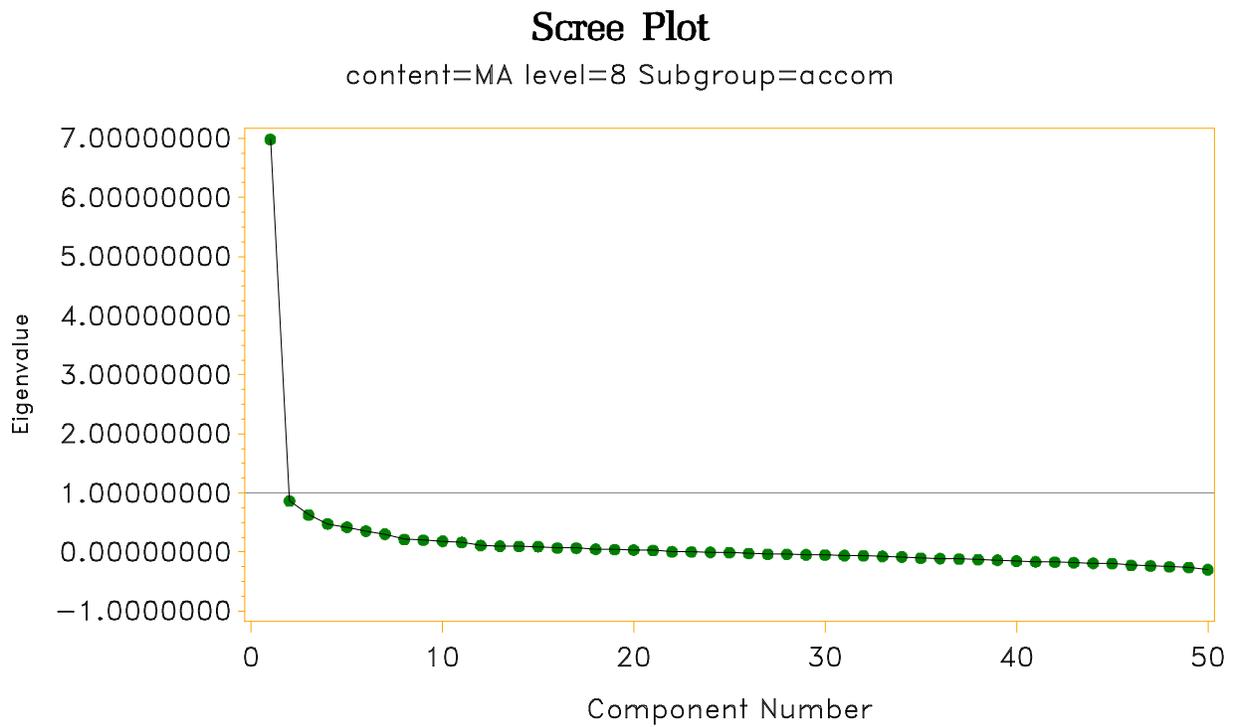


Figure 60. Spring 2014 Grade 8 Mathematics scree plot: English Language Learner

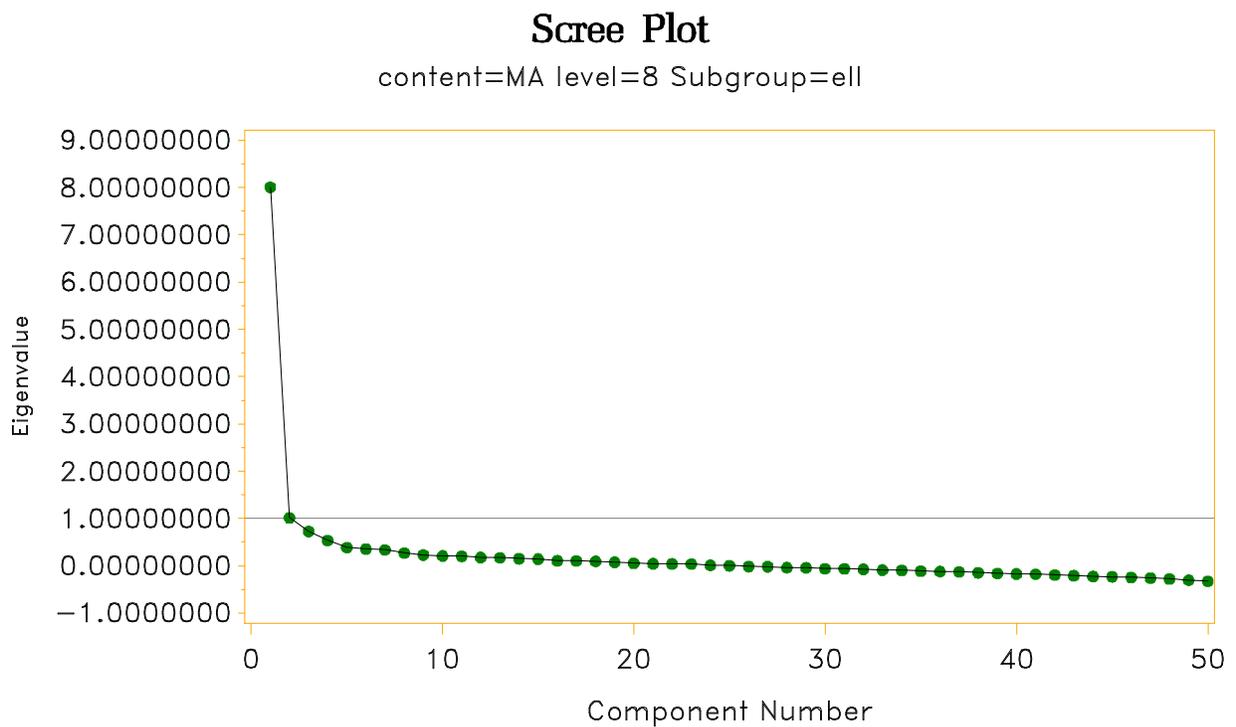


Figure 61. Spring 2014 Grade 8 Mathematics scree plot: Free Lunch

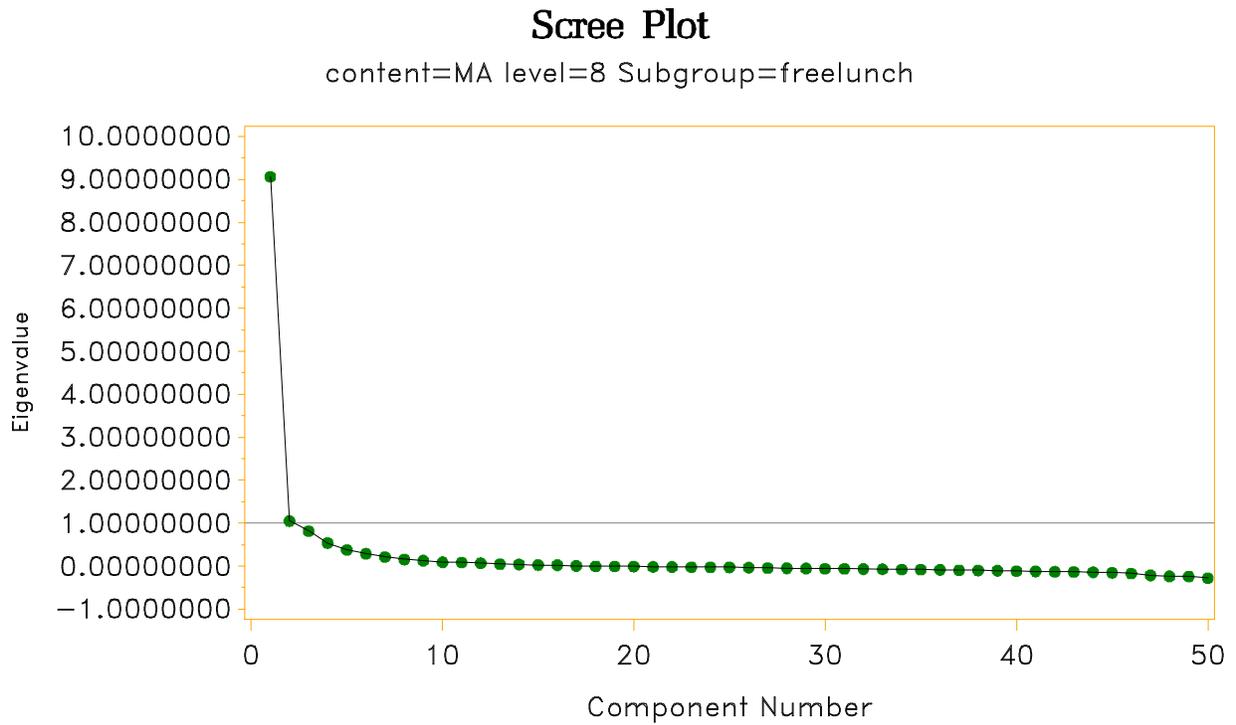


Figure 62. Spring 2014 Grade 8 Mathematics scree plot: Individualized Education Program

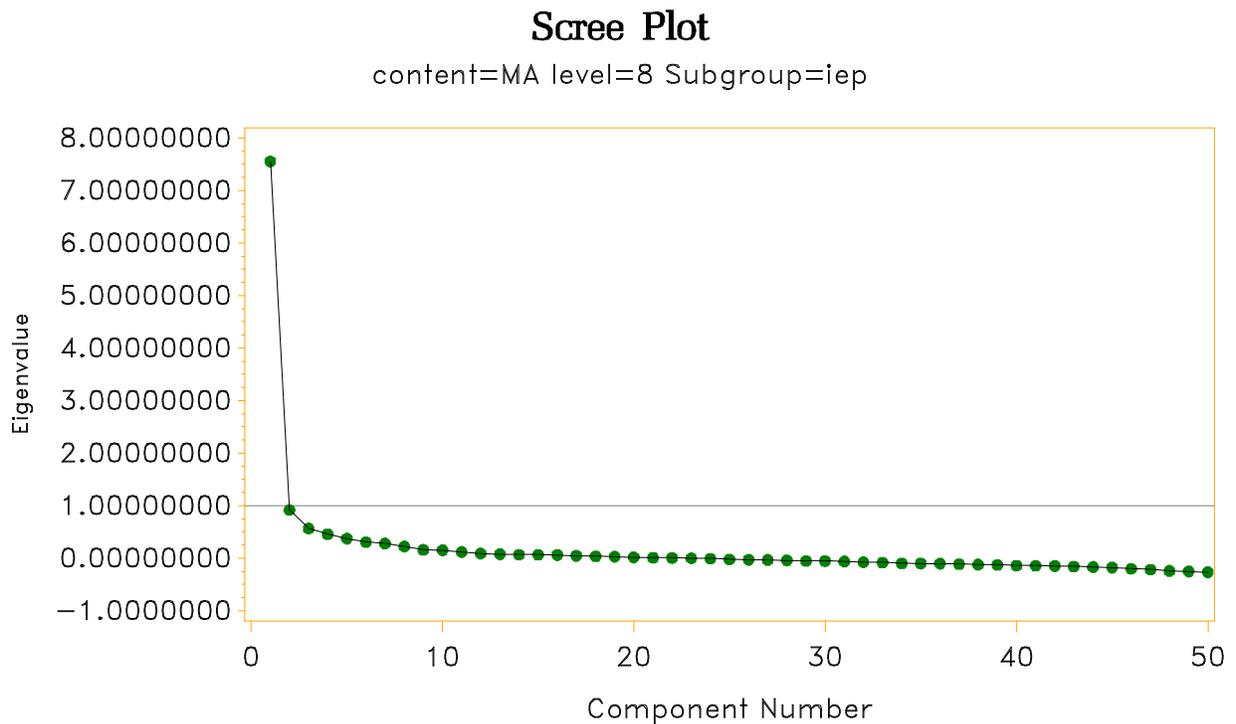


Figure 63. Spring 2014 Grade 3 Reading scree plot: Total

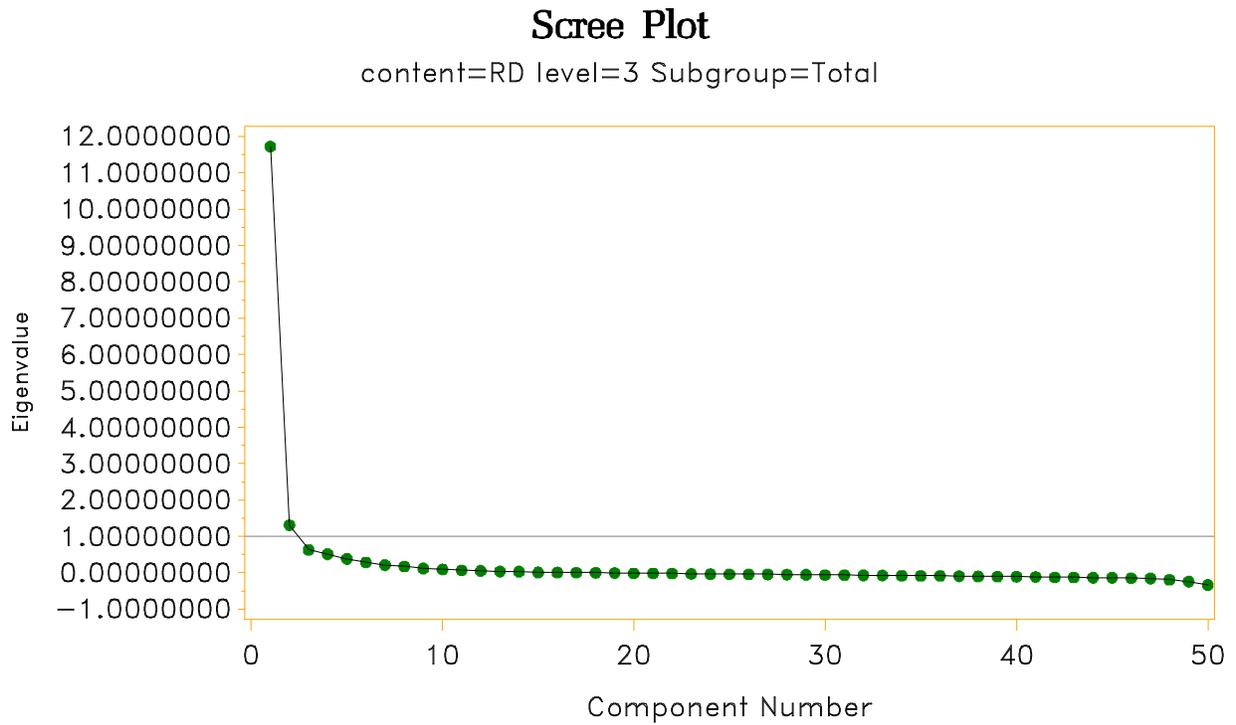


Figure 64. Spring 2014 Grade 3 Reading scree plot: Accommodated

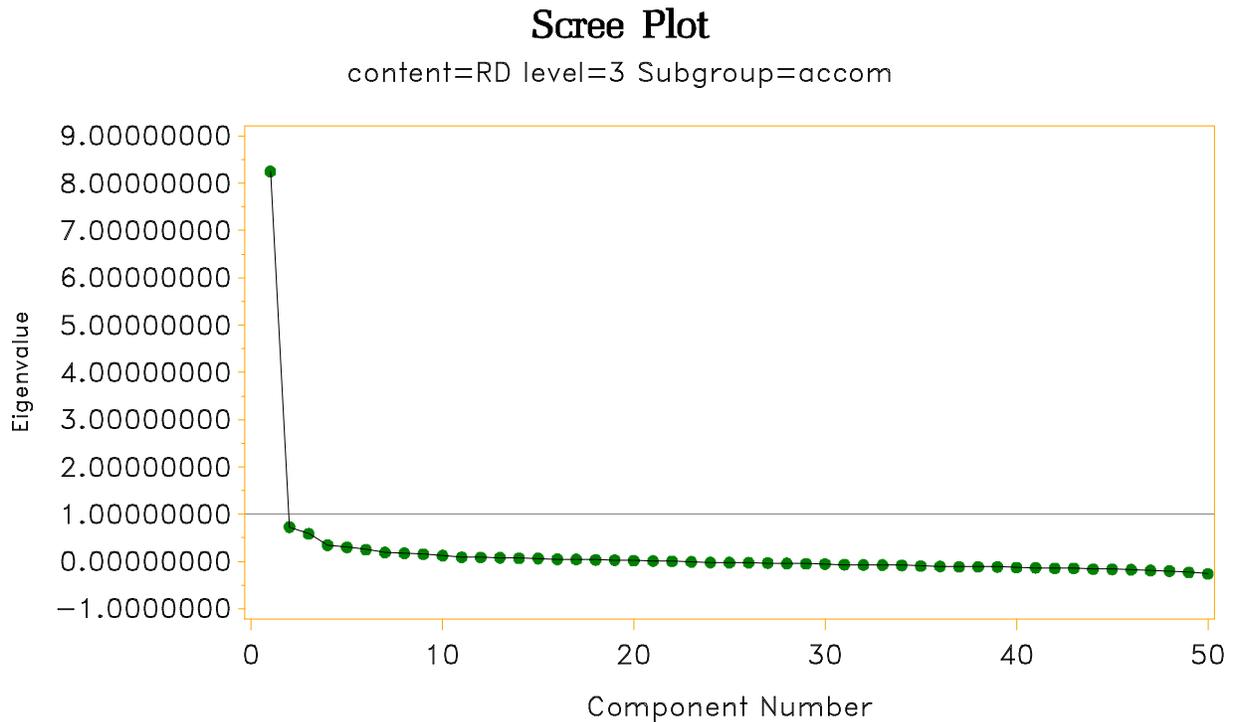


Figure 65. Spring 2014 Grade 3 Reading scree plot: English Language Learner

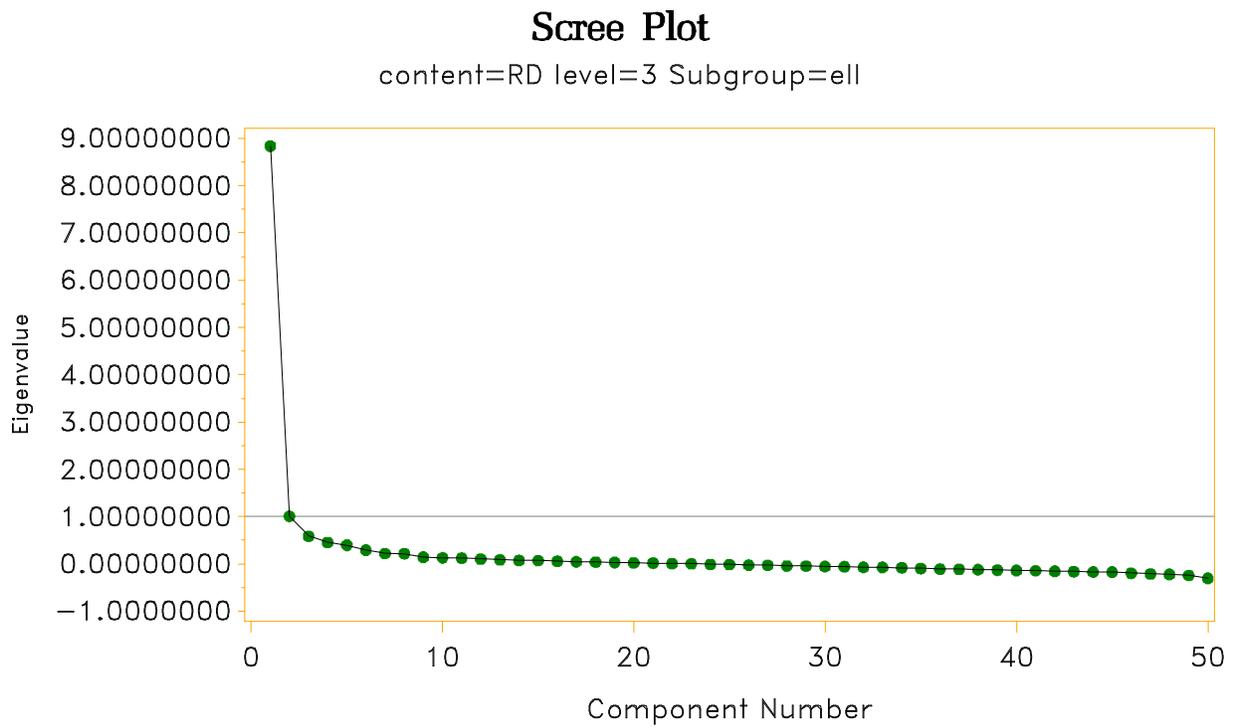


Figure 66. Spring 2014 Grade 3 Reading scree plot: Free Lunch

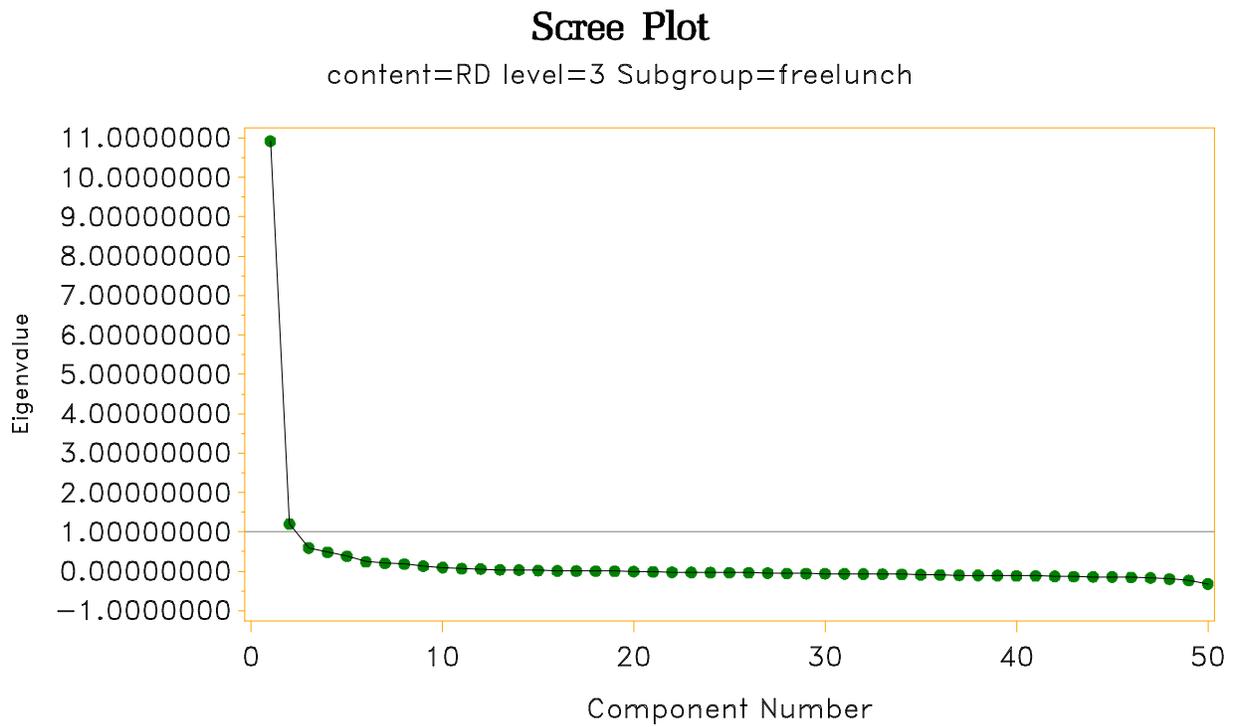


Figure 67. Spring 2014 Grade 3 Reading scree plot: Individualized Education Program

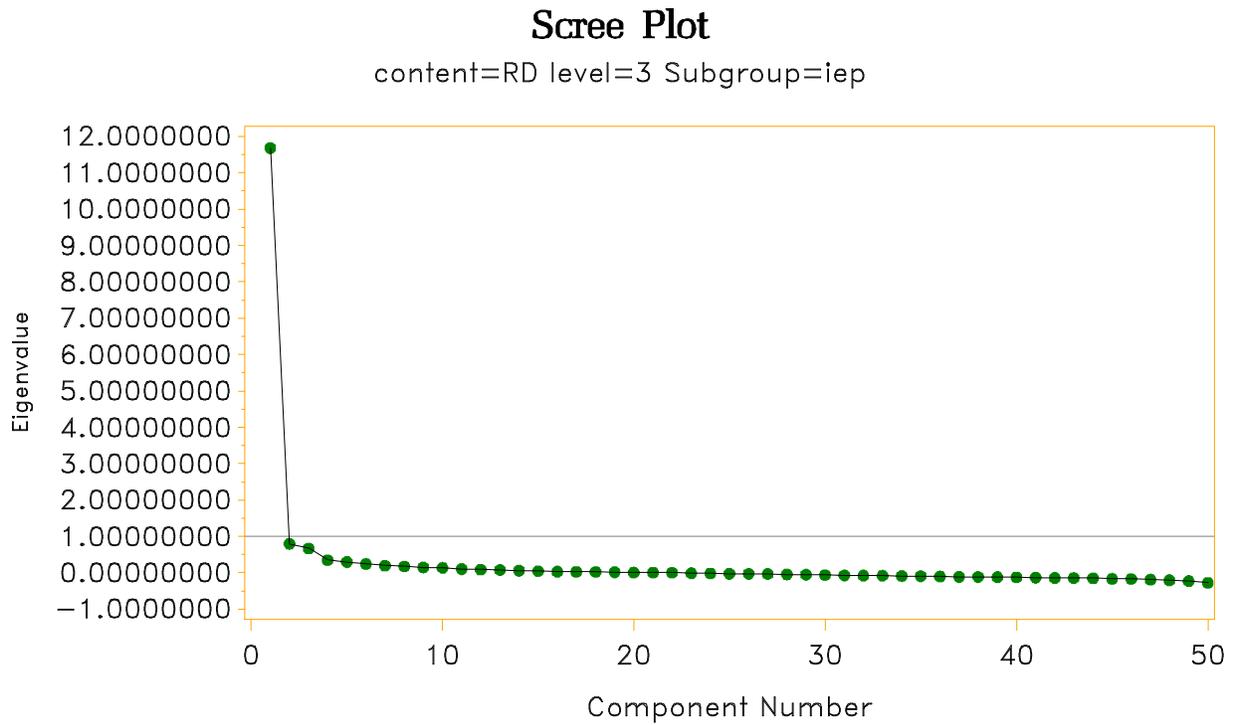


Figure 68. Spring 2014 Grade 4 Reading scree plot: Total

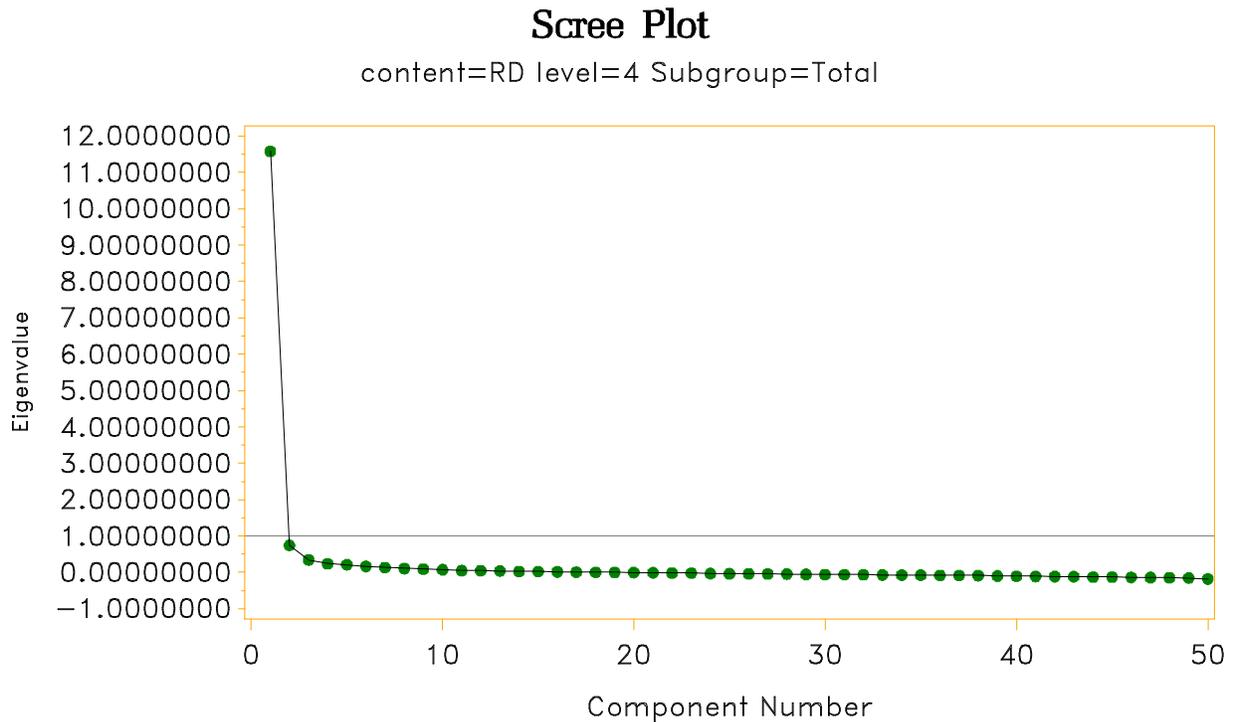


Figure 69. Spring 2014 Grade 4 Reading scree plot: Accommodated

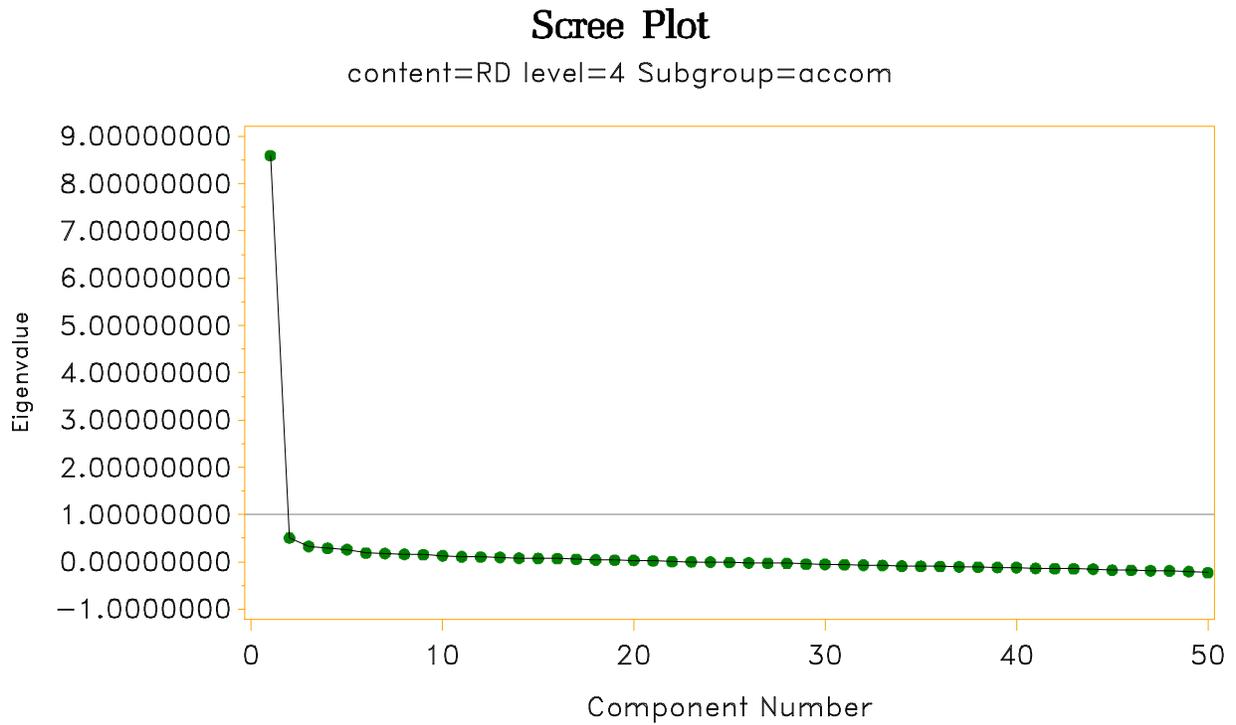


Figure 70. Spring 2014 Grade 4 Reading scree plot: English Language Learner

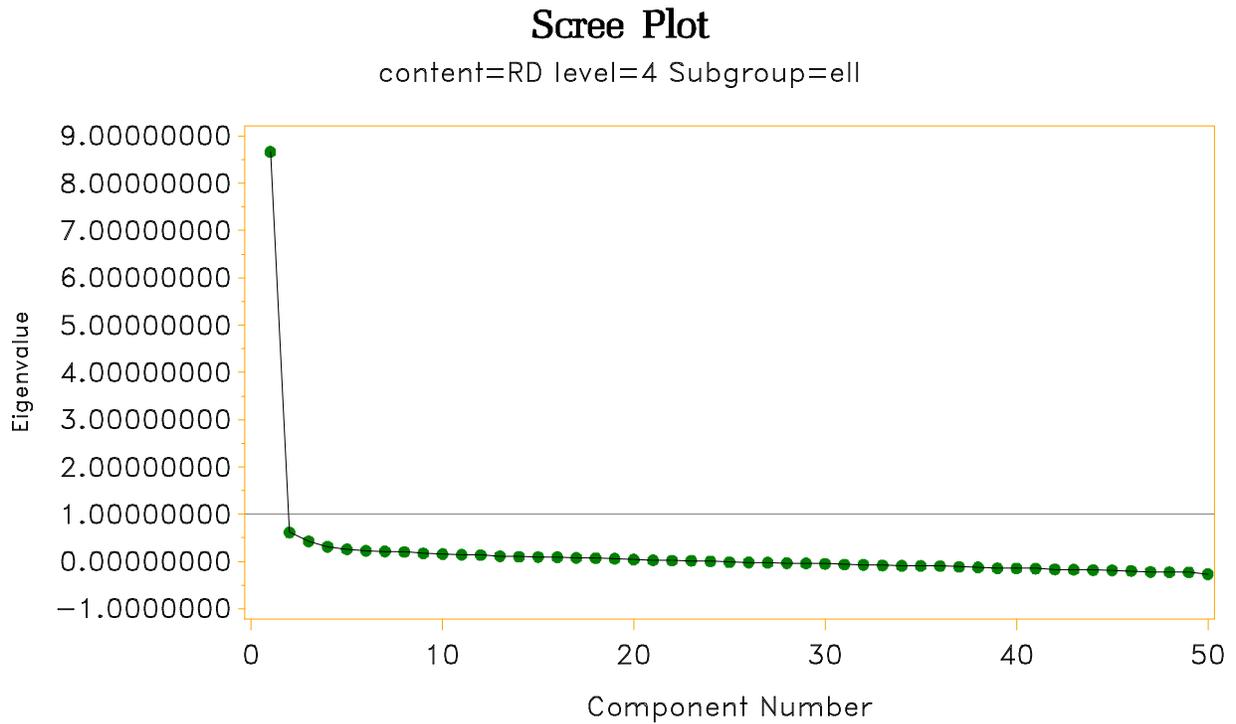


Figure 71. Spring 2014 Grade 4 Reading scree plot: Free Lunch

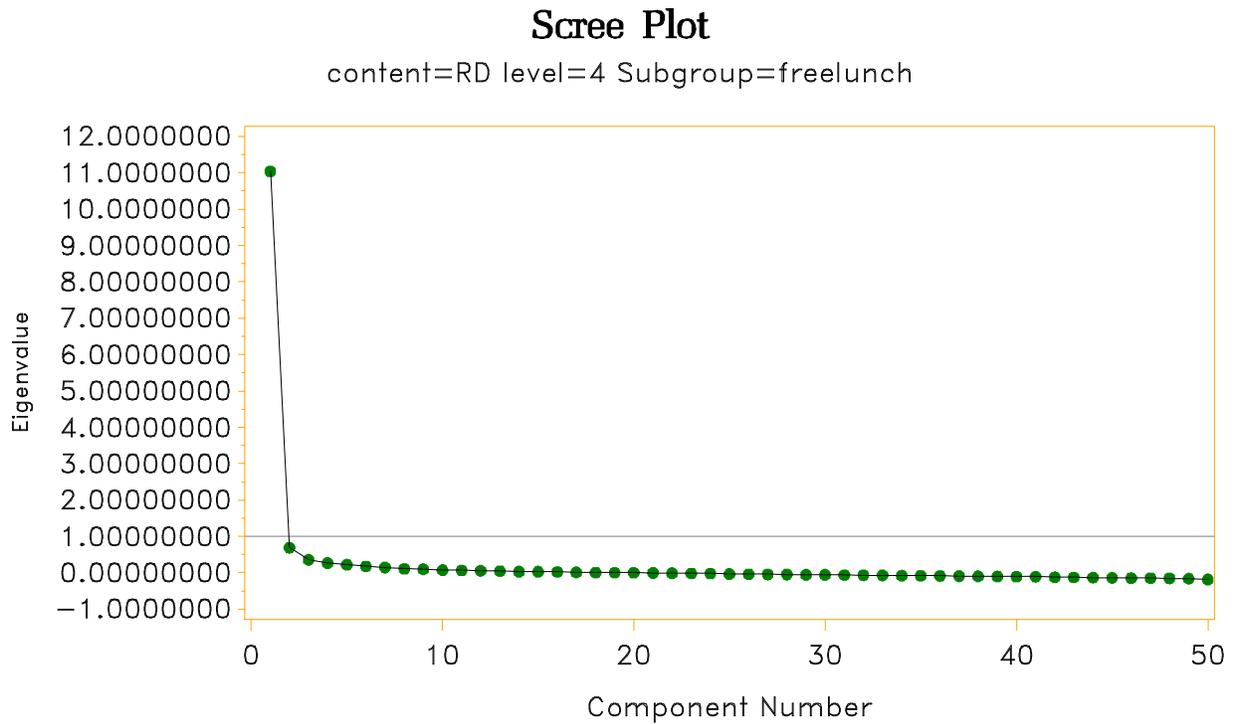


Figure 72. Spring 2014 Grade 4 Reading scree plot: Individualized Education Program

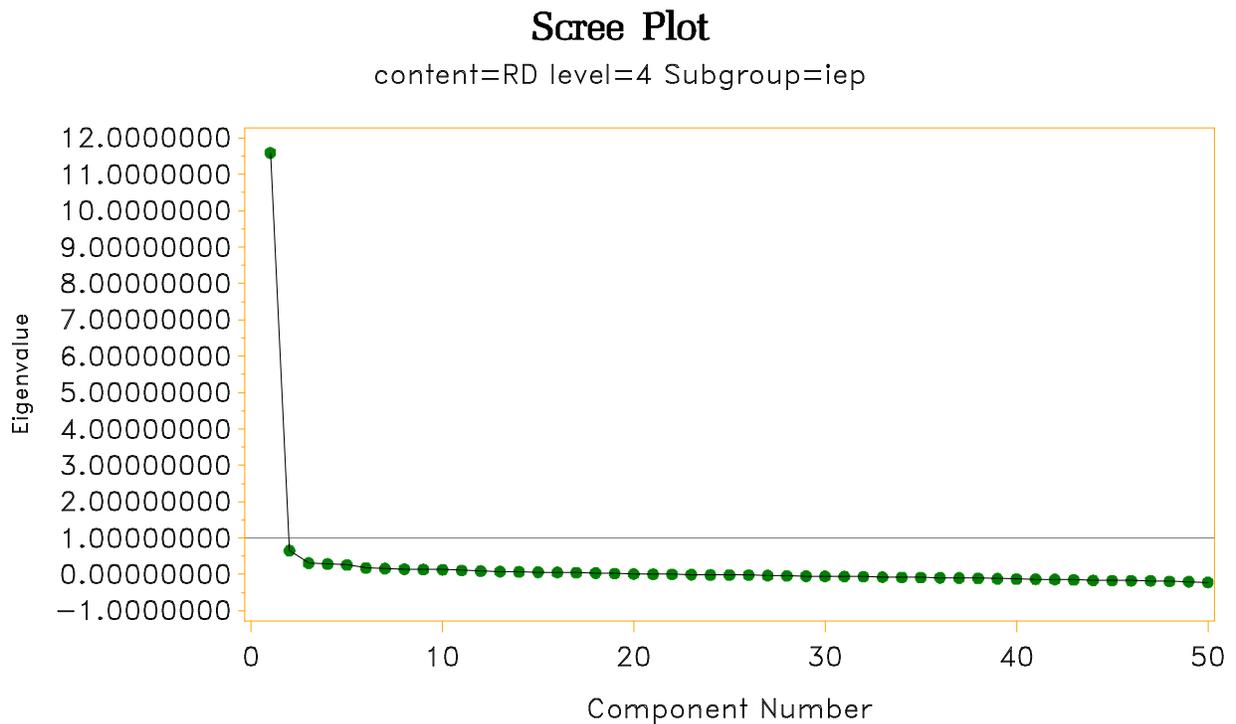


Figure 73. Spring 2014 Grade 5 Reading scree plot: Total

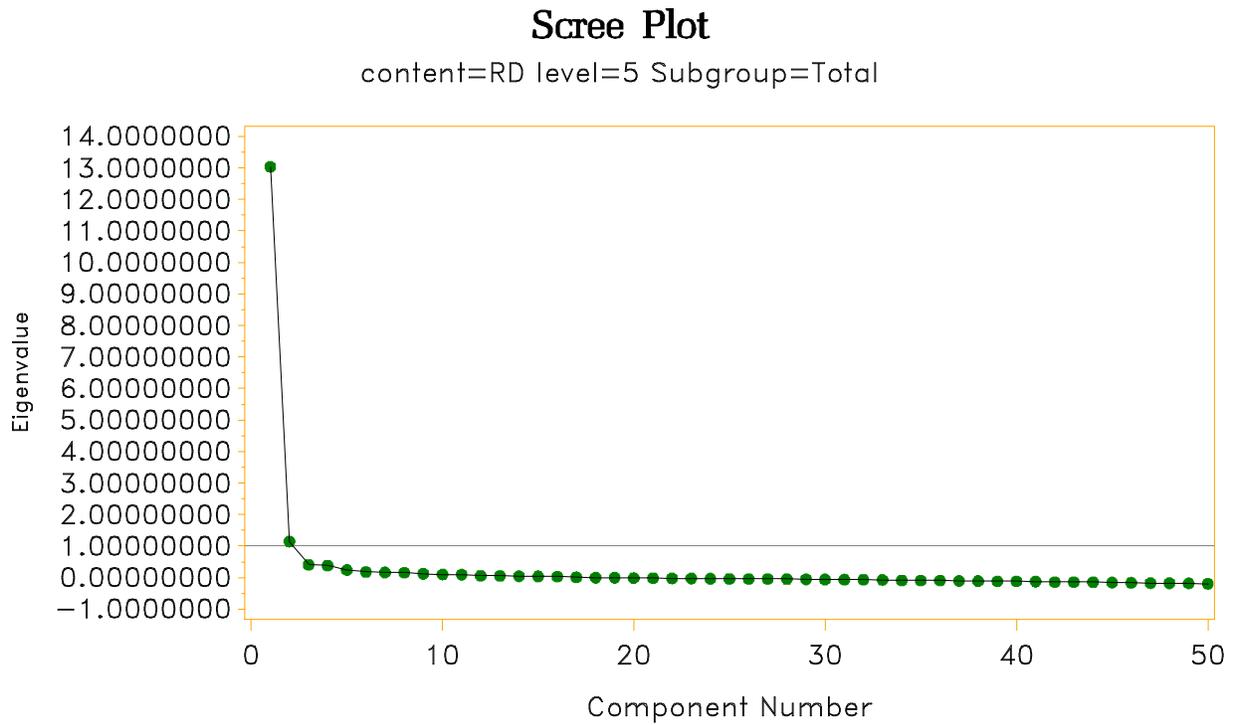


Figure 74. Spring 2014 Grade 5 Reading scree plot: Accommodated

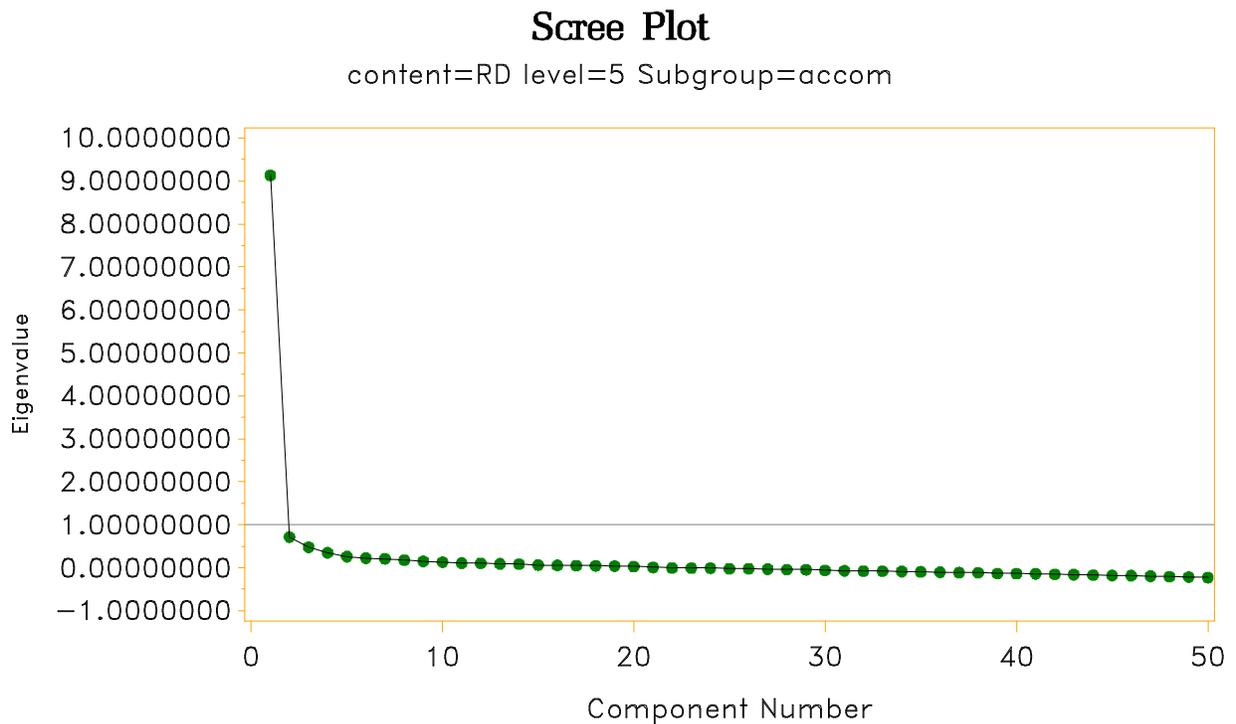


Figure 75. Spring 2014 Grade 5 Reading scree plot: English Language Learner

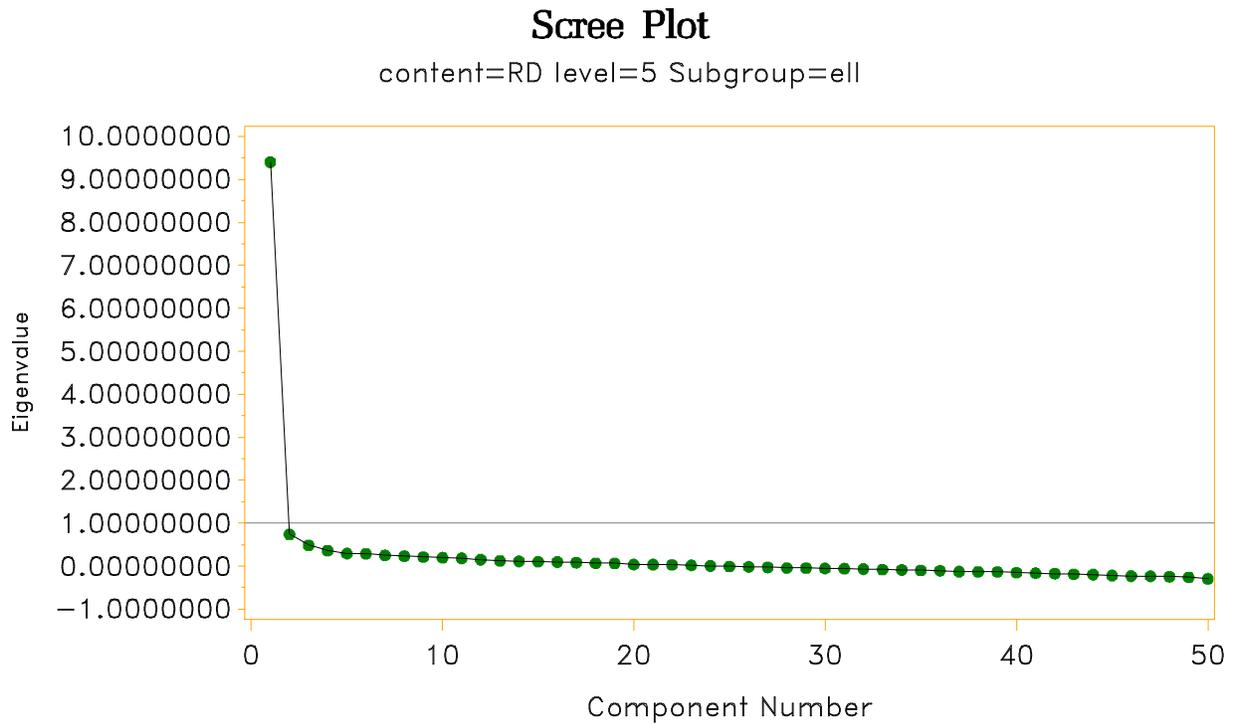


Figure 76. Spring 2014 Grade 5 Reading scree plot: Free Lunch

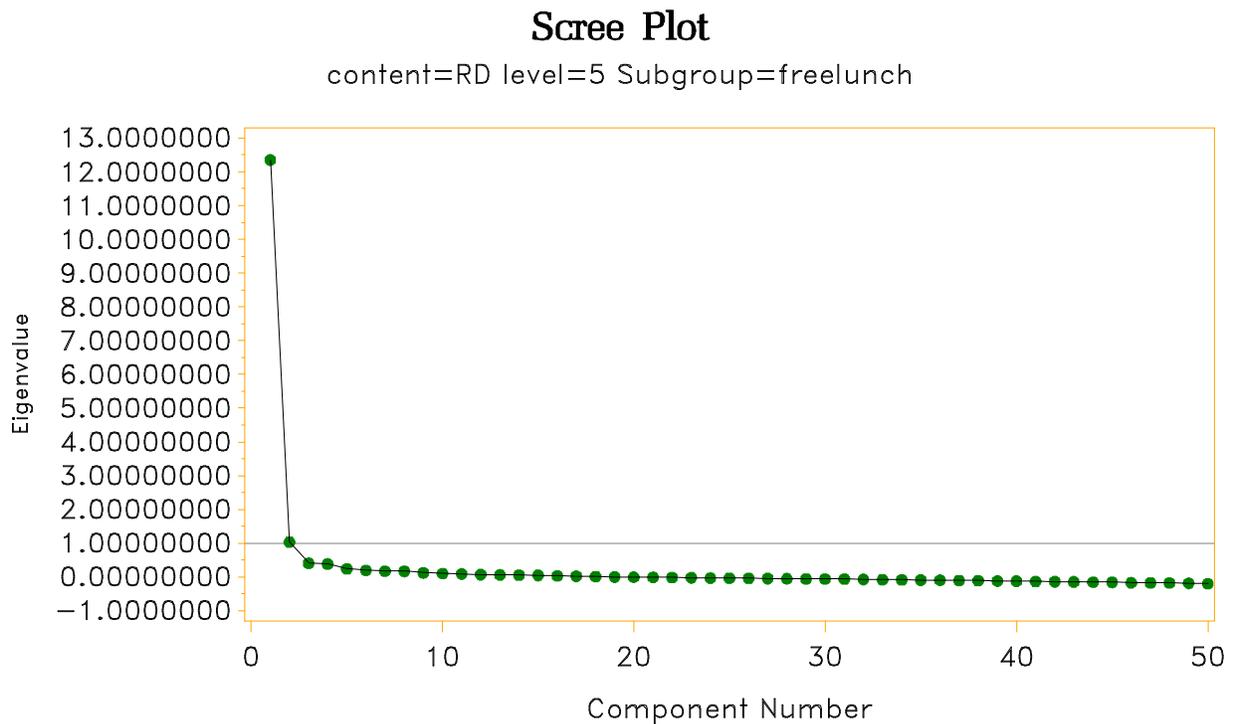


Figure 77. Spring 2014 Grade 5 Reading scree plot: Individualized Education Program

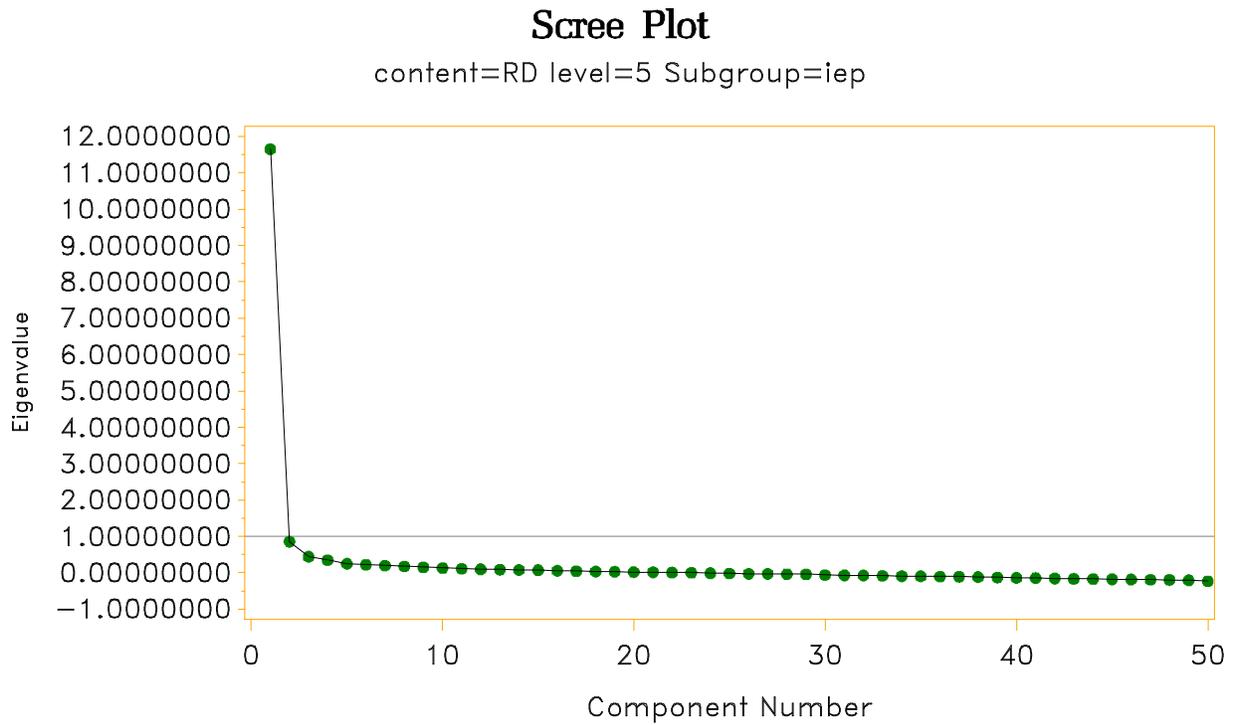


Figure 78. Spring 2014 Grade 6 Reading scree plot: Total

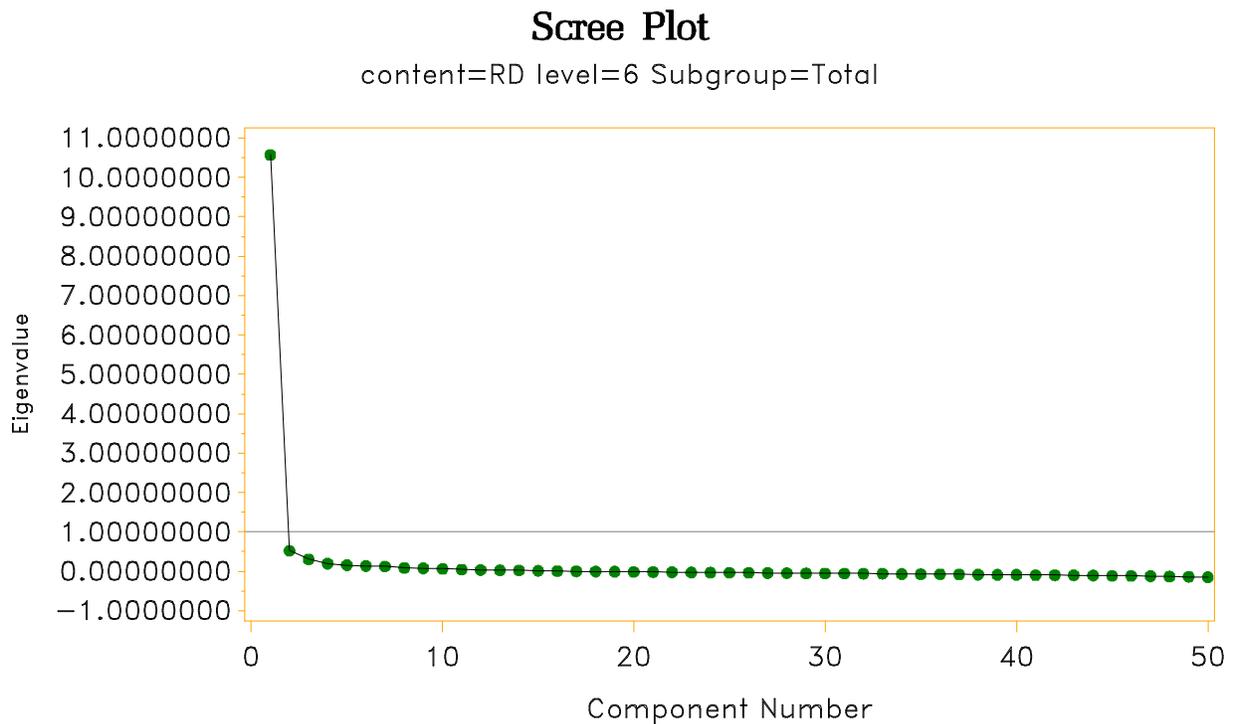


Figure 79. Spring 2014 Grade 6 Reading scree plot: Accommodated

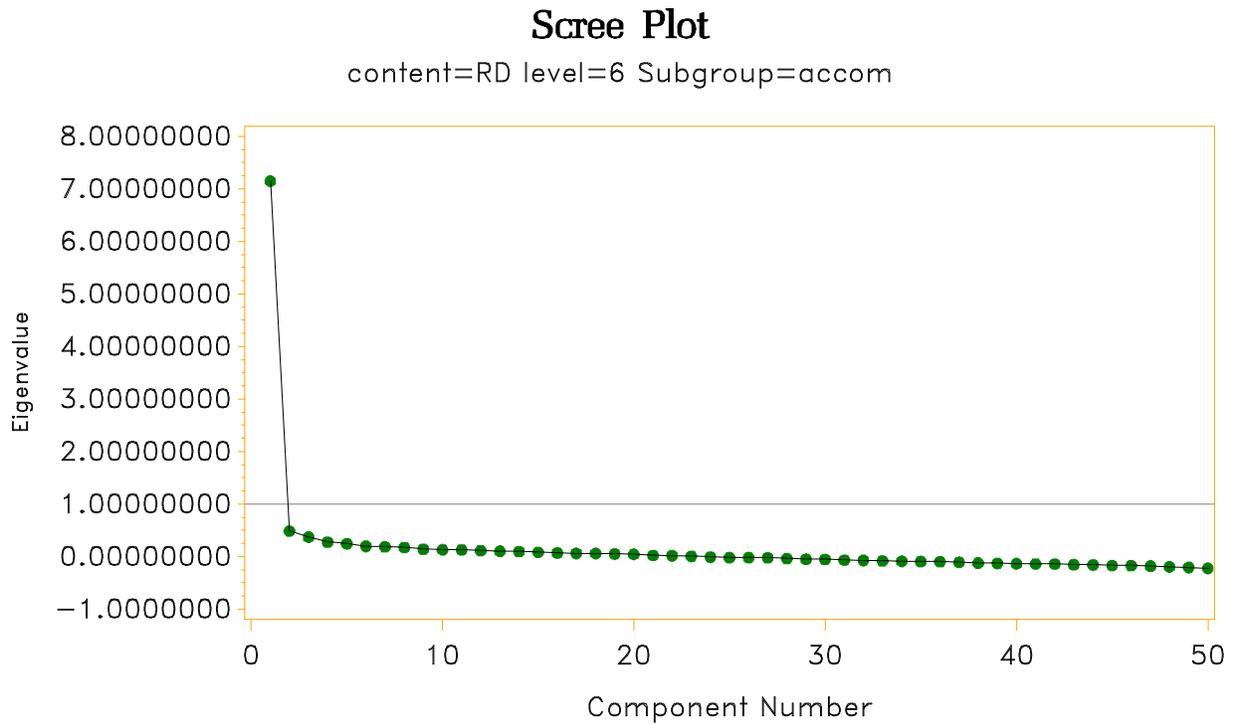


Figure 80. Spring 2014 Grade 6 Reading scree plot: English Language Learner

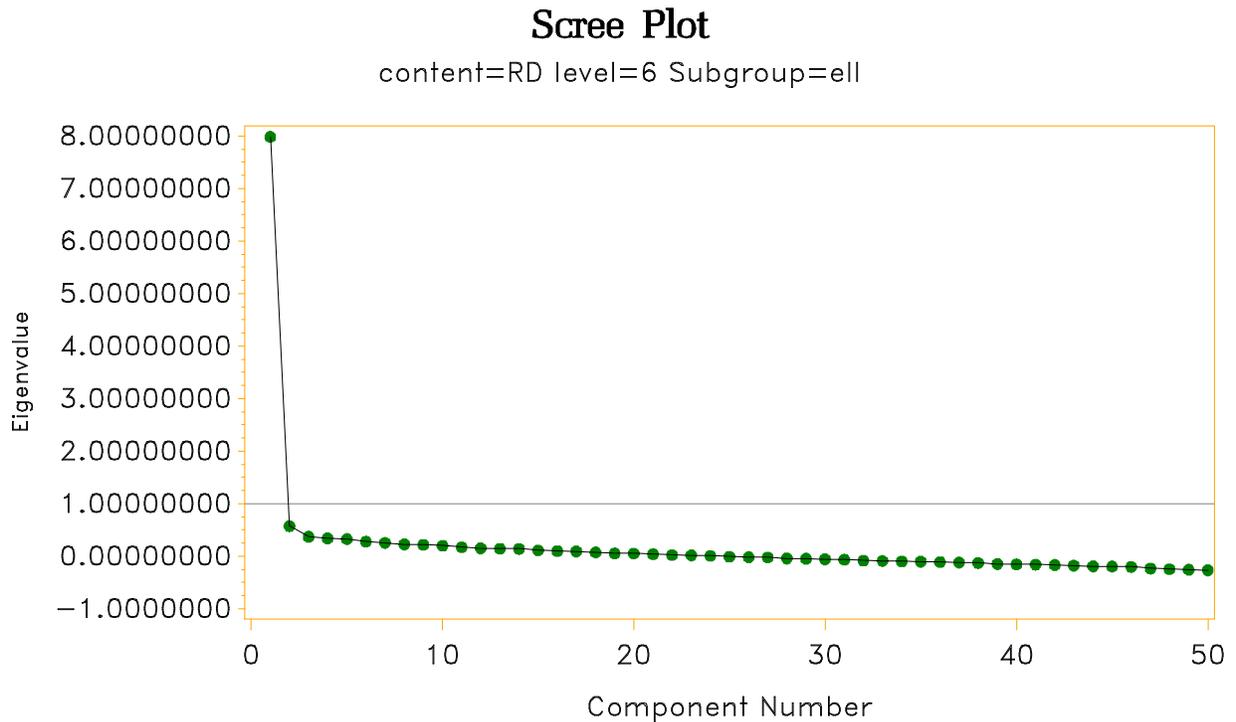


Figure 81. Spring 2014 Grade 6 Reading scree plot: Free Lunch

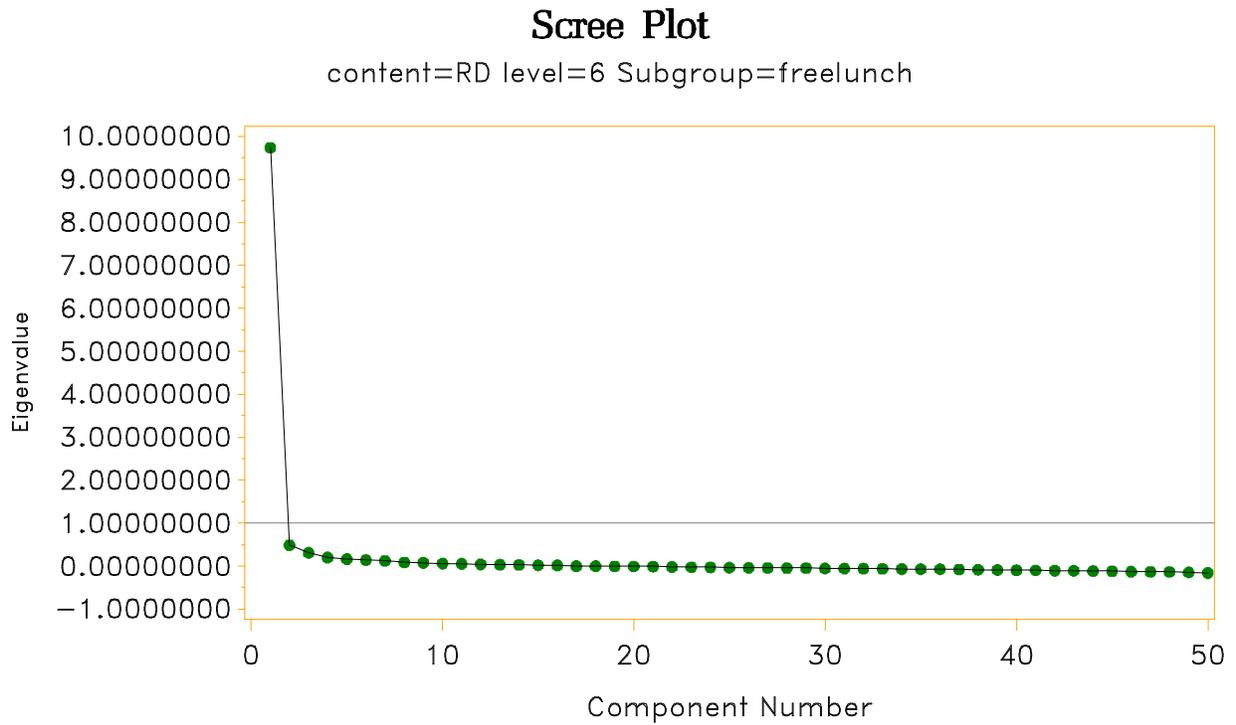


Figure 82. Spring 2014 Grade 6 Reading scree plot: Individualized Education Program

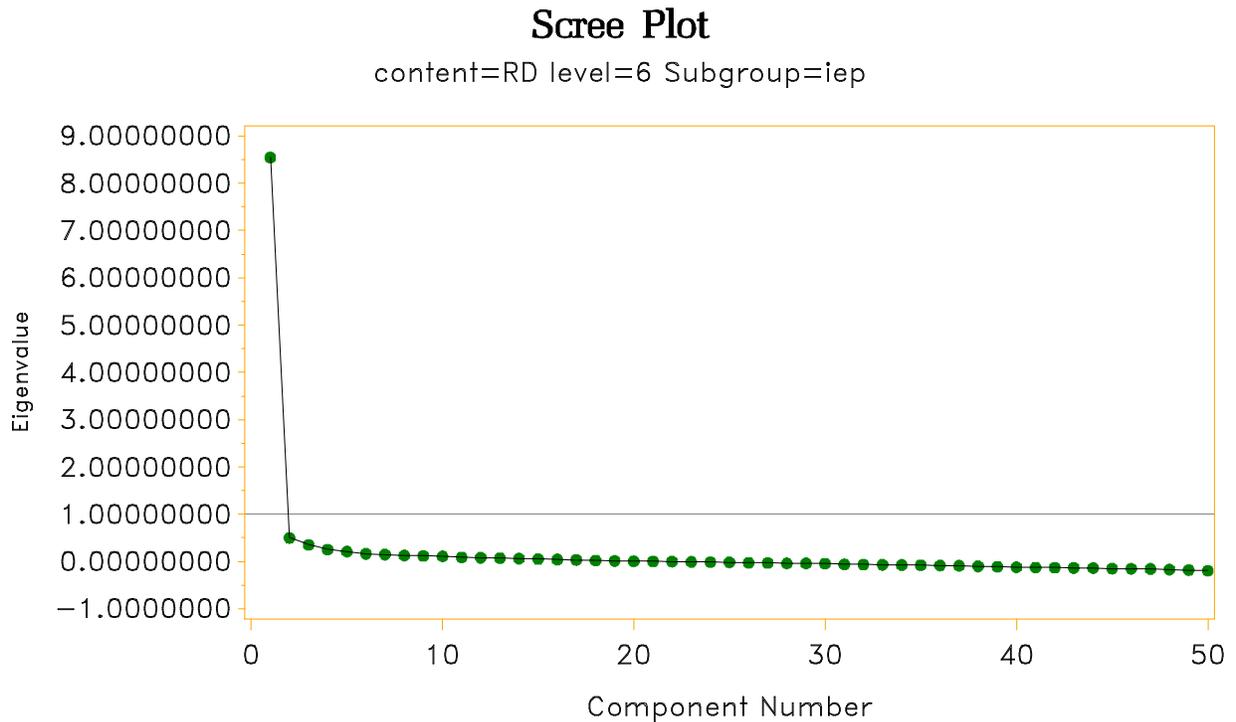


Figure 83. Spring 2014 Grade 7 Reading scree plot: Total

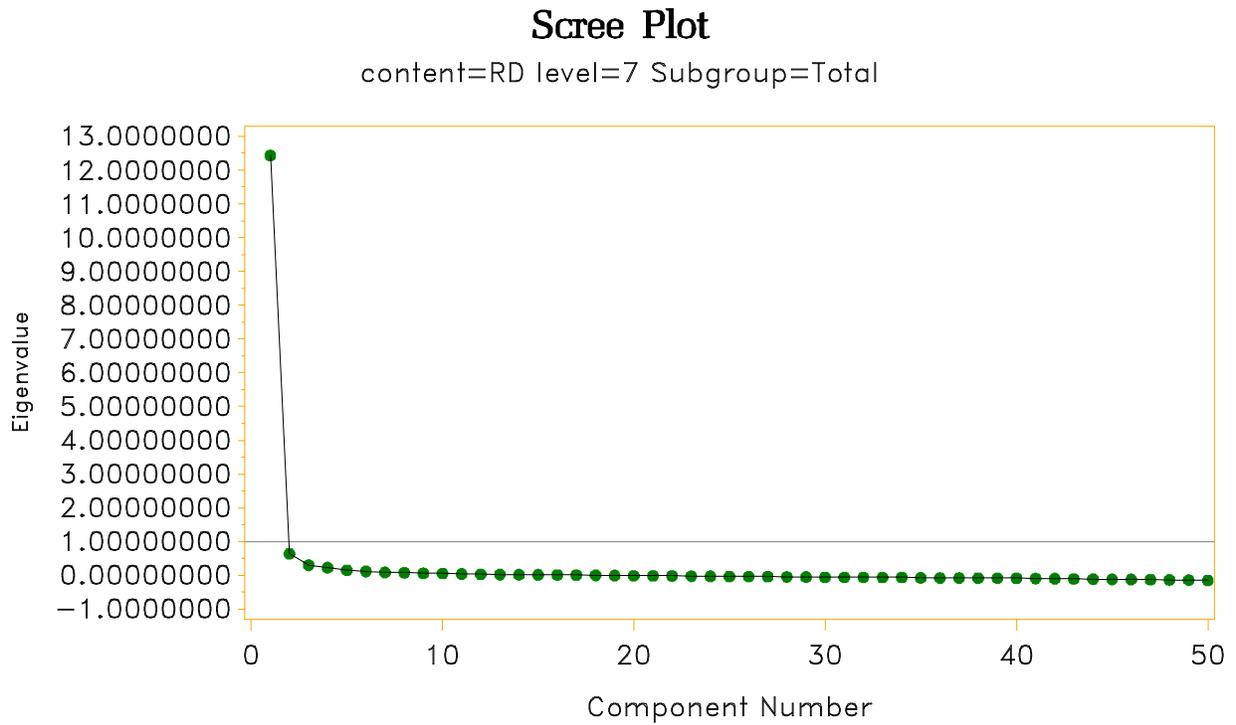


Figure 84. Spring 2014 Grade 7 Reading scree plot: Accommodated

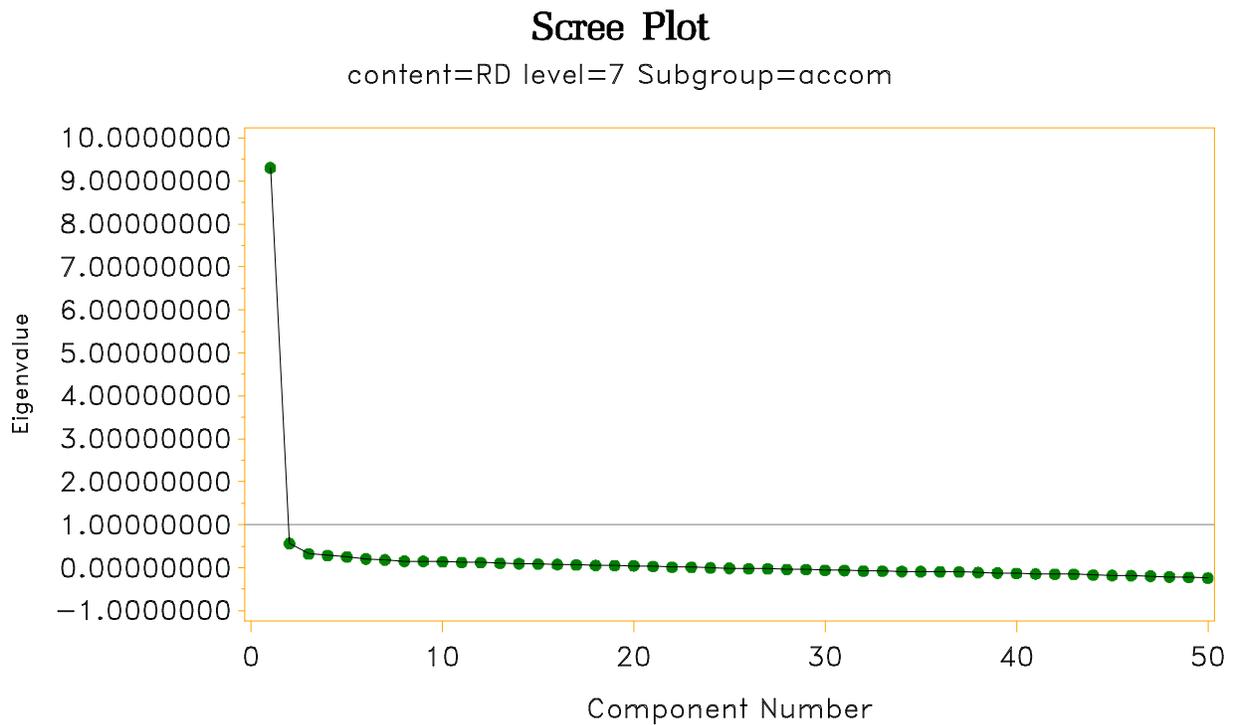


Figure 85. Spring 2014 Grade 7 Reading scree plot: English Language Learner

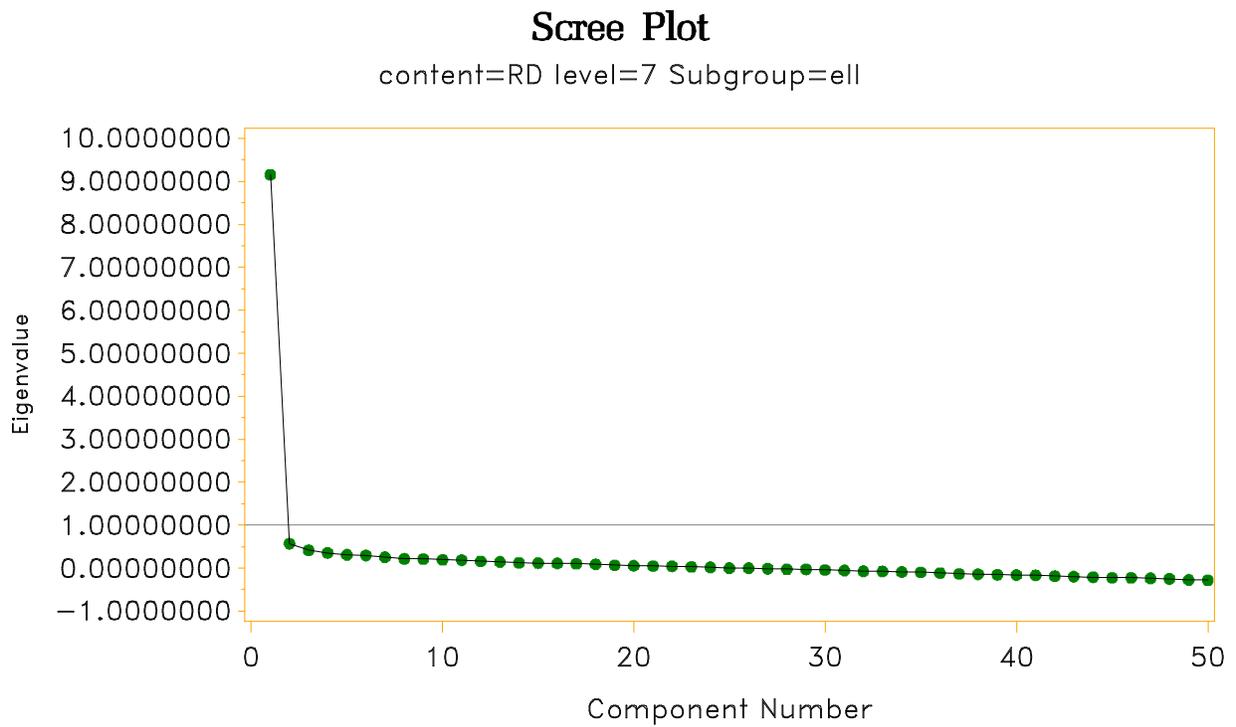


Figure 86. Spring 2014 Grade 7 Reading scree plot: Free Lunch

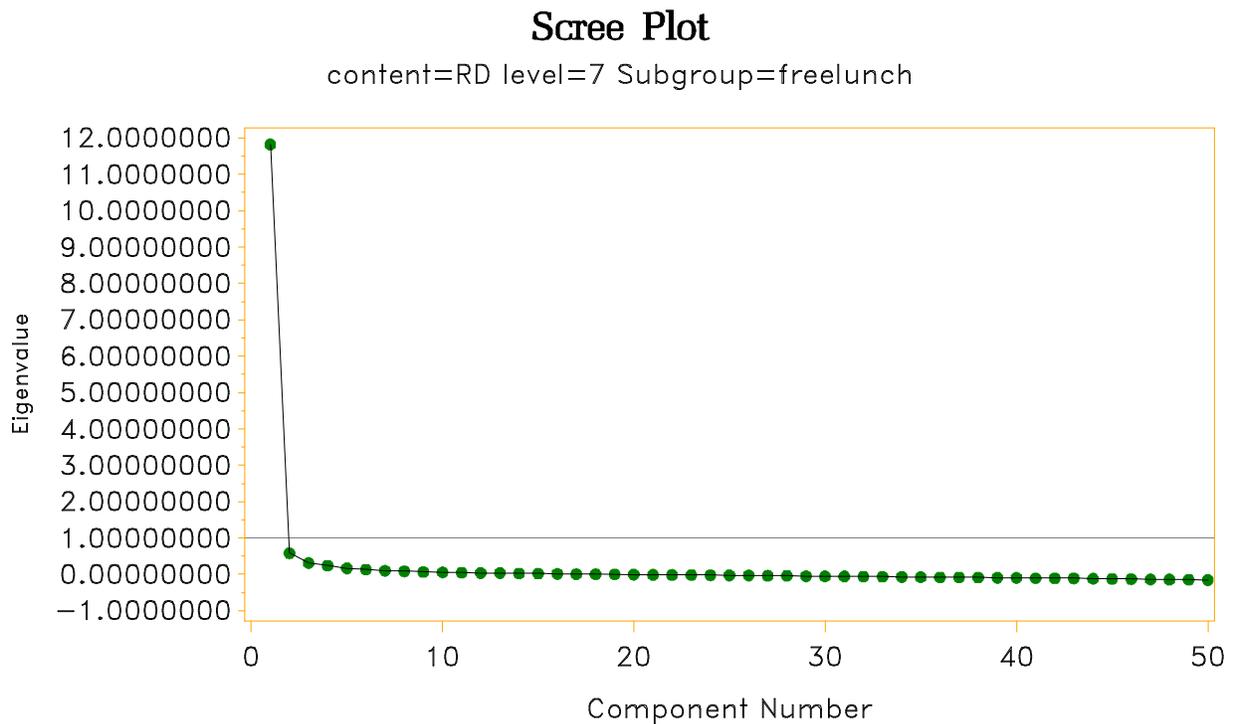


Figure 87. Spring 2014 Grade 7 Reading scree plot: Individualized Education Program

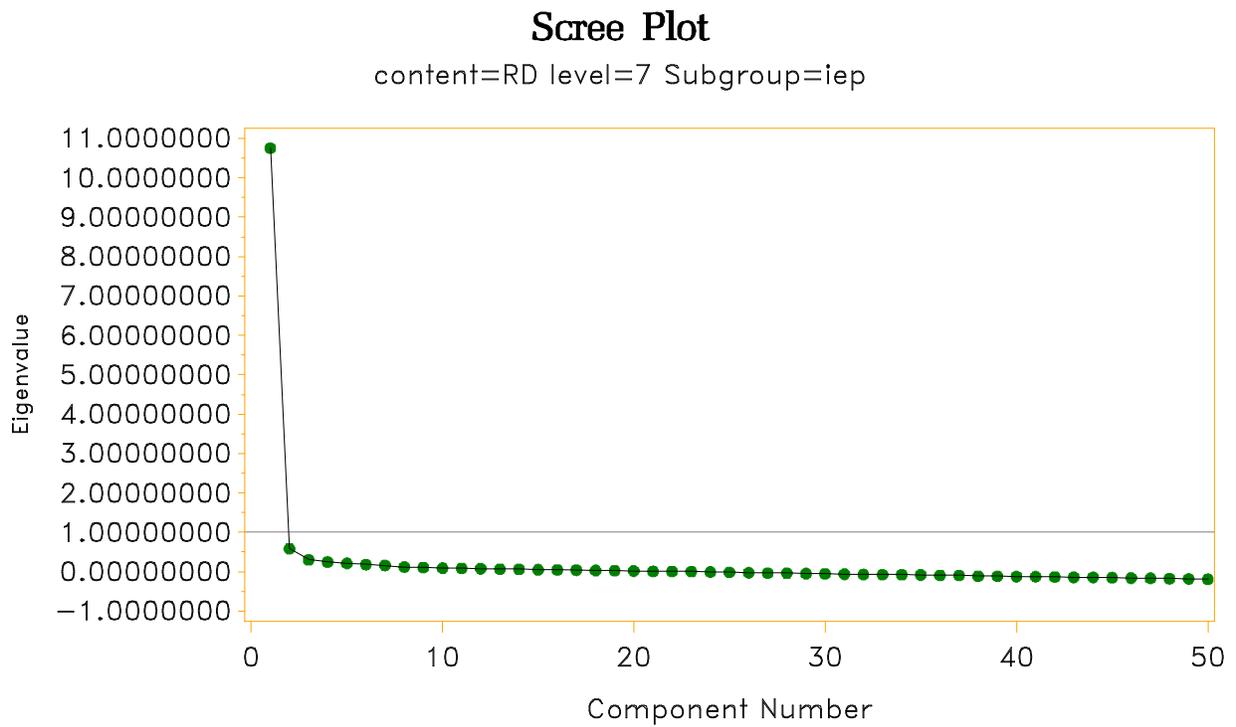


Figure 88. Spring 2014 Grade 8 Reading scree plot: Total

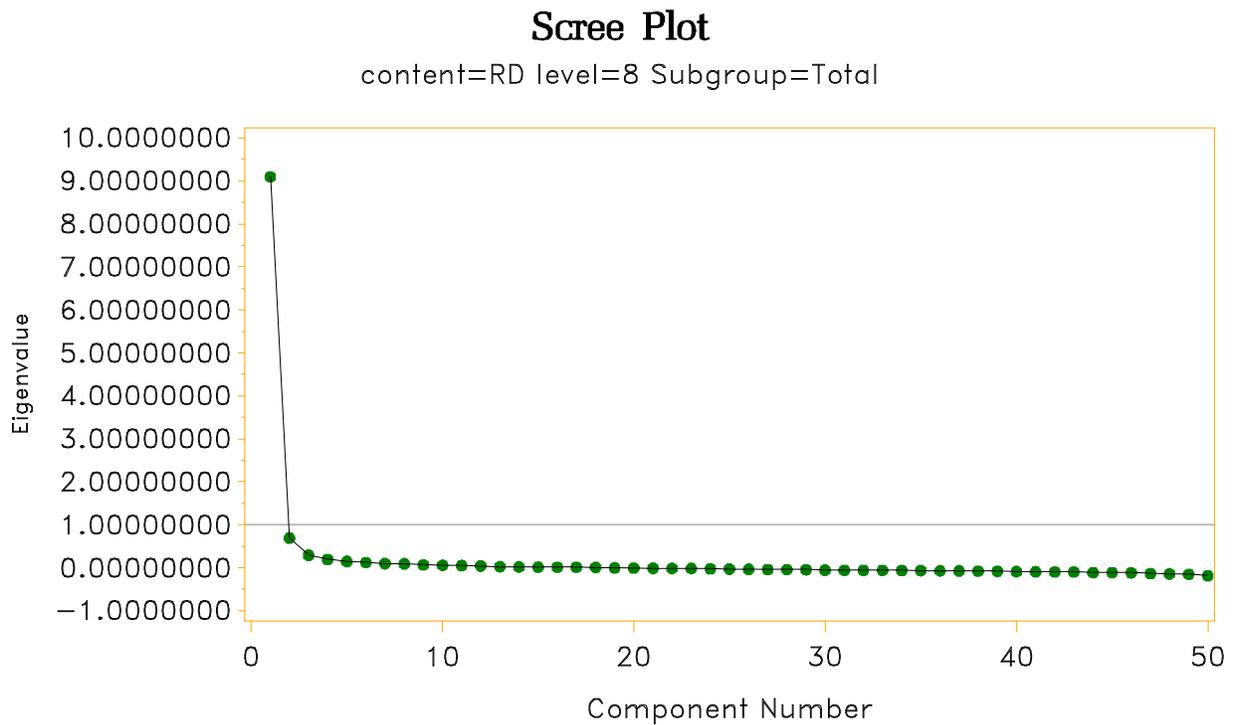


Figure 89. Spring 2014 Grade 8 Reading scree plot: Accommodated

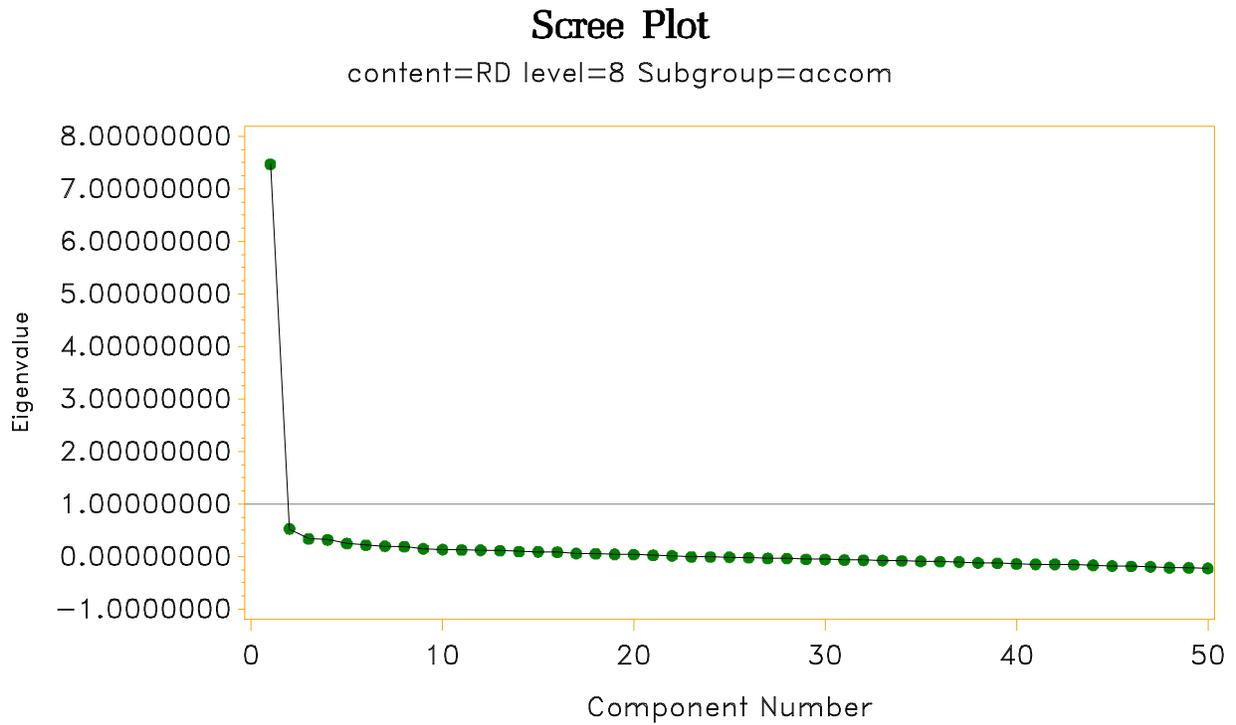


Figure 90. Spring 2014 Grade 8 Reading scree plot: English Language Learner

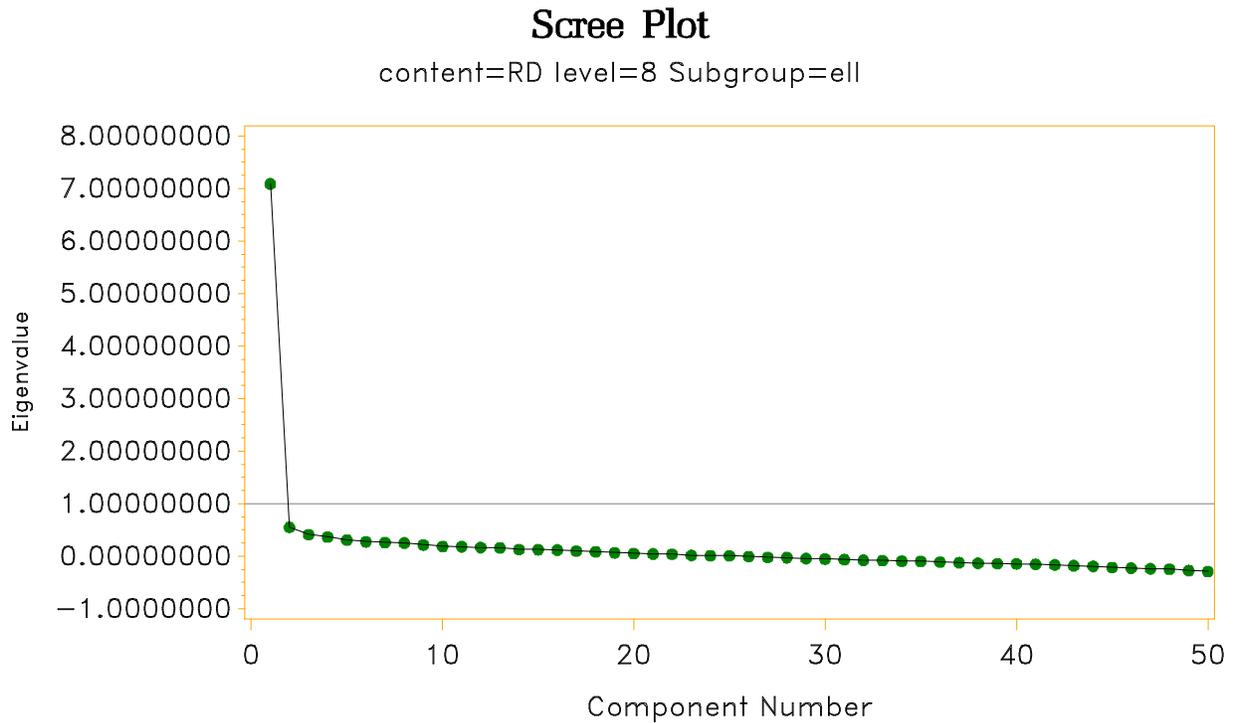


Figure 91. Spring 2014 Grade 8 Reading scree plot: Free Lunch

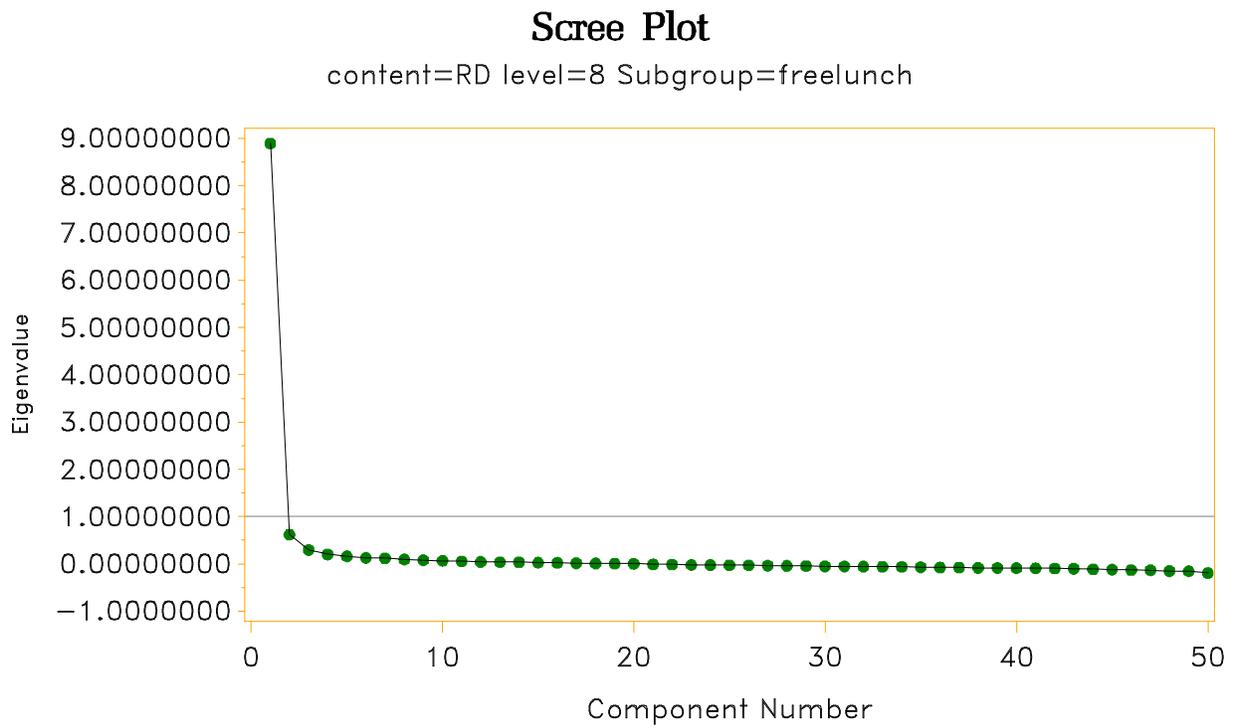


Figure 92. Spring 2014 Grade 8 Reading scree plot: Individualized Education Program

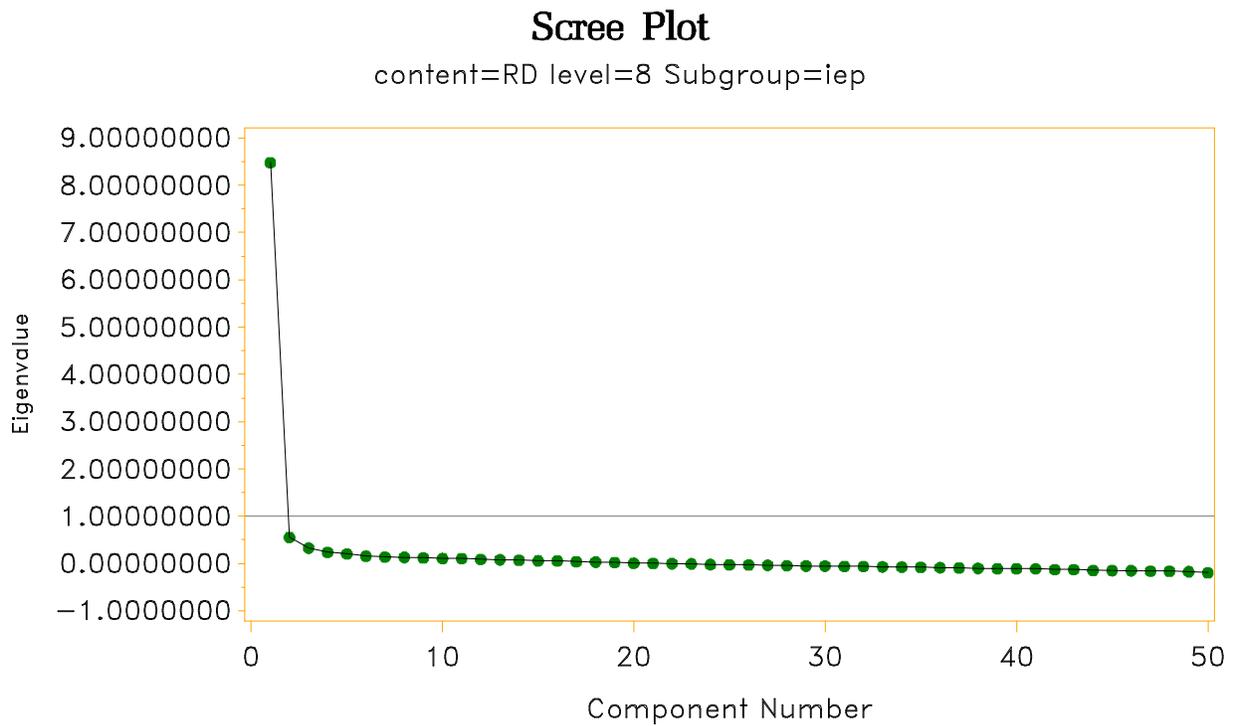


Figure 93. Spring 2014 Grade 5 Science scree plot: Total

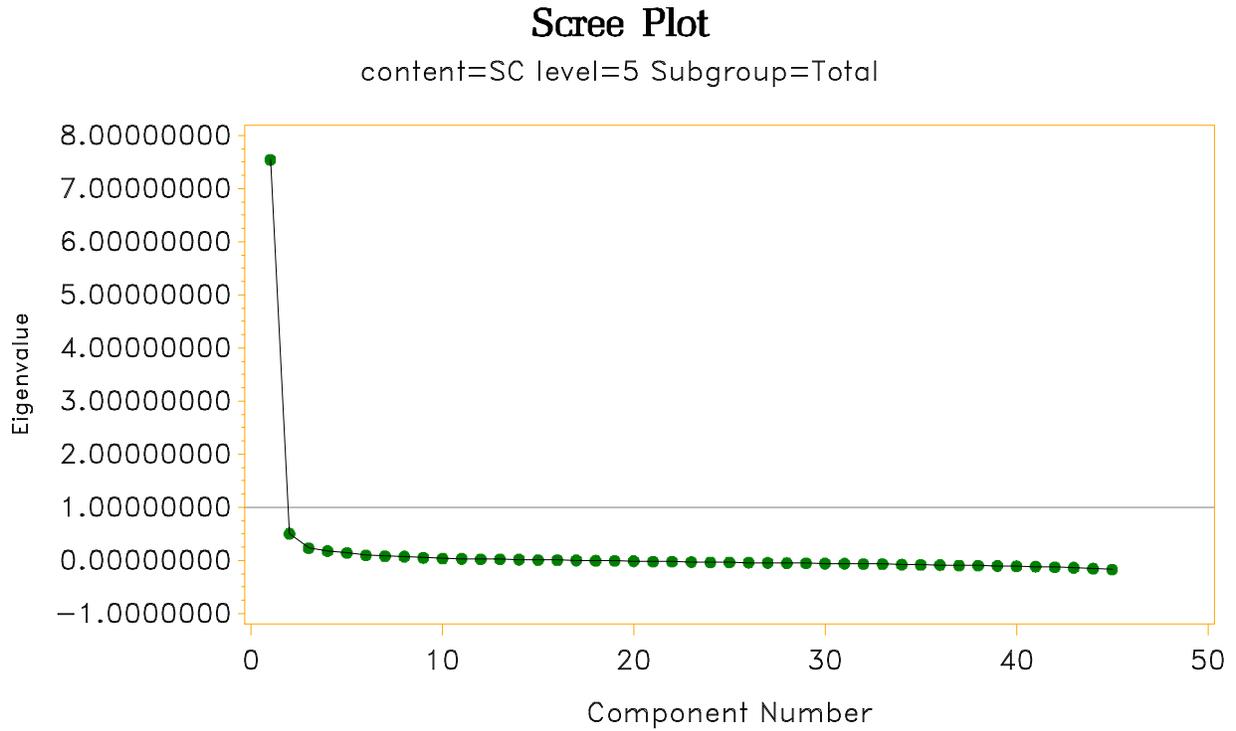


Figure 94. Spring 2014 Grade 5 Science scree plot: Accommodated

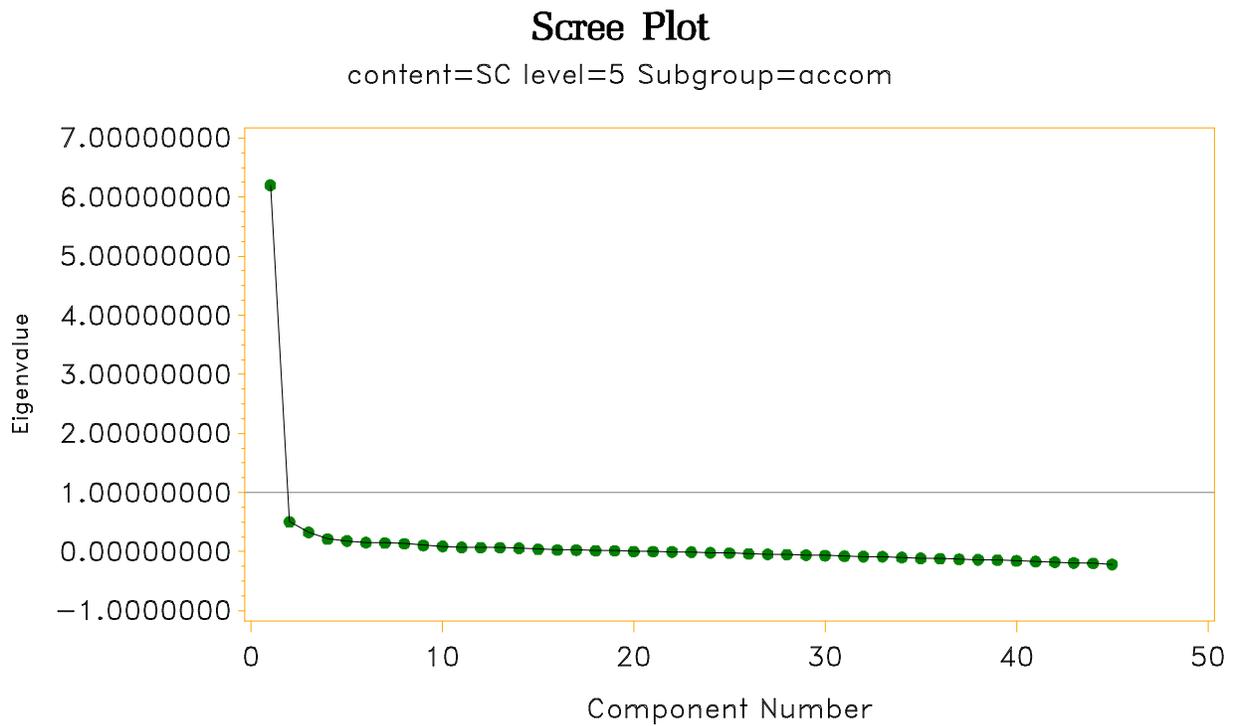


Figure 95. Spring 2014 Grade 5 Science scree plot: English Language Learner

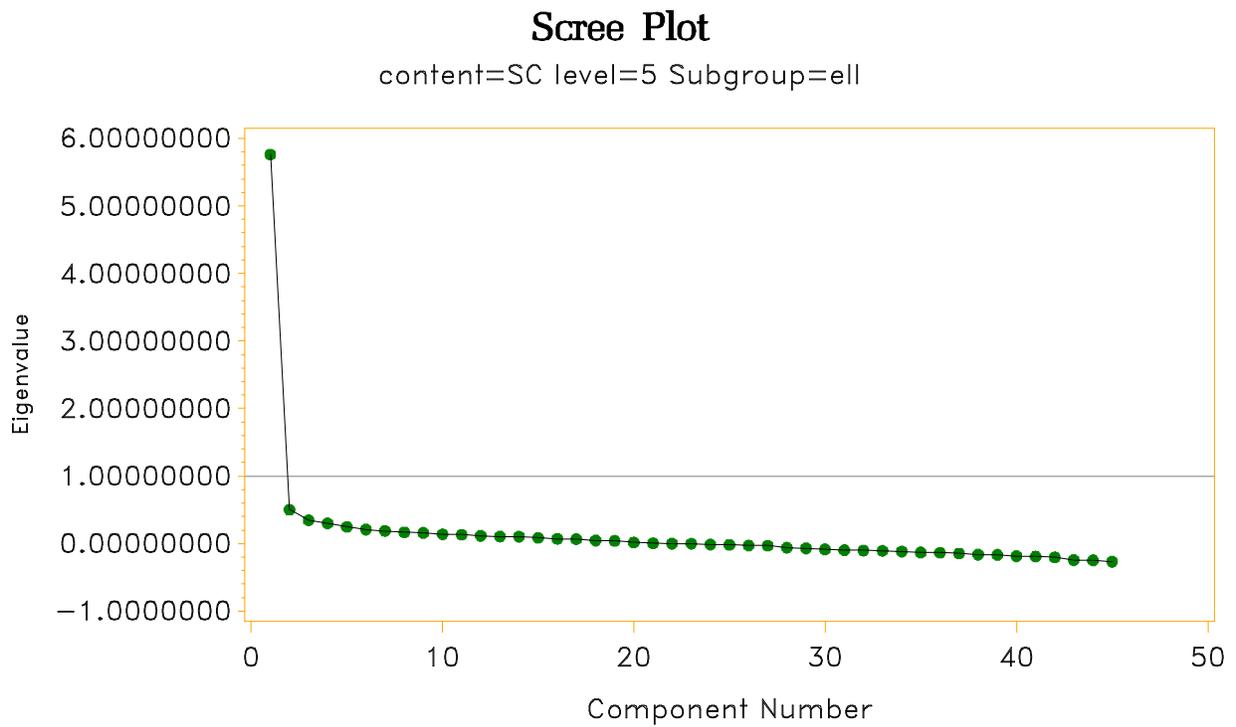


Figure 96. Spring 2014 Grade 5 Science scree plot: Free Lunch

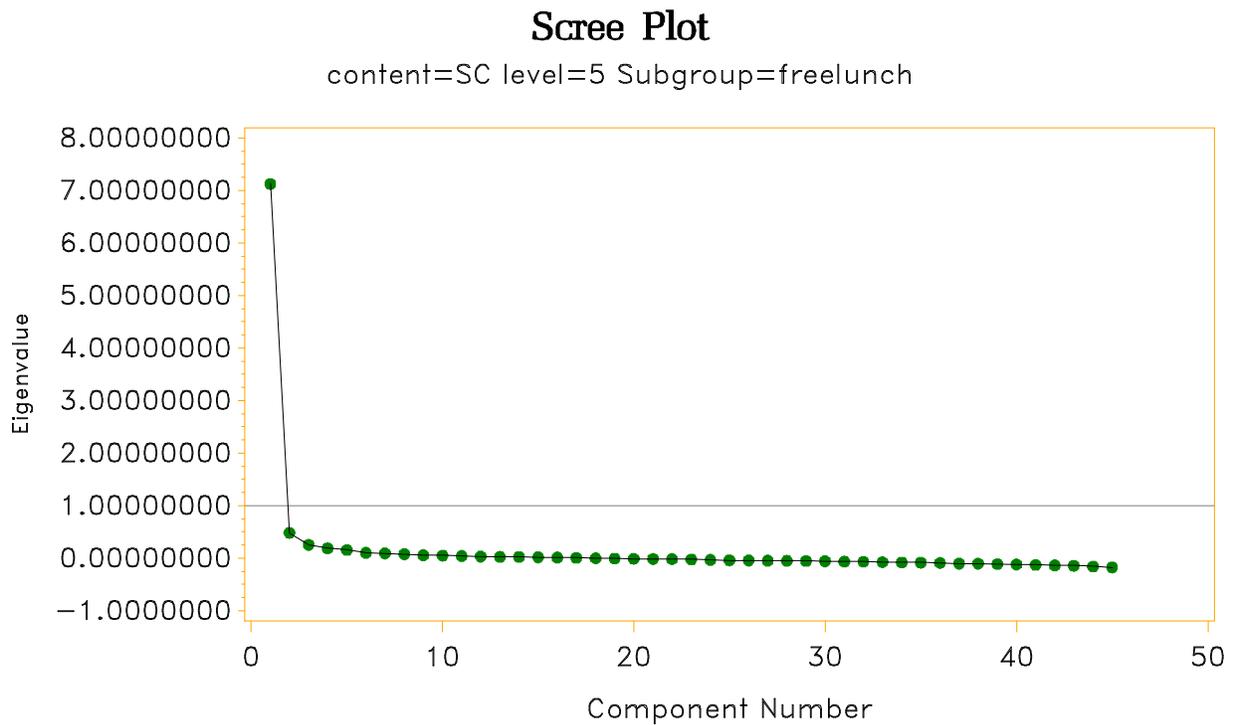


Figure 97. Spring 2014 Grade 5 Science scree plot: Individualized Education Program

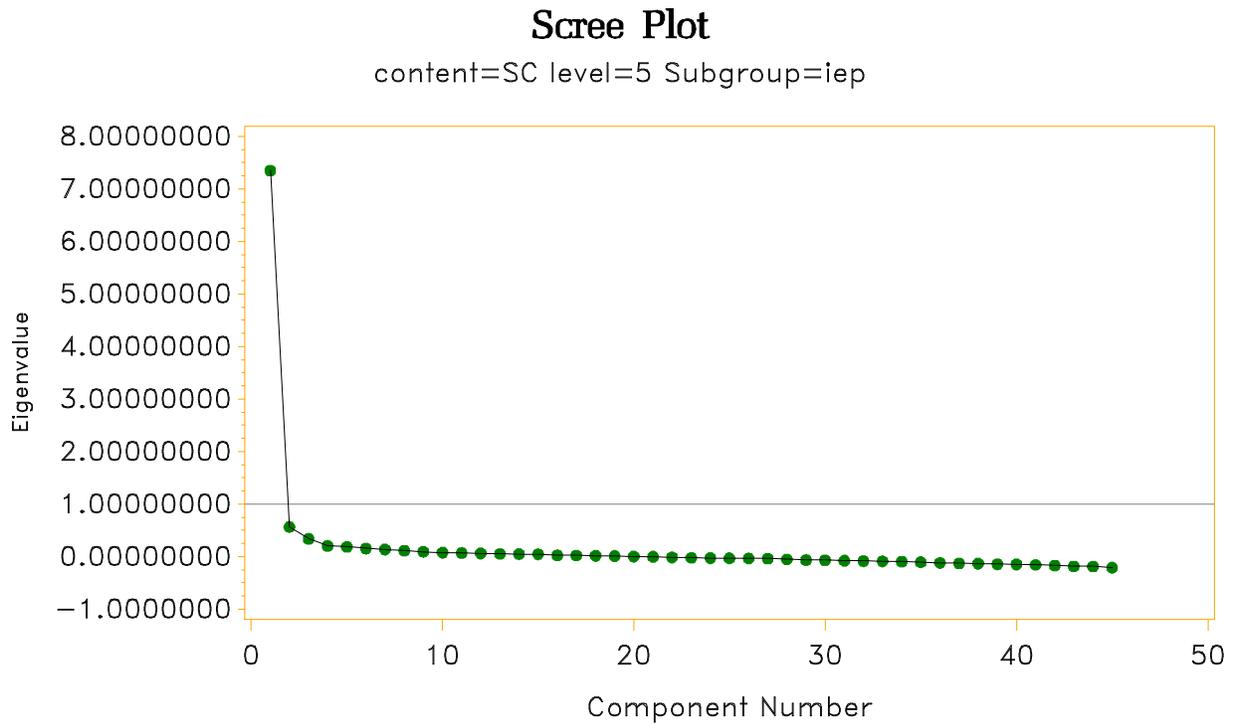


Figure 98. Spring 2014 Grade 8 Science scree plot: Total

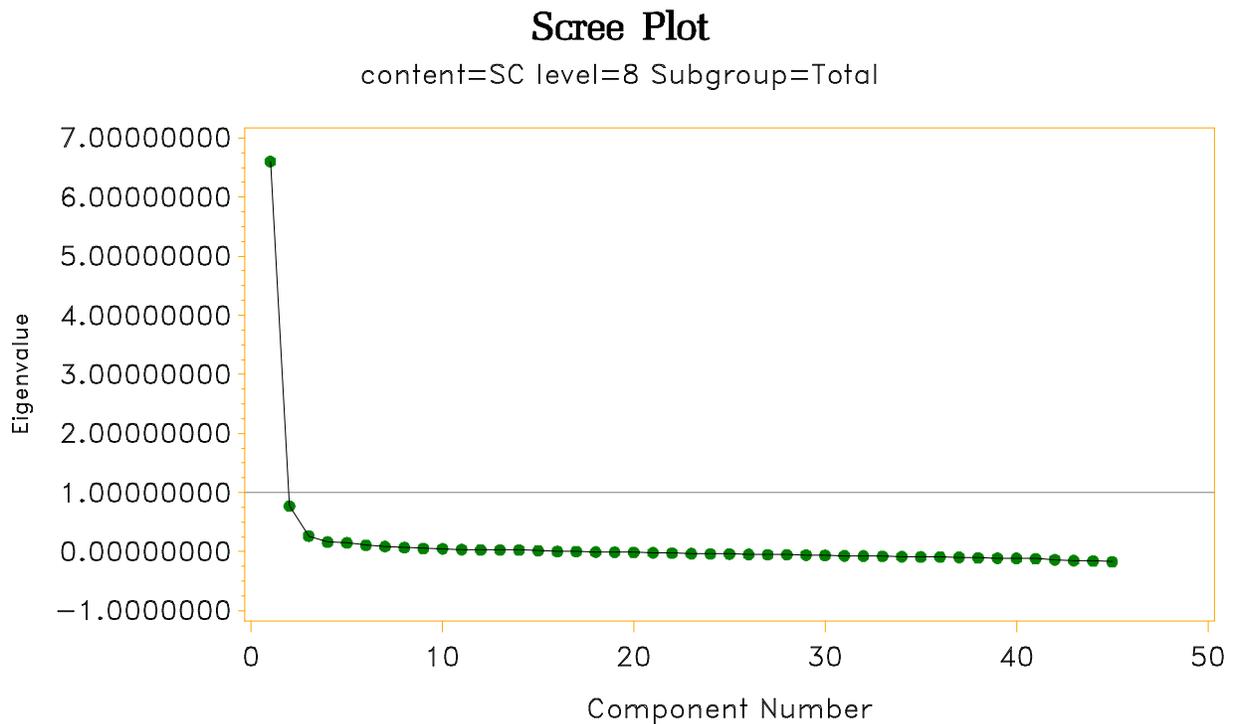


Figure 99. Spring 2014 Grade 8 Science scree plot: Accommodated

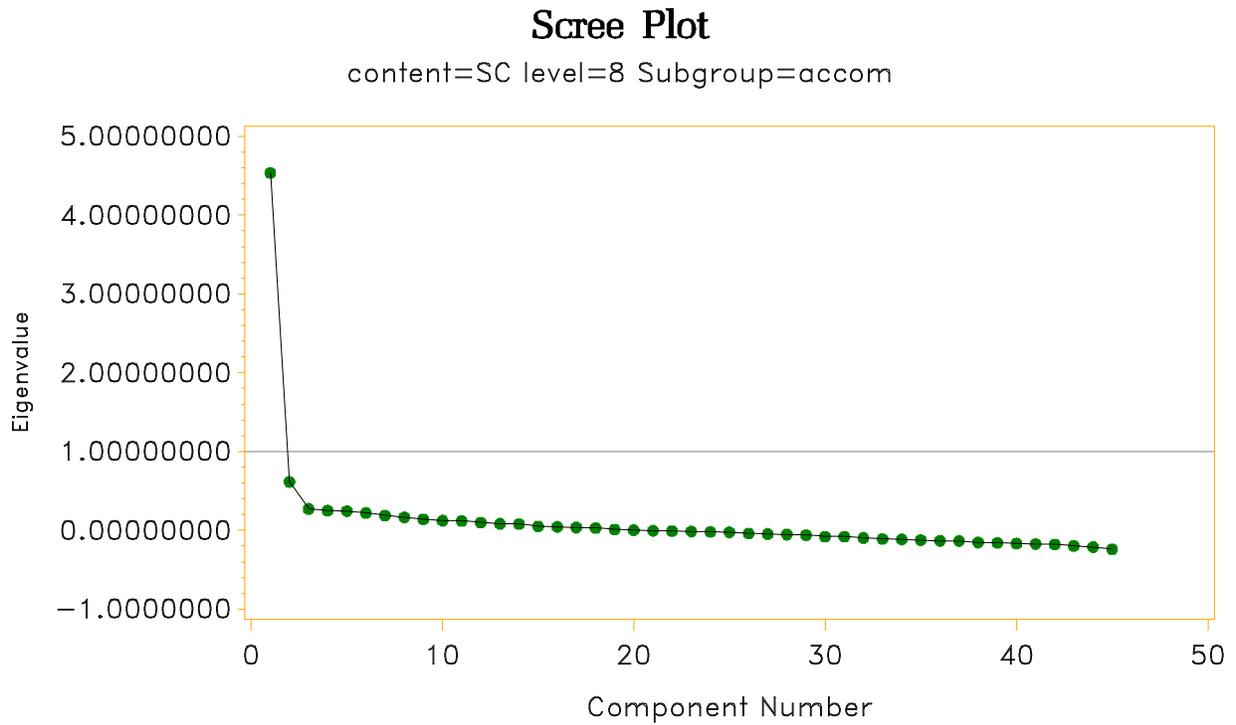


Figure 100. Spring 2014 Grade 8 Science scree plot: English Language Learner

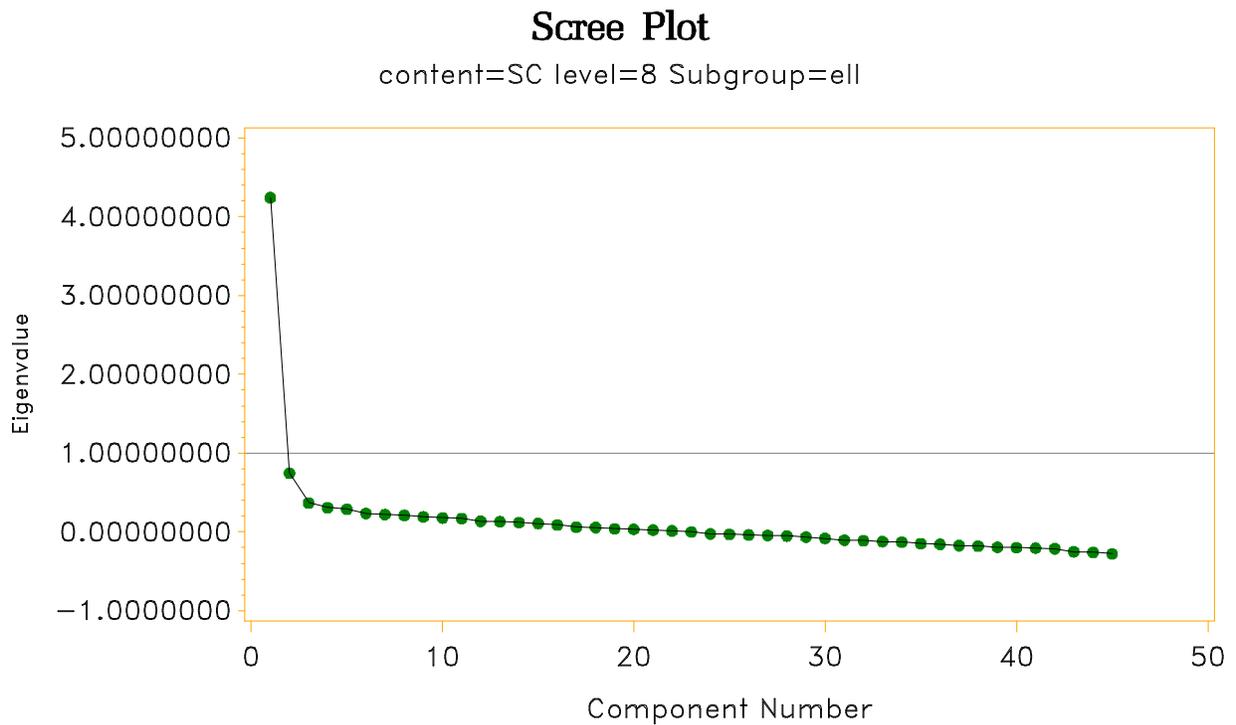


Figure 101. Spring 2014 Grade 8 Science scree plot: Free Lunch

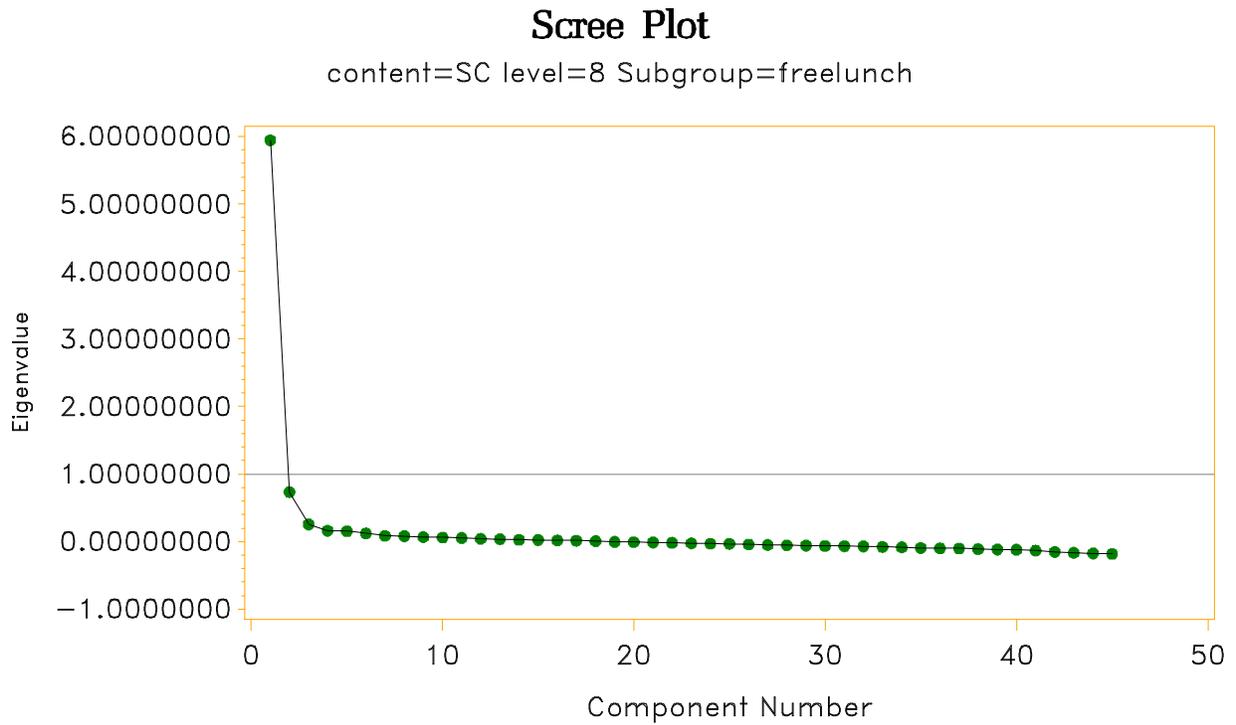


Figure 102. Spring 2014 Grade 8 Science scree plot: Individualized Education Program

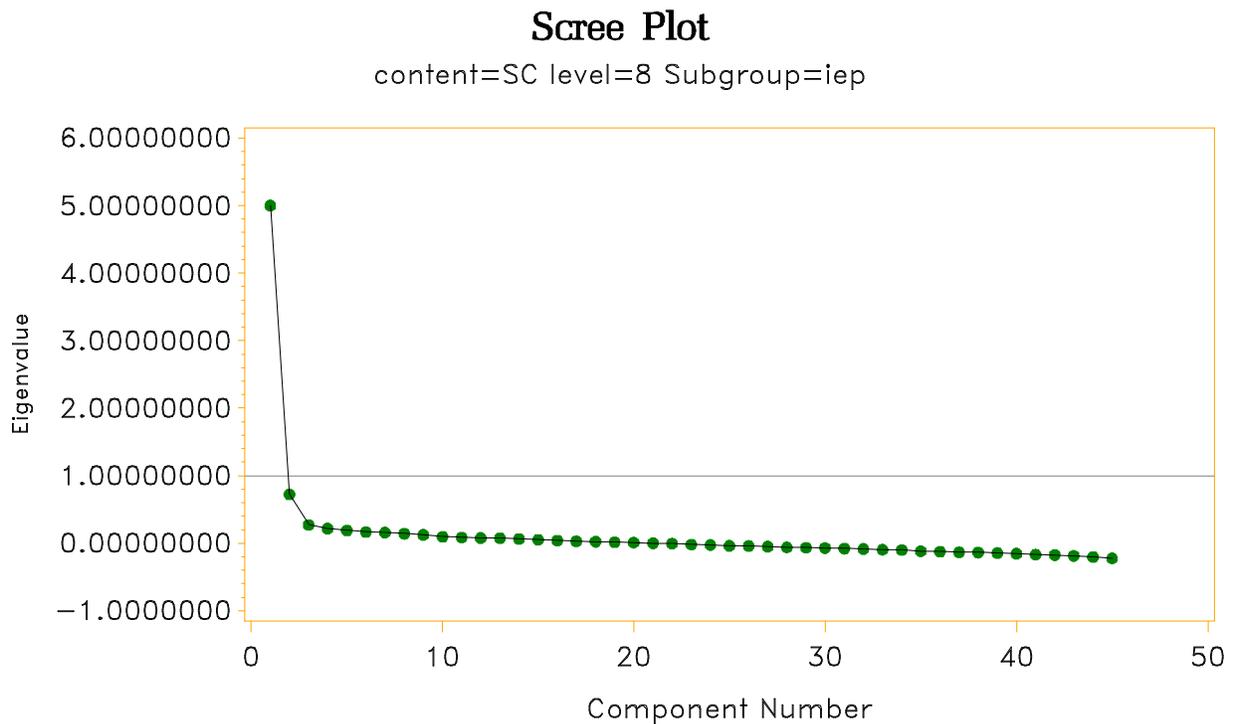


Figure 103. Spring 2014 Grade 5 Social Studies scree plot: Total

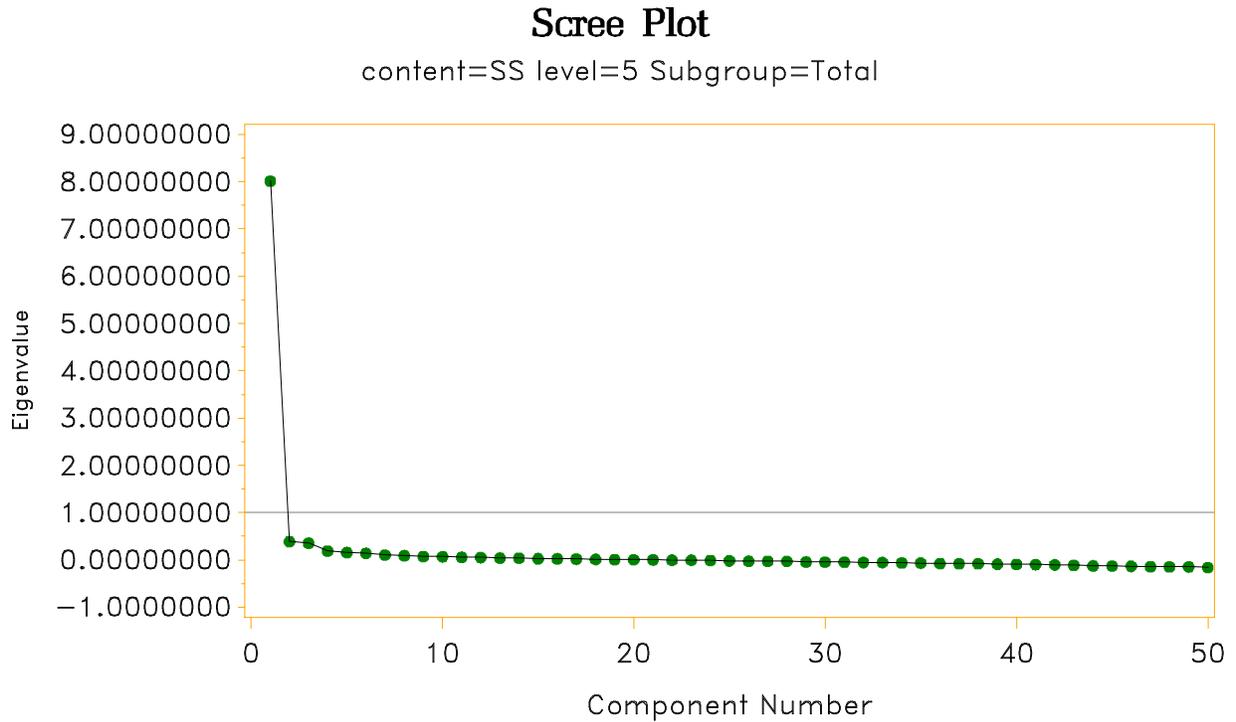


Figure 104. Spring 2014 Grade 5 Social Studies scree plot: Accommodated

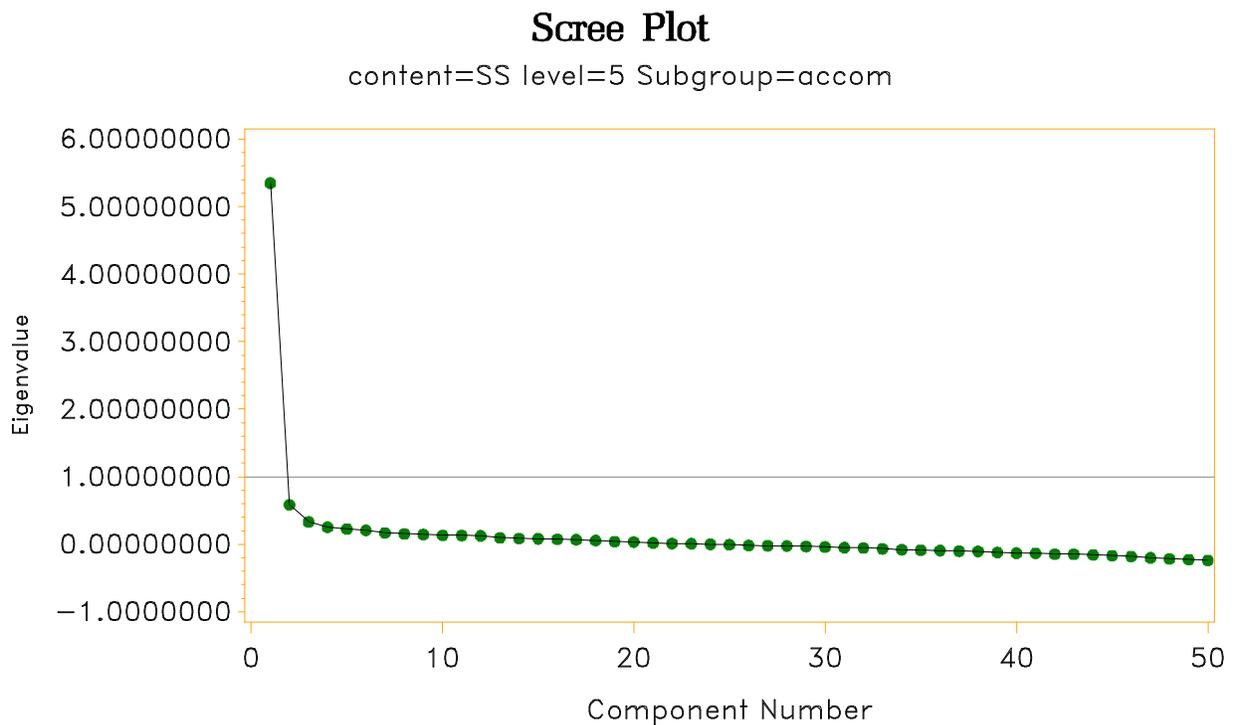


Figure 105. Spring 2014 Grade 5 Social Studies scree plot: English Language Learner

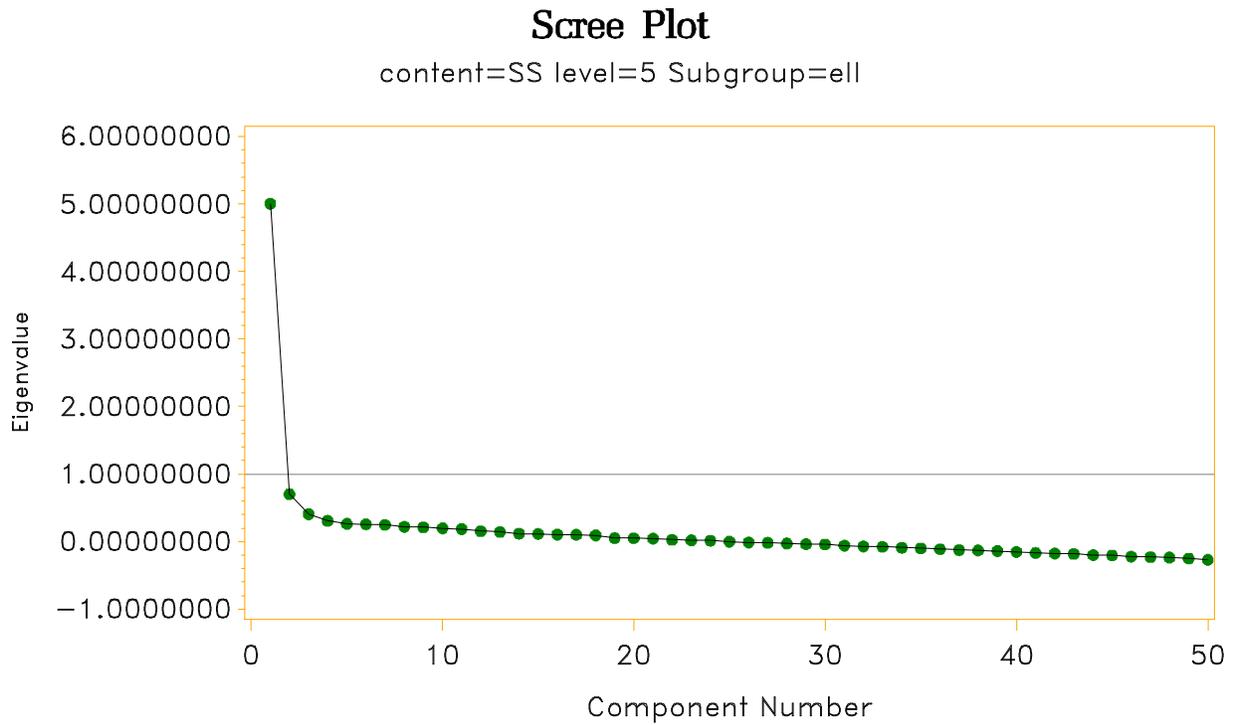


Figure 106. Spring 2014 Grade 5 Social Studies scree plot: Free Lunch

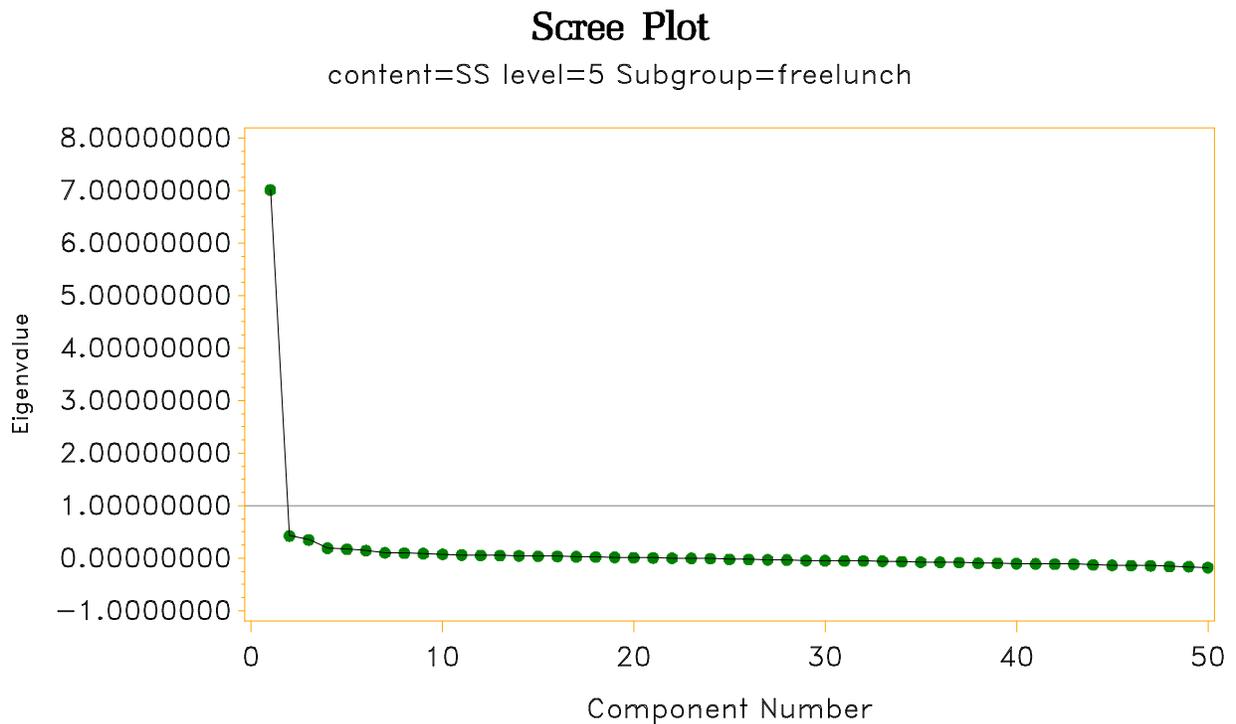


Figure 107. Spring 2014 Grade 5 Social Studies scree plot: Individualized Education Program

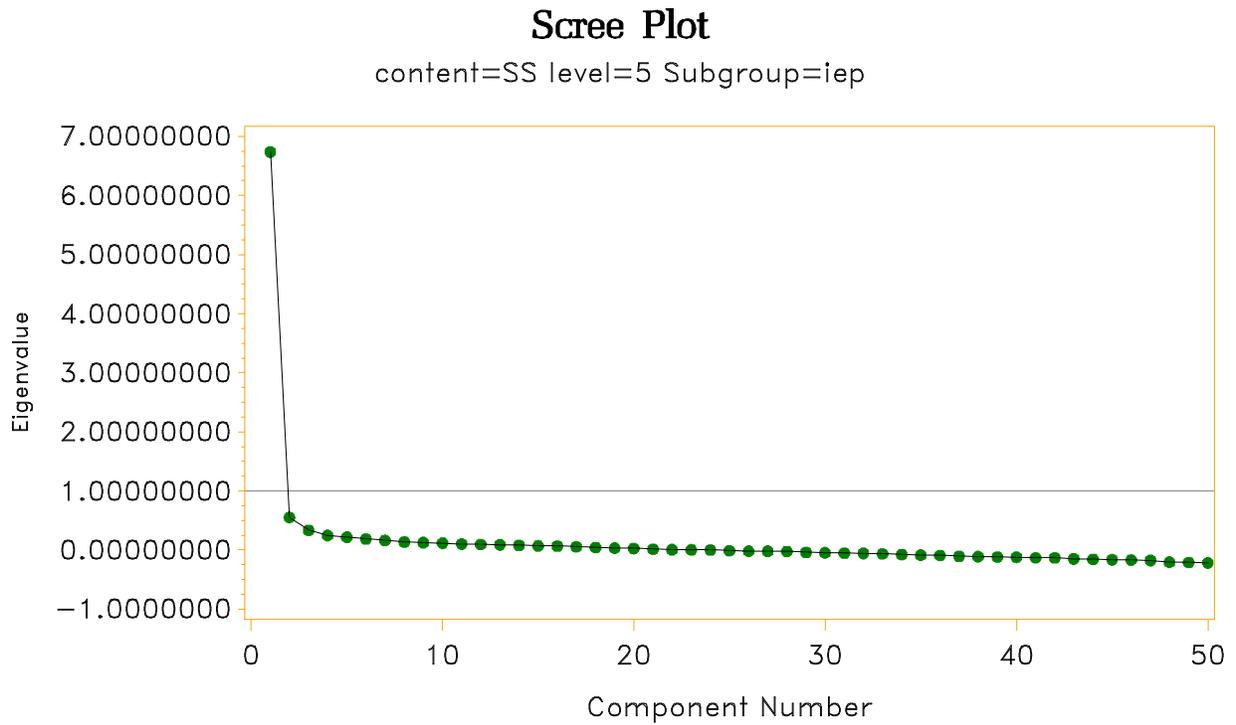


Figure 108. Spring 2014 Grade 8 U.S. History scree plot: Total

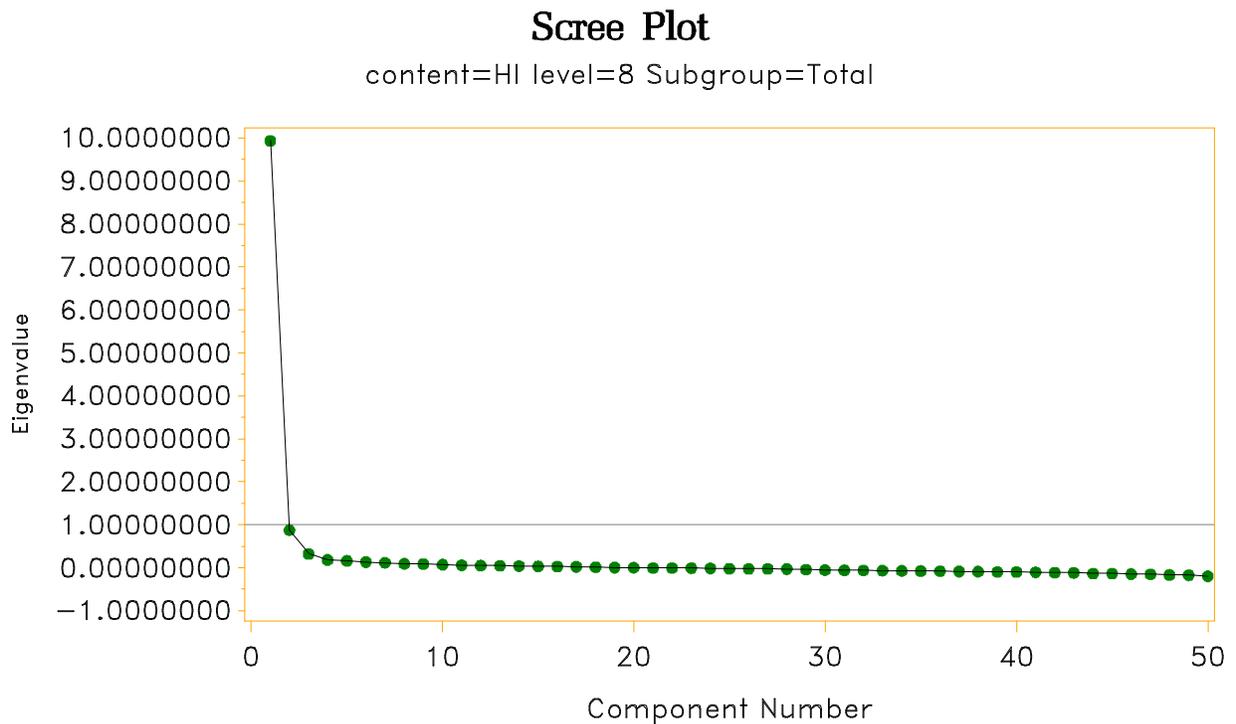


Figure 109. Spring 2014 Grade 8 U.S. History scree plot: Accommodated

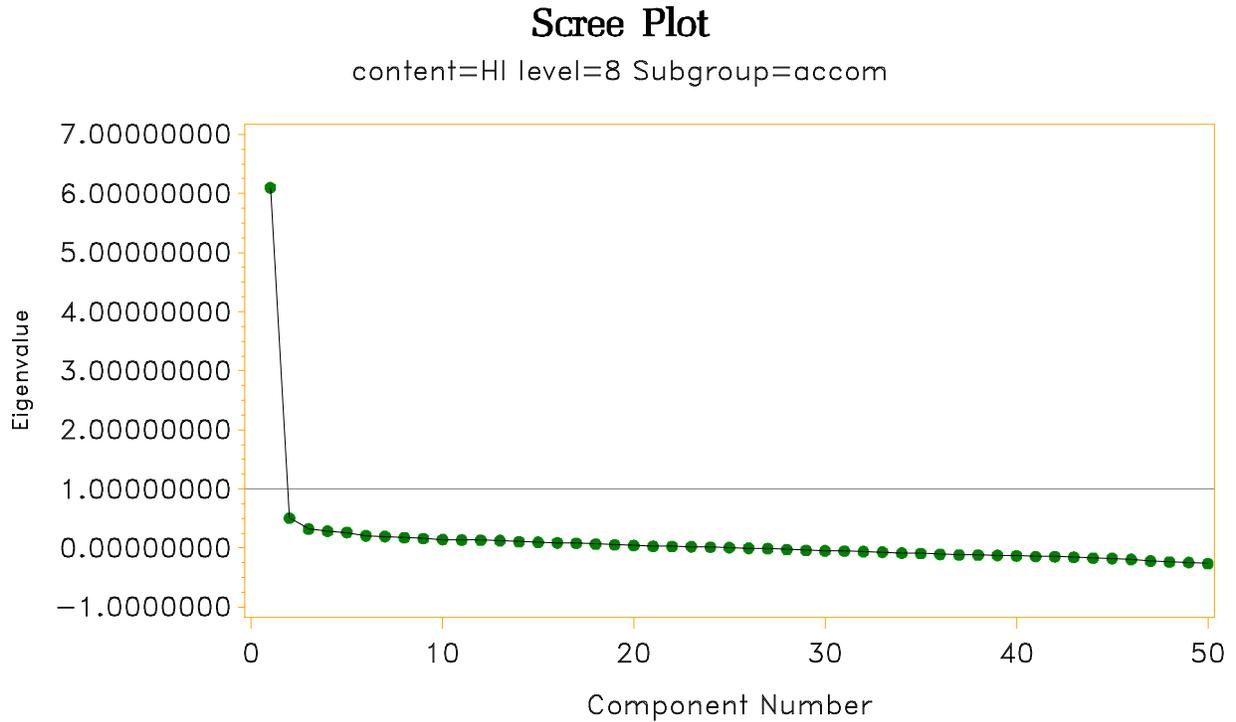


Figure 110. Spring 2014 Grade 8 U.S. History scree plot: English Language Learner

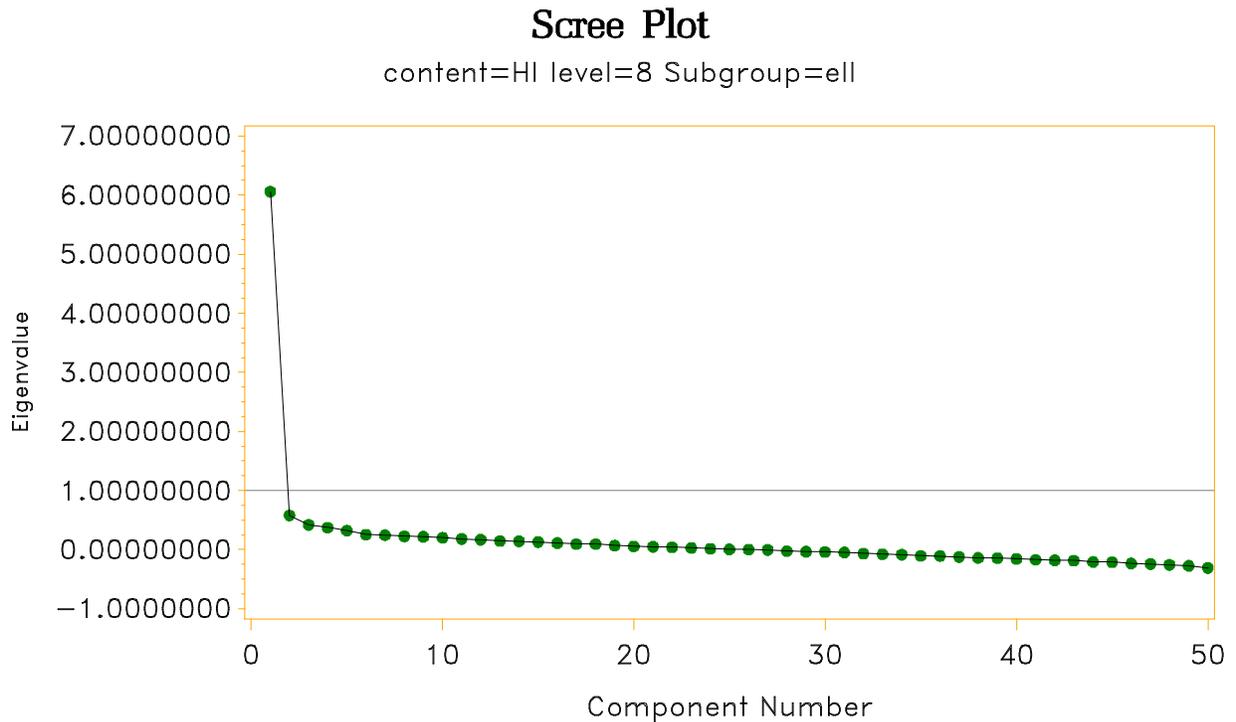


Figure 111. Spring 2014 Grade 8 U.S. History scree plot: Free Lunch

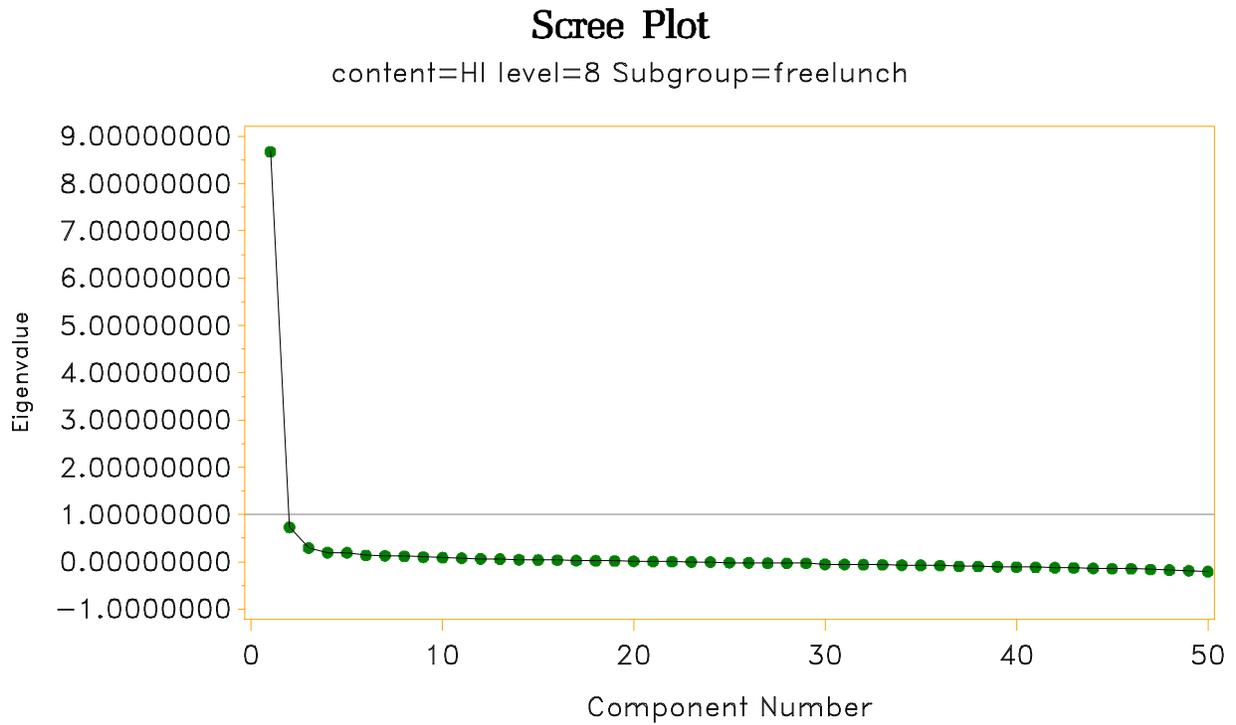
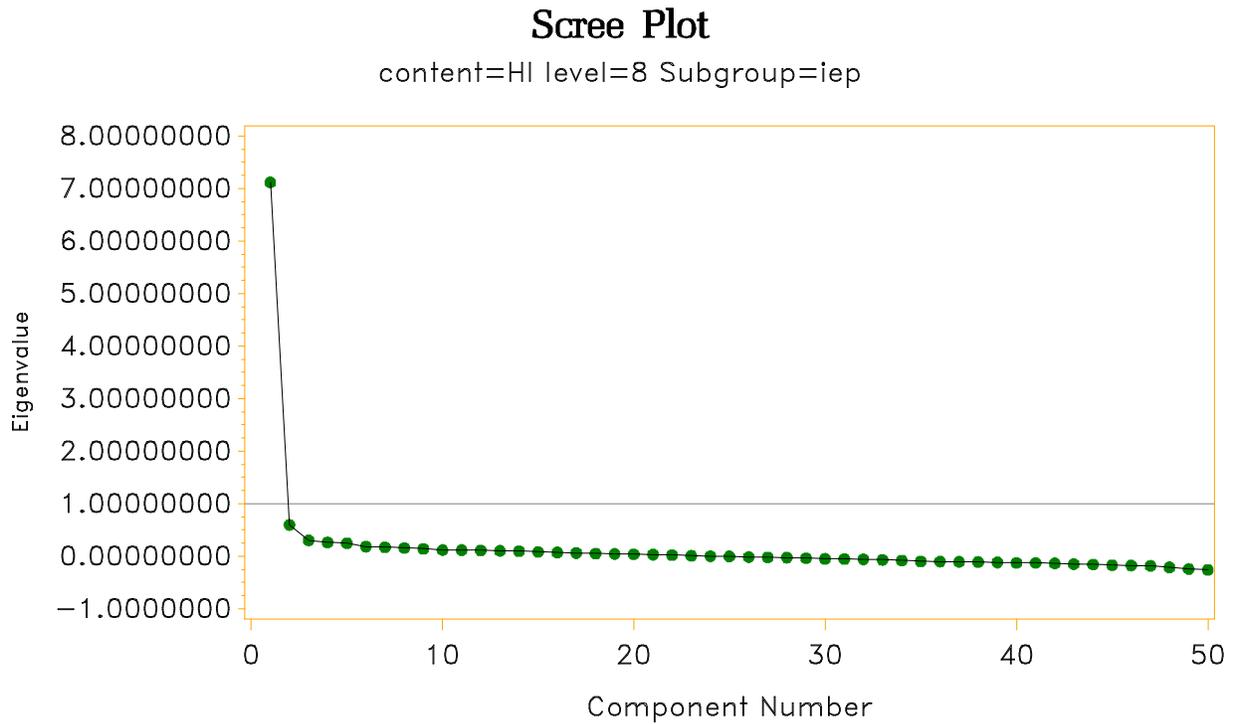


Figure 112. Spring 2014 Grade 8 U.S. History scree plot: Individualized Education Program



Appendices

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

Table A1. OCCT Test Blueprint and Actual Item Counts: Grade 3 Mathematics

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Algebraic Reasoning: Patterns and Relationships	7	7	0
1.1 Algebra Patterns, Equations	2	2	0
1.2 Equations	2	2	0
1.3 Number Properties	3	3	0
Number Sense and Operation	20	20	0
2.1 Number Sense	10	10	0
2.2 Number Operations	10	10	0
Geometry	7	7	0
3.1 Properties of shapes	3	3	0
3.2 Spatial Reasoning	2	2	0
3.3 Coordinate Geometry	2	2	0
Measurement	9	9	0
4.1 Measurement	4	4	0
4.2 Time and Temperature	2	2	0
4.3 Money	3	3	0
Data Analysis	7	7	0
5.1 Data Analysis	4	3	0
5.2 Probability	3	4	0
Total Test	50	50	0

Table A2. OCCT Test Blueprint and Actual Item Counts: Grade 4 Mathematics

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Algebraic Reasoning: Patterns and Relationships	7	7	0
1.1 Algebra Patterns	3	3	0
1.2 Equations	2	2	0
1.3 Number Properties	2	2	0
Number Sense and Operation	18	18	0
2.1 Number Sense	8	8	0
2.2 Number Operations	10	10	0
Geometry	9	9	0
3.1 Lines	2	2	0
3.2 Angles	2	2	0
3.3 Polygons	3	3	0
3.4 Transformations	2	2	0
Measurement	9	9	0
4.1 Measurement	5	5	0
4.2 Time and Temperature	2	2	0
4.3 Money	2	2	0
Data Analysis	7	7	0
5.1 Data Analysis	2	2	0
5.2 Probability	2	2	0
5.3 Central Tendency	3	3	0
Total Test	50	50	0

Table A3. OCCT Test Blueprint and Actual Item Counts: Grade 5 Mathematics

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Algebraic Reasoning: Patterns and Relationships	13	13	0
1.1 Algebra Patterns	5	5	0
1.2 Equations	4	3	0
1.3 Number Properties	4	5	0
Number Sense and Operation	16	16	0
2.1 Number Sense	8	8	0
2.2 Number Operations	8	8	0
Geometry	7	7	0
3.1 Circles and Polygons	4	4	0
3.2 Angles	3	3	0
Measurement	7	7	0
4.1 Measurement	5	5	0
4.2 Money	2	2	0
Data Analysis	7	7	0
5.1 Data Analysis	3	3	0
5.2 Probability	2	2	0
5.3 Central Tendency	2	2	0
Total Test	50	50	0

Table A4. OCCT Test Blueprint and Actual Item Counts: Grade 6 Mathematics

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Algebraic Reasoning: Patterns and Relationships	13	13	0
1.1 Algebra Patterns	4	3	0
1.2 Expressions and Equations	4	5	0
1.3 Number Properties	3	3	0
1.4 Solving Equations	2	2	0
Number Sense and Operation	15	15	0
2.1 Number Sense	5	4	0
2.2 Number Operations	10	11	0
Geometry	8	8	0
3.1 Three Dimensional Figures	2	2	0
3.2 Congruent and Similar Figures	2	2	0
3.3 Coordinate Geometry	4	4	0
Measurement	7	7	0
4.1 Circles	4	4	0
4.2 Conversions	3	3	0
Data Analysis	7	7	0
5.1 Data Analysis	3	3	0
5.2 Probability	2	2	0
5.3 Central Tendency	2	2	0
Total Test	50	50	0

Table A5. OCCT Test Blueprint and Actual Item Counts: Grade 7 Mathematics

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Algebraic Reasoning: Patterns and Relationships	15	15	0
1.1 Linear Relationships	5	4	0
1.2 Solving Equations	5	6	0
1.3 Solving and Graphing Inequalities	5	5	0
Number Sense and Operation	11	11	0
2.1 Number Sense	5	6	0
2.2 Number Operations	6	5	0
Geometry	8	7	0
3.1 Classifying Figures	1–3	1	0
3.2 Lines and Angles	1–3	2	0
3.3 Transformations	4	4	0
Measurement	9	9	0
4.1 Perimeter and Area	5	6	0
4.2 Circles	2	2	0
4.3 Composite Figures	2	1	0
Data Analysis	7	8	0
5.1 Data Analysis	2	3	0
5.2 Probability	2	2	0
5.3 Central Tendency	3	3	0
Total Test	50	50	0

Table A6. OCCT Test Blueprint and Actual Item Counts: Grade 8 Mathematics

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Algebraic Reasoning: Patterns and Relationships	16	16	0
1.1 Equations	10–12	9	0
1.2 Inequalities	4–6	7	0
Number Sense and Operation	11	11	0
2.1 Number Sense	3–4	4	0
2.2 Number Operations	7–8	7	0
Geometry	9	9	0
3.1 Three Dimensional Figures	5	5	0
3.2 Pythagorean Theorem	4	4	0
Measurement	7	7	0
4.1 Surface Area and Volume	3	3	0
4.2 Ratio and Proportions	2	2	0
4.3 Composite Figures	2	2	0
Data Analysis	7	7	0
5.1 Data Analysis	3	3	0
5.3 Central Tendency	4	4	0
Total Test	50	50	0

Table A7. OCCT Test Blueprint and Actual Item Counts: Grade 3 Reading

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Vocabulary	12	12	0
2.1 Words in Context	2–4	3	0
2.2 Affixes, Roots, and Stems	2–4	3	0
2.3 Synonyms, Antonyms, and Homonyms	2–4	2	0
2.4 Using Resource Materials	2–4	4	0
Comprehension/Critical Literacy	24	26	0
4.1 Literal Understanding	5	5	0
4.2 Inferences and Interpretation	7	5	0
4.3 Summary and Generalization	6	8	0
4.4 Analysis and Evaluation	6	8	0
Literature	8	7	0
5.2 Literary Elements	3–4	3	0
5.3 Figurative Language/Sound Devices	4–5	4	0
Research and Information	6	5	0
6.1 Accessing Information	6	5	0
Total Test	50	50	0

Table A8. OCCT Test Blueprint and Actual Item Counts: Grade 4 Reading

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Vocabulary	12	11	0
1.1 Words in Context	4	2	0
1.2 Affixes, Roots, and Stems	4	5	0
1.3 Synonyms, Antonyms and Homonyms	4	4	0
Comprehension/Critical Literacy	23	24	0
3.1 Literal Understanding	4	5	0
3.2 Inferences and Interpretation	6	5	0
3.3 Summary and Generalization	7	5	0
3.4 Analysis and Evaluation	6	9	0
Literature	9	9	0
4.2 Literary Elements	5	4	0
4.3 Figurative Language/Sound Devices	4	5	0
Research and Information	6	6	0
5.1 Accessing Information	6	6	0
Total Test	50	50	0

Table A9. OCCT Test Blueprint and Actual Item Counts: Grade 5 Reading

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Vocabulary	12	11	0
1.1 Words in Context	4	4	0
1.2 Affixes, Roots, and Stems	4	3	0
1.3 Synonyms, Antonyms, and Homonyms	4	4	0
Comprehension/Critical Literacy	20	19	0
3.1 Literal Understanding	4	4	0
3.2 Inferences and Interpretation	4–6	6	0
3.3 Summary and Generalization	4–6	6	0
3.4 Analysis and Evaluation	4–6	3	0
Literature	12	13	0
4.1 Literary Genre	4	5	0
4.2 Literary Elements	4	5	0
4.3 Figurative Language/Sound Devices	4	3	0
Research and Information	6	7	0
5.1 Accessing Information	2–4	2	0
5.2 Interpreting Information	2–4	5	0
Total Test	50	50	0

Table A10. OCCT Test Blueprint and Actual Item Counts: Grade 6 Reading

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Vocabulary	8	8	0
1.1 Words in Context	4	2	0
1.2 Word Origins	4	4	0
Comprehension/Critical Literacy	20	21	0
3.1 Literal Understanding	4	9	0
3.2 Inferences and Interpretation	4–6	2	0
3.3 Summary and Generalization	4–6	7	0
3.4 Analysis and Evaluation	4–6	2	0
Literature	14	14	0
4.1 Literary Genres	4	3	0
4.2 Literary Elements	4–6	6	0
4.3 Figurative Language/Sound Devices	4–6	5	0
Research and Information	8	7	0
5.1 Accessing Information	4	4	0
5.2 Interpreting Information	4	3	0
Total Test	50	50	0

Table A11. OCCT Test Blueprint and Actual Item Counts: Grade 7 Reading

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Vocabulary	10	11	0
1.1 Words in Context	3–4	4	0
1.2 Word Origins	3–4	4	0
1.3 Idioms and Comparisons	3–4	3	0
Comprehension/Critical Literacy	20	20	0
3.1 Literal Understanding	4–5	6	0
3.2 Inferences and Interpretation	4–6	5	0
3.3 Summary and Generalization	4–6	4	0
3.4 Analysis and Evaluation	4–6	5	0
Literature	12	12	0
4.1 Literary Genres	4	4	0
4.2 Literary Elements	4	3	0
4.3 Figurative Language/Sound Devices	4	5	0
Research and Information	8	7	0
5.1 Accessing Information	4	3	0
5.2 Interpreting Information	4	4	0
Total Test	50	50	0

Table A12. OCCT Test Blueprint and Actual Item Counts: Grade 8 Reading

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Vocabulary	6	4	0
1.1 Words in Context	2	1	0
1.2 Word Origins	2	1	0
1.3 Idioms and Comparisons	2	2	0
Comprehension/Critical Literacy	21	22	0
3.1 Literal Understanding	4–5	4	0
3.2 Inferences and Interpretation	4–6	5	0
3.3 Summary and Generalization	5–7	5	0
3.4 Analysis and Evaluation	6–8	8	0
Literature	15	14	0
4.1 Literary Genre	4–5	4	0
4.2 Literary Elements	5–7	3	0
4.3 Figurative Language/Sound Devices	4–6	7	0
Research and Information	8	10	0
5.1 Accessing Information	4	5	0
5.2 Interpreting Information	4	5	0
Total Test	50	50	0

Table A13. OCCT Test Blueprint and Actual Item Counts: Grade 5 Science

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS*	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Process Standards			
Observe and Measure	10	9	14
P1.1 SI Metric	4–6	4	6
P1.2 Similar/different characteristics	4	5	8
Classify	10	9	13
P2.1 Observable properties	4–6	4	7
P2.2 Serial order	4–5	5	6
Experiment	13–15	13	8
P3.2 Experimental design	9–11	9	6
P3.4 Hazards/practice safety	4	4	2
Interpret and Communicate	12–14	14	25
P4.2 Data tables/line/bar/trend and circle graphs	4–6	46	8
P4.3 Prediction based on data	4–6	4	11
P4.4 Explanations based on data	4–6	4	6
Total Test	45	45	60
Content Standards			
Properties of Matter and Energy	16–18	18	23
1.1 Matter has physical properties	4–5	5	6
1.2 Physical properties can be measured	4–5	5	4
1.3 Energy can be transferred	4–5	4	5
1.4 Potential/Kinetic Energy	4–5	4	8
Organisms and Environments	10–13	10	17
2.1 Organisms dependence	5–7	5	10
2.2 Individual organism and species survival	5–7	5	7
Structures of the Earth and the Solar System	12–15	13	18
3.1 Properties of Soils	4–6	4	6
3.2 Weather patterns	4–6	5	6
3.3 Earth as a planet	4	4	6
Total Test	41	41	58

* Items from the Safety Objective (P3.4) are not dual aligned to a content standard

Table A14. OCCT Test Blueprint and Actual Item Counts: Grade 8 Science

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS*	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Process Standards			
Observe and Measure	8–11	11	11
P1.1 Qualitative/quantitative observations/changes	4–6	5	4
P1.2 and P1.3 SI (metrics) units/appropriate tools	4–5	6	7
Classify	7–9	7	12
P2.1 Classification system	4–6	4	5
P2.2 Properties ordered	3–5	3	7
Experiment	15–17	14	18
P3.2 Experimental design	6–7	5	7
P3.3 Identify variables	6–7	6	6
P3.6 Hazards/practice safety	3–4	3	5
Interpret and Communicate	12–14	13	19
P4.2 Data tables/line/bar/trend and circle graphs	6–7	6	8
P4.3 Explanations/prediction	6–7	7	11
Total Test	45	45	60
Content Standards			
Properties and Chemical Changes in Matter	8	8	11
1.1 Chemical reactions	4	4	5
1.2 Conservation of matter	4	4	6
Motion and Forces	8	8	9
2.1 Motion of an object	4	4	4
2.2 Object subjected to a force	4	4	5
Diversity and Adaptations of Organisms	7	7	10
3.1 Classification	3	3	5
3.2 Internal and external structures	4	4	5
Structures/Forces of the Earth/Solar System	11	11	15
4.1 Landforms result from constructive and destructive forces	4	4	6
4.2 Rock cycle	3–4	4	4
4.3 Global Weather Patterns	3–4	3	5
Earth’s History	7–8	8	10
5.1 Catastrophic events	3–4	4	6
5.2 Fossil evidence	3–4	4	4
Total Test	41–42	42	55

* Items from the Safety Objective (P3.6) are not dual aligned to a content standard

Table A15. OCCT Test Blueprint and Actual Item Counts: Grade 5 Social Studies

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS*	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
James Towne Settlement and Plimoth Plantation Exploration	8	8	12
1.1, 1.2, 1.3, 1.4 James Towne Settlement	4	4	8
1.5 Plimoth Plantation	4	4	4
Colonial America	10	10	12
2.1, 2.3, 2.6 Colonial economics, trade/migration, perspectives	4–6	5	7
Self-government, role of religion, leaders, and 2.2, 2.4, 2.5 British and Native American Relationships	4–6	5	5
American Revolution	18	18	19
3.1 Causes and effects of American Revolution	4–6	5	7
3.2, 3.3, 3.4 Founding Documents of the Revolutionary Era	4–5	5	8
3.5 Events of the Revolutionary War	4–5	3	1
3.6 Key individuals of the Revolutionary Era	4–5	5	3
Early Federal Period	14	14	17
4.1, 4.2 Causes, leaders, and issues of the Constitutional Convention	4–5	7	8
4.3 Purposes and principles of the U.S. Constitution	4–6	3	5
4.4, 4.5 Ratification of the U.S. Constitution and the Bill of Rights	4–5	4	4
Total Test	50	50	60

Table A16. OCCT Test Blueprint and Actual Item Counts: Grade 7 Social Studies (Geography)

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Geographic Tools/Geography Skills	6	0	8
1.1,1.2,1.3,1.4,1.5 Geographic tools and skills	4–5	0	7
1.6 Freedom Week	1–2	0	1
Human and Physical Characteristics of Regions	12	0	16
2.1, 2.2 Political and Physical/Cultural Regions	4–6	0	8
2.3, 2.5 Physical and Human Characteristics Linking/Dividing Regions	4–6	0	3
2.4 Conflict and Cooperation	4–6	0	5
Physical Systems of the Earth	6	0	9
3.1 Visual Information, Landforms and Bodies of Water	2–4	0	3
3.2 Impact of Natural Disasters on Human Populations	4–5	0	6
Human Systems: People and Cultures	16	0	17
4.1, 4.2, 4.5 Cultural Traits, Major World Religions, and Major Political Systems	6–8	0	7
4.4, 4.6 Economic Systems, Economic Interdependence and Trade	4–5	0	5
4.3, 4.7 Human Characteristics of Developing and Developed Countries and Population Issues	4–5	0	5
Human Interaction with the Environment	10	0	10
5.1 Distribution of Resources	4–6	0	2
5.2, 5.3 Human Modification and Regional Problems	4–6	0	8
Total Test	50	0	60

Table A17. OCCT Test Blueprint and Actual Item Counts: Grade 8 Social Studies (U.S. History)

OAS Standard and Objective	Ideal Number of Items for Alignment to OAS	Actual Number of Items on 2014 Test	Number of Items Field Tested in 2014
Causes and Events of the American Revolution	8	8	6
1.1, 1.2 Consequences of the French and Indian War, British Imperial Policies	4	4	3
1.3, 1.4, 1.5 Ideological War, Declaration of Independence’s Grievances, Ideals, and Social Contract Selection	4	4	3
The Revolutionary Era (2.0)	6	6	6
2.1, 2.2, 2.3 Articles of Confederation, Motivations & Choices, Key Military & Diplomatic Events	6	6	6
Developing the American Government System (3.0)	10	10	18
3.1, 3.2, 3.3 Causes for the Constitutional Convention, and Ratification	4–6	5	6
3.4, 3.5 Constitutional Principles and the Bill of Rights	4–6	5	12
The Transformation of the United States to the Mid-1800s	16	16	19
4.1 Major Events and Issues of Early Presidential Administrations	4–6	6	3
4.2, 4.6 Jacksonian Era and Westward Expansion	4–6	5	8
4.3, 4.4, 4.5 Sectional Economic Systems, African American Experiences, and Reform Movements/Leaders	4–6	5	8
Causes, Events, and Leadership in the Civil War	10	10	11
5.1, 5.2 Causes of the Civil War: 1850s through the 1860s Presidential Elections	4–6	5	4
5.3, 5.4 Advantages/Disadvantages, Leadership, Major Turning Points of the War	4–6	5	7
Total Test	50	50	60

Appendix B
Spring 2014 OK Grades 3 to 8 Writing Linking Study

SPRING 2014 OK GRADES 3 TO 8 WRITING LINKING STUDY

TO: LISA CHANDLER, JOYCE DEFEHR
FROM: DONG-IN KIM, LITONG ZHANG
SUBJECT: WRITING EQUATING FOR GRADES 5 AND 8
DATE: MAY 4, 2014
CC: TRACY KEITH, LINDY WIENAND, AMANDA BREITMAIER,
 SHIVA DORESWAMY, DAVID COSIO

This document serves to describe the analyses conducted for equating of the OK Grades 5 and 8 Writing tests. All analyses results can be made available upon request.

In Spring 2013, one operational Writing prompt and four Field Test (FT) prompts were administered in grades 5 and 8. An analytic scoring was applied to the Writing prompts. Each Writing prompt consisted of 5 traits, and each trait score ranged from 1.0 to 4.0. Condition codes were assigned to a 0 score point (Please see **Appendix A. Spring 2014 OCCT Grades 5 & 8 Writing Composite Scoring**). Final Writing composite scores ranged from 15 to 60. Equipercentile linking method was performed among 5 prompts: one operational Writing prompt and four FT Writing prompts, to align the 5 Writing prompts to the same scale in Spring 2013.

In Summer 2013, standard setting was performed using the Spring 2013 operational Writing prompt, and final cut scores were decided.

Table 1. OCCT Grades 5 & 8 Writing Cut Score Ranges and Impact Data

Grade	Unsatisfactory	Limited Knowledge	Proficient	Advanced
5	15-22(6)	23-35(36)	36-47(56)	48-60(2)
8	15-24(9)	25-35(34)	36-49(50)	50-60(7)

*values in parenthesis are percent (%) for each Performance Level

Table 2 shows that Information Writing prompts were administered as 2013 operational Writing prompts and opinion/argument Writing prompts were administered as 2014 operational Writing prompts. According to CTB Content, 2014 operational Writing prompts are not intact 2013

Writing prompts, which were used for Equipercentile linking. 2013 Writing prompts were vastly modified to be used for the 2014 administration. Therefore, the 2013 FT prompts and 2014 operational prompts are not the same, and the 2013 linking study results cannot be used.

Table 2. Genre of Writing Prompts for 2013 and 2014

Grade	2013 OP	2013 FT
5	Info WP (camouflage and vision of animals was the topic)	Narrative WP (visiting a farm) Opinion WP (screen-free day) Narrative WP (Hershey) Opinion WP (animal helpers)
8	Info WP (brain power)	Info WP (volunteering) Argument WP (animal protection) Argument WP (Olympics) Narrative WP (cowboy life)

*2013 and 2014 Operational items are bolded in red

To apply the 2013 cut scores from 2013 standard setting to 2014 Writing scores, 2014 Writing prompts need to be placed on the scale of the 2013 operational Writing prompt. Since there are no anchor items, such as Writing multiple-choice items, for OCCT Grades 5 & 8, there is no way to apply a common item linking design. Also mentioned above, the 2013 linking study results cannot be applied because 2013 Writing prompts were modified after the linking study. So, SDE and CTB have discussed and agreed to apply equipercentile linking under the assumption that there is flat growth of Oklahoma students between 2013 and 2014 Writing performance.

Equating Sample

2014 Census data was used for equating analysis, such that no special sampling was conducted. All students who took Braille form were excluded in this study. The following exclusion rules were applied to clean the data set:

1. A second time test-taker
2. Had an invalidated condition code
3. Marked NFAY in school, district, and state
4. Attended a private schools

Table 3 shows the number of students considered and the number of valid students included in the analysis. Invalid student count in each exclusion category is also presented here.

Table 3. Case Counts Summary for Linking Study

Year	Grade	Total	Valid	Exclusion				
				Total (1+2+3+4)	1	2	3	4
2013	5	49198	43504	5694	0	2002	3758	0
2014	5	50794	45242	5552	149	3109	2496	0
2013	8	48115	42851	5264	0	1731	3708	0
2014	8	50166	44702	5464	142	2860	2696	0

A, B, C, D and U are 5 invalidation conditional codes used when a score cannot be assigned:

- A: Blank/No responses
- B: Illegible
- C: Other language
- D: Off topic
- U: Unclear Image

Their frequencies are presented in Table 4.

Table 4. Case Counts for Invalidation Condition Codes

Year	Grade	A	B	C	D	U	Total
2013	5	1609	5	7	379	2	2002
2014	5	2834	19	12	244	0	3109
2013	8	1645	2	2	82	0	1731
2014	8	2843	4	2	11	0	2860

Note that all students with condition cords were excluded in this study.

Linking Result

A concordance table from equipercentile linking includes two scores, Writing composite scores and their corresponding equivalent scores. Equipercentile equating determines the equating relationship as one where a score could have an equivalent percentile on either form. To keep the same range of scores 15~60, a minimum possible score of 15 and maximum possible score of 60 were assigned as equivalent scores of composite scores 15 and 60.

Tables 5.1 and 5.2 show the final concordance tables for grades 5 and 8 respectively. In these tables the first column shows 2014 Writing composite scores, which were calculated using Writing scoring rules in Appendix A, and the second column shows their equivalent (concordance) scores from equipercentile linking. For example, let's suppose a student gets a 2014 Writing composite score of 39 in grade 5. Then, the student's performance level is "Proficient" because the Proficient cut score is at 36 and his equivalent score is 43, which is above the cut score. For grade 5, score points of 17, 31, and 53 are cut scores for Limited Knowledge, Proficient, and Advanced, respectively. For grade 8, score points of 19, 35, and 53 are cut scores for Limited Knowledge, Proficient, and Advanced, respectively.

Tables 6.1 and 6.2 compare the writing score frequency distributions between Spring 2013 and Spring 2014 administrations. Table 7 shows the percentages of each performance level for these administrations. As can be seen in Table 7, pass rates for Spring 2013 and Spring 2014 grade 5 are 58.01% and 47.82%, and those for Spring 2013 and Spring 2014 grade 8 are 57.43% and 57.42%. Grade 5 shows that there are about 10% passing rate decrease between 2013 and 2014. This happens because about 24% of 2014 students got a composite score of 30, and a composite score of 31 is the pass cut score as can be seen in Table 6.1. If we choose a composite score of 30, one score lower than the current pass cut, pass rate for Spring 2014 becomes 72%. This means there is about 14% (=72%-58%) pass rate increase between 2013 and 2014. Please note that Out of about 11,000 students with composite score point of 30, most students received 2 trait score points across the 5 traits from both raters.

Table 5.1 Concordance Table between 2013 and 2014 Grade 5 Writing Scores

2014 Writing Composite Score	Equivalent (Concordance) Score	Comment
15	15	
16	23	
17	23	Limited Knowledge
18	24	
19	24	
20	25	
21	26	
22	26	
23	28	
24	30	
25	30	
26	30	
27	31	
28	31	
29	32	
30	35	
31	38	Proficient
32	39	
33	39	
34	40	
35	40	
36	41	
37	41	
38	42	
39	43	
40	44	
41	44	
42	44	
43	45	
44	45	
45	45	
46	46	
47	46	
48	46	
49	46	
50	46	
51	46	

2014 Writing Composite Score	Equivalent (Concordance) Score	Comment
52	46	
53	48	Advanced
54	50	
55	51	
56	51	
57	52	
58	52	
59	53	
60	60	

Table 5.2 Concordance Table between 2013 and 2014 Grade 8 Writing Scores

2014 Writing Composite Score	Equivalent (Concordance) Score	Comment
15	15	
16	24	
17	24	
18	24	
19	25	Limited Knowledge
20	25	
21	26	
22	26	
23	29	
24	30	
25	30	
26	30	
27	30	
28	30	
29	30	
30	32	
31	34	
32	35	
33	35	
34	35	
35	36	Proficient
36	36	
37	37	
38	38	

2014 Writing Composite Score	Equivalent (Concordance) Score	Comment
39	41	
40	41	
41	41	
42	42	
43	42	
44	42	
45	45	
46	47	
47	47	
48	47	
49	48	
50	48	
51	49	
52	49	
53	52	Advanced
54	54	
55	54	
56	54	
57	54	
58	55	
59	56	
60	60	

Table 6.1 Frequency Comparison between 2013 and 2014 for Grade 5

Composite Score	Spring 2013				Spring 2014			
	FD	%FD	CFD	%CFD	FD	%FD	CFD	%CFD
15	1442	3.31	1442	3.32	3348	7.40	3348	7.40
16	178	0.41	1620	3.72	126	0.28	3474	7.68
17	188	0.43	1808	4.16	416	0.92	3890	8.60
18	250	0.57	2058	4.73	260	0.57	4150	9.17
19	40	0.09	2098	4.82	126	0.28	4276	9.45
20	232	0.53	2330	5.36	450	0.99	4726	10.45
21	279	0.64	2609	6.00	321	0.71	5047	11.16
22	131	0.30	2740	6.30	54	0.12	5101	11.28
23	958	2.20	3698	8.50	3424	7.57	8525	18.84
24	481	1.11	4179	9.61	421	0.93	8946	19.77
25	470	1.08	4649	10.69	631	1.39	9577	21.17
26	646	1.48	5295	12.17	650	1.44	10227	22.61
27	503	1.16	5798	13.33	537	1.19	10764	23.79
28	862	1.98	6660	15.31	1255	2.77	12019	26.57
29	632	1.45	7292	16.76	725	1.60	12744	28.17
30	2950	6.78	10242	23.54	10862	24.01	23606	52.18
31	1448	3.33	11690	26.87	1031	2.28	24637	54.46
32	1845	4.24	13535	31.11	1912	4.23	26549	58.68
33	2119	4.87	15654	35.98	1308	2.89	27857	61.57
34	458	1.05	16112	37.04	804	1.78	28661	63.35
35	2154	4.95	18266	41.99	1757	3.88	30418	67.23
36	2364	5.43	20630	47.42	1228	2.71	31646	69.95
37	1315	3.02	21945	50.44	233	0.52	31879	70.46
38	2613	6.01	24558	56.45	4460	9.86	36339	80.32
39	2464	5.66	27022	62.11	740	1.64	37079	81.96
40	1569	3.61	28591	65.72	675	1.49	37754	83.45
41	2964	6.81	31555	72.53	737	1.63	38491	85.08
42	1245	2.86	32800	75.40	608	1.34	39099	86.42
43	3057	7.03	35857	82.42	892	1.97	39991	88.39
44	981	2.26	36838	84.68	448	0.99	40439	89.38
45	5016	11.53	41854	96.21	2348	5.19	42787	94.57
46	331	0.76	42185	96.97	175	0.39	42962	94.96
47	334	0.77	42519	97.74	249	0.55	43211	95.51
48	223	0.51	42742	98.25	133	0.29	43344	95.81

Spring 2013					Spring 2014			
Composite Score	FD	%FD	CFD	%CFD	FD	%FD	CFD	%CFD
49	90	0.21	42832	98.46	111	0.25	43455	96.05
50	162	0.37	42994	98.83	210	0.46	43665	96.51
51	171	0.39	43165	99.22	191	0.42	43856	96.94
52	12	0.03	43177	99.25	10	0.02	43866	96.96
53	222	0.51	43399	99.76	787	1.74	44653	98.70
54	27	0.06	43426	99.82	81	0.18	44734	98.88
55	9	0.02	43435	99.84	82	0.18	44816	99.06
56	24	0.06	43459	99.90	44	0.10	44860	99.16
57	8	0.02	43467	99.92	49	0.11	44909	99.26
58	14	0.03	43481	99.95	65	0.14	44974	99.41
59	5	0.01	43486	99.96	48	0.11	45022	99.51
60	18	0.04	43504	100.00	220	0.49	45242	100.00

Table 6.2 Frequency Comparison between 2013 and 2014 for Grade 8

Spring 2013					Spring 2014			
Composite Score	FD	%FD	CFD	%CFD	FD	%FD	CFD	%CFD
15	1161	2.71	1161	2.71	3560	7.96	3560	7.96
16	97	0.23	1258	2.94	92	0.21	3652	8.17
17	218	0.51	1476	3.44	198	0.44	3850	8.61
18	188	0.44	1664	3.88	136	0.30	3986	8.92
19	67	0.16	1731	4.04	121	0.27	4107	9.19
20	189	0.44	1920	4.48	242	0.54	4349	9.73
21	182	0.42	2102	4.91	206	0.46	4555	10.19
22	71	0.17	2173	5.07	24	0.05	4579	10.24
23	1338	3.12	3511	8.19	3485	7.80	8064	18.04
24	265	0.62	3776	8.81	151	0.34	8215	18.38
25	430	1.00	4206	9.82	258	0.58	8473	18.95
26	558	1.30	4764	11.12	346	0.77	8819	19.73
27	431	1.01	5195	12.12	386	0.86	9205	20.59
28	625	1.46	5820	13.58	554	1.24	9759	21.83
29	438	1.02	6258	14.60	413	0.92	10172	22.76
30	5711	13.33	11969	27.93	7208	16.13	17380	38.88
31	913	2.13	12882	30.06	346	0.77	17726	39.65
32	1542	3.60	14424	33.66	472	1.06	18198	40.71

Composite Score	Spring 2013				Spring 2014			
	FD	%FD	CFD	%CFD	FD	%FD	CFD	%CFD
33	1606	3.75	16030	37.41	401	0.90	18599	41.61
34	694	1.62	16724	39.03	434	0.97	19033	42.58
35	1519	3.54	18243	42.57	706	1.58	19739	44.16
36	1524	3.56	19767	46.13	758	1.70	20497	45.85
37	325	0.76	20092	46.89	86	0.19	20583	46.05
38	4669	10.90	24761	57.78	7537	16.86	28120	62.91
39	1091	2.55	25852	60.33	397	0.89	28517	63.79
40	1162	2.71	27014	63.04	386	0.86	28903	64.66
41	1580	3.69	28594	66.73	587	1.31	29490	65.97
42	1245	2.91	29839	69.63	512	1.15	30002	67.12
43	1606	3.75	31445	73.38	807	1.81	30809	68.92
44	955	2.23	32400	75.61	643	1.44	31452	70.36
45	5960	13.91	38360	89.52	8968	20.06	40420	90.42
46	363	0.85	38723	90.37	122	0.27	40542	90.69
47	470	1.10	39193	91.46	164	0.37	40706	91.06
48	352	0.82	39545	92.29	95	0.21	40801	91.27
49	338	0.79	39883	93.07	182	0.41	40983	91.68
50	396	0.92	40279	94.00	241	0.54	41224	92.22
51	386	0.90	40665	94.90	269	0.60	41493	92.82
52	14	0.03	40679	94.93	13	0.03	41506	92.85
53	1367	3.19	42046	98.12	2223	4.97	43729	97.82
54	77	0.18	42123	98.30	59	0.13	43788	97.96
55	71	0.17	42194	98.47	40	0.09	43828	98.05
56	88	0.21	42282	98.67	52	0.12	43880	98.16
57	100	0.23	42382	98.91	52	0.12	43932	98.28
58	98	0.23	42480	99.13	89	0.20	44021	98.48
59	60	0.14	42540	99.27	72	0.16	44093	98.64
60	311	0.73	42851	100.00	609	1.36	44702	100.00

Table 7 Performance Level Percentage Comparison between 2013 and 2014

Grade	Spring 2013					Spring 2014				
	1	2	3	4	Passing	1	2	3	4	Passing
5	6.30	35.69	55.75	2.26	58.01	7.68	44.50	44.78	3.05	47.82
5	6.30	35.69	55.75	2.26	58.01	7.68	20.49	68.79	3.05	71.84
8	8.81	33.76	50.50	6.93	57.43	8.92	33.66	50.27	7.15	57.42

- 1: Unsatisfactory; 2: Limited Knowledge; 3: Proficient; 4: Advanced
- Passing = 3 + 4

Appendix A. Spring 2014 OCCT Grades 5 & 8 Writing Composite Scoring

- The writing prompts are read for 5 analytic traits.
- Each analytic trait has a score range of 1.0 – 4.0. A ‘0’ is not a valid score.

Condition Codes

Condition codes are assigned when a score cannot be assigned. The Description listed below is the actual text that will appear on the screen for the rater/reader to see when assigning a condition code.

The order of the condition codes shown below is how they will appear on reports.

Code	EHS Code Description	Description	Comments
B	Illegible/Incomprehensible	Response is illegible or incomprehensible	Will be translated to “I” for reporting
C	Language Other than English	Written predominantly in a language other than English	Will be translated to “L” for reporting
A	Blank/No Response/Refusal	Blank/no response/refusal to answer	Will be translated to “N” for reporting
D	Off Topic	Response off the topic of the writing task	Will be translated to “O” for reporting
U	Unclear image		Not reported

Writing Composite Score (Grades 5 and 8):

Composite Score

A student’s composite score on the Writing assessment, in part, is derived by assigning various weights to the five analytic traits. The averaged analytic score for each category is multiplied by the appropriate weight (percentage) and summed. The sum is then multiplied by 15 to place the score on the appropriate scale. The score is then rounded up to the nearest whole number. Each student’s composite score will range from 60 (the highest score) to 15 (the lowest score). The weights attributed to each analytic score are given in the table below.

Composite Score

Percentage	Analytic Score Category
30%	Ideas and Development
25%	Organization, Unity, and Coherence
15%	Word Choice
15%	Sentences and Paragraphs
15%	Grammar, Usage, and Mechanics

Steps to Calculate the Writing Composite Score

The steps outlined below show how Writing scores are calculated based on the trait scores in one writing prompt. The table gives an example of how Writing scores will be calculated.

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, and write the results in Column F.

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

STEP 4: Multiply the sum from Column F by 15. This is the Raw Composite Score.

STEP 5: Round this transformed Writing score to the nearest whole number to obtain the final Writing score. After calculation, the final Writing score value will range from 15 to 60.

Calculating Writing Composite scores

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 × E)
Ideas and Development	.30	3	2	$(3+2)/2=2.5$	$.30 \times 2.5 = 0.75$
Organization, Unity, and Coherence	.25	3	3	$(3+3)/2=3.0$	$.25 \times 3.0 = 0.75$
Word Choice	.15	3	2	$(3+2)/2=2.5$	$.15 \times 2.5 = 0.375$
Sentences and Paragraphs	.15	2	3	$(2+3)/2=2.5$	$.15 \times 2.5 = 0.375$
Grammar/Usage and Mechanics	.15	3	2	$(3+2)/2=2.5$	$.15 \times 2.5 = 0.375$
					Sum Above
					= 2.625

$2.625 \times 15 = 39.375$

Writing Composite Score = 39