

**Technical Report**  
**Proficiency Assessments for Wyoming Students (PAWS) and Student**  
**Assessment of Writing Skills (SAWS)**

Reading and Mathematics: Grades 3–8

Science: Grades 4 and 8

Writing: Grades 3, 5, and 7

2013-2014 Administration

Submitted to Wyoming Department of Education

Prepared by  
Educational Testing Service

December 1, 2014

Final

Copyright © 2014 by Educational Testing Service.

All rights reserved

The Wyoming Department of Education does not discriminate on the basis of race, color, national origin, sex, age, or disability in admission or access to, or treatment of employment in its educational programs or activities. Inquiries concerning Title VI, Title IX, Section 504, and the Americans with Disabilities Act may be referred to the Wyoming Department of Education, Office for Civil Rights Coordinator, 2nd floor, Hathaway Building, Cheyenne, Wyoming 82002-0050 or (307) 777-7673, or the Office for Civil Rights, Region VIII, U.S. Department of Education, Federal Building, Suite 320, 1244 Speer Boulevard, Denver, CO 80204-3582, or (303) 844-5695 or TDD (303) 844-3417.

This report is available on the World Wide Web at <http://edu.wyoming.gov/default.aspx>

***Wyoming Department of Education***

Cindy Hill

Superintendent of Wyoming Department of Education

***Wyoming Department of Education Assessment Team***

Deb Lindsay

Assessment Division Administrator

Jessica Steinbrenner

Assessment Division Supervisor

Bailey Anthony and Michelle Carroll

Assessment Consultants

***Educational Testing Service***

Lauren Hebert

Wyoming Program Manager

Rebecca Rabin-Reed

SAWS Program Manager

Valerie Link

Senior Psychometrician

Junhui Liu

Psychometrician

Hyeonjoo Oh

Psychometric Manager

Mark Batleman

Principal Statistical Assoc - Level 1

William Bennett

Statistical Assistant-Senior

Su Liu

Data Analyst Manager

## Executive Summary

This report presents the technical results for the Proficiency Assessments for Wyoming Students (PAWS) 2013–2014 school year. The document covers the Grades 3 to 8 Reading and Mathematics administration and the Grades 4 and 8 Science administration. Additionally, the Grades 3, 5, and 7 administrations of the Student Assessment of Writing Skills (SAWS) are included. The populations are composed of on-grade students at each grade level. There were approximately 6,760 to 7,370 students in the total Wyoming student population at grades 3 to 8.

### Structure of This Report

The initial chapter provides relevant policy decisions regarding the PAWS and SAWS programs, followed by brief descriptions of the PAWS and SAWS as it was administered during the 2013–2014 school year. The reliability and validity chapters present the evidence gathered to support the intended uses and interpretations of scores for the PAWS and SAWS assessment programs. In short, the validity process began with test design and continued through the entire assessment process, including item development and field testing, analyses of item and test data, test scaling, scoring, and score reporting. Each of these processes is described in detail in the starting with Test Design and Development chapter, and concluded in the Historical Comparisons chapter. Operational aspects of the program are discussed in the remaining chapters.

### New Reading and Mathematics Assessments

The PAWS Reading and Mathematics assessments experienced significant changes in 2014. The 2014 PAWS Reading and Mathematics assessments utilized the Common Core Strands as reporting categories and becomes the new scale measuring students' academic performance on 2012 Wyoming Content Performance Standards (2012 WyCPS).

The Reading vertical scale allows for direct comparisons of student test scores across grade levels within a content area. More details regarding the Reading Vertical scale development can be found in the separate *Proficiency Assessments for Wyoming Students (PAWS) 2014 Calibration and Vertical Scale Report Reading Grades 3–8* document.

The Mathematics vertical span scale was designed to address the Wyoming mathematics blueprint having three content continuums across grades. This scale divides grades 3–8 mathematics tests into 3 spans (Span I: grades 3-5; Span II: grades 6-7; and Span III: grade 8). As shown in the section of PAWS test design, the reporting categories of the PAWS Mathematics assessments are the same within each span but different across spans.

### Other Program Changes

The grade 8 science administration moved to a non-consumable test booklet and consumable answer document. The WDE made the end-users aware of this change and an opportunity was provided to practice bubbling in answer documents to the students. The test design and processes were not changed in Science grade 4 2014 administration .

The SAWS program is currently in transition, with the expansion from a single writing prompt to two writing prompts in grade 3. There has also been an addition of the 4 + 8 point set for grades 5 and 7 for 2014.

A standard setting meeting was held in July to set the cut scores for different performance levels for Reading, Mathematics, and SAWS. PAWS and SAWS standard setting activities are described in detail in the *2014 Standard Setting Summary Proficiency Assessments for Wyoming Students (PAWS for Reading, Mathematics) and Student Assessment of Writing Skills (SAWS)* (Baron, 2014) report.

### Conclusion

The technical efforts conducted in 2013–14, described later in this report, demonstrated the fidelity of the program to its long standing levels of validity and reliability. This quality was retained through diligent compliance to procedures and high caliber judgment and evaluation of numerous national professionals from the field of test measurement.

## Table of Contents

Executive Summary .....	v
Structure of This Report.....	v
New Reading and Mathematics Assessments.....	v
Other Program Changes .....	vi
Conclusion .....	vi
LIST OF TABLES .....	xvi
LIST OF FIGURES .....	xxv
1. OVERVIEW OF THE 2014 PAWS AND SAWS .....	1
1.1 Introduction.....	1
1.2 Background of PAWS and SAWS.....	1
1.3 Overview of PAWS and SAWS Test Components .....	5
1.4 Overview of the PAWS and SAWS Design .....	5
1.5 State Policy on Student Participation.....	6
1.5.1 Students with Disabilities, 504 Plans, and English Language Learners .....	7
2. VALIDITY .....	12
2.1 Overview.....	12
2.1.1. Purpose of the PAWS and SAWS .....	13
2.1.2. The Constructs to Be Measured.....	13
2.1.3. The Interpretations and Uses of the Scores Generated .....	13
2.1.4. Intended Test Population(s) .....	14
2.2 Evidence Based on Content-related Validity .....	14
2.2.1. Description of the state standards .....	15
2.2.2. Specifications and Blueprints .....	15
2.2.3. Item development process.....	15
2.2.4. Item review process .....	15
2.2.5. Form construction process .....	15
2.2.6. Alignment study.....	16
2.3 Evidence Based on Relations to Other Variables .....	16

2.3.1. Correlations between Content Areas .....	16
2.3.2. Differential Item Functioning Analyses.....	17
2.4 Evidence Based on Response Processes .....	17
2.4.1. Evidence of Interrater Agreement.....	17
2.4.1.1. Interrater Reliability.....	17
2.4.1.2. Interrater Agreement.....	18
2.5 Evidence Based on Internal Structure.....	18
2.5.1. Classical Statistics.....	18
2.5.2. Reliability.....	19
2.5.3. Dimensionality.....	19
2.6 Evidence Based on Consequences of Testing.....	20
3. PAWS AND SAWS TEST DESIGN AND DEVELOPMENT .....	21
3.1 Overview.....	21
3.2 Test Design and Blueprints.....	21
3.2.1. Purpose.....	21
3.2.2. Plan .....	22
3.3 Types of Items Used in PAWS and SAWS .....	22
3.3.1. PAWS Reading Tests.....	23
3.3.2. PAWS Mathematics Tests .....	26
3.3.3. PAWS Science Tests.....	29
3.3.4. SAWS Tests .....	32
12-point prompt .....	34
8-point Prompt, Grade 3 .....	34
4- and 8-point Response-to-Text Prompt Sets, Grades 5 and 7 .....	35
3.4 PAWS Test Development Process.....	35
3.5 Item and Test Form Development .....	36
3.5.1. Item Specifications.....	37
3.5.2. Item Difficulty Requirements .....	37
3.5.3. Item Graphics Requirements.....	37
3.6 Item Review .....	38



3.7 Field Testing .....	39
3.7.1. Classical Item Statistics .....	40
3.7.1.1. Item Difficulty .....	41
3.7.1.2. Item Discrimination .....	41
3.7.2. Differential Item Functioning .....	41
3.7.3. Item Response Theory (IRT) Analysis .....	43
3.7.4. Data Review Procedures .....	44
3.8 Test Form Construction .....	45
3.8.1. Construction of the Reading and Mathematics Forms.....	46
3.8.2. Construction of the Science Forms .....	46
3.8.3. Construction of the SAWS Forms .....	47
3.8.4. Final Review of Assembled Operational Tests.....	48
4. TEST ADMINISTRATION .....	49
4.2 Materials Return.....	49
4.3 Directions for Administration .....	49
4.3.1. Allowed Student Manipulatives.....	50
4.3.2. Test Security .....	50
4.4 Student Participation.....	51
4.5 PAWS and SAWS Standard Accommodations .....	51
4.5.1. Students Eligible for Test Accommodations .....	53
4.5.2. Requirements for Use of Test Accommodations .....	53
4.5.3. Description of Standard Accommodations for Students with Disabilities .....	55
4.5.3.1. Presentation Accommodations.....	55
4.5.3.2. Response Accommodations .....	56
4.5.3.3. Setting Accommodations .....	57
4.5.3.4. Timing and Scheduling Accommodations.....	57
4.5.4. Description of Standard Accommodations for English Language Learners (ELL) ...	57
4.5.4.1. Presentation Accommodations.....	58
4.5.4.2. Setting Accommodations .....	58
4.5.4.3. Timing and Scheduling Accommodations.....	58

4.5.5. PAWS 2014 Monitoring of Appropriate Accommodations .....	58
4.5.5.1. Empirical Analysis of Accommodations .....	59
4.5.6. Selection and Administration of Accommodations .....	59
5. PROCESSING AND SCORING OF PAWS AND SAWS ITEMS .....	61
5.1 Overview .....	61
5.1.1 Multiple Choice Items.....	61
5.1.2 SAWS Prompts .....	61
5.2 Receipt Control, Processing, Scanning, Editing .....	61
5.2.1 Receipt Control .....	62
5.2.2 Processing .....	62
5.2.3 Scanning.....	62
5.2.4 Editing.....	63
5.3 Scorer Recruitment and Qualifications .....	64
5.3.1 Eligibility .....	64
5.3.2 Rater Qualifications .....	64
5.3.2.1. Training.....	64
5.3.2.2 Certification Procedures.....	65
5.4 Topic Notes and Annotations Meeting and Operational Rangefinding Meeting .....	65
5.4.1. Operational Rangefinding Meeting.....	65
5.4.2. Pre-Rangefinding Procedures .....	65
5.4.3. Rangefinding Procedures .....	66
5.5: Hand-Scoring Process.....	67
5.5.1. Scoring Responses to SAWS Writing Prompts .....	67
5.5.2. About SAWS Scoring.....	67
5.5.3. Reference Materials for Scoring .....	67
5.5.4. Guidance to Raters: Points to Remember about SAWS Scoring.....	67
5.6: Procedures for Maintaining and Retrieving Individual Scores.....	68
5.7 Interrater Reliability.....	68
5.8 Accuracy Monitoring.....	69
5.9 Blanks and Invalid Responses for SAWS.....	70

5.9.1. Blank (BB) .....	70
5.9.2. Copy of Prompt (CP) .....	70
5.9.3. Foreign Language (NL) .....	70
5.9.4. Illegible (IL) .....	71
5.9.5. Incomprehensible (IN) .....	71
5.9.6. Insufficient to Score (IS) .....	71
5.9.7. Refusal (XX) .....	71
5.10 Reporting of PSC Alerts .....	71
5.10.1. Policy on the Reporting of Alerts .....	71
5.10.2. Reporting Procedure .....	72
6. EQUATING AND SCALING PROCEDURES .....	73
6.1 Overview .....	73
6.2 Item and Forms Development .....	73
6.3 IRT Models and Calibrations .....	74
6.4 Fit Statistics for the Rasch Model .....	75
6.5 Reading and Mathematics Vertical Scales .....	76
6.6 Science Equating Analyses .....	77
6.6.1. Calibration and Equating Process for the 2014 Science Administration .....	78
6.7 Translating Raw Scores to Scaled Scores and Performance Levels .....	80
7. PAWS AND SAWS REPORTING .....	82
7.1 Overview .....	82
7.2 Scaled Scores .....	82
7.3 Performance Levels .....	83
7.4 Content Standard-Level Raw and Scale Scores .....	84
7.5 SAWS Raw Scores .....	84
7.6 Production of Printed Score Reports for PAWS and SAWS .....	85
7.7 Assessment Score Reports: Supplement Guide for Districts and Schools for PAWS and SAWS .....	86
8. RELIABILITY .....	87
8.1 Overview .....	87

8.2 Internal Consistency Reliability.....	87
8.3 Interrater Reliability.....	91
8.4 Weighted Kappa.....	93
8.5 Classical and Conditional Standard Errors of Measurement .....	94
8.6 Accuracy and Consistency of Classifications .....	95
9. QUALITY CONTROL PROCEDURES .....	100
9.1 Quality Control of Item Development .....	100
9.1.1. Item and Prompt Specifications .....	100
9.1.2. Item Writers .....	100
9.1.3. Internal Contractor Reviews .....	100
9.1.4. Content Review.....	101
9.1.5. Editorial Review .....	101
9.1.6. Fairness Review .....	101
9.1.7. Assessment Director Review .....	101
9.1.8. Data Review of Field Tested Items.....	101
9.1.9. Quality Control of the Item Bank .....	101
9.2 Quality Control of Test Materials .....	102
9.2.1. Collecting Test Materials.....	102
9.2.2. Processing Test Materials .....	102
9.3 Quality Control of Scanning .....	103
9.4 Quality Control of Psychometric Analyses.....	103
9.5 Quality Control of Reporting .....	103
9.5.1. Excluding Student Scores from Summary Reports .....	104
10. HISTORICAL COMPARISONS .....	105
11. REFERENCES .....	110
12. GLOSSARY OF TERMS .....	113
Appendix A: PAWS 2014 Reading, Math, Science, and SAWS Blueprints .....	119
Reading .....	119
Mathematics .....	137
Science .....	143

SAWS .....	147
Appendix B: Sample PAWS/SAWS Student Reports: Grade 3 Reading, Mathematics, and SAWS (exemplar for Grades 5 and 7) .....	149
Appendix C: Sample PAWS/SAWS Student Reports: Grade 4 Reading, Mathematics, and SAWS (exemplar for Grade 8) .....	152
Appendix D: Sample PAWS/SAWS Student Reports: Grade 6 Reading and Mathematics .....	155
Appendix E: SAWS Field Test Subscale Correlations .....	158
Appendix F: PAWS and SAWS Operational Subscale Correlations.....	161
PAWS .....	161
SAWS .....	167
Appendix G: DIF Results for Field Test 2014 Items .....	169
Reading .....	169
Mathematics .....	173
Science .....	176
SAWS .....	177
Appendix H: SAWS Field Test Rater Reliability .....	179
Appendix I: SAWS Operational Rater Reliability .....	183
Appendix J: Classical Item Statistics for 2014 Field Test Items .....	185
Reading .....	185
Mathematics .....	211
Science .....	234
SAWS .....	242
Appendix K: Classical Item Statistics for 2014 Operational Items .....	253
Reading .....	253
Mathematics .....	265
Science .....	277
SAWS .....	281
Appendix L: Rasch Difficulty, Standard Error, Fit Statistics, and <i>N</i> -counts for 2014 Field Test Items.....	286
Reading .....	286

Mathematics .....	306
Science .....	324
Appendix M: Rasch Difficulty, Standard Error, and Fit Statistics for 2014 Operational Items. ....	332
Reading .....	332
Mathematics .....	344
Science .....	356
Appendix N: Frequency of Individual Accommodations for 2014 PAWS and SAWS Tests....	360
Reading .....	360
Mathematics .....	363
Science .....	366
SAWS .....	369
Appendix O: Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status .....	372
Reading .....	372
Mathematics .....	372
Science .....	373
SAWS .....	373
Appendix P: Scaled Score Descriptive Statistics by Demographic Subgroup .....	374
Reading .....	375
Mathematics .....	378
Science .....	381
SAWS .....	382
Appendix Q: Raw Score to Scaled Score Tables.....	392
Reading .....	392
Mathematics .....	404
Science .....	418
SAWS .....	422
Appendix R: Performance Level Percentages by Demographic Subgroup .....	425
Reading .....	426
Mathematics .....	429

Science .....	432
SAWS .....	433
Appendix S: SAWS Field Test Demographic Performance .....	435

## LIST OF TABLES

Table 1. Statewide Participation in Reading PAWS.....	8
Table 2. Statewide Participation in Mathematics PAWS .....	9
Table 3. Statewide Participation in Science PAWS.....	10
Table 4. Statewide Participation in SAWS .....	11
Table 5. Reporting Strands Design for Grade 3 Reading .....	25
Table 6. Reporting Strands Design for Grade 4 Reading .....	25
Table 7. Reporting Strands Design for Grade 5 Reading .....	25
Table 8. Reporting Strands Design for Grade 6 Reading .....	25
Table 9. Reporting Strands Design for Grade 7 Reading .....	26
Table 10. Reporting Strands Design for Grade 8 Reading .....	26
Table 11. Reporting Strands Design for Grade 3 Mathematics .....	27
Table 12. Reporting Strands Design for Grade 4 Mathematics .....	28
Table 13. Reporting Strands Design for Grade 5 Mathematics .....	28
Table 14. Reporting Strands Design for Grade 6 Mathematics .....	28
Table 15. Reporting Strands Design for Grade 7 Mathematics .....	28
Table 16. Reporting Strands Design for Grade 8 Mathematics .....	29
Table 17. Reporting Strands Design for Grade 4 Science .....	31
Table 18. Reporting Strands Design for Grade 8 Science .....	32
Table 19. SAWS Standards.....	33
Table 20. SAWS Grades and Content.....	33
Table 21. SAWS Grades, Prompt Type, and Field Test Design.....	34
Table 22. DIF Categories for Multiple-Choice Items .....	43
Table 23. DIF Categories for Constructed-Response Items .....	43
Table 24. Criteria to Evaluate Mean-Square Fit Statistics.....	76
Table 25. PAWS Reading, Mathematics, and Science Scaling Constants, Lowest Obtainable Scale Scores, and Highest Obtainable Scale Scores .....	81
Table 26. Proficiency Level Ranges for Grades 3–8 Reading.....	83
Table 27. Proficiency Level Ranges for Grades 3–8 Mathematics .....	83
Table 28. Proficiency Level Ranges for Grades 4 and 8 Science.....	83
Table 29. Proficiency Level Ranges for Grades 3, 5, and 7 SAWS .....	83
Table 30. SAWS State Level Normative Percentage Distributions.....	85
Table 31. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade.....	88
Table 32. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade and Reading Domain.....	89
Table 33. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade and Mathematics Domain .....	90



Table 34. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade and Science Domain.....	91
Table 35. SAWS 2014 Overall Interrater Reliability for 12-point Prompt and Trait Rater Agreement.....	92
Table 36. SAWS 2014 Overall Interrater Reliability for 4-point Prompt and Trait Rater Agreement.....	93
Table 37. SAWS 2014 Overall Interrater Reliability for 8-point Prompt and Trait Rater Agreement.....	93
Table 38. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 3 .....	98
Table 39. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 4 .....	98
Table 40. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 5 .....	98
Table 41. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 6 .....	98
Table 42. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 7 .....	99
Table 43. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 8 .....	99
Table 44. Scaled Scores Descriptive Statistics for the PAWS Reading Tests.....	106
Table 45. Percentage Proficient and Advanced for the PAWS Reading Tests.....	106
Table 46. Scaled Scores Descriptive Statistics for the Mathematics Tests.....	106
Table 47. Percentage of Proficient + Advanced Students for the Mathematics Tests.....	106
Table 48. Scaled Scores Descriptive Statistics for the Science Tests.....	107
Table 49. Percentage of Proficient + Advanced Students for the Science Tests .....	107
Table 50. Raw Scores Descriptive Statistics for the SAWS Tests .....	107
Table 51. Percentage Proficient and Advanced for the SAWS Tests .....	107
Table A1. PAWS 2014 Grade 3 Reading Blueprint .....	119
Table A2. PAWS 2014 Grade 4 Reading Blueprint .....	122
Table A3. PAWS 2014 Grade 5 Reading Blueprint .....	125
Table A4. PAWS 2014 Grade 6 Reading Blueprint .....	128
Table A5. PAWS 2014 Grade 7 Reading Blueprint .....	131
Table A6. PAWS 2014 Grade 8 Reading Blueprint .....	134
Table A7. PAWS 2014 Grade 3 Math Blueprint .....	137
Table A8. PAWS 2014 Grade 4 Math Blueprint .....	138
Table A9. PAWS 2014 Grade 5 Math Blueprint .....	139
Table A10. PAWS 2014 Grade 6 Math Blueprint .....	140
Table A11. PAWS 2014 Grade 7 Math Blueprint .....	141
Table A12. PAWS 2014 Grade 8 Math Blueprint .....	142
Table A13. PAWS 2014 Science Grade 4 Blueprint .....	143
Table A14. PAWS 2014 Science Grade 8 Blueprint .....	145
Table A15. SAWS 2014 Writing Grade 3 Blueprint .....	147
Table A16. SAWS 2014 Writing Grade 5 Blueprint .....	147
Table A17. SAWS 2014 Writing Grade 7 Blueprint .....	148

Table E1. Grade 5 SAWS 2014 Field Test Subscale Correlations – 12-point prompt.....	158
Table E2. Grade 5 SAWS 2014 Field Test Subscale Correlations – 4-point prompt.....	158
Table E3. Grade 5 SAWS 2014 Field Test Subscale Correlations – 8-point prompt.....	159
Table E4. Grade 7 SAWS 2014 Field Test Subscale Correlations – 12-point prompt.....	159
Table E5. Grade 7 SAWS 2014 Field Test Subscale Correlations – 4-point prompt.....	160
Table E6. Grade 7 SAWS 2014 Field Test Subscale Correlations – 8-point prompt.....	160
Table F1. Grade 3 PAWS Total Test and Subscale Correlations .....	161
Table F2. Grade 4 PAWS Total Test and Subscale Correlations .....	162
Table F3. Grade 5 PAWS Total Test and Subscale Correlations .....	163
Table F4. Grade 6 PAWS Total Test and Subscale Correlations .....	164
Table F5. Grade 7 PAWS Total Test and Subscale Correlations .....	165
Table F6. Grade 8 PAWS Total Test and Subscale Correlations .....	166
Table F7. SAWS 2014 Operational Test Subscale Correlations – 12-point Prompt.....	167
Table F8. SAWS 2014 Operational Test Subscale Correlations – 4-point Prompt.....	167
Table F9. SAWS 2014 Operational Test Subscale Correlations –8-point Prompt.....	168
Table G1. Grade 3 Reading DIF Summary Statistics for Embedded Field Test Items .....	169
Table G2. Grade 4 Reading DIF Summary Statistics for Embedded Field Test Items .....	170
Table G3. Grade 5 Reading DIF Summary Statistics for Embedded Field Test Items .....	170
Table G4. Grade 6 Reading DIF Summary Statistics for Embedded Field Test Items .....	171
Table G5. Grade 7 Reading DIF Summary Statistics for Embedded Field Test Items .....	171
Table G6. Grade 8 Reading DIF Summary Statistics for Embedded Field Test Items .....	172
Table G7. Grade 3 Mathematics DIF Summary Statistics for Embedded Field Test Items .....	173
Table G8. Grade 4 Mathematics DIF Summary Statistics for Embedded Field Test Items .....	173
Table G9. Grade 5 Mathematics DIF Summary Statistics for Embedded Field Test Items .....	174
Table G10. Grade 6 Mathematics DIF Summary Statistics for Embedded Field Test Items .....	174
Table G11. Grade 7 Mathematics DIF Summary Statistics for Embedded Field Test Items .....	175
Table G12. Grade 8 Mathematics DIF Summary Statistics for Embedded Field Test Items .....	175
Table G13. Grade 4 Science DIF Summary Statistics for Embedded Field Test Items .....	176
Table G14. Grade 8 Science DIF Summary Statistics for Embedded Field Test Items .....	176
Table G15. Grade 3 SAWS DIF Summary Statistics for Embedded Field Test Items .....	177
Table G16. Grade 5 SAWS DIF Summary Statistics for Embedded Field Test Items .....	177
Table G17. Grade 7 SAWS DIF Summary Statistics for Embedded Field Test Items .....	178
Table H1. Grade 3 Field Test Trait Rater Agreement and Weighted Kappa – 8-point Prompt .	179
Table H2. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 12-point prompts .....	179
Table H3. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 4-point prompts	180
Table H4. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 8-point prompts	180
Table H5. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 12-point prompts .....	181

Table H6. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 4-point prompts	181
Table H7. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 8-point prompts	182
Table I1. SAWS 2014 Overall Interrater Reliability for 12-point Prompt and Trait Rater Agreement.....	183
Table I2. SAWS 2014 Overall Interrater Reliability for 4-point Prompt and Trait Rater Agreement.....	184
Table I3. SAWS 2014 Overall Interrater Reliability for 8-point Prompt and Trait Rater Agreement.....	184
Table J1. Reading Grade 3 Classical Statistics for Field Test Items .....	185
Table J2. Reading Grade 4 Classical Statistics for Field Test Items .....	190
Table J3. Reading Grade 5 Classical Statistics for Field Test Items .....	194
Table J4. Reading Grade 6 Classical Statistics for Field Test Items .....	198
Table J5. Reading Grade 7 Classical Statistics for Field Test Items .....	202
Table J6. Reading Grade 8 Classical Statistics for Field Test Items .....	206
Table J7. Mathematics Grade 3 Classical Statistics for Field Test Items .....	211
Table J8. Mathematics Grade 4 Classical Statistics for Field Test Items .....	215
Table J9. Mathematics Grade 5 Classical Statistics for Field Test Items .....	219
Table J10. Mathematics Grade 6 Classical Statistics for Field Test Items .....	223
Table J11. Mathematics Grade 7 Classical Statistics for Field Test Items .....	227
Table J12. Mathematics Grade 8 Classical Statistics for Field Test Items .....	231
Table J13. Science Grade 4 Classical Statistics for Field Test Items .....	234
Table J14. Science Grade 8 Classical Statistics for Field Test Items .....	238
Table J15. Distributions of SAWS Grade 3 Field Test –8-point Prompt .....	242
Table J16. Distributions of Scores for SAWS Grade 5 – 12-point Prompt .....	242
Table J17. Distributions of Scores for SAWS Grade 5 by Trait – 12-point Prompt .....	243
Table J18. Distributions of Scores for SAWS Grade 5 – 4-point Prompt .....	243
Table J19. Distributions of Scores for SAWS Grade 5 by Trait – 4-point Prompt .....	244
Table J20. Distributions of Rater 1 Scores for SAWS Grade 5 – 8-point Prompt .....	245
Table J21. Distributions of Scores for SAWS Grade 5 by Trait – 8-point Prompt .....	246
Table J22. Distributions of Rater 1 Scores for SAWS Grade 7 – 12-point Prompt .....	247
Table J23. Distributions of Rater 1 Scores for SAWS Grade 7 by Trait – 12-point Prompt.....	248
Table J24. Distributions of Scores for SAWS Grade 7 – 4-point Prompt .....	248
Table J25. Distributions of Scores for SAWS Grade 7 by Trait – 4-point Prompt .....	249
Table J26. Distributions of Scores for SAWS Grade 7 – 8-point Prompts.....	250
Table J27. Distributions of Scores for SAWS Grade 7 by Trait – 8-point Prompt .....	251
Table K1. Reading Grade 3 Classical Statistics for Operational Items .....	253
Table K2. Reading Grade 4 Classical Statistics for Operational Items .....	255
Table K3. Reading Grade 5 Classical Statistics for Operational Items .....	257
Table K4. Reading Grade 6 Classical Statistics for Operational Items .....	259

Table K5. Reading Grade 7 Classical Statistics for Operational Items .....	261
Table K6. Reading Grade 8 Classical Statistics for Operational Items .....	263
Table K7. Mathematics Grade 3 Classical Statistics for Operational Items .....	265
Table K8. Mathematics Grade 4 Classical Statistics for Operational Items .....	267
Table K9. Mathematics Grade 5 Classical Statistics for Operational Items .....	269
Table K10. Mathematics Grade 6 Classical Statistics for Operational Items .....	271
Table K11. Mathematics Grade 7 Classical Statistics for Operational Items .....	273
Table K12. Mathematics Grade 8 Classical Statistics for Operational Items .....	275
Table K13. Science Grade 4 Classical Statistics for Operational Items .....	277
Table K14. Science Grade 8 Classical Statistics for Operational Items .....	279
Table K15. Distributions of Rating 1 Scores for SAWS .....	281
Table K16. Distributions of Trait Scores for SAWS – Grade 3 – 12-point Prompt 1 .....	282
Table K17. Distributions of Trait Scores for SAWS - Grade 3 - 12-point Prompt 2 .....	282
Table K18. Distributions of Trait Scores for SAWS - Grade 5 - 12-point Prompt .....	282
Table K19. Distributions of Trait Scores for SAWS - Grade 7 - 12-point Prompt .....	282
Table K20. Distributions of Scores for SAWS Grade 5 - 4-point Prompt .....	283
Table K21. Distributions of Trait Scores for SAWS - Grade 5 – 4-point Prompt.....	283
Table K22. Distributions of Rater 1 Scores for SAWS Grade 5 - 8-point Prompt.....	283
Table K23. Distributions of Scores for SAWS Grade 5 by Trait – 8-point Prompt .....	284
Table K24. Distributions of Scores for SAWS Grade 7 - 4-point Prompt .....	284
Table K25. Distributions of Trait Scores for SAWS - Grade 5 – 4-point Prompt.....	284
Table K26. Distributions of Rater 1 Scores for SAWS Grade 7 - 8-point Prompt.....	285
Table K27. Distributions of Scores for SAWS Grade 7 by Trait – 8-point Prompt.....	285
Table L1. Reading Grade 3 IRT Statistics for Field Test Items .....	286
Table L2. Reading Grade 4 IRT Statistics for Field Test Items .....	290
Table L3. Reading Grade 5 IRT Statistics for Field Test Items .....	293
Table L4. Reading Grade 6 IRT Statistics for Field Test Items .....	296
Table L5. Reading Grade 7 IRT Statistics for Field Test Items .....	299
Table L6. Reading Grade 8 IRT Statistics for Field Test Items .....	302
Table L7. Mathematics Grade 3 IRT Statistics for Field Test Items .....	306
Table L8. Mathematics Grade 4 IRT Statistics for Field Test Items .....	309
Table L9. Mathematics Grade 5 IRT Statistics for Field Test Items .....	312
Table L10. Mathematics Grade 6 IRT Statistics for Field Test Items .....	315
Table L11. Mathematics Grade 7 IRT Statistics for Field Test Items .....	318
Table L12. Mathematics Grade 8 IRT Statistics for Field Test Items .....	321
Table L13. Science Grade 4 IRT Statistics for Field Test Items .....	324
Table L14. Science Grade 8 IRT Statistics for Field Test Items .....	328
Table M1. Reading Grade 3 IRT Statistics for Operational Items .....	332
Table M2. Reading Grade 4 IRT Statistics for Operational Items .....	334

Table M3. Reading Grade 5 IRT Statistics for Operational Items .....	336
Table M4. Reading Grade 6 IRT Statistics for Operational Items .....	338
Table M5. Reading Grade 7 IRT Statistics for Operational Items .....	340
Table M6. Reading Grade 8 IRT Statistics for Operational Items .....	342
Table M7. Mathematics Grade 3 IRT Statistics for Operational Items .....	344
Table M8. Mathematics Grade 4 IRT Statistics for Operational Items .....	346
Table M9. Mathematics Grade 5 IRT Statistics for Operational Items .....	348
Table M10. Mathematics Grade 6 IRT Statistics for Operational Items .....	350
Table M11. Mathematics Grade 7 IRT Statistics for Operational Items .....	352
Table M12. Mathematics Grade 8 IRT Statistics for Operational Items .....	354
Table M13. Science Grade 4 IRT Statistics for Operational Items .....	356
Table M14. Science Grade 8 IRT Statistics for Operational Items .....	358
Table N1. Frequency of IEP Student's Standard Accommodations: Presentation Accommodations .....	360
Table N2. Frequency of IEP Student's Standard Accommodations: Response Accommodations .....	361
Table N3. Frequency of IEP Student's Standard Accommodations: Setting Accommodations	361
Table N4. Frequency of IEP Student's Standard Accommodations: Timing and Scheduling Accommodations .....	362
Table N5. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations .....	362
Table N6. Frequency of English Language Learners Standard Accommodations: Setting Accommodations .....	362
Table N7. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations .....	362
Table N8. Frequency of IEP Student's Standard Accommodations: Presentation Accommodations .....	363
Table N9. Frequency of IEP Student's Standard Accommodations: Response Accommodations .....	364
Table N10. Frequency of IEP Student's Standard Accommodations: Setting Accommodations .....	365
Table N11. Frequency of IEP Student's Standard Accommodations: Timing and Scheduling Accommodations .....	365
Table N12. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations .....	365
Table N13. Frequency of English Language Learners Standard Accommodations: Setting Accommodations .....	365
Table N14. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations .....	366

Table N15. Frequency of IEP Student’s Standard Accommodations: Presentation Accommodations .....	366
Table N16. Frequency of IEP Student’s Standard Accommodations: Response Accommodations .....	367
Table N17. Frequency of IEP Student’s Standard Accommodations: Setting Accommodations .....	367
Table N18. Frequency of IEP Student’s Standard Accommodations: Timing and Scheduling Accommodations .....	368
Table N19. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations .....	368
Table N20. Frequency of English Language Learners Standard Accommodations: Setting Accommodations .....	368
Table N21. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations .....	368
Table N22. Frequency of IEP Student’s Standard Accommodations: Presentation Accommodations .....	369
Table N23. Frequency of IEP Student’s Standard Accommodations: Response Accommodations .....	370
Table N24. Frequency of IEP Student’s Standard Accommodations: Setting Accommodations .....	370
Table N25. Frequency of IEP Student’s Standard Accommodations: Timing and Scheduling Accommodations .....	371
Table N26. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations .....	371
Table N27. Frequency of English Language Learners Standard Accommodations: Setting Accommodations .....	371
Table N28. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations .....	371
Table O1. Reading Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status .....	372
Table O2. Mathematics Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status .....	372
Table O3. Science Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status .....	373
Table O4. SAWS Mean Raw Scores, Counts, and Raw Score Standard Deviations for IEP Students by Accommodation Status .....	373
Table P1. Summary Statistics of Reading, Mathematics, and Science Scale Score by Grade ...	374
Table P2. Summary Statistics of Reading Grade 3 Scale Score .....	375
Table P3. Summary Statistics of Reading Grade 4 Scale Score .....	375

Table P4. Summary Statistics of Reading Grade 5 Scale Score .....	376
Table P5. Summary Statistics of Reading Grade 6 Scale Score .....	376
Table P6. Summary Statistics of Reading Grade 7 Scale Score .....	377
Table P7. Summary Statistics of Reading Grade 8 Scale Score .....	377
Table P8. Summary Statistics of Mathematics Grade 3 Scale Score.....	378
Table P9. Summary Statistics of Mathematics Grade 4 Scale Score.....	378
Table P10. Summary Statistics of Mathematics Grade 5 Scale Score.....	379
Table P11. Summary Statistics of Mathematics Grade 6 Scale Score.....	379
Table P12. Summary Statistics of Mathematics Grade 7 Scale Score.....	380
Table P13. Summary Statistics of Mathematics Grade 8 Scale Score.....	380
Table P14. Summary Statistics of Science Grade 4 Scale Score .....	381
Table P15. Summary Statistics of Science Grade 8 Scale Score .....	381
Table P16. Summary Statistics of SAWS Grade 3 Total Raw Score .....	382
Table P17. Summary Statistics of SAWS Grade 5 Total Raw Score .....	382
Table P18. Summary Statistics of SAWS Grade 7 Total Raw Score .....	383
Table P19. Gender and Ethnicity Performance by SAWS Prompt and Trait – Grade 3 Prompt 1 .....	384
Table P20. Gender and Ethnicity Performance by SAWS Prompt and Trait – Grade 3 Prompt 2 .....	385
Table P21. Gender and Ethnicity Performance by SAWS 12-point Prompt and Trait – Grade 5 .....	386
Table P22. Gender and Ethnicity Performance by SAWS 4-point Prompt and Trait – Grade 5	387
Table P23. Gender and Ethnicity Performance by SAWS 8-point Prompt and Trait – Grade 5	388
Table P24. Gender and Ethnicity Performance by SAWS Prompt and Trait – Grade 7 .....	389
Table P25. Gender and Ethnicity Performance by SAWS 4-point Prompt and Trait – Grade 7	390
Table P26. Gender and Ethnicity Performance by SAWS 8-point Prompt and Trait – Grade 7	391
Table Q1. Reading Grade 3 Raw Score to Scaled Score .....	392
Table Q2. Reading Grade 4 Raw Score to Scaled Score .....	394
Table Q3. Reading Grade 5 Raw Score to Scaled Score .....	396
Table Q4. Reading Grade 6 Raw Score to Scaled Score .....	398
Table Q5. Reading Grade 7 Raw Score to Scaled Score .....	400
Table Q6. Reading Grade 8 Raw Score to Scaled Score .....	402
Table Q7. Mathematics Grade 3 Raw Score to Scaled Score.....	404
Table Q8. Mathematics Grade 4 Raw Score to Scaled Score.....	406
Table Q9. Mathematics Grade 5 Raw Score to Scaled Score.....	409
Table Q10. Mathematics Grade 6 Raw Score to Scaled Score.....	411
Table Q11. Mathematics Grade 7 Raw Score to Scaled Score.....	413
Table Q12. Mathematics Grade 8 Raw Score to Scaled Score.....	415
Table Q13. Science Grade 4 Raw Score to Scaled Score .....	418

Table Q14. Science Grade 8 Raw Score to Scaled Score .....	420
Table Q15. SAWS Grade 3 Raw Score to Performance Level.....	422
Table Q16. SAWS Grade 5 Raw Score to Performance Level.....	423
Table Q17. SAWS Grade 7 Raw Score to Performance Level.....	424
Table R1. Performance Levels of Reading, Mathematics, and Science by Grade Revised .....	425
Table R2. Performance Levels of Reading by Grade 3 Demographic Subgroup.....	426
Table R3. Performance Levels of Reading by Grade 4 Demographic Subgroup.....	426
Table R4. Performance Levels of Reading by Grade 5 Demographic Subgroup.....	427
Table R5. Performance Levels of Reading by Grade 6 Demographic Subgroup.....	427
Table R6. Performance Levels of Reading by Grade 7 Demographic Subgroup.....	428
Table R7. Performance Levels of Reading by Grade 8 Demographic Subgroup.....	428
Table R8. Performance Levels of Mathematics by Grade 3 Demographic Subgroup.....	429
Table R9. Performance Levels of Mathematics by Grade 4 Demographic Subgroup.....	429
Table R10. Performance Levels of Mathematics by Grade 5 Demographic Subgroup.....	430
Table R11. Performance Levels of Mathematics by Grade 6 Demographic Subgroup.....	430
Table R12. Performance Levels of Mathematics by Grade 7 Demographic Subgroup.....	431
Table R13. Performance Levels of Mathematics by Grade 8 Demographic Subgroup.....	431
Table R14. Performance Levels of Science by Grade 4 Demographic Subgroup.....	432
Table R15. Performance Levels of Science by Grade 8 Demographic Subgroup.....	432
Table R16. Performance Levels of SAWS by Grade 3 Demographic Subgroup .....	433
Table R17. Performance Levels of SAWS by Grade 5 Demographic Subgroup .....	433
Table R18. Performance Levels of SAWS by Grade 7 Demographic Subgroup .....	434
Table S1. SAWS Field Test Gender and Ethnicity Performance by 8-point Prompt – Grade 3	435
Table S2. SAWS Field Test Gender and Ethnicity Performance by 12-point Prompt – Grade 5	436
Table S3. SAWS Field Test Gender and Ethnicity Performance by 4-point Prompt – Grade 5	437
Table S4. SAWS Field Test Gender and Ethnicity Performance by 8-point Prompt – Grade 5	438
Table S5. SAWS Field Test Gender and Ethnicity Performance by 12-point Prompt – Grade 7	439
Table S6. SAWS Field Test Gender and Ethnicity Performance by 4-point Prompt – Grade 7	440
Table S7. SAWS Field Test Gender and Ethnicity Performance by 8-point Prompt – Grade 7	441



## LIST OF FIGURES

Figure 1. Sample item characteristic curve.....	75
Figure 2. Accuracy or Consistency around Critical Cut Point.....	97
Figure 3. Percentage of Proficient and Advanced Students for Grade 4 Science.....	108
Figure 4. Percentage of Proficient and Advanced Students for Grade 8 Science.....	109

# 1. OVERVIEW OF THE 2014 PAWS AND SAWS

## 1.1 Introduction

This report describes the technical characteristics of the Proficiency Assessments for Wyoming Students (PAWS) and Student Assessment of Writing Skills (SAWS) for the 2013–2014 school year. Primary purposes of the PAWS and SAWS include improving teaching and learning, fostering school and program improvement, and measuring student performance indicators under the Wyoming Accountability in Education Act.

Beginning with the spring 2006 administration, PAWS became the official statewide assessment used to measure individual student achievement against the Wyoming Content and Performance Standards in English Language Arts and Mathematics at grades 3–8 and 11. The PAWS Reading and Mathematics tests meet all requirements of the No Child Left Behind Act of 2001 (NCLB). In 2008, a Science assessment was implemented at grades 4, 8, and 11.

The PAWS writing test was discontinued beginning with the 2012 administration and was administered separately from the PAWS in 2013. The newly renamed SAWS program received further revision through the Select Committee on Education during 2011–2012. Current legislation (version C3 of EA90) required the state board to “establish a separate writing and language assessment to be implemented and administered statewide in school year 2013–2014 and each school year thereafter” (Section 3(a)).

Further legislative action, 2013 Wyoming State Enrolled Act 65, removed grade 11 from the 2013 PAWS and SAWS future administrations.

In 2012, the Wyoming State Board adopted the 2012 Wyoming Content Performance Standards (2012 WyCPS) in English Language Arts and Mathematics. ETS developed and field tested items in 2013 aligned to the 2012 WyCPS. The 2014 PAWS assessments utilize the 2012 WyCPS as reporting categories and have become the new scale measuring students’ academic performance.

## 1.2 Background of PAWS and SAWS

In the spring of 2006, the Proficiency Assessments for Wyoming Students (PAWS) in reading, writing, and mathematics were administered for the first time to Wyoming students in grades 3–8 and 11. Wyoming statutes require that a statewide assessment system shall be substantially aligned with the uniform education program and student content and performance standards imposed by law and by board rule and regulation (§21-2-304 (a)(v)(A)).

In early 2003, the Wyoming State Legislature established the Wyoming Statewide Task Force on Student Assessment and Education Accountability and provided two central charges to this group. The legislature asked that the Task Force:

1. Recommend modifications, if necessary, to Wyoming's statewide assessment system to improve teaching and learning and foster school improvement; and
2. Recommend an accountability system with consequences assisting in meeting NCLB's accountability requirements while maintaining uniformity and quality of state standards.

Staff of the Wyoming Department of Education (WDE) served in an advisory capacity to this group. The 13-member Task Force included one district superintendent; five administrators; two members of the Wyoming legislature; two teachers; a parent; and the editor of the *Casper Star Tribune* newspaper. In October 2003, *The Wyoming Statewide Task Force on Student Assessment and Education Accountability Report and Recommendations* set forth various suggestions to the WDE for consideration as the new assessment system was designed.

The task force recommended a statewide assessment system that would include, among other things, the following:

- A summative assessment that would maintain some, but not all, of the features of the Wyoming Comprehensive Assessment System (WyCAS) and that would satisfy the core requirements of the NCLB related to standards, assessments, and accountability;
- Comparability of scores *across* grades to allow for meaningful evaluation of individual student performance and progress as that student moves from grade to grade while also allowing for meaningful within-grade comparisons from year to year;
- Embedded tools and assessments in Reading, Writing, and Mathematics (and possibly Science) that would: be developed and implemented over time; be based on ongoing research and evaluation; fit within existing district assessment systems; be administered periodically during the school year preceding the summative assessment; inform instructional strategies; assist in improving student learning during the year; and supplement summative assessment results;
- Use of the *National Assessment of Educational Progress* (NAEP) results for the state to provide national comparison data; and
- Timely and meaningful feedback to educators, parents, and students regarding student, school, district, and state performance, which could improve teaching and learning over the course of the school year.

As a result, PAWS replaced WyCAS as the statewide accountability assessment. The WyCAS was initially designed to comply with the provisions of the 1994 reauthorization of the ESEA, the *Improving America's Schools Act (IASA)*. With the introduction of the PAWS, the WDE has not only implemented an assessment system that meets the accountability requirements of NCLB, but one that also provides the data necessary to inform instructional decision-making by Wyoming classroom teachers to address the specific academic needs of students.

In 2012, Wyoming adopted the Common Core standards for English Language Arts and Mathematics, henceforth called 2012 WyCPS. The statutes read:

“W.S. 21-2-304(a)(iii) - By rule and regulation and in consultation with local school districts, prescribe uniform student content and performance standards for the common core of knowledge and the common core of skills specified under W.S. 21-9-101(b), and promulgate uniform standards . . .

W.S.21-2-304(c) - The state board shall perform an ongoing review of state board duties prescribed by law and may make recommendations to the legislature on board duties. In addition and not less than once every five (5) years, the board shall evaluate and review the uniformity and quality of the content and performance standards imposed under W.S. 21-9-101 and 21-9-102 and the student content and performance standards promulgated under paragraph (a)(iii) of this section . . .”

To comply with this legislative action, the WDE piloted 2012 WyCPS-aligned items in spring 2013. These embedded field test items were administered across the state in grades 3 through 8 for reading and mathematics. New vertical scales for Reading and Math were established and approved by the WDE in May 2014. A standard setting for 2014 PAWS Reading and Mathematics assessments was conducted in July 2014, establishing the performance standards.

In spring 2006, the Proficiency Assessments for Wyoming Students (PAWS) writing test was administered for the first time to Wyoming students in grades 3 through 8 and 11. PAWS writing was designed to provide information for federal, state, and local indicators of student academic performance requirements in writing. From 2006 until 2011, two 12-point writing prompts were administered to participating students.

During the 2011 Wyoming legislative session, the state legislature passed Senate File 70 (SF 70), also known as Enrolled Act 90 (EA 90). This legislation reads:

“Section 5(b)(i) requires the state board shall, through the state superintendent and the department, develop an authentic statewide assessment of student writing skills which is:

(i) Limited to one (1) writing prompt in school year 2011–2012, the initial year of implementation statewide as a pilot assessment; . . .

(iv) Administered separate and at different times from the statewide summative assessment in other subject areas; . . .

(v) Fully implemented in the 2013–2014 school year and each year thereafter.”

The newly renamed SAWS program received further revision through the Select Committee on Education during 2011–2012. Current legislation (version C3 of EA90) required the state board to “establish a separate writing and language assessment to be implemented and administered statewide in school year 2013–2014 and each school year thereafter” (Section 3(a)).

To comply with this legislative action, the WDE piloted SAWS in spring 2012. This stand-alone, one-prompt writing assessment was administered across the state in grades 3 through 8 and 11. Further legislative action, 2013 Wyoming State Enrolled Act 65, removed grade 11 from the 2013 SAWS and future administrations. During the 2012 Wyoming legislative session, the state legislature passed Senate File 65 (SF 65), also known as Enrolled Act 65 (EA 65). Section 1(c) requires that:

*(c) School level performance shall be determined by measurement of performance indicators and attainment of student performance as specified by this section. To the extent applicable, each measure shall be aggregated to the school level based upon those grades served inclusive to each school as reported by the respective school district to the department of education. . . .*

*The indicators of school level performance shall be: (ii) Student academic achievement in reading, mathematics, science, and writing and language as measured by: (A) The statewide assessment administered under W.S. 21-2-304(a)(v) in: . . .*

*(III) Writing and language in grades three (3), five (5) and seven (7).*

Therefore, in 2014, every eligible child in grades 3, 5, and 7 are eligible to participate in the SAWS assessment and were administered multiple operational writing prompts.

In response to the statutory and regulatory requirements and the recommendations of the task force, the PAWS and SAWS state-level assessments are aligned with the 2012 WyCPS in English Language Arts and Mathematics in grades 3 through 8 and 11, Wyoming Content and

Performance<sup>1</sup> Standards Science at grades 4, 8, and 11 and SAWS in grades 3, 5, and 7. PAWS and SAWS are designed to provide information for use as federal, state, and local indicators of the extent to which students satisfy academic performance requirements. PAWS and SAWS results provide reliable information that can be used as a basis for drawing valid inferences that enable:

- Students to know the extent to which they have mastered expected knowledge and skills in the Standards;
- Parents to know if their children are acquiring the knowledge and skills aligned with the Wyoming Content and Performance Standards;
- Teachers to know if their students have mastered grade-level knowledge and skills in the Standards and, if not, what weaknesses need to be addressed; and
- Community leaders and lawmakers to know if students in Wyoming schools are improving their performance over time.

### 1.3 Overview of PAWS and SAWS Test Components

The entire assessment program administered in 2013–2014 consisted of the following components:

- PAWS Reading, Mathematics, and Science assessments
- SAWS Writing assessments
- PAWS Alternate Assessment Reading, Mathematics, and Science assessments
- SAWS Alternate Assessment writing assessments

The test design for the spring 2014 administration of the PAWS and SAWS included content area assessments in reading, mathematics, science, and writing. For reading, mathematics, science, and writing, each test had two to three sessions. Multiple choice items and writing prompts were administered via pencil and paper in a consumable test booklet for students in grades 3-5 and via a separate answer sheet for students in grades 6-8.

### 1.4 Overview of the PAWS and SAWS Design

As stated above, the intent of the PAWS and SAWS assessment is not only to meet the accountability requirements of NCLB and the Wyoming Accountability in Education Act, but also to inform teaching, learning, and school improvement activities.

---

<sup>1</sup> Wyoming uses the term “performance” to describe the characteristics of student achievement of mastery of the content of Wyoming’s Standards, whereas NCLB describes this measure as “achievement.”

Therefore, PAWS and SAWS were conceptually constructed around an instructionally supportive design to include clear targets for instruction and informative reporting categories.

The PAWS and SAWS assessment are used to measure individual student achievement against the newly adopted 2012 Wyoming Content and Performance Standards (2012 WyCPS) in English Language Arts and Mathematics. In addition, SAWS writing blueprints were revised and test items developed to better align these assessments with the 2012 WyCPS. The Wyoming Content and Performance Standards outline knowledge and skills students are expected to acquire at each grade in order to succeed in school and at work.

The PAWS Science provides additional skill-level reporting categories aligned to the Wyoming Content and Performance Standards as organized by the Wyoming Assessment Descriptions to assist teachers in interpreting and addressing specific academic needs of students.

Assessment results provide important information to all facets of the school community. Policymakers, administrators, teachers, students, and parents all use assessment information for a variety of purposes. Collectively, these users make decisions about how well students are achieving, whether schools are functioning effectively for each child, and whether they are functioning well for all children collectively.

PAWS and SAWS results are particularly intended to help educators make informed decisions about curriculum and instruction. Since PAWS and SAWS are aligned to academic content and student performance standards, its results can reveal weaknesses and strengths in curricula or instructional methodology. Thus, they can also help educators target specific areas necessary for school and district improvement.

### *1.5 State Policy on Student Participation*

With two exceptions, all students in grades 3 through 8 must participate in the regular PAWS and grades 3, 5, and 7 in SAWS tests if they receive any instruction on Wyoming state academic standards. The only exceptions are for students with significant cognitive disabilities who meet Wyoming Alternate Assessment participation guidelines and ELL students who have been in the United States for less than a full year. The exemption for ELL students is only for the reading component of PAWS. They are required to take the mathematics and science portions of PAWS and SAWS, but may take the Wyoming ELL assessment as a substitution for the ELA portions of PAWS.

Students with significant cognitive disabilities were required to take the Proficiency Assessments for Wyoming Students–Alternate (PAWS-ALT) and Student Assessment of Writing Skills–Alternate (SAWS-ALT). All students will participate in the state accountability assessment program in one of three ways:

- Participation in PAWS and SAWS regular assessment without accommodation
- Participation in PAWS and SAWS regular assessment with accommodation
- Participation in PAWS-ALT and SAWS-ALT

#### *1.5.1 Students with Disabilities, 504 Plans, and English Language Learners*

Following are procedures and practices related to the participation in the statewide assessments of students with disabilities, students who have 504 Plans, and students with limited English proficiency in the statewide assessments:

Students with disabilities participate with appropriate accommodations based on each student's Individualized Education Program (IEP) committee's recommendation. Students with 504 Plans and English Language Learners (ELL) also take the PAWS.

Some students with disabilities, for whom even the PAWS with accommodations is inappropriate, participate in the PAWS-ALT and SAWS-ALT as provided for by a student's IEP. The PAWS and SAWS are intended to include all of the public school students in Wyoming. However, students with the most significant cognitive disabilities are exempted from the PAWS and SAWS under the Individuals with Disabilities Education Act or Section 504 of the Rehabilitation Act. These students are assessed using the PAWS-ALT and SAWS-ALT. The decision for exemption from the PAWS and SAWS is made on an individual basis according to professional judgments of the IEP team. Corresponding documentation for any exemption is required.

School districts may not exempt ELL students from the assessment, except for students who are in their first year in the United States. Only students who are in their first year may take the Wyoming ELL assessment instead of the Reading component of PAWS and SAWS, but they are not exempt from the Mathematics and Science tests. The Wyoming ELL assessment measures English language academic readiness.

Tables 1 through 4 provide data on the numbers of students tested in 2014. Additional information can be found on the WDE website: <http://edu.wyoming.gov/default.aspx>.



Table 1. Statewide Participation in Reading PAWS

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	N	%	N	%	N	%	N	%	N	%	N	%
Total	7365	100.0	7022	100.0	7075	100.0	6758	100.0	6796	100.0	6781	100.0
Male	3709	50.4	3639	51.8	3635	51.4	3509	51.9	3498	51.5	3558	52.5
Female	3636	49.4	3377	48.1	3430	48.5	3244	48.0	3294	48.5	3220	47.5
Unknown	20	0.3	6	0.1	10	0.1	5	0.1	4	0.1	3	0.0
American Indian/Alaska Native	309	4.2	267	3.8	268	3.8	269	4.0	245	3.6	222	3.3
Asian	52	0.7	72	1.0	74	1.0	53	0.8	58	0.9	57	0.8
African American	92	1.2	85	1.2	70	1.0	68	1.0	83	1.2	98	1.4
Native Hawaiian or other/Pacific Islander	8	0.1	6	0.1	7	0.1	10	0.1	15	0.2	11	0.2
Hispanic/Latino	1068	14.5	968	13.8	933	13.2	912	13.5	916	13.5	891	13.1
White	5662	76.9	5480	78.0	5578	78.8	5305	78.5	5345	78.6	5363	79.1
Multiracial	146	2.0	133	1.9	133	1.9	135	2.0	128	1.9	134	2.0
Unknown	28	0.4	11	0.2	12	0.2	6	0.1	6	0.1	5	0.1
Free/Reduced Lunch	2363	32.1	2195	31.3	2151	30.4	2009	29.7	2009	29.6	1832	27.0
Not Free/Reduced Lunch	5002	67.9	4827	68.7	4924	69.6	4749	70.3	4787	70.4	4949	73.0
Special Education	1083	14.7	1084	15.4	1037	14.7	920	13.6	843	12.4	868	12.8
Not Special Education	6282	85.3	5938	84.6	6038	85.3	5838	86.4	5953	87.6	5913	87.2
English Language Learner	371	5.0	216	3.1	149	2.1	155	2.3	163	2.4	145	2.1
Not English Language Learner	6994	95.0	6806	96.9	6926	97.9	6603	97.7	6633	97.6	6636	97.9

Table 2. Statewide Participation in Mathematics PAWS

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	N	%	N	%	N	%	N	%	N	%	N	%
Total	7369	100.0	7026	100.0	7077	100.0	6760	100.0	6799	100.0	6784	100.0
Male	3718	50.5	3641	51.8	3634	51.3	3511	51.9	3504	51.5	3561	52.5
Female	3639	49.4	3370	48.0	3431	48.5	3244	48.0	3291	48.4	3220	47.5
Unknown	12	0.2	15	0.2	12	0.2	5	0.1	4	0.1	3	0.0
American Indian/Alaska Native	307	4.2	268	3.8	267	3.8	268	4.0	244	3.6	222	3.3
Asian	52	0.7	74	1.1	75	1.1	55	0.8	58	0.9	57	0.8
African American	94	1.3	86	1.2	71	1.0	69	1.0	84	1.2	98	1.4
Native Hawaiian or other/Pacific Islander	9	0.1	6	0.1	7	0.1	10	0.1	15	0.2	11	0.2
Hispanic/Latino	1076	14.6	969	13.8	933	13.2	920	13.6	920	13.5	895	13.2
White	5666	76.9	5474	77.9	5573	78.7	5298	78.4	5345	78.6	5362	79.0
Multiracial	143	1.9	132	1.9	134	1.9	134	2.0	127	1.9	134	2.0
Unknown	22	0.3	17	0.2	17	0.2	6	0.1	6	0.1	5	0.1
Free/Reduced Lunch	2364	32.1	2194	31.2	2146	30.3	2011	29.7	2011	29.6	1836	27.1
Not Free/Reduced Lunch	5005	67.9	4832	68.8	4931	69.7	4749	70.3	4788	70.4	4948	72.9
Special Education	1081	14.7	1085	15.4	1037	14.7	919	13.6	840	12.4	866	12.8
Not Special Education	6288	85.3	5941	84.6	6040	85.3	5841	86.4	5959	87.6	5918	87.2
English Language Learner	380	5.2	223	3.2	155	2.2	167	2.5	174	2.6	149	2.2
Not English Language Learner	6989	94.8	6803	96.8	6922	97.8	6593	97.5	6625	97.4	6635	97.8

Table 3. Statewide Participation in Science PAWS

	Grade 4		Grade 8	
	<i>N</i>	%	<i>N</i>	%
Total	7022	100.0	6770	100.0
Male	3639	51.8	3555	52.5
Female	3371	48.0	3212	47.4
Unknown	12	0.2	3	0.0
American Indian/Alaska Native	267	3.8	221	3.3
Asian	74	1.1	56	0.8
African American	86	1.2	98	1.4
Native Hawaiian or other/Pacific Islander	6	0.1	11	0.2
Hispanic/Latino	973	13.9	892	13.2
White	5467	77.9	5353	79.1
Multiracial	133	1.9	134	2.0
Unknown	16	0.2	5	0.1
Free/Reduced Lunch	2192	31.2	1830	27.0
Not Free/Reduced Lunch	4830	68.8	4940	73.0
Special Education	1081	15.4	863	12.7
Not Special Education	5941	84.6	5907	87.3
English Language Learner	224	3.2	149	2.2
Not English Language Learner	6798	96.8	6621	97.8

Table 4. Statewide Participation in SAWS

	<u>Grade 3</u>		<u>Grade 5</u>		<u>Grade 7</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Total	7315	100.0	7033	100.0	6763	100.0
Male	3690	50.4	3622	51.5	3474	51.4
Female	3621	49.5	3400	48.3	3276	48.4
Unknown	4	0.1	11	0.2	13	0.2
American Indian/Alaska Native	291	4.0	254	3.6	231	3.4
Asian	52	0.7	75	1.1	57	0.8
African American	89	1.2	69	1.0	81	1.2
Native Hawaiian or other/Pacific Islander	7	0.1	6	0.1	14	0.2
Hispanic/Latino	1065	14.6	925	13.2	913	13.5
White	5651	77.3	5556	79.0	5318	78.6
Multiracial	144	2.0	131	1.9	126	1.9
Unknown	16	0.2	17	0.2	23	0.3
Free/Reduced Lunch	2315	31.6	2116	30.1	1965	29.1
Not Free/Reduced Lunch	5000	68.4	4917	69.9	4798	70.9
Special Education	1066	14.6	1029	14.6	834	12.3
Not Special Education	6249	85.4	6004	85.4	5929	87.7
English Language Learner	367	5.0	148	2.1	162	2.4
Not English Language Learner	6948	95.0	6885	97.9	6601	97.6

## 2. VALIDITY

### 2.1 Overview

Validity refers to the degree to which each interpretation or use of a test score is supported by evidence that is gathered (American Educational Research Association [AERA], American Psychological Association [APA], and National Council on Measurement in Education [NCME], 2014; ETS, 2002). It is a central concern underlying the development, administration, and scoring of a test and the uses and interpretations of test scores.

Validation is the process of accumulating evidence to support each proposed score interpretation or use. It does not involve a single study or gathering one particular kind of evidence. Validation involves multiple investigations and various kinds of evidence (AERA, APA, and NCME, 2014; Cronbach, 1971; ETS, 2002; Kane, 2006). The process begins with test design and continues through the entire assessment process including item development and field testing, analyses of item and test data, test scaling, scoring, and score reporting.

In this section, the evidence gathered is presented to support the intended uses and interpretations of scores for the PAWS and SAWS assessment programs. The description is organized in the manner prescribed by *The Standards for Educational and Psychological Testing* (AERA, APA, and NCME, 2014). These standards require a clear definition of the purpose of the test, which includes a description of the qualities called constructs that are to be assessed by a test, the population to be assessed, as well as how the scores are to be interpreted and used.

In addition, the *Standards* identify five kinds of evidence that can provide support for score interpretations and uses, which are as follows:

- Evidence based on test content;
- Evidence based on relations to other variables;
- Evidence based on response processes;
- Evidence based on internal structure, and;
- Evidence based on the consequences of testing.

These kinds of evidence are also defined as important elements of validity information in documents developed by the U.S. Department of Education for the peer review of testing programs administered by states in response to the Elementary and Secondary Education Act (USDOE, 2001).

The next section defines the purpose of the PAWS and SAWS assessments, followed by a description and discussion of the kinds of validity evidence that have been gathered.

### *2.1.1. Purpose of the PAWS and SAWS*

The purposes of the PAWS and SAWS are multifold, as outlined in Chapters 1 and 3. The assessment is intended to comply with federal mandates, to inform ongoing instruction, and to help teachers plan instruction for the following year. Additionally, the PAWS in grades 3 through 8 are used in determining AYP that applies toward meeting the requirement of the federal No Child Left Behind (NCLB) Act of 2001.

### *2.1.2. The Constructs to Be Measured*

The PAWS and SAWS are designed to show how well students perform relative to the Wyoming content standards. These content standards describe what students should know and be able to do at each grade level.

Test blueprints and specifications define the procedures used to measure the content standards. These documents also provide an operational definition of the construct to which each set of standards refers. That is, they define, for each subject area the content to be assessed, the tasks to be presented, the administration instructions to be given, and the rules used to score examinee responses. They control as many aspects of the measurement procedure as possible so that the testing conditions will remain the same over test administrations (Cronbach, 1971; Cronbach, Gleser, Nanda, and Rajaratnam, 1972) in order to minimize construct irrelevant score variance (Messick, 1989). The content blueprints for the PAWS and SAWS can be found in Chapter 3, Appendix A, and on the WDE Web page at <http://edu.wyoming.gov/educators/assessment/paws/>. ETS has developed all PAWS and SAWS test items to conform to the Wyoming content standards and test blueprints.

### *2.1.3. The Interpretations and Uses of the Scores Generated*

Total scores expressed as scale scores, student performance levels, and subscores for each reporting cluster are generated for each subject area test. Based on a student's total score, an inference is drawn about how much knowledge and skill in the subject area the student has. The total score is also used to classify students in terms of their level of knowledge and skill in the subject area. These levels are called performance levels and are as follows: advanced, proficient, basic, and below basic.

Subscores are used to draw inferences about a student's achievement in each of several specific knowledge or skill areas covered by each test. Subscore results compare an individual student's scale score to the average scale score for the state as a whole. A detailed description of the uses and applications of PAWS and SAWS scores is presented in Chapter 8. Examples of individual student reports are provided in Appendix B showing the report for reading, mathematics, and SAWS students in grades 3, 5, and 7, Appendix C demonstrating the reading, mathematics, and science for grades 4 and 8, and Appendix D for grade 6 reading and mathematics.

The tests that make up the PAWS and SAWS assessment provide results or score summaries that are used for different purposes. The four major purposes are:

1. Communicating with parents and guardians;
2. Informing decisions needed to support student achievement;
3. Evaluating school programs; and
4. Providing data for state and federal accountability programs for schools.

These are the only uses and interpretations of scores for which validity evidence has been gathered. If the user wishes to interpret or use the scores in other ways, the user is cautioned that the validity of doing so has not been established. The user is advised to gather evidence to support these additional interpretations or uses (AERA, APA, and NCME, 2014, Standard 1.4).

#### *2.1.4. Intended Test Population(s)*

Wyoming public school students are the intended test population for the PAWS and SAWS. Students in grades 3–8 are tested in reading and mathematics, and grades 3, 5, and 7 in writing. In addition, students in grades 4 and 8 take a grade-level science test. Section 1.5 provides details regarding state policy for student participation. Further details regarding student participation and accommodations can be found in chapter 4.

### *2.2 Evidence Based on Content-related Validity*

According to the AERA, APA, and NCME's *Standards for Educational and Psychological Testing* (2014), analyses that demonstrate a strong relationship between a test's content and the construct that the test was designed to measure can provide important evidence of validity. In current K–12 testing, the construct of interest usually is operationally defined by state content standards and the test blueprints that specify the content, format, and scoring of items that are admissible measures of the knowledge and skills described in the content standards. Evidence that the items meet these specifications and represent the domain of knowledge and skills referenced by the standards supports the inference that students' scores on these items can be appropriately regarded as measures of the intended construct.

As noted in the AERA, APA, and NCME's *Standards for Educational and Psychological Testing* (2014), evidence based on test content may involve logical analyses of test content in which experts judge the adequacy with which the test content conforms to the test specifications and represents the intended domain of content. Such reviews can also be used to determine whether the test content contains material that is not relevant to the construct of interest. Analyses of test content may also involve the use of empirical evidence of item quality.

The procedures used for test administration and test scoring are also to be considered in evaluating test content. As Kane (2006, p. 29) has noted, although evidence that appropriate administration and scoring procedures have been used does not provide compelling evidence to

support a particular score interpretation or use, such evidence may prove useful in refuting rival explanations of test results. Evidence based on content includes the following:

### *2.2.1. Description of the state standards*

As was noted in Chapter 1, Wyoming adopted rigorous content standards in 2008 for science. In 2012, the Wyoming State Legislature adopted the 2012 Wyoming Content Performance Standards (2012 WyCPS), for the PAWS assessment program for reading, mathematics, and SAWS. These standards were designed to guide instruction and learning for all students in the state and to bring Wyoming students to world-class levels of achievement.

### *2.2.2. Specifications and Blueprints*

ETS maintains item development specifications for each PAWS and SAWS assessment. The item specifications describe the characteristics of the items that should be written to measure each content standard. A thorough description of the specifications can be found in Chapter 3. Once the items are developed, ETS selects all PAWS items and SAWS prompts to conform to the Wyoming content standards and test blueprints. Test blueprints for the components of the PAWS and SAWS assessments were proposed by ETS and reviewed and approved by the WDE. There has been only one recent change in the blueprints for the PAWS with the removal of constructed response items. The content blueprints for the PAWS and SAWS can be found in Chapter 3, Appendix A, and on the WDE Web page at <http://edu.wyoming.gov/educators/assessment/paws/>.

### *2.2.3. Item development process*

A detailed description of the content and psychometric criteria applicable to the construction of the 2014 PAWS and SAWS is presented in Chapter 3.

### *2.2.4. Item review process*

Chapter 3 explains in detail the extensive item review process applied to items written for use in the PAWS and SAWS. In brief, items written for the PAWS and SAWS go through multiple review cycles and involve multiple groups of reviewers.

### *2.2.5. Form construction process*

For each test, the content standards, blueprints, and test specifications are used as the basis for choosing items. Additional targets for item difficulty that are used for test construction were defined in light of what are desirable statistical characteristics in test items and statistical evaluations of the PAWS items and SAWS prompts. Guidelines for test construction were established with the goal of maintaining parallel forms to the greatest extent possible from year to year. Details can be found in Chapter 3.



#### *2.2.6. Alignment study*

Strong alignment between standards and assessments is fundamental to meaningful measurement of student achievement and instructional effectiveness. Alignment results should demonstrate that the assessments represent the full range of the content standards and that these assessments measure student knowledge in the same manner and at the same level of complexity as expected in the content standards. The alignment study for the PAWS Science assessment was completed in previous years to the present administration and recommendations from those studies incorporated into current item and test development processes (for details please refer to past years' editions of the PAWS technical reports). Alignment studies for PAWS reading and mathematics are expected to be completed in the next two years.

### *2.3 Evidence Based on Relations to Other Variables*

Empirical results concerning the relationships between scores on a test and measures of other variables external to the test can also provide evidence of validity when these relationships are found to be consistent with the definition of the construct that the test is intended to measure. As indicated in the *Test Standards* (AERA, APA, and NCME, 2014), the variables investigated can include other tests that measure the same construct and different constructs, criterion measures that scores on the test are expected to predict, as well as demographic characteristics of examinees that are expected to be related and unrelated to test performance.

#### *2.3.1. Correlations between Content Areas*

To the degree that students' content area scores correlate as expected, evidence of the validity in regarding those scores as measures of the intended constructs is provided. Appendixes E (field test) and F (operational) provide the intercorrelations for SAWS. Correlations among SAWS 12-point prompt operational traits were lower (ranging from 0.59 to 0.78) than the correlations between traits and total scores, with the majority of the correlations in the 0.65 range. This indicated that the four traits previously mentioned are distinctive features of writing. Correlations of traits with the total scores are necessarily higher than those between the components.

PAWS Reading, Mathematics, and Science tests and subscale intercorrelations are presented in Appendix F. There are strong relationships between the PAWS Reading, Mathematics, and Science scores. In the grades where Science was tested, it tended to be more strongly related to both Reading and Mathematics than Reading was to Mathematics, though the average difference was small. For Reading, this was probably because the Science items were tied to common passages rather than being discrete and independent items, thus requiring more reading ability. For Mathematics, Science items often involve mathematical functions or terms, thus giving students with higher levels of mathematical ability an advantage in answering them. The strong relationships between the scaled scores for Reading, Mathematics, and Science support the validity of the PAWS assessments. Taken together, they can be seen as measuring scholarship or

academic achievement, and they tend to covary together as would be expected. All can be seen to have strong relationships with the other subscales within each of the subjects, indicating that the subscales are measuring different yet related areas of knowledge.

### *2.3.2. Differential Item Functioning Analyses*

Analyses of DIF can provide evidence of the degree to which a score interpretation or use is valid for individuals who differ in particular demographic characteristics. For PAWS and SAWS assessments, DIF analyses were performed on all field-test items for which sufficient student samples were available.

The results of the DIF analyses are presented in Appendix G. The vast majority of the items exhibited little or no significant DIF, suggesting that, in general, scores based on the PAWS/SAWS items would have the same meaning for individuals who differed in their demographic characteristics. Due to small case counts, DIF analyses for ethnicities were not performed.

## *2.4 Evidence Based on Response Processes*

As noted in the AERA, APA, and NCME's *Standards for Educational and Psychological Testing* (2014) additional support for a particular score interpretation or use can be provided by theoretical and empirical evidence indicating that examinees are using the intended response processes when responding to the items in a test. This evidence may be gathered from interacting with examinees in order to understand what processes underlie their item responses. Finally, evidence may also be derived from feedback provided by observers or judges involved in the scoring of examinee responses.

### *2.4.1. Evidence of Interrater Agreement*

Rater consistency for the writing prompt is critical to the SAWS scores and their interpretations. These findings provide evidence of the degree to which raters agree in their observations about the qualities evident in students' responses. In order to evaluate the reliability of the student scores on the writing prompts, two raters scored approximately 25% of the examinee responses. The data collected were used to evaluate interrater reliability and interrater agreement.

#### *2.4.1.1. Interrater Reliability*

Cohen's Kappa statistics findings provide evidence of the degree to which a student's score may vary from rater to rater. Without explicit criteria to guide the rating process, two independent raters might not assign the same score to a given response. The results showed moderate levels of agreement between raters' scores on examinees' written responses to the prompts administered in grades 3, 5, and 7. Appendix I contains the weighted kappas for all SAWS prompts scoring.

#### *2.4.1.2. Interrater Agreement*

As noted previously, 25% of the responses to the SAWS prompts were scored by two raters. The total score of 12-point prompts showed approximately 24.0% exact agreement and 54.0% exact + adjacent agreement. Tables 32–34, in Chapter 8, provide the agreement rates for the SAWS operational prompts. The traits scores, having only 3 points, resulted in an exact score agreement from 55.4% to 63.0%. Exact + adjacent score agreement ranged from 98.0% to 98.7%. The grades 5 and 7 prompts worth a total of 4-points showed approximately 62.0% exact agreement and 95.0% exact + adjacent agreement. The traits scores, having only 2 points, resulted in an exact score agreement from 73.9% to 81.8%. Exact + adjacent score agreement ranged from 99.4% to 99.9%. The 8-points prompt total score showed approximately 35.3% exact agreement and 82.1% exact + adjacent agreement. The 8-point response to text traits scores, having only 2 points, resulted in an exact score agreement from 67.1% to 72.7%. Exact + adjacent score agreement ranged from 97.5% to 99.3%. The 8-point holistic traits score, having 6 points, resulted in an exact score agreement from 43.8% to 46.2%. Exact + adjacent score agreement ranged from 89.2% to 91.1%. Appendices H and I present the results of the SAWS field test and operational interrater reliability.

### *2.5 Evidence Based on Internal Structure*

As suggested by the *Standards* (AERA, APA, and NCME, 2014), evidence of validity can also be obtained from studies of the properties of the scores and the relationship between these scores and scores on components of the test. To the extent that the score properties and relationships found are consistent with the definition of the construct measured by test, support is gained for interpreting these scores as measures of the construct.

For the PAWS and SAWS, it is assumed that a single construct underlies the total scores obtained on each test. Evidence to support this assumption can be gathered from the results of item analyses, evaluations of internal consistency, and studies of model-data fit, dimensionality, and reliability.

With respect to the subscores that are reported, these scores are intended to reflect examinees' knowledge and/or skill in an area that is part of the construct underlying the total test. Analyses of the intercorrelations among the subscores themselves and between the subscores and total test score can be used for this purpose. Information about the internal consistency of the items on which each subscore is based is also useful and is provided in Section 8.2.

#### *2.5.1. Classical Statistics*

Point biserial correlations calculated for the items in a test show the degree to which the items discriminate between students with low and high scores on a test. To the degree that the correlations are high, evidence that the items assess the same construct is provided. The point

biserials for the items in the PAWS are presented in Appendices J (field test) and K (operational).

Also germane to the validity of a score interpretation are the ranges of item difficulty for the items on which a test score will be based. The finding that items have difficulties that span the range of examinee ability provides evidence that examinees at all levels of ability are adequately measured by the items. Information on average item score (i.e.,  $p$ -values) is given in Appendices J (field test) and K (operational); the distributions of item  $b$ -values are given in Appendices L (field test) and M (operational). A description of  $p$ -values and item means can be found in Section 3.7.1.1. Item Difficulty. Section 3.7.3. Item Response Theory (IRT) Analysis and Chapter 6 provide details about  $b$ -values.

SAWS classical statistics are also provided at the end of the Appendices J (field test) and K (operational) and in Tables J15–J28 and K15–K27. The prompt distributions show the ranges of student scores across the score levels and the prompt means across the prompt types.

### 2.5.2. Reliability

Reliability is a prerequisite for *validity*. The finding of reliability in student scores supports the validity of the inference that the scores reflect a stable construct. This section will describe briefly findings concerning the total test reliability, as well as reliability results for the reporting clusters.

**Overall reliability**—The reliability analyses on each of the PAWS assessments are presented in Chapter 8. The results indicate that the reliabilities for all PAWS were medium to high, ranging from 0.81 to 0.93.

**Reliability of performance classifications**—The methodology used for estimating the reliability of classification decisions is described in section 8.6, Accuracy and Consistency of Classifications. These levels of accuracy and consistency are high, and they are consistent with levels seen in previous years.

### 2.5.3. Dimensionality

Measurement using IRT implies order and magnitude on a single dimension (Andrich, 1989). However, unidimensionality cannot be strictly met in a real testing situation because students' cognitive, personality, and test-taking factors usually have a unique influence on their test performance to some level (Andrich, 1988; Hambleton, Swaminathan, and Rogers, 1991). Consequently, what is required for unidimensionality to be met is an investigation of the presence of a dominant factor that influences test performance. If present, this dominant factor can be considered to be the ability measured by the test (Andrich, 1988; Hambleton et al., 1991; Ryan, 1983). The results of science dimensionality studies are provided in the 2012 PAWS

Technical Report. The PAWS Reading and Mathematics dimensionality is planned for 2014–2015 administration.

### 2.6 Evidence Based on Consequences of Testing

As observed in the AERA, APA, and NCME’s *Standards for Educational and Psychological Testing* (2014), tests are usually administered “with the expectation that some benefit will be realized from the intended use of the scores” (p. 18). When this is the case, expected benefits evidence will provide support for intended use of the scores. The WDE and ETS are in the process of determining what kinds of information can be gathered to assess the consequences of administration of the PAWS.

### 3. PAWS AND SAWS TEST DESIGN AND DEVELOPMENT

#### 3.1 Overview

The Wyoming PAWS and SAWS statewide assessments adhere to the principles of sound and ethical test construction set forth in the *Standards for Educational and Psychological Testing* (AERA, 2014). These assessments comply with the requirements of NCLB (P.L. 107–110) and were designed to provide teachers with information to improve instruction based on the Wyoming Content and Performance Standards.

#### 3.2 Test Design and Blueprints

##### *3.2.1. Purpose*

Standards-based educational reform began in Wyoming in 1997–98, with adoption of rigorous academic content standards in language arts,<sup>2</sup> mathematics, science, and social studies.<sup>3</sup> Wyoming educators have continued the other earlier efforts to implement standards-based curriculum and assessment to meet the goals of improving teaching and the academic achievement of all of our students.

In 2004, the Wyoming Legislature passed a law describing the purpose and implementation of a statewide assessment system (§21-2-304) in order to meet the requirements of NCLB.<sup>4</sup> As a result, PAWS became the official instrument for measuring individual student achievement. Results of student achievement are reported at the student level and aggregated at the classroom, school, district, and state levels. As previously noted, the primary purpose of the PAWS is to foster program improvement at the school, district, and state levels that supports the teaching and learning that takes place in Wyoming public classrooms. The construction of PAWS also ensures that it meets NCLB requirements. Improvement of teaching and learning in schools and fostering school program improvement are the primary purposes of statewide assessment of student performance in Wyoming.

To achieve these goals, the first step taken by the WDE in early 2004 was to contract Dr. Robert Marzano to evaluate the Wyoming Content and Performance Standards with the intent of developing an organizing framework for reading, writing, and mathematics content. The second step was to empanel content experts from around the state to review and revise Dr. Marzano's

---

<sup>2</sup> As previously noted, Wyoming tests only the Reading Language Arts Standards.

<sup>3</sup> Social studies is not presently tested in the PAWS assessments.

<sup>4</sup> The decision is made based on the recommendations of the Wyoming Statewide Task Force on Student Assessment and Education Accountability.

work. The major purpose of this exercise was the support of an assessment design that measured integrated concepts and skills. The WDE undertook this challenging task in order to better promote student learning of clear and rigorous content.

The documents were open to public comment during the fall of 2004 and again in 2012 for the 2012 WyCPS. From these documents arose the guiding principle of the design of PAWS and SAWS as an assessment focused on powerful, content-subsuming cognitive skills and not on isolated collections of information. Thus, the knowledge, skills, and the expectation of Wyoming student performance as envisioned by Wyoming teachers and the Wyoming Content and Performance Standards led to the development of the PAWS and SAWS blueprints and specifications.

### *3.2.2. Plan*

The first step in test development is to create item and test specifications. WDE's test specifications reflect skill expectations that are outlined in Wyoming's Content and Performance Standards. These item specifications established guidelines for selecting test content and writing test items. For PAWS, the specifications determined both the composition of the item pool and the rules for item selection.

The academic content and skills measured by a test and distributions of emphasis are set forth in the test blueprints and test specifications along with the number of points possible in each category. The test blueprints and test specifications were developed by content specialists of the Wyoming Department of Education and staff at ETS, based on the Wyoming Content and Performance Standards.

Wyoming considers a test blueprint to be a detailed plan for building test forms. The blueprint and specifications include:

- Knowledge and skills as specified in the reading, mathematics, science, and writing standards to be tested
- Number of items and points per test form
- Percentage and/or number of items and points per content standard
- Distribution of multiple item types (multiple choice and constructed response)
- Proposed distribution of items by cognitive complexity, i.e., percentage of items with low, moderate, or high levels of cognitive complexity

Approximate time requirements for each assessment

### *3.3 Types of Items Used in PAWS and SAWS*

Consistent with Wyoming state law, legislation passed in 2013 [Enrolled Act 90, The Wyoming Accountability in Education Act] modified this requirement; beginning with the 2014

administration, the PAWS assessments were composed solely of multiple choice items for the PAWS assessments. Each item measures a single skill-reporting category within a content standard. Multiple-choice items have four response options and do not use “none of the above” or “all of the above” as response options. Reading and science items are grouped together into item sets that refer to a common passage. For SAWS tests, each participating grade 3 student responded to two 12-point prompts, one Narrative and the other Informative. The students participating in grades 5 and 7 had a single 12-point Narrative writing prompt and a 4 + 8-point set. A student was provided a single passage to read and then answered a 4-point response followed by an 8-point response.

The PAWS assessment is used to measure individual student achievement against the newly adopted Wyoming Content and Performance Standards (2012 WyCPS) in Reading and Mathematics. For Science, the WyCPS adopted in 2008 remains in place. From 2012 to 2014, PAWS Reading and Mathematics blueprints were revised and test items developed to better align these assessments with the 2012 WyCPS. The newly developed blueprints are detailed in Appendix A.

The Wyoming Content and Performance Standards identify knowledge and skills students are expected to acquire at each grade in order to succeed in school and at work. It is important to develop items that elicit the complexity of knowledge required to meet these objectives. The degree of challenge on PAWS items is categorized based on Dr. Norman Webb’s work with Depth of Knowledge levels. The categories *low complexity*, *moderate complexity*, and *high complexity* form an ordered description of the cognitive load involved in responding to the item.

### 3.3.1. PAWS Reading Tests

The Wyoming Language Arts Content and Performance Standards include an expectation that all students will become effective raters, writers, listeners, and speakers. However, due to the limitations of large-scale testing and the desire to minimize student time spent on testing, the Wyoming Legislature determined that only reading will be assessed by PAWS (Beginning in 2014, a reconfigured writing test was administered apart from the PAWS assessment). The WDE provides ongoing technical support and guidance for schools and districts to include instruction and monitoring of student achievement in the areas of listening, speaking, and writing, but these measures are not included in the state’s determinations of adequate yearly progress (AYP) under NCLB.

The PAWS reading assessment is designed to measure the reading content standard requiring that students use the reading process to apply a variety of comprehension strategies and demonstrate an understanding of literary and informational text. Testing of Wyoming students’ reading comprehension skills relative to the reading proficiency goals required to meet the standards is one component of the PAWS. Students were tested in reading at grades 3 through 8.



Reading concepts were measured by requiring students to examine texts with accuracy, to make relevant connections, and to support their inferences.

The structure of the operational 2014 PAWS reading test was based on the 2014 PAWS Reading Blueprint (see Appendix A). The content of the test is aligned to the reading content standards of the Wyoming Language Arts Content and Performance Standards. The PAWS assessment is designed to assess overall literacy skills in the following skill-reporting categories:

- Determine information's relevance and importance, and select and apply information for a task within a functional text;

Understand main points and supporting details, recognize expository organization and its use, and see relationship of text's content to broader issues/topics within an expository text; and

- Identify the development of basic story elements, understand a story's plot development, and identify a story's theme(s) and its (their) development within a narrative text.

Three Content Standards are assessed for each grade for grades 3–8: *Reading Literature*, *Reading Informational Text*, and *Language*.

Within the Content Standards of *Reading Literature* and *Reading Informational Text*, there are four benchmarks for each grade for grades 3–8: Key Ideas and Details, Craft and Structure, Integration of Knowledge and Ideas, and Range of Reading and Level of Text Complexity. There are no additional benchmarks for the Content Standard of *Language*.

The 2013–2014 PAWS Reading blueprints and reporting categories for each of the grade levels are provided in Appendix A. As noted in the blueprints, the percentage of assessment coverage of text type reflects the emphasis of instruction in Wyoming classrooms across grades. Tables 5–10 provide the number of items for each reading assessment by reporting strand, for the assessment overall and for the vertical scale set. Note: Integration of Knowledge and Ideas reporting category items were still being field tested and therefore not reported in 2014. In general, the proportions of vertical scale items in each strand reflect the overall reporting category distribution in the assessment.

Table 5. Reporting Strands Design for Grade 3 Reading

Reporting Category Name	No. of Items	% of Items
Lit: Key Ideas and Details	20	40
Lit: Craft and Structure	6	12
Inf.: Key Ideas and Details	10	20
Inf.: Craft and Structure	7	14
Integration of Knowledge and Ideas*	0	0
Language	7	14
Total	50	

Table 6. Reporting Strands Design for Grade 4 Reading

Reporting Category Name	No. of Items	% of Items
Lit: Key Ideas and Details	15	30
Lit: Craft and Structure	6	12
Inf.: Key Ideas and Details	15	30
Inf.: Craft and Structure	8	16
Integration of Knowledge and Ideas*	0	0
Language	6	12
Total	50	

Table 7. Reporting Strands Design for Grade 5 Reading

Reporting Category Name	No. of Items	% of Items
Lit: Key Ideas and Details	14	26
Lit: Craft and Structure	7	13
Inf.: Key Ideas and Details	17	31
Inf.: Craft and Structure	8	15
Integration of Knowledge and Ideas*	0	0
Language	8	15
Total	54	

Table 8. Reporting Strands Design for Grade 6 Reading

Reporting Category Name	No. of Items	% of Items
Lit: Key Ideas and Details	15	27
Lit: Craft and Structure	9	16
Inf.: Key Ideas and Details	15	27
Inf.: Craft and Structure	9	16
Integration of Knowledge and Ideas*	0	0
Language	8	14
Total	56	

Table 9. Reporting Strands Design for Grade 7 Reading

Reporting Category Name	No. of Items	% of Items
Lit: Key Ideas and Details	13	23
Lit: Craft and Structure	9	16
Inf.: Key Ideas and Details	19	34
Inf.: Craft and Structure	8	14
Integration of Knowledge and Ideas*	0	0
Language	7	13
Total	56	

Table 10. Reporting Strands Design for Grade 8 Reading

Reporting Category Name	No. of Items	% of Items
Lit: Key Ideas and Details	12	21
Lit: Craft and Structure	7	13
Inf.: Key Ideas and Details	20	36
Inf.: Craft and Structure	9	16
Integration of Knowledge and Ideas*	0	0
Language	8	14
Total	56	

### 3.3.2. PAWS Mathematics Tests

In the area of mathematics, the focus is on the ability of students to demonstrate basic computational skills along with the higher-level thinking skills of reasoning and problem solving. To achieve this end, the PAWS mathematics assessment is designed to measure whether students have acquired the skills to analyze, reason, and communicate ideas effectively as they pose, formulate, solve, and interpret mathematical problems in a variety of real-world situations. Because of this, Wyoming’s framework for assessing Mathematics is based upon mathematical problem solving.

The structure of the operational 2014 PAWS Mathematics test is explicated in the 2014 PAWS Mathematics Blueprints (see Appendix A). The content of the test is aligned to the five content standards within the Wyoming Mathematics Content and Performance Standards for grades 3–5:

- Operations and Algebraic Thinking
- Number and Operations – Base Ten
- Number and Operations – Fractions
- Measurement and Data
- Geometry

Five content standards within the Wyoming Mathematics Content and Performance Standards for grades 6 and 7:

- Ratios and Proportional Relationships
- The Number System
- Expressions and Equations
- Geometry
- Statistics and Probability

Five content standards within the Wyoming Mathematics Content and Performance Standards for grade 8:

- The Number System
- Expressions and Equations
- Functions
- Geometry
- Statistics and Probability

The 2014 PAWS Mathematics blueprints and reporting categories for each of the grade levels are provided in Appendix A. As noted in the tables below, the percentage of assessment coverage of each content standard reflects the emphasis of instruction in Wyoming classrooms across grades. For example, at grade 8 the emphasis is placed upon Expressions and Equations. Calculator use is not permitted for the grades 3–5 assessments. Tables 11–16 provide the number of items for each Mathematics assessment by reporting strand, for the assessment overall.

Table 11. Reporting Strands Design for Grade 3 Mathematics

Reporting Category Name	No. of Items	% of Items
Operations and Algebraic Thinking	20	40
Number Operations–Base Ten	6	12
Number Operations–Fractions	6	12
Measurement and Data	12	24
Geometry	6	12
Total	50	

Table 12. Reporting Strands Design for Grade 4 Mathematics

Reporting Category Name	No. of Items	% of Items
Operations and Algebraic Thinking	13	22
Number Operations–Base Ten	10	17
Number Operations–Fractions	20	34
Measurement and Data	10	17
Geometry	6	10
Total	59	

Table 13. Reporting Strands Design for Grade 5 Mathematics

Reporting Category Name	No. of Items	% of Items
Operations and Algebraic Thinking	6	10
Number Operations–Base Ten	16	27
Number Operations–Fractions	19	32
Measurement and Data	12	20
Geometry	6	10
Total	59	

Table 14. Reporting Strands Design for Grade 6 Mathematics

Reporting Category Name	No. of Items	% of Items
Geometry	6	10
Ratios and Proportional Relationships	10	17
The Number System	15	25
Expressions and Equations	20	34
Statistics and Probability	8	14
Total	59	

Table 15. Reporting Strands Design for Grade 7 Mathematics

Reporting Category Name	No. of Items	% of Items
Geometry	9	15
Ratios and Proportional Relationships	13	22
The Number System	10	17
Expressions and Equations	18	31
Statistics and Probability	9	15
Total	59	

Table 16. Reporting Strands Design for Grade 8 Mathematics

Reporting Category Name	No. of Items	% of Items
Geometry	16	25
Ratios and Proportional Relationships		
The Number System	6	9
Expressions and Equations	23	35
Statistics and Probability	6	9
Functions	14	22
Total	65	

### 3.3.3. PAWS Science Tests

The Wyoming Science Content and Performance Standards specify that all students should understand science concepts and processes, scientific inquiry, and the history and nature of science. Because of the constraints of space available on the assessment and the desire to limit testing time, the WDE determined that only the skills of science concepts and processes and scientific inquiry would be assessed by PAWS, as these skills allow students to process, apply, and effectively communicate scientific knowledge. The WDE provides support and guidance for schools and districts to ensure that instruction and monitoring of student achievement in the areas of the history and nature of science take place at the local level, but these measures are not assessed by the PAWS at present.

In order to accurately reflect the expectations of the Wyoming Science Content and Performance Standards, the PAWS science assessments for grades 4 and 8 are designed to measure students' abilities to connect science knowledge with science process. The Wyoming Performance Standards instruct teachers to judge where students are performing in relation to the benchmarks, and ultimately, the standards. To evaluate students' mastery against the Wyoming Performance Level Descriptors, teachers are required to measure each student's ability to make connections among concepts and processes and apply scientific information as the criteria for determining performance levels (advanced, proficient, basic, and below basic). As stated in the Wyoming Science Content and Performance Standards, students develop an understanding of scientific content through inquiry. Therefore, when considering the appropriateness of the PAWS science tests, careful consideration was given to the relevant criterion intended to be measured and the alignment to the intent of the Wyoming Science Content and Performance Standards, notably, the science performance inferences to be drawn from the results.

Based on this design, the PAWS science assessment items are written to measure students' mastery of science inquiry skills within the context of the benchmarks from Standard I: Concepts and Processes. The items are distributed equally among the physical science, life science, and earth/space science benchmarks. Over the course of a two-year cycle, each of the inquiry skills is assessed within the context of each benchmark in Standard I Concepts and Processes. All too

often, students' understanding of core concepts and scientific theories is measured without careful attention to how students internalize core assumptions, apply important ideas, or make connections to relevant everyday experiences. Without measurement of such epistemological standards, teachers will not know whether students have a firm foundation on which to base scientific arguments.

The design of both the Wyoming Science Content and Performance Standards and the PAWS science assessments is based on a view of proficiency in science that values students' understanding of science concepts and their ability to think critically and apply scientific logic and reasoning, rather than simply memorizing and recalling science facts. Students were tested in science at grades 4 and 8. Science concepts and inquiry skills were measured by requiring students to examine scientific investigations accurately, to make relevant connections, and to support their inferences.

The structure of the operational 2014 PAWS science test was based on the 2013–2014 PAWS science Blueprint. The content of the test is aligned to the Science as Inquiry content standard of the Wyoming Science Content and Performance Standards. Because scientific inquiry involves many processes, the PAWS assessment is designed to assess inquiry skills overall in the following skill reporting categories:

- Use observation to pose questions that can be addressed through a scientific investigation;
- Design and conduct a scientific investigation;
- Organize and represent data; and
- Draw conclusions and make connections with concepts and knowledge.

The content of the test is aligned to the three content areas within the Wyoming Science Content and Performance Standard I: Concepts and Processes, and a score analysis is reported in each of the following areas:

- Life science;
- Physical science; and
- Earth/Space science

The number of items assessing each skill-reporting category and content standard is constant across all grade levels, they are provided in Appendix A. Tables 17, and 18 provide the number of items (and points) for each Science assessment, by reporting strand for the assessment overall, and for the anchor item set. The anchor item set is utilized for year to year equating. These tables include similar information for 2013 and 2014 for comparison purposes. There is some fluctuation between the percentage of the total raw score represented by the reporting strand and the percentage of anchor item points for the strand. Across all grades, the representation of the

anchor sets remained stable although the number of anchor items was reduced to ensure equal representation.

Table 17. Reporting Strands Design for Grade 4 Science

Strand	Number of Items			
	<u>2013</u>		<u>2014</u>	
	Total Test	Anchors	Total Test	Anchors
	Items	Items	Items	Items
LIFE	16	9	16	6
PHYS	18	6	18	7
ESCI	16	6	16	6
Totals	50	21	50	19

Strand	Percentages of Items			
	<u>2013</u>		<u>2014</u>	
	Total Test	Anchors	Total Test	Anchors
	Items	Items	Items	Items
LIFE	32	43	32	32
PHYS	36	29	36	37
ESCI	32	29	32	32



Table 18. Reporting Strands Design for Grade 8 Science

Strand	Number of Items			
	<u>2013</u>		<u>2014</u>	
	Total Test	Anchors	Total Test	Anchors
	Items	Items	Items	Items
LIFE	16	3	16	7
PHYS	18	6	18	8
ESCI	16	12	12	4
Totals	50	21	46 <sup>5</sup>	19

Strand	Percentages of Items			
	<u>2013</u>		<u>2014</u>	
	Total Test	Anchors	Total Test	Anchors
	Items	Items	Items	Items
LIFE	32	14	35	37
PHYS	34	29	39	42
ESCI	32	57	26	21

### 3.3.4. SAWS Tests

The SAWS assessments are open-ended, requiring students to write responses to multiple prompts at grades 3, 5, and 7. The writing assessment is designed to measure current (2012) Wyoming writing content standards as shown in Table 19. Table 20 provides the grades, prompt types, and operational standards measured in the 2013–2014 administration. In addition to the operational prompts, students were also administered additional field test prompts in the same testing window, as presented in Table 21. The multiple writing prompts required three sessions to administer, each of approximately 1.5 hours.

All texts and prompts were chosen and developed to be free of age, gender, geographic, ethnic, socioeconomic, religious, or physical disability stereotypes. Committees of Wyoming educators have confirmed there is no apparent bias in any text or prompt. Stand-alone prompts address experiences and interests common to the student’s age level. Although they offer the opportunity

---

<sup>5</sup> A passage in Grade 8 Science Test Booklet also appeared in the Released Test Questions posted earlier this year on the WDE website. After reviewing the consequences, ETS decided to remove the four questions altogether from the scorable bank of operational items and treat them as a “Do Not Score.” This reduced the number of maximum raw points for Science in Grade 8, but did not bear any negative consequences to content or statistical reliability.

to write from experience, prompts were designed not to intrude on the student’s personal feelings or to require a student to discuss personal values.

Table 19. SAWS Standards

Grade	Standard 1 – Opinion Response to Text	Standards 1 , 2, or 3	Standard 9 – Response to Literary or Informational Text (Set Items)	
Number of Score Points	8	12	4	8
3	✓	2 – Narrative or 3 - Informational	-	-
5	-	1 – Opinion, 2 – Narrative, or 3 - Informational	✓	✓
7	-	1 – Argument, 2 – Narrative, or 3 - Informational	✓	✓

Table 20. SAWS Grades and Content

<u>Operational</u>				
Grade	Operational Test	Operational Standard	Number of Field Test Forms	Field Test
3	2 12-point Prompts <sup>6</sup>	Informational and Narrative	6	8-point Prompt
5	12-point Prompt and 4 + 8 set <sup>7</sup>	Informational and Std 9	8	12-point Prompt Or 4 + 8 set
7	12-point Prompt and 4 + 8 set	Informational and Std 9	8	12-point Prompt Or 4 + 8 set

<sup>6</sup> All 12-point prompts are scored analytically.

<sup>7</sup> All 4- and 8-point prompts are scored holistically.

Table 21. SAWS Grades, Prompt Type, and Field Test Design

Form	Grade 3		Grade 5		Grade 7	
	Operational	Field Test	Operational	Field Test	Operational	Field Test
1	12 and 12	8	12 and 4 + 8	12 Std 1	12 and 4 + 8	12 Std 1
2	12 and 12	8	12 and 4 + 8	12 Std 2	12 and 4 + 8	12 Std 2
3	12 and 12	8	12 and 4 + 8	12 Std 3	12 and 4 + 8	12 Std 3
4	12 and 12	8	12 and 4 + 8	4 + 8	12 and 4 + 8	4 + 8
5	12 and 12	8	12 and 4 + 8	4 + 8	12 and 4 + 8	4 + 8
6	12 and 12	8	12 and 4 + 8	4 + 8	12 and 4 + 8	4 + 8
7	-	-	12 and 4 + 8	4 + 8	12 and 4 + 8	4 + 8
8	-	-	12 and 4 + 8	4 + 8	12 and 4 + 8	4 + 8

### *12-point prompt*

The 12-point prompt for grades 3, 5, and 7 assessed one of the following standards:

**Standard 1:** Opinion/Argument Essay: Grades 3 and 5 students wrote opinion pieces on topics or texts, supporting a point of view with reasons and information. Grade 7 students wrote arguments to support claims with clear reasons and relevant evidence. This prompt required students to write an essay scored using a 12-point analytic scoring guide.

**Standard 2:** Informative/Explanatory Essay: Students in grades 3, 5, and 7 wrote informative/explanatory texts to examine a topic and convey ideas and information clearly. Please see each grade level standard for a more detailed description of the writing skills measured. This prompt required students to write an essay scored using a 12-point analytic scoring guide. Scoring guides for each grade can be found at the links below.

**Standard 3:** Narrative Essay: Students in grades 3, 5, and 7 wrote narratives to develop real or imagined experiences or events using effective technique, descriptive details, and clear event sequences. Please see each grade level standard for a more detailed description of the writing skills measured.

Each student response was scored against the following Wyoming benchmark writing skills: Idea Development, Organization, Voice, and Conventions. Each skill received 0 to 3 points, with blank and invalid responses receiving the same score as a minimal response (zero points). Off-topic responses received zero points for the Idea Development skill, with the scores for the remaining three skills determined according to the rubric. Response to each prompt received up to twelve points on the SAWS writing assessment at all grade levels.

### *8-point Prompt, Grade 3*

The grade 3 response-to-text/opinion field-test prompts assessed Standard 1 and included a brief informational, functional, or literary text that offered context for the prompt that is intended to elicit an opinion essay. These essays were assigned a holistic writing score of 2, 4, 6, or 8 points based on the scoring rubric.

#### *4- and 8-point Response-to-Text Prompt Sets, Grades 5 and 7*

For these items, students read either an informational or a literary text and then responded to a 4-point prompt and an 8-point prompt related to the text. Students received up to 2 points per item/prompt for specifically citing relevant information from the text, and up to an additional 2 or 6 points for an overall holistic writing score. The 4-point and 8-point Response-to-Text Prompts, assessed at grades 5 and 7 addressed the following standard:

**Standard 9:** Students addressed a pair of prompts after they read one passage, each prompt drew evidence from literary or informational texts to support analysis, reflection, and research. Please see grade level standards for a more detailed description of the writing skills measured. This prompt type required a short 4-point response and a longer 8-point response; each response was scored using a holistic scoring guide. There were a total of 12 points possible for the prompts addressing Standard 9.

(Released prompts and analytic Scoring Guides for each grade and Standard are available at <http://edu.wyoming.gov/educators/assessment/saws/>).

### *3.4 PAWS Test Development Process*

A state committee, consisting of regional representatives, utilized national and regional documents to establish that the rigor of the Wyoming Language Arts standards are consistent with these documents, and adjustments were made as deemed appropriate by the state committees.<sup>8</sup> The Wyoming Language Arts Content and Performance Standards address three content standards: (1) Reading, (2) Writing, and (3) Speaking and Listening. Content standard 3, (Speaking and Listening), is not currently assessed by PAWS. Content standard 2 (Writing) is currently assessed by SAWS through a single writing prompt.

Multiple choice items were used on the PAWS reading, mathematics, and science portions. The SAWS writing prompts were classified as Stand Alone prompts (SA), Short Response (SR) and Extended Response (ER). Stand Alone response templates were 3 pages long and had a maximum possible score of 12 points. Short Response item templates were ½ page long and had

---

<sup>8</sup> These documents included the following publications:

- National Council of Teachers of English and International Reading Program
- Standards for the English Language Arts;
- National Center on Education and the Economy,
- New Standards Performance Standards; Speech Communication Association, Speaking, Listening, and Media Literacy Standards for K through 12 Education, and
- Guidelines for Assessing Communication in Primary and Secondary Education; the Colorado Model Content Standards for Reading and writing; and the Standards of Learning for Virginia Public Schools.

a maximum possible score of 4 points. Extended Response item templates were 2 pages long and had a maximum possible score of 8 points.

The Wyoming Mathematics Content and Performance Standards are consistent with those of the National Council of Teachers of Mathematics (NCTM) as they are written in *Principles and Standards for School Mathematics* (April 2000). The Wyoming Mathematics Standards address five content standards: (1) Number Operations and Concepts, (2) Geometry, (3) Measurement, (4) Algebraic Concepts and Relationships, and (5) Data Analysis and Probability. Multiple choice items were used on the Mathematics portions of the PAWS.

The Wyoming Science Content and Performance Standards address three content standards: (1) Concepts and Processes, (2) Science as Inquiry, and (3) History and Nature of Science in Personal and Social Decisions. Content standard 3, History and Nature of Science in Personal and Social Decisions is not assessed by PAWS. Multiple choice items were used on the Science portions of the PAWS.

Initial creation of blueprints, item and passage specifications, and assessment descriptions took place in the fall of 2004. Development of these documents has been an ongoing process, and they guided the development, review, and field testing of items for use on the PAWS assessments.

Item development was a cooperative effort involving WDE and ETS content staff as well as Wyoming teachers. All items were authored by ETS content staff and reviewed by and revised at the direction of WDE content staff. After items were approved by WDE, they were then reviewed by committees of Wyoming educators (see Section 3.6 Item Review). Items approved at item review then became eligible for field testing, after which they were evaluated in light of their statistics from field testing (see Section 3.7.4, Data Review). Items approved at data review then were eligible for use as operational items.

The PAWS tests were constructed to produce assessments that are psychometrically sound, measure the academic content outlined in Wyoming's grade-level content standards and described in the test specifications, and to interest and engage students. WDE content staff and ETS content specialists and psychometricians collaborated to choose items for use on the 2014 forms considering both the content and psychometric properties of each item selected.

### 3.5 Item and Test Form Development

In this section, the general process for item development is described. Using the Wyoming Content and Performance Standards as a foundation, test blueprints were developed by the WDE setting forth the number of items for each Reading, Mathematics, or Science content standard. These blueprints were initially developed in the fall of 2006 for Science and 2012 for Reading and Mathematics. They have been refined during the course of the program, balancing the need

to provide a high level of information about student ability to inform instruction against the desire to impinge upon instructional time as little as possible.

Wyoming's item development procedures are consistent with industry practice and take approximately two years, including writing, review, and field-testing before an item is eligible for inclusion in the item pool.

#### *3.5.1. Item Specifications*

Test items were created by ETS item writers (Wyoming educators are involved in the item review process) who are selected for their academic content and grade-level experience and who are experienced in the development of statewide assessments. Item writers selected to write items for the PAWS were then trained on PAWS specific requirements, including the WY Content and Performance Standards for their specific grade and subject and style guidelines for the PAWS. These PAWS specific requirements were collected in an Item Specifications document. All items were written to measure specific content standards at a variety of specified levels of cognitive complexity as developed from Webb's Depth of Knowledge levels.

For example, the Mathematics Item Specifications were intended to accomplish two purposes: (1) to provide both general and specific guidelines for development of all test items at the grade levels assessed by PAWS Mathematics, and (2) to describe the test items and prompt types to be developed for the PAWS Mathematics assessments. Within the specifications document are sections dedicated to information about item contexts, cognitive task levels, use of graphics, item style and format, and general content limits by grade. Comparable information was provided for PAWS Reading and Science items in Reading and Science Item Specifications.

#### *3.5.2. Item Difficulty Requirements*

The Rasch measurement model was used to develop the scale for each of the PAWS Reading, Science, and Mathematics assessments. The Rasch model is robust and is used for many large-scale, high stakes assessment programs. In general, the Rasch model assumes that the probability that a student will answer an item correctly is a function of the latent trait that underlies performance on the assessment and the difficulty of the item. This underlying trait, usually referred to as ability, is nothing more than what the assessment is designed to measure (e.g., Mathematics, Reading, or Science). See chapter 5 for further detail on the Rasch model.

#### *3.5.3. Item Graphics Requirements*

Many items contain graphics. For example, mathematics items frequently contain charts, spinners, box-and-whisker plots, line graphics, clocks, and geometric shapes. WDE reviewed all test items and forms to ensure an appropriate use and balance of these types of graphics.

### 3.6 Item Review

Items accepted from ETS item writers for consideration by the PAWS program are reviewed against WDE-established criteria (i.e., alignment with Wyoming Content Standards, grade-level appropriateness, cognitive demand, appropriate item type, and bias) by ETS assessment specialists and content specialists at the WDE. ETS and the WDE collaborate to consider and implement WDE-proposed revisions to the items. Items passing this review phase become eligible for external review by Wyoming teachers.

Annually, an external review of items is completed by a panel of experienced teachers at each grade level selected by the WDE. Each panel has approximately 10–15 members. Panel members committed up to two weeks of service during the summer and were compensated for their service.

Most members of these panels are classroom teachers. University of Wyoming and district curriculum personnel have also participated. Criteria for the panel selection include the following:

- Knowledge of the Wyoming Content and Performance Standards and expertise in the subject area
- Teaching experience at the grade level to which the individual will be assigned
- Geographical location to ensure all regions of Wyoming are represented

All reviewers first received training in how to effectively evaluate items, including strategies for examining the overall technical qualities of all items, such as language clarity, readability, plausibility of options, parallel structure of response options, significance and suitability of subject content, lack of bias, veracity of the correct answer, proper level of difficulty, and alignment to Wyoming Content and Performance Standards.

The evaluations and recommendations of the educators for each item were evaluated by ETS and WDE. All of the feedback generated by the reviewers was utilized to make final decisions on which items to accept and what revisions to include in the version of the item that was field tested. Only the items that measure grade-level expectations are carried forward to the field-test stage. The criteria used for item review are listed below.

1. Conceptual criteria:

- Grade-level appropriateness
- Thinking skill match
- Lack of bias
- Clear statement
- One best answer
- Each distractor credible
- Meets all technical criteria for item parameters

2. Language criteria:

- Appropriate for age
- Correct punctuation
- Spelling and grammar
- Lack of excess words
- No stem/foil clues

3. Format criteria:

- Logical order of distractors
- Familiar presentation style, print size, and type
- Correct mechanics and appearance
- Equal-length distractors

4. Graphic stimuli criteria:

- Necessary
- Clean
- Relevant
- Unbiased

The item review panel also provided input on potential bias and/or sensitivity in the test content. With regard to fairness and content, panelists suggested revision or deletion of items as they deemed necessary. Any items that survived this rigorous examination became part of the pool of items eligible for field testing.

### 3.7 Field Testing

During the 2014 PAWS and SAWS administration, reading, mathematics, science, and writing field test items were embedded within the operational forms, respectively. In 2014, there were



ten field test forms for PAWS and six for SAWS for each grade and content area. Since field test items could appear on multiple forms within a grade level and the numbers of students per grade varied, the numbers of examinees attempting each field test item also varied. The items were responded to by between 600 and 2000 students, depending on subject, grade level, and number of field test forms the item appeared on. Student responses to the field test items did not affect their operational test scores. Data on the field test items were used only in data review as an aid in determining whether the item was suitable for use.

Field test forms were created to have the same length and same item types in the same relative positions across forms. They were spiraled within classroom and school in order that randomly equivalent samples of students would receive each of the forms. The WDE reviewed the assembled field test forms for clarity, correctness, potential bias, and curricular appropriateness. Field test items were indistinguishable from operational items so that the students' motivation in responding to them would be at the same level as their motivation in responding to operational items.

All field test items underwent comprehensive statistical analysis to provide the WDE with the information necessary to make informed decisions about the likelihood of each item providing reliable information that could be used in drawing valid inferences concerning student performance. The following analyses were conducted on the field test items (processes and findings are discussed below):

#### Classical item analyses

- Differential Item Functioning (DIF) analyses
- Rasch Item Response Theory (IRT) analyses

##### *3.7.1. Classical Item Statistics*

Classical item statistics were computed for all field test items in Mathematics, Reading, Science, and SAWS. The field test classical analysis results appear in Appendix J. For each item, the following statistics were computed:

- *N*-counts for each statistic;
- Item difficulty (or average item score);

#### Item discrimination (or point biserial correlation);

- Multiple choice item distractor discrimination for PAWS only;
- Multiple choice item response and constructed response score distributions (total and broken out by trait by form); and
- DIF statistics (Mantel and Haenszel, 1959) and standardized mean difference (SMD) by gender and ethnicity.

#### *3.7.1.1. Item Difficulty*

Item difficulty is typically defined as the average of scores for a given item. For multiple choice items, this value (commonly referred to as a  $p$ -value) ranged from 0 to 1. For the SAWS prompts, this value was expressed as item mean.

#### *3.7.1.2. Item Discrimination*

Item discrimination is defined here as the correlation between a score on a given test question and the overall operational raw test score. For multiple-choice items, it is also known as the point biserial correlation. The discrimination for multiple choice distractors (incorrect answer options) was also computed. The operational test score used in calculating this coefficient did not include field test item scores.

#### *3.7.2. Differential Item Functioning*

In addition to classical item analyses, Differential Item Functioning (DIF) analyses are conducted on the field test items. DIF statistics are not computed on operational items. DIF analyses are used to identify those items that identifiable groups of students (e.g., males, females) with the same underlying level of ability have different probabilities of answering correctly. Examinees are separated into relevant subgroups based on ethnicity or gender for analysis. Then examinees in each subgroup are ranked relative to their total test score (conditioning on ability). Examinees in the focal group (e.g., females) are compared to examinees in the reference group (e.g., males) relative to their performance on individual items.

If the item is differentially more difficult for an identifiable subgroup when conditioned on ability, it may be measuring something different from the intended construct. However, it is important to recognize that DIF-flagged items might be related to actual differences in relevant knowledge or skills (item impact) or statistical Type I error. As a result, DIF statistics are used to identify items that are potentially functioning differentially. Subsequent review by content experts and bias/sensitivity committees are required to determine the source and meaning of performance differences. For the spring 2014 PAWS Reading, Mathematics, and Science tests, DIF analyses were conducted for gender groups (male/female) and ethnicity groups (White/Asian, White/African American, White/Hispanic/Latino, and White/Native American) where sample size was sufficient.

Statistics from two DIF detection methods were computed: the Mantel-Haenszel procedure (Mantel and Haenszel, 1959) for multiple choice items and the standardization procedure (Dorans and Kulick, 1983, 1986) for writing prompts. As part of the Mantel-Haenszel procedure, the statistic described by Holland and Thayer (1988), known as MH D-DIF, was used.

The formula for the estimate of constant odds ratio is:

$$\alpha_{MH} = \frac{\left( \sum_m \frac{R_{rm} W_{fm}}{N_m} \right)}{\left( \sum_m \frac{R_{fm} W_{rm}}{N_m} \right)},$$

where

$R_{rm}$  = number in reference group at ability level  $m$  answering the item right,

$W_{fm}$  = number in focal group at ability level  $m$  answering the item wrong,

$R_{fm}$  = number in focal group at ability level  $m$  answering the item right,

$W_{rm}$  = number in reference group at ability level  $m$  answering the item wrong,

$N_m$  = total group at ability level  $m$ .

This statistic is expressed as the differences between members of the “focal group” (female, Asian, African American, Hispanic/Latino, and Native American) and members of the “reference group” (male and White) after conditioning on total operational test score. This statistic is reported on the ETS delta scale, which is a normalized transformation of item difficulty ( $p$ -value) with a mean of 13 and a standard deviation of 4. Negative MH D-DIF statistics favor the reference group and positive values favor the focal group. The classification logic used for flagging items is based on a combination of absolute differences and significance testing. Items that are not statistically significantly different based on the MH D-DIF ( $p > 0.05$ ) are considered to have similar performance between the two studied groups; these items are considered to be functioning appropriately. For items where the statistical test indicates significant differences ( $p < 0.05$ ), the effect size is used to determine the direction and severity of the DIF.

SMD is the Standardized Mean Difference index, and  $SD$  is the total group standard deviation of the item scores (in its original metric). A negative SMD value shows that the question is more difficult for the focal group, whereas a positive value indicates that it is more difficult for the reference group.

DIF analyses were not conducted if the sample size for either the reference group or focal group was less than 100 and the sample size for the two groups combined was less than 400. Items are classified into one of three categories and assigned values of A, B, or C based on these DIF statistics. Category A items contain negligible DIF. Category B items exhibit slight or moderate DIF. Category C items have moderate to large values of DIF. Negative values imply that, conditional on the matching variable, the focal group has a lower mean item score than the

reference group. In contrast, a positive value implies that, conditional on total test score, the reference group has lower mean item score than the focal group. The flagging criteria for multiple-choice items and writing prompts are provided in Tables 22 and 23 respectively.

Table 22. DIF Categories for Multiple-Choice Items

DIF Category	Definition
A (negligible)	Absolute value of the MH D-DIF is not significantly different from zero, or is less than one.
B (slight to moderate)	1. Absolute value of the MH D-DIF is significantly different from zero but not from one, and is at least one; OR 2. Absolute value of the MH D-DIF is significantly different from one, but is less than 1.5. Positive values are classified as “B+” and negative values as “B-”.
C (moderate to large)	Absolute value of the MH D-DIF is significantly different from one, and is at least 1.5. Positive values are classified as “C+” and negative values as “C-.”

Table 23. DIF Categories for Constructed-Response Items

DIF Category	Definition
A (negligible)	Mantel Chi-square $p$ -value $>0.05$ and $ SMD/SD  \leq 0.17$
B (slight to moderate)	Mantel Chi-square $p$ -value $<0.05$ and $0.17 \leq  SMD/SD  \leq 0.25$
C (moderate to large)	Mantel Chi-square $p$ -value $<0.05$ and $ SMD/SD  > 0.25$

DIF statistics are computed for all field test items and reviewed at Data Review as part of the evaluation process for inclusion into the active item pool. Appendix G summarizes the number and percentage of items by DIF category from the 2014 field test items for each grade and content area. The 2014 operational tests are composed of items that were piloted in years prior to 2014, which were reviewed and approved by Content Review, Bias and Fairness Review, and Data Review Committees.

### 3.7.3. Item Response Theory (IRT) Analysis

Rasch IRT was used to scale the PAWS<sup>9</sup>. IRT is widely used because it allows for invariant estimation of item and ability parameters. Regardless of the distribution of the sample, the parameter estimates will be linearly related to the parameters estimated from another sample drawn from the same population apart from random measurement error. IRT allows the comparison of two students’ levels of ability even though they may have taken different sets of

<sup>9</sup> SAWS remains in a classical metric per decision of the WDE and Wyoming TAC.

items. An important characteristic of IRT is its item-level orientation. IRT expresses the probability of a student answering a particular item correctly in terms of the student's ability (i.e., the student's level of achievement) and the item difficulty (*b*-value). The probability of a correct response to an item increases as the student's ability increases. See Chapter 6 for further details on the Rasch model. The results of the Rasch IRT analyses of the field test items can be found in Appendix L.

#### *3.7.4. Data Review Procedures*

Following the spring 2014 PAWS administration the statistics discussed above were computed for each item field tested. These statistics will be compiled into books along with images of the items for use in data review meetings. Each item will appear on one page of the data review book with its statistics on the opposite page. An item with any statistics outside pre-established limits will have an appropriate annotation.

Field test items are evaluated by panels of Wyoming state educators selected by the WDE. Each data review panel consists of 8–12 educators with experience in the target grade and subject. Items field tested during the 2014 administration were reviewed in July 2014 by a panel in Laramie, Wyoming.

In addition to judgments of content relevance, panelists evaluate the technical quality of items, checking each field test item (including those with appropriate statistics) for such flaws as:

1. inappropriate readability level
2. ambiguities in the questions or answer options
3. clueing within the body of the item
4. keyed answers that were partially or wholly incorrect
5. distractors that were partially or wholly correct
6. unclear instructions
7. factual inaccuracy
8. any other concrete and material flaws

All items, statistics, and comments were reviewed by the WDE determining the final disposition of all field test items. Items found by the WDE to be inappropriate for curricular or psychometric reasons were removed from the pool of items eligible for use in future PAWS assessments.

The data review meetings begin with a training session led by an ETS assessment lead and psychometrician. This session covers the statistics that the panelists will be using as they evaluate each item, the meaning of each in the context of evaluating item quality and suitability

for use on future operational exam forms, and the role of the panelists' expertise in the data review process.

Panelists were provided with measures of item difficulty (item mean score) and discrimination (item score-test score correlation). They were also given response or score distributions for all examinees. In addition, for multiple choice items they received distractor discrimination values. This information was presented in tabular format. Items with low or negative discrimination and/or with distractors with positive discriminations were culled out, along with items flagged for possible DIF.

Panelists were instructed that the statistics and notes were supplemental to their experience as Wyoming educators in recommending acceptance or rejection of the items being reviewed. That is, they could indicate possible locations of flaws in the item (for example, a distractor with a positive discrimination could indicate that an item actually has two correct options). However, panelists were asked to use their professional experience in educating and working with Wyoming students when deciding to recommend that an item should be rejected.

Items that appear to be bad based on their statistics may actually address areas about which students had misconceptions or in which they had not received effective or sufficient instruction. Such items could be helpful in highlighting areas where instruction can be improved. Similarly, good items may contain flaws and might need to be rejected. Panelists were asked not to blindly recommend acceptance or rejection based solely on an item's statistics, but rather to carefully consider each item in light of their expertise, using the statistical information to supplement their professional judgment. Only items with concrete and identifiable flaws should be recommended for rejection. Panelists were reminded in particular that items should not be rejected simply because they are deemed to be too hard or too easy, and that items of all difficulty levels are needed to effectively assess the entire range of student abilities within Wyoming.

The results of the Rasch IRT analyses of the field test items can be found in Appendix L, the PAWS and SAWS classical analysis results appear in Appendix J, and PAWS and SAWS DIF in Appendix G. Items accepted at data review from the 2014 administration are eligible for use as operational items beginning with the spring 2015 administration.

### 3.8 Test Form Construction

After each administration, analyses were conducted by the ETS psychometrician to determine the statistical properties of all items that were present on any of the forms (both operational items and field test items). This includes estimation of Rasch difficulty parameters on the current scale for all items. Thus, all items that have been field tested or used operationally were equated to the original scales and have known Rasch difficulty. Therefore, when forms were constructed for the 2014 administration it was possible to create test forms that were targeted to not only meet

content and blueprint specifications, but also to match statistical characteristics of the 2013 base PAWS Science tests, as test characteristic curves (TCCs), information, and standard error curves could be evaluated to help ensure statistical comparability.

### *3.8.1. Construction of the Reading and Mathematics Forms*

ETS utilized proprietary test construction software for the construction of the 2014 forms for the new base scale for reading and mathematics. The ETS psychometrician utilized the test content blueprint and the preliminary statistical targets in a configuration file for each grade and subject test being constructed. The blueprints were new for 2014 and can be found in Appendix A.

In addition, the targets for key balance (for multiple choice items, approximately 25% for each of options A–D) were used.

The assessment development leads assembled a draft form conforming to the blueprint and tentative statistical targets were then reviewed by the psychometrician. The test construction software provided real-time feedback on the psychometric properties of the form, allowing the psychometrician and content staff to immediately see the results of a proposed change in the items on the form.

Assessment development leads focused on the content of the form, including checking that the items conformed to the blueprint, that there was balance across the items and passages (for example, there should be a balance in gender and ethnic representation across items and passages. A Reading test where all passages were about females playing sports would lack balance, as would a Mathematics test where all the items referenced Cartesian graphs), that the items did not provide clues to the correct answers of other items, and other similar content-based issues.

The psychometrician and the assessment development leads checked the conformance of the test to its statistical targets and blueprint, key balance (i.e., that approximately the same number of multiple choice items were keyed to each of the possible answer options [A, B, C, and D]). Other checks were to determine that the same key occurred no more than three times in a row) and that the other statistical properties of the items and forms were within desired limits.

Changes in the composition of the forms (either in the items themselves or the ordering of the items) by either the assessment development leads or psychometrician had to be approved by the opposite party. Once a form had been approved by both the assessment development leads and the psychometrician it was sent to the WDE for their review and approval.

### *3.8.2. Construction of the Science Forms*

ETS utilized proprietary test construction software for the construction of the 2014 forms. The ETS psychometrician utilized the test content blueprint and the statistical targets in a configuration file for each grade and subject test being constructed. The blueprints were

unchanged from 2013 and can be found in Appendix A. The TCCs, information, and standard error curves from the 2013 administration constituted the statistical targets for the 2014 science forms.

Additional check of forms includes the targets for key balance (for multiple choice items, approximately 25% for each of options A–D), proportion of items from the 2013 operational forms (approximately 30% of the test), and proportion of items that had previously been used operationally versus those that had only been field tested (between 40% and 60% of each) were used. Moreover, limits were set on the year an item had been field tested to maximize the use of newer items as much as possible.

The assessment development leads assembled a draft form conforming to the blueprint and statistical targets; then it was reviewed and edited by the psychometrician. The test construction software provided real-time feedback on the psychometric properties of the form, allowing the psychometrician and content staff to immediately see the results of a proposed change in the items on the form. Finally, the software noted the items' positions, used to minimize the difference between the items' position on the 2014 form and its position on the form from its most recent use.

Assessment development leads focused on the content of the form, including checking that the items conformed to the blueprint, that there was balance across the items and passages (for example, there should be a balance in gender and ethnic representation across items and scenarios), that the items did not provide clues to the correct answers of other items, and other similar content-based issues.

The psychometrician and the assessment development leads checked the conformance of the test against its statistical targets and blueprint, key balance (i.e., that approximately the same number of multiple choice items were keyed to each of the possible answer options [A, B, C, and D] and that the same key occurred no more than three times in a row), and that the other statistical properties of the items and forms were within desired limits.

Changes in the composition of the forms (either in the items themselves or the ordering of the items) by either the assessment development leads or psychometrician had to be approved by the opposite party. Once a form had been approved by both the assessment development leads and the psychometrician it was sent to the WDE for their review and approval.

### *3.8.3. Construction of the SAWS Forms*

The test design for the grade 3 spring 2014 SAWS assessment was two 12-point prompts, one Narrative, and the other Informative. Grades 5 and 7 had a single 12-point Narrative Writing prompt and a 4 + 8 point set. ETS content specialists and psychometricians jointly selected prompts according to test build specifications and test blueprints for the 2014 administration. A



number of factors were considered during the test construction process. Prompts were selected to satisfy the test design, meet target test difficulty, and represent an overall test with balanced content. A test development checklist was used to review the initial test assembled during the test build. Test build was an iterative process to balance test content and its statistical properties. The 2014 operational prompt was selected to ensure the writing prompts across administrations have difficulties that are as similar as possible. The selected prompts were provided to the WDE content specialists for approval.

#### *3.8.4. Final Review of Assembled Operational Tests*

Once the forms were assembled to meet test specifications and statistical targets, WDE content specialists reviewed the assembled forms. The criteria for evaluating each group of forms included the following:

- The content of the test forms should reflect the goals and objectives of the Wyoming Content and Performance Standards (curricular validity);
- The content of test forms should reflect the knowledge and skills as taught in Wyoming Schools (instructional validity);
- Items should be clearly and concisely written and the vocabulary appropriate to the target age level (item quality); and
- Content of the test forms should be balanced in relation to ethnicity, gender, socioeconomic status, and geographic district of the state (free from test/item bias).

After any changes from the WDE review had been completed, ETS staff (test development staff members, content specialists and editors) conducted a final review including a content and grammar check. The WDE then completed their final review and provided approval and sign-off for each PAWS operational test form.

## 4. TEST ADMINISTRATION

### 4.1 Test Materials

Test materials were sent to each Wyoming PAWS and SAWS Building Coordinator(s) in shrink-wrapped packages within boxes that included school inventories. All students in grades 3–5 received scorable test booklets. Students in grades 6–8 received answer documents to record responses to questions in the test booklets.

Building Coordinators were responsible for distributing the materials to test administrators. Materials were color-coded by grade. Coordinators applied Pre-ID labels with student identification and demographic information to test books or answer documents. Materials distributed each day were limited to those needed for testing on that particular day. When not in use, materials were locked in secure storage.

### 4.2 Materials Return

Once test administrations were completed, materials were collected and tabulated by Building Coordinators. In addition, the demographic information was hand gridded on the Test and Answer books or answer documents if the student did not have a Pre-ID label. The documents were then packaged together and locked in secure storage until their shipment to ETS. Each box was labeled with a unique tracking number by the shipping carrier.

### 4.3 Directions for Administration

The *PAWS and/or SAWS Directions for Administration* and *PAWS and SAWS Building Coordinator's Manual* provided the guidelines for planning and managing the PAWS and SAWS administration for district and school administrators. The *PAWS and SAWS Directions for Administration* provided specific directions for test administrators, from scheduling and timing for sessions and preparing students to testing students from special populations. Two half-day comprehensive training sessions conducted jointly by the WDE and ETS were held in January 2014 prior to the 2014 testing window. All test administrators around the state were required to view the Test Administrator Training Video before the test window opened. Building principals required test administrators as well as anyone handling test materials to sign off after viewing the training video. These certification documents were retained in the school and were available to the WDE upon request.

The PAWS tests were administered under untimed testing conditions. Grades 3–6 Reading were administered in four untimed sessions. Grades 7 and 8 Reading were administered in three untimed sessions. Grade 3–5 Mathematics was administered in three untimed sessions (this was the only grade which did not have separate calculator and non-calculator sessions). Grades 6–8

of Mathematics were administered in three untimed sessions, one non-calculator and two calculator sessions.

All grades of Science (4 and 8) were administered in two untimed sessions. The expected time for testing was provided by grade and content area in the *PAWS Directions for Administration Manual*, but students could take more time if needed.

The SAWS assessment was administered in three untimed sessions.

#### *4.3.1. Allowed Student Manipulatives*

Calculators were not allowed on the PAWS Mathematics test in grades 3–5. Calculators were permitted for students in grades 6–8 on two sections of the assessment. In addition, a *PAWS 2014 Allowable Resources* document was posted to the WDE webpage to assist test administrators in administering PAWS in a standardized manner.

#### *4.3.2. Test Security*

PAWS and SAWS test security guidelines strictly prohibit the photocopying of all or any part of a test booklet, and require that all violations of the Wyoming Department of Education’s regulations be reported to the WDE immediately. Under the state law, violations are dealt with at the school district level. The reporting of violations to the WDE ensured that test scores could be invalidated if necessary. All test booklets were considered secure materials. The PAWS Building Coordinators were required to document the receipt of secure materials, check the lists of students, and return all test materials to ETS for scoring.

The specific procedures that were to be followed during any test administration and used in the handling of documentation were outlined in the 2014 *PAWS and SAWS Directions for Administration*. Persons designated to administer the PAWS and SAWS tests were expected to:

- Keep all test materials in locked storage.
- Not reproduce any test materials in any manner.
- Not disclose any actual test items to students prior to and after testing.
- Not provide answers to any test items to any students.
- Not change or otherwise alter a student’s answer.
- Follow the suggested time periods as closely as possible in order to maintain uniformity in the test administration. (Note: PAWS is an untimed test.)
- Follow the *Directions for Administration* manual explicitly.
- Follow all Ethics and Security Requirements as outlined in the 2014 *PAWS and SAWS Directions for Administration*. If there is a violation, the students’ materials *will not be scored* and the school will not be able to count the student(s) for participation.

In addition, financial rewards related to test performance were strongly discouraged.

PAWS and SAWS test administrators (teachers) were instructed to immediately report any loss of test materials or other testing irregularities to the school principal or Building Coordinator. The PAWS and SAWS District Coordinator subsequently reported all irregularities to the WDE Standards, Assessment, and Accountability Unit.

#### 4.4 Student Participation

As noted previously, all Wyoming students in grades 3 through 8 were required to participate in the regular PAWS and SAWS tests, the PAWS and SAWS with appropriate accommodations, or the PAWS-ALT (for students with the most significant cognitive disabilities). Federal and state law (i.e., the Individuals with Disabilities Act of 1997 and W.S. 21-9-101 (c)(i)) did not exempt any student from participating in the statewide assessments. Students with disabilities, who were on a 504 Plan, or who were English Language Learners (ELL) were allowed to be provided with standard accommodations during the administration of PAWS consistent with guidance provided by the Wyoming Department of Education. Students with significant cognitive disabilities were required to take the Proficiency Assessments for Wyoming Students-Alternate (PAWS-ALT) as determined by their IEP teams.

All students participated in the state accountability assessment program in one of three ways:

- Participation in PAWS and SAWS regular assessment without accommodation;
- Participation in PAWS and SAWS regular assessment with standard accommodation;
- Participation in PAWS-ALT and SAWS-ALT.

#### 4.5 PAWS and SAWS Standard Accommodations

Accommodations are practices and procedures in the areas of presentation, response, setting, and timing/scheduling that provide equitable access for students during instruction and assessment. Accommodations changed the way a test was administered or the way a student responded to test questions to reduce or eliminate the effects of a student's disability or lack of proficiency in English, but did not reduce learning expectations. Allowable accommodations on PAWS did not change the construct being tested nor did they affect the psychometric characteristics of the assessment.

Standard accommodations were allowed on the PAWS and SAWS for students with disabilities, for students on a 504 Plan, and English Language Learners (ELL). The WDE recognizes that the proper administration of standard accommodations allows these students access to the test, resulting in the students' ability to demonstrate their knowledge and skills consistent with the measured test constructs in each content area. Often the conditions under which the test was

standardized differ from those present when accommodations were used. These differences, in some cases like reading the reading passages, reached a level sufficient to jeopardize the validity of interpretations. However, based on available evidence, the standard accommodations allowed for PAWS and SAWS were considered incidental to the construct intended to be measured by the test (Standards for Educational and Psychological Testing, 1999, p.101) by the WDE. Thus, students using accommodations received scores on PAWS and SAWS that are considered valid and were aggregated with those of other students. WDE and ETS staff paid careful attention to the potential effects of testing conditions on test score interpretations and adhered to the Standards for Educational and Psychological Testing (2014).

The administration of standard accommodations during PAWS and SAWS has potential implications for the validity of resulting scores. Therefore, it was necessary for test administrators and access assistants to be trained annually and to be familiar with updated standard accommodations documents related to the selection, administration, and evaluation of standard accommodations.

In January 2006, the *Wyoming Accommodations Manual for Instruction and Assessment: How to Select, Administer, and Evaluate Use of Accommodations for Instruction and Assessment of Students with Disabilities* was developed by the Wyoming Department of Education in conjunction with the CCSSO State Collaborative on Assessment and Student Standards Assessing Special Education Students (SCASS-ASES). Information in the manual guides the selection, administration, and evaluation of accommodations to ensure that the validity and comparability of resulting scores are preserved. It is available along with other documents related to PAWS and SAWS standard accommodations on the WDE website.

In November 2006, the Standards, Assessment, and Accountability and Special Programs Units provided state-wide training for school district personnel representing every school district in the state on the selection, administration, and evaluation of accommodations to further standardize the use of accommodations in the PAWS administration. Training materials provided by CCSSO/SCASS-ASES were adapted, utilized, and distributed. Training materials were made available on CD and were sent to all districts that were not able to attend the training. Additionally, a presentation was made by the Wyoming Institute for Disabilities (WIND) of the University of Wyoming on assistive technology and augmentative devices. Based on feedback provided during the 2005–2006 administration and the November 2006 training and recommendations made by the Wyoming Technical Advisory Committee, revisions were made and are reflected in the approved list of PAWS Standard Accommodations (see 2014 *PAWS Directions for Administration*) to improve clarity and ensure the standard use of accommodations.

Additionally each year, a required Standard Accommodations Online Training is provided and notice of this training is provided through a Superintendent's Memo. The purpose of the Standard Accommodations Online Training is to ensure that test administrators and access assistants are trained on the guidelines and requirements to select, administer, and evaluate standard accommodations for the current administration to all three eligible student groups.

This required training provides information regarding the following topics: students eligible to receive standard accommodations, persons eligible to administer standard accommodations, standard and nonstandard accommodations, 2014 PAWS and SAWS standard accommodations, English Language Learners (ELL) standard accommodations, the selection, administration, and evaluation of accommodations, special test forms, documentation of accommodations, and participation exemption from state assessment. Verification of completion of this training by Test Administrators and Access Assistants must be provided to the building principal or the District or Building PAWS Coordinator using the 2014 Proficiency Assessments for Wyoming Students Test Administrators Verification Form found at the WDE website.

Two addendums related to the administration of standard accommodations were distributed through Superintendent's Memo and postings on the WDE website including the *Wyoming Statewide Assessment System 2014 PAWS and SAWS Standard Accommodations* and the *2014 PAWS and SAWS Standard Accommodations Frequently Asked Questions (FAQ)*. The *Wyoming Statewide Assessment System 2014 PAWS and SAWS Standard Accommodations* document provides information about the administration of standard accommodations and also identifies the allowable standard accommodations, divided into four categories (presentation, response, setting, and timing and scheduling). The FAQ document provides information about the administration and documentation of standard accommodations as well as detailed information regarding specific accommodations including the administration of standard accommodations for ELL students, best practices associated with the selection and administration of accommodations, and a specific list of standard accommodations for ELL students.

#### *4.5.1. Students Eligible for Test Accommodations*

The right to receive accommodations on state assessment is guaranteed by law to a student with a disability. The process of making decisions about accommodations is one in which members of the IEP team facilitate the participation of students with disabilities in general state assessments. Students eligible for accommodations also include those students with a 504 Plan and English Language Learners (ELL).

#### *4.5.2. Requirements for Use of Test Accommodations*

For students with disabilities, the selection of accommodations for the general assessment was the responsibility of a student's IEP team or 504 Plan committee. Guidance was provided in the *Wyoming Accommodations Manual for Instruction and Assessment: How to Select, Administer,*

*and Evaluate Use of Accommodations for Instruction and Assessment of Students with Disabilities* (January 2006). Currently permitted standard accommodations for students with disabilities, 504 Plans, or who were ELL were listed in the 2014 *Directions for Administration* (DFA). Accommodations were matched to an individual student's needs and were only provided when all of the following conditions were met:

1. The accommodations were documented on the student's IEP or 504 Plan.
2. The accommodations for ELL were determined at the local level.
3. The selection and administration of accommodations were consistent with the 2014 PAWS standard accommodations.
4. Standard accommodations were administered as described in the *Wyoming Statewide Assessment System 2014 PAWS Standard Accommodations* and the *Wyoming Accommodations Manual for Instruction and Assessment*.
5. The accommodations provided were effective in providing access to the test and had been regularly used by the student during instruction and classroom assessment.
6. The accommodations were administered by a trained Test Administrator or access assistant who was familiar to the student.

Accommodations could not:

1. Result in adverse consequences;
2. Alter the construct being tested; or
3. Provide additional information, prompting, or clueing to suggest or support the selection of correct answers.

Standard accommodations must have been used consistently for instruction and assessment prior to the test administration. Accommodations were not allowed for non-ELL students or for any students without an IEP or 504 Plan. Accommodations were administered by a trained certified teacher, certified staff member, or access assistant. A certified teacher, certified staff member, or access assistant was qualified to administer accommodations if that teacher:

1. Understood the procedures for administering standard accommodations; and
2. Has effectively administered the accommodation(s) to the student during instruction and/or assessment; and
3. Has attended a 2014 PAWS Training or has viewed the 2014 PAWS Training online and submitted record of the training to the building principal; and

4. Has completed the 2014 PAWS Accommodations Training online and a submitted record of the training to the building principal.

PAWS administrations were untimed for all students. Large print, audio, and Braille versions of PAWS were available for all grade levels and content areas.

#### *4.5.3. Description of Standard Accommodations for Students with Disabilities*

As mentioned above, the types of standard and allowable accommodations used with PAWS and SAWS were grouped into four categories:

- Presentation (visual, tactile, auditory, and multisensory),
- Response,
- Setting, and
- Timing/scheduling.

Appropriate documentation and monitoring of the standardized use of accommodations was required of test administrators, test coordinators, and/or principals. Monitoring of the selection, administration, and evaluation of accommodations by school personnel was provided by the Wyoming Department of Education and occurred during the administration of the tests as well as following the administration of the PAWS and SAWS. Additionally, the Special Programs Unit reviewed documentation of accommodations during on-site monitoring visits. The following assessment accommodations were allowable for students with an IEP or 504 Plan.

##### *4.5.3.1. Presentation Accommodations*

1. Student uses a Braille Special Test Form.
2. Student uses a Large Print Special Test Form.
3. Student uses an Audio Special Test Form.
4. Student uses magnification devices.
5. Student uses color overlays to reduce glare or enhance text.
6. Student uses templates to reduce the amount of visible print.
7. Student uses tactile graphics.
8. Sign language interpreter signs directions in all content areas and/or signs test questions as written in all content areas EXCEPT Reading. The interpreter may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Interpreters need to be familiar with the terminology and symbols specific to the content. It is recommended that one interpreter be provided for each individual student.



9. A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.
10. A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.
11. Student asks for clarification of directions (not test questions or answer choices).
12. Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.
13. Student uses text-to-speech software in all content areas EXCEPT Reading.

#### *4.5.3.2. Response Accommodations*

14. A certified staff member or access assistant scribes what a student dictates through alternate augmentative communications (AAC), pointing, sign language, or speech. The scribe may not edit or alter the student's work in any way and must record, word for word, exactly what the student has dictated. A scribe must allow the student to review and edit what that student has written. The student's final response must be transcribed by a certified staff member or access assistant into the Student Test and Answer Book on the pages in which the student's response is to be written.
15. A student types responses using a word processor. Dictionary and synonym/thesaurus devices MUST be disabled. The margins for word-processed documents should match the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages in which the student's response is to be written.
16. Student uses speech-to-text conversion or voice recognition in all content areas. The margins for this document should match as closely as possible the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages in which the student's response is to be written.
17. Student uses a Braille. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book or answer document on the pages in which the student's response is to be written.
18. Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the

Student Test and Answer Book or answer document on the pages in which the student's response is to be written.

19. A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book or answer document.
20. Student uses visual organizers including graph paper, place markers, and templates. Student uses a pencil to underline text. Highlighters CANNOT be used in the Student Test and Answer Book or answer document.

#### *4.5.3.3. Setting Accommodations*

21. Student takes the test in a different building location, in a small group, or individually. Changes can also be made to a student's location within a room to reduce distractions to the student or to other students, to increase physical access, or enable the use of special equipment. Students must be monitored by a certified staff member.

#### *4.5.3.4. Timing and Scheduling Accommodations*

22. Student is provided with extended time to complete the assessment.
23. Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.
24. Student takes the tests at the time of day when that student is most likely to demonstrate peak performance.

#### *4.5.4. Description of Standard Accommodations for English Language Learners (ELL)*

Schools could not exempt ELL students from the PAWS and SAWS content assessments. The only exception to this policy was that students who were enrolled in U.S. schools for less than one year as of March 31, 2014, could be waived from taking the Reading PAWS content assessments with an exemption approved by the Wyoming Department of Education. Students who received this exemption took the Wyoming ELL assessment instead of the Reading portion of PAWS, but were not exempted from the mathematics and science portions of PAWS or SAWS.

ELL students could be provided with accommodations during PAWS and SAWS as long as they met eligibility criteria. In addition, students who no longer meet the eligibility criteria as ELL and were identified as proficient or transitional could also receive standard accommodations for a period of up to two academic years when appropriate. These accommodations have been demonstrated to be effective in providing access to the test and should have been used regularly by the student during instruction and assessment prior to the 2014 administration.

#### *4.5.4.1. Presentation Accommodations*

25. A certified staff member or access assistant translates written directions to the student.
26. A certified staff member or access assistant re-reads, simplifies, or clarifies directions in English or in the student's primary language (NOT test questions or answer choices) without clueing correct responses.
27. A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one reader be provided for each individual student.
28. Student uses a bilingual dictionary provided by the school.

#### *4.5.4.2. Setting Accommodations*

29. Student takes the test in a different building location, in a small group, or individually. Changes can also be made to a student's location within a room to reduce distractions to the student or to other students, to increase physical access, or enable the use of special equipment. Students must be monitored by a certified staff member.

#### *4.5.4.3. Timing and Scheduling Accommodations*

30. Student is provided with multiple, individual breaks as needed.
31. Student is allowed to complete the test over multiple days.

#### *4.5.5. PAWS 2014 Monitoring of Appropriate Accommodations*

Through its Continuous Improvement Focused Monitoring process, the WDE Special Programs Division monitors the appropriate selection and use of accommodations for both instruction and assessment. Each school year, Special Programs staff members visit at least 16% of Wyoming districts to investigate potential noncompliance within the priority areas of Free and Appropriate Public Education in the Least Restrictive Environment (FAPE in the LRE), Postsecondary Transition, Child Find, Disproportionality, and other procedural areas.

While on-site in school districts, WDE staff members review Individual Education Program (IEP) files looking for evidence that IEP teams have made sound accommodations decisions to enable students with disabilities to gain access to instructional content and assessment measures. In addition, general and special education teachers, administrators, and service providers are interviewed to provide further information about school and district practices regarding accommodations. Failure to provide accommodations listed in a student's IEP or failure to thoughtfully consider accommodations for a student or students may contribute to a finding of noncompliance, thus requiring the district to address the issue through the creation and

implementations of a Corrective Action Plan (CAP). Monitoring of standard accommodations for ELL's was provided by the Local Education Agency.

#### *4.5.5.1. Empirical Analysis of Accommodations*

IEP and 504 Plan students comprised approximately 11%–14% of students at each grade level, with between 60%–80% of those IEP and 504 students receiving testing accommodations (depending on grade and subject). While Wyoming allows 31 specific accommodations on PAWS as described herein, the overwhelming majority across all content areas were provided as auditory presentations (e.g., Reading directions, Reading questions, clarifying directions, or the audio form), setting accommodations (i.e., testing in a separate location), or an accommodation in timing/scheduling (e.g., extended time, multiple breaks, test over multiple days). This breakdown by specific accommodation also provides a baseline for monitoring accommodations in future years. Frequency tables for accommodations provided during the 2014 PAWS for Mathematics, Reading, and Science for all grades are presented in Appendix N. In general, IEP students who did not receive accommodations had higher mean scale scores. Mean scale scores for IEP and 504 Plan students broken down by accommodation status are presented in Appendix O.

#### *4.5.6. Selection and Administration of Accommodations*

An important question regarding the use of accommodations in large-scale assessment is whether the resultant student scores mean the same thing as scores resulting from non-accommodated assessment (Kim, Wang, Zhao, and Li, 2006). In other words, do the accommodations yield meaningful, valid scores of the level of a student's subject mastery? It is also imperative to know the effect of including scores of accommodated students in test calibration<sup>10</sup>, specifically in terms of item parameters and resulting test scores (Karkee, Lewis, and Barton, 2005). Wyoming recognizes the need to examine the data associated with the administration of standard accommodations for students with disabilities, students with 504 Plans, and English Language Learners, and for the continued evaluation of the standard accommodations with regard to current research.

Standard accommodations were implemented for students with disabilities, students with 504 Plans, and the English Language Learners (ELLs) participating in the PAWS testing. In providing for the use of accommodations, the state recognized that it is important to ensure that

---

<sup>10</sup> Note that responses to Braille, audio, and large print forms were excluded from calibration, scaling, and equating analyses, but are included in all descriptive statistics reported in this technical report except those that come directly from the calibration, scaling, and equating analyses (such as Rasch item difficulties). Responses to the regular forms from students who received accommodations were included in the calibration, scaling, and equating analyses.

accommodated testing conditions did not change the construct being tested nor affect the psychometric characteristics of the assessments. ETS and WDE will continue to monitor the appropriate use of accommodations for students that require them. Special attention will be given to ensure that the use of accommodations does not negatively affect the validity of the test results for such students or for students who did not require accommodations.

To ensure the appropriate selection and administration of standard accommodations for the 2014 PAWS administration, the Standards and Assessment Division provided training required of all test administrators and access assistants responsible for administering accommodations. Additionally, updated guidance on the 2014 PAWS and SAWS Standard Accommodations and 2014 PAWS and SAWS Standard Accommodations FAQ were distributed via a Superintendent's Memo.

The training provided critical information regarding students eligible to receive standard accommodations, persons eligible to be administered standard accommodations, standard and nonstandard accommodations, 2014 PAWS and SAWS standard accommodations, ELL standard accommodations, the selection, administration, and evaluation of accommodations, special test forms (Braille, Large Print, Audio), documentation of accommodations, and participation exemption from state assessment. Verification of completion of this training was required by Test Administrators and Access Assistants and was provided to the building principal or the District or Building PAWS Coordinator using the 2014 PAWS Test Administrator Training Verification Form. All training materials and documents were available on the WDE website.

## 5. PROCESSING AND SCORING OF PAWS AND SAWS ITEMS

### 5.1 Overview

This chapter describes the receipt control, scanning, and scoring procedures used at ETS for the 2014 PAWS including details of the hand-scoring of the SAWS prompts.

At the close of testing, the PAWS and SAWS Student Test and Answer Books and answer documents (PAWS grades 6–8) were returned to ETS. Upon receipt, they were scanned into ETS’s electronic imaging system. Subsequent processing of student responses necessary to score those responses and to produce reports used these images rather than the paper documents. After scanning, the physical documents were put into archival storage.

Student responses to PAWS multiple-choice test items were machine-scored. Student responses to SAWS prompts were individually read and evaluated by raters employed by ETS. The WDE had upfront oversight and control of training materials and audited scorer trainings at their discretion. 2014 operational SAWS hand-scored items were scored in Concord, CA.

#### *5.1.1 Multiple Choice Items*

Multiple-choice items were used in all tests. Correct answers were assigned a score of one point and incorrect answers were assigned a score of zero points.

#### *5.1.2 SAWS Prompts*

The scanning of student test and answer books into the electronic imaging system allowed student responses to constructed-response items to be scored online at all scoring sites while maintaining the original documents at a central facility. The imaging system randomly distributed responses, ensuring that no one reader scored a disproportionate number of responses from any one school. The online scoring system maintained a database of actual student responses and the scores associated with those responses. The system also provided continuous up-to-date monitoring of all scoring activities.

### 5.2 Receipt Control, Processing, Scanning, Editing

ETS’s Operations Center was responsible for the processing of documents received from Wyoming for each individual student’s work. The team consisted of software and process engineers, management professionals, systems and requirements analysts, and customer service specialists. The receiving staff accepted and counted PAWS and SAWS cartons that were returned to ETS, confirming shipments from schools. The editing staff captured and verified customer information via the Header Sheet to compare number of documents scanned to number indicated as being returned on the Header Sheet. The Document Staging department ensured that box contents matched the information provided on the Header Sheet. This step linked every document to the proper scannable scoring order number (batch number) that was utilized

throughout the remaining steps of the scanning and scoring process. The scanning process captured data from student test and answer books, answer documents, and school headers.

Within each functional area, specific tasks were accomplished and quality checks were performed both within and across functional areas. The quality checks performed were documented in the custom program specifications.

#### *5.2.1 Receipt Control*

Receipt control began when the receiving staff accepted and counted cartons as they were delivered, sorting them by district into scorable and non-scorable queues. The first quality checkpoint was a comparison of what was received against what was expected to be received. This check was performed utilizing the tracking system to flag any anomalies in the shipment and to begin immediate investigation of any such. The process was utilized to produce a daily report listing districts for which materials had not arrived.

ETS and WDE have established rules for handling issues encountered while processing the answer documents. These are located in the program specifications.

#### *5.2.2 Processing*

ETS used Header Sheets to capture and verify customer information to ensure that complete results were delivered to the proper location. The information that was verified included the returned scorable document *N*-count, grade, and subject for each returned scorable document, building name and number, and district name.

To minimize or eliminate student coding errors on the student demographic page, ETS provided a pre-identification service to the WDE. This service was utilized to provide student demographic data that were printed on pre-ID labels that were scanned during processing.

During the staging process, ETS staff removed the documents from the boxes and arranged them on carts. A preprinted scannable scoring order number (batch number) was matched to each cart. Each Header Sheet was matched to a specific batch number that was placed with the documents so that when it was scanned the batch number was associated with those documents. This step is important because it linked every individual document to the proper order number throughout the remaining steps in the scoring and reporting process.

#### *5.2.3 Scanning*

In the scanning stage, ETS captured all the data from the student response forms and school headers created during the staging process. All scannable documents were processed in a temperature-controlled environment. This allowed the paper to normalize and eliminated paper distortion caused by the environment. Properly stabilized paper improved scan reliability and quality. Prior to scanning, the spines of multiple-page documents were cut to create single sheets that were then scanned.

ETS utilized image-scanning technology to capture information from all scannable documents. A scanner diagnostic test was executed prior to scanning the documents on each cart, and a calibration check was performed to validate that the scanner was imaging properly. The calibration check ensured that the scanner was accurately capturing the range of darkness of the written and gridded responses. This was critical to the post-processing that occurred in editing and scoring.

The images produced by the scanner included document identification and all information gridded by the test-taker and were stored as 8-bit (256 level) grayscale images. The scanning program checked the validity of the document identification using optical mark recognition (OMR), skunk codes, and optical character recognition (OCR) module codes to ensure that the booklet that was being scanned was the correct booklet. The scanning program also compared the actual number of pages scanned to the number of pages expected for the document according to its identification. These two checks ensured that the correct document was being imaged and that the entire document was imaged. Finally, the skunk and module codes acted as reference points indicating the orientation of the document as it moved through the scanner.

Scanned documents were sent to databases where images were distributed to editors and/or raters based upon rules established for the program. The data collected from the image scanners were stored in a scan file, which was used to generate an edit report. When this was completed, the cart containing the scanned documents was logged out of the scanning workstation.

Constructed-response image files were distributed to ETS's Performance Assessment and Scoring Service (PASS) for human scoring, while images of selected responses and demographic data were made available to scoring editing for human review. PSC was responsible for all activities related to the scoring of constructed-response assessments. The PSC maintains a large pool of qualified, trained, professional raters who are experienced in scoring a wide range of open-ended assessments in writing. Raters for the SAWS were drawn from this pool and received additional SAWS-specific training prior to their scoring the assessment.

#### *5.2.4 Editing*

The first step in the editing process was to electronically compare each student's scanned data to the business rules established by WDE for processing the student's information. The results of this comparison were used to generate an edit report listing documents requiring correction or validation. This report included all documents with a data field that did not match program specifications. A scoring editor reviewed every flag by referencing the source document and validating or correcting the field. Data items edited included the student id, name, and date of birth. The edits that were applied to the student's scanned data were also applied when registering the student online. In the online system, edits were applied immediately and data were not accepted into the system if invalid.



Another step in the paper editing process is *N*-count verification. The number of documents scanned was compared to the number of documents recorded on the Header Sheet and collected in the structure definition. When the *N*-counts did not match, the paper documents for that batch were manually counted, and based on the business rule variance, an alert was issued for document *N*-count discrepancies.

When all edits were resolved any corrections were incorporated into the file containing student records. Once all corrections were made, the edit routine was rerun to ensure data validity. When no fields were flagged as suspect, all the records for that order were considered clean and the tracking system moved the order to job submission. The physical documents were no longer needed in the scoring process and were moved to the archiving workstation.

### 5.3 Scorer Recruitment and Qualifications

#### *5.3.1 Eligibility*

Applicants were required to possess a Bachelor's degree from an accredited institution and have computers meeting the minimum Hardware and Software Requirements. Current or prior teaching experience in Wyoming was highly desirable. The applicant must have resided in the United States, and be a U.S. citizen, a resident alien, or authorized to work for remuneration in the United States.

Prospective raters were self trained to apply scoring criteria for the program and were then invited to take a certification test. Upon passing certification, prospective raters were placed into the rater pool. Invitations to score were based on volume and program needs. Not all successful applicants were certified; not all certified raters were invited to score at each administration. If raters were invited to score, all raters were to score online and received an hourly honorarium.

#### *5.3.2 Rater Qualifications*

##### *5.3.2.1. Training*

ETS raters were an integral part of the SAWS program, as they were responsible for evaluating the student written responses. By completion of the interactive tutorial raters had to:

- Learn about the Writing prompt of the SAWS.
- Understand the principles by which this assessment is scored.
- Practice scoring papers accurately.
- Be introduced to Online Network for Evaluation (ONE), the software used for online scoring.
- Familiarize themselves with the logistics and policies by which online scoring is conducted.
- Take a practice certification test in preparation for the rater qualification test.

#### *5.3.2.2 Certification Procedures*

After completing ETS's Online Scoring Network (OSN)<sup>TM</sup> general training and SAWS specific training provided in the "LearnOSN" self tutorial website, prospective raters made arrangements with the Performance Assessment and Scoring Services (PASS) to schedule a certification test. Prospective raters did not have any content-related assistance while taking the certification tests and did not discuss them with anyone.

Prospective raters had to pass a certification test in order to become certified SAWS raters. The scoring center coordinator notified the prospective rater regarding whether or not that rater had successfully certified. No more than two attempts were permitted.

### *5.4 Topic Notes and Annotations Meeting and Operational Rangefinding Meeting*

In preparation for scoring field-test responses for the 2014 new-item type (Grade 3, 8-point opinion), ETS staff selected approximately 60 responses/samples for each of three field-test prompts. These responses had been "prescored" internally and were intended to represent the range of responses at each score point. These, in addition to draft Scoring Notes, were reproduced in binders and forwarded to the WDE. ETS Test Developer/Content Specialist and WDE representatives then met via conference call (June 12, 2014) and discussed these responses as a "mini-rangefinding" session to ensure ETS staff had clear guidance and consensus from the WDE to move forward with creating benchmark/anchor paper sets, seeded feedback responses, and other training materials for scoring the field-test responses (immediately following the scoring of all SAWS operational responses).

#### *5.4.1. Operational Rangefinding Meeting*

Rangefinding for operational responses was divided into two parts: pre-rangefinding and rangefinding. Rangefinding facilitators, representatives of the WDE, Chief Scoring Leaders (CSLs), and ETS Test Developer/Content Specialist participated in both steps of the process.

The object of the pre-rangefinding and rangefinding sessions was to identify anchor papers, produce annotations for each of these papers, and supply topic notes/scoring summaries for each prompt. Anchor papers were typical examples of each skill for each score point. These identified papers were selected to teach a lesson about scoring a particular topic or to demonstrate the range of types of papers possibly found at a given score point.

#### *5.4.2. Pre-Rangefinding Procedures*

Facilitators read a large number of responses, looking for responses with a teaching point, or those making a particularly good anchor paper. After facilitators finished reading a number of responses, they looked for several responses at each score point to give a range for discussion during the rangefinding session. These resulting responses were then assembled into sets for rangefinding team discussion.

About fifty responses were needed for the first two days of operational scoring as a pool for calibration sets and monitors. Therefore, it was necessary to discuss all responses selected for rangefinding. Any responses not used as anchor responses were used for calibration and monitor (validity) responses.

#### *5.4.3. Rangefinding Procedures*

The PASS department prepared copies of all five sets for each participant. Facilitators then explained the purpose of and procedure for the rangefinding process. Raters (representatives from WDE, AD, and CSL (Chief Scoring Leaders)) were then directed to read and score the first set responses. They were provided approximately twenty-five minutes per set to mimic the minimum reading rate required by raters. Each reader was provided a score sheet to record rater's scores with room on the score sheet to record additional scores. After scoring the first set, facilitators collected and recorded each reader's verdict on each paper. The order in which scores were solicited varied.

Their object was twofold: they facilitated discussion about some of the responses and encouraged the group to reach a consensus about the utility of responses for use as anchors or rangefinders at the various score points. One facilitator led the discussion as a moderator, leading everyone to consensus, and not forcing the discussion toward a specific score.

This discussion also served as material for the annotations facilitators subsequently wrote for each of the chosen responses. Because of this, the facilitators made careful notes of the discussion on each paper. Facilitators began with any responses where there was consensus, or a near consensus. If most of the table agreed a given paper was a 2 or 2+, the paper was selected as the "solid" or "high" anchor 2 point paper. Once an anchor was chosen for a specific score, the specific score point was not revisited, even if another paper met the criteria for that score point.

If one or two raters disagreed with the score, in order to fill as many slots as possible through the first set, the facilitators initiated a discussion between a reader or raters in the majority and one of those in the minority, reading the responses aloud. After discussion, raters were permitted to modify their scores, thus producing a consensus. On the other hand, if a reader or raters had sound and serious objections to the use of a given paper as an anchor or rangefinder, the paper was thrown out. This process was followed until all prompts in the sets were reviewed.

The same procedure was followed for each subsequent set until all anchor responses were selected. Once all anchor sets were selected for each topic by score and trait, the sample sets were provided to the WDE for approval. Any changes were communicated and fixed.

### 5.5: Hand-Scoring Process

All student responses to the writing prompts were scored in the ONE system, a distributed, Web-based scoring system that enables a large number of raters to view and score assigned responses from remote locations. All identifying information from the responses sent to raters was removed so that neither the identity of the student nor the student's school was revealed to the rater; the rater saw only the student response.

#### *5.5.1. Scoring Responses to SAWS Writing Prompts*

Raters scored writing prompts online after they were trained in SAWS scoring and certified. The system they used was ONE. As they scored the responses, raters referred to the appropriate SAWS Scoring Guide and sample responses ("benchmarks" and "rangefinders"). A scoring leader guided and monitored the process to further ensure accuracy in scoring.

#### *5.5.2. About SAWS Scoring*

The SAWS writing prompt responses were scored in relation to the SAWS Scoring Guides, which were based on the Wyoming Content and Performance Standards. The rater used the Scoring Guide to evaluate each trait of the response separately: Idea Development, Voice, Organization, and Conventions. The rater did not respond to the overall quality of the response. Quality was defined for each of the traits described by the scoring guides for each score point, and was illustrated by sample responses exemplifying each trait and each score point. In determining a score for a trait, the rater made an assessment of how well the paper reflected the characteristics of the score point.

#### *5.5.3. Reference Materials for Scoring*

The following reference materials were easily accessible in ONE:

- Each grade had its own **Scoring Guide**, which explained the criteria for each score point.
- Each prompt had **Benchmarks**, which were responses intended to provide a solid example illustrating each score point for each of the four skills.
- Each Benchmark had an **Annotation** explaining how the paper fits the Scoring Guide criteria.
- All prompt support materials, other than the SAWS Scoring Guides, were confidential documents not to be shared by a rater with anyone else.

#### *5.5.4. Guidance to Raters: Points to Remember about SAWS Scoring*

The following information was provided:

- The SAWS Scoring Guide is based on the Wyoming Content and Performance Standards, so it is important to score each response according to the discrete skills described in the Scoring Guide.
- Raters must recognize that responses must meet a standard in writing conventions at each score point. The conventions include grammar, usage, punctuation, capitalization, and spelling.

- Read the entire response; the writing sometimes changes dramatically after the beginning of the response.
- Do not take notes on the response.
- Do not penalize an unfinished but developed response for lacking a conclusion.
- Do not judge a response by its length; some short responses are very good and some long ones deserve low scores.
- Remember that some responses of slightly different quality earn the same score. Each score point represents a range (a high 3, a middle 3, and a low 3, for example).
- Use the full scoring scale; match the quality of each response to the standards described in the Scoring Guides and illustrated in the Benchmarks.
- Remember that the SAWS standards must determine your scoring decisions. In fairness to the students, you must accept and apply the SAWS Scoring Guide.

#### 5.6: Procedures for Maintaining and Retrieving Individual Scores

All Wyoming SAWS student responses are contractually stored for the “lifetime.” All scanned images of responses will be stored through Information Management System (IMS) for six years. To retrieve a student’s responses, a student’s scanned document from IMS can be pulled with their student ID number. (ONE Constructed Response ID number) and have a PDF image within the day it was requested. Requests can come only from WDE.

#### 5.7 Interrater Reliability

ETS’s online scoring system generated many different kinds of internal monitoring reports that enabled ETS and WDE Content Specialists, Scoring Directors, and Scoring Supervisors to monitor the accuracy of scoring. These reports listed all of a team’s raters and provided the results of their scoring on an ongoing basis. Information on these reports included the number of responses read by the raters, the number and percent of invalid (blank, foreign language, etc.) responses scored, and the number of responses that received second scores.

The second scores provided data on the percent of perfect agreement between first and second raters, percent of responses on which the first scorer was a point higher or lower than the second scorer, and the number and percent of responses differing by more than one point (non-adjacent scores).

All SAWS operational writing prompts received a single reading used for reporting. 25% of the responses were randomly routed by ETS’s on-line system for a reading by a second scorer to monitor interrater reliability. Non-adjacent scores received a third score or resolution score performed by a Supervisor, Scoring Director, or Content Specialist that was used as the operational score for the student’s response (i.e., the resolution score overrode both the initial and read behind scores).

Using the results of online monitoring, including the 25% second reads, qualified raters were expected to maintain a minimum cumulative perfect agreement rate of 70% agreement for extended-response items. Raters who fell below this standard were targeted for additional training and check sets.

Section 8.3 presents the overall interrater reliability information for the 2014 SAWS prompts. These are presented in terms of the percentage of responses scored that were exact matches, the percentage that were adjacent (+/- one score point), and the percentage of responses that received non-adjacent scores for the prompt total and by trait.

PAWS field test items received a single score with 25% of the daily scoring output randomly routed by ETS's on-line system for a second score. Non-adjacent scores received a third score or resolution score performed by a Supervisor, Scoring Director, or Content Specialist that was used as the operational score for that item. The second scoring was used for interrater monitoring purposes only.

### 5.8 Accuracy Monitoring

The monitoring functions of the ONE provided a useful method for overseeing the accuracy of scoring and the performance of individual topics. The ONE produced a variety of reports with extensive data on both raters and topics, as well as an overview of the progress and accuracy of the overall scoring process. Most reader performance data were available immediately. A content specialist or a scoring leader was able to view statistical tabulations of reader performance within any given time period. Scoring leaders had the capability of monitoring raters while they were actively scoring a group of responses.

The ONE produced reports showing the degree to which raters were consistent in scores they assign. In addition, the overall mean and the percentage of scores awarded at each score point revealed whether the reader fulfilled the performance standard of using the full range, or whether the reader was scoring consistently low, consistently high, or too exclusively in the middle. If a reader's rate of agreement began to decline, the reader was retrained by a scoring leader and closely monitored thereafter. If the reader's performance did not improve, the reader was released.

In addition to a statistical depiction of reader performance, the ONE monitoring function also provided a statistical portrait of topic performance. Test development staff were able to see over time whether a given SAWS trait was performing well by considering:

- The average rate at which responses are read
- The mean score overall
- The percentage of scores awarded at each point

Analytic evaluation was a procedure for scoring student work samples in which the evaluator made a single judgment of the response, awarding points separately for each trait. Trained evaluators used a scoring guide to describe a typical response at each score level, along with exemplar responses to serve as illustrations of each score level. This was calibrated with continual monitoring of scoring and interrater reliability calculations.

### 5.9 Blanks and Invalid Responses for SAWS

The WDE and ETS developed rules concerning writing prompts that should be scored as blank or invalid. For purposes of scoring and item and test statistics, blank and invalid responses were treated as zeroes.

Available condition codes for blank and invalid responses included Blank (BB), Copy of Prompt (CP), Foreign Language (NL), Illegible (IL), Incomprehensible (IN), Insufficient to Score (IS), and Refusal (XX).

Condition codes could only be assigned by a Scoring Supervisor or Director (with the exception that a condition code of Blank could only be assigned by a Scoring Director and required a second reading to confirm it as such). Raters forwarded responses that they identified as blank or invalid to the review queue for review by a Scoring Supervisor or Director. If the Supervisor or Director determined that a condition code was appropriate then that supervisor or director scored it as such. If the Supervisor or Director determined that a condition code was not appropriate, the paper was returned to the scoring queue.

#### *5.9.1. Blank (BB)*

A response that:

- has no writing or marks of any sort.
- is completely erased (erasure marks are apparent, but no words are discernible).
- is incompletely erased (some words may be readable, but the student clearly intended to erase the response).
- has been crossed-out, even if parts or all of the response is readable.

#### *5.9.2. Copy of Prompt (CP)*

A response that:

- consists solely of a copy of most or all of the prompt.

#### *5.9.3. Foreign Language (NL)*

A response that:

- is *entirely* in a language other than English. (If a portion of the response is in English, Raters should score that portion according to the Scoring Guide.)

#### *5.9.4. Illegible (IL)*

A response that:

- is illegible (i.e., all or a substantial portion of the response is illegible to the point that a score cannot be applied). NOTE: any responses may appear illegible at first, but experienced scoring staff can often read these responses.

#### *5.9.5. Incomprehensible (IN)*

A response that:

- has decipherable words or letters, but no sense/meaning can be determined.

#### *5.9.6. Insufficient to Score (IS)*

A response that:

- has legible writing, is in English, and is not a refusal. (e.g., the prompt asks for a student to write a letter to the principal about needing new lockers and the response is: “Dear Principal,” or “New lockers? Yes”).

#### *5.9.7. Refusal (XX)*

A response that:

- indicates a refusal to respond in writing (e.g., I don’t know/care; I don’t understand; I hate this)
- indicates a refusal to respond not in writing (e.g., an X across the page or a question mark).
- includes *only* drawings or doodles (i.e., there is no accompanying scorable writing).

### *5.10 Reporting of PSC Alerts*

Students’ responses occasionally contained what is termed a PSC Alert, that is, some responses stated or implied threats of violence to self or others or possible cases of abuse or neglect.

Copies of responses demonstrating potential irregularities (i.e., writings on suicide, abuse, neglect, or possibly indicating teacher interference) were provided to the WDE by ETS. PSC staff forwarded copies of responses to the Program Manager who forwarded the copies to the WDE.

#### *5.10.1. Policy on the Reporting of Alerts*

ETS’s raters were instructed to forward student responses that contain one or more of the following elements to a Review queue.

- 1. Statement of intent to inflict serious and imminent physical harm to self.
- 2. Statement of intent to inflict serious and imminent physical harm to others.



- 3. Statement reporting past or current child abuse or neglect.

The raters were not instructed to flag and report any statements beyond the above three categories. The raters were instructed, however, that they could at their discretion flag and report any other material that they believed may reflect a serious situation requiring action.

#### *5.10.2. Reporting Procedure*

When a scorer identified a response containing a PSC Alert in one or more of the categories listed above, the following procedure was followed:

The scorer forwarded the response to Review. The Scoring Director reviewed the response to determine whether it fits the criteria of an alert. The WY PSC Content Specialist was consulted if needed. If the determination was that the response did not contain alert content, no report was made. If the response contained content of a possible alert, a copy of the student's response with a completed project alert form was posted to the ETS State Services Program Team who contacted the WDE.

If ETS referred a student's test to WDE, it did so without making any assessment or recommendation other than to make note of the PSC Alert. Due to the nature of the material and lack of appropriate context, ETS was not in a position to determine whether threats or other statements contained in test responses were serious or joking, real or imaginary.

## 6. EQUATING AND SCALING PROCEDURES

### 6.1 Overview

This chapter covers:

- Creation of vertical scale for PAWS Reading and Mathematics,
- The equating of the 2014 PAWS Science assessments, and
- Translation of raw scores to scale scores along with descriptive statistics for all of the 2014 PAWS scales.

ETS Statistical Analysis team for Wyoming program conducted and quality checked all analyses for the WY PAWS and SAWS assessment, and documented the primary analyses results in this chapter. After all analyses were concluded and documented, preliminary results were calculated over the student data sample used for equating (all grades and subjects had responses from > 95% of the population of Wyoming students in the sample). Documentation of the analyses and the preliminary statewide results were presented to the WDE assessment leadership team for their review and approval by ETS's lead psychometrician via conference call and WebEx prior to the release of the scoring tables for production of reports. Scoring tables were released and production activities commenced after ETS received written approval of the results by the WDE assessment leadership.

### 6.2 Item and Forms Development

Kolen and Brennan (2014) state that equating adjusts for differences in difficulty, not for differences in content. Properly, then, a discussion of the creation of the Reading and mathematics vertical scale and equating of the PAWS Science assessment begins by noting that the development of the items and forms began in 2005 and has been an ongoing process. Reading and mathematics have new 2012 style guide based on the 2012 WyCPS. Science items have been developed to the same style guide since 2005 (with minor updates throughout), and tests have used comparable blueprints since the first operational administration in 2008. Science test blueprints did change between the 2013 and 2014 administrations with the removal of constructed response items from the assessments. It should be noted that the SAWS writing assessment was discontinued as a component of PAWS and became an independent assessment beginning with the 2013 administration.

The numbers of items in the grade 8 Science assessment were reduced since a passage in 2014 Grade 8 Science Test Booklet also appeared in the Released Test Questions posted earlier this year on the WDE website. After reviewing the consequences, ETS decided to remove the four questions altogether from the scorable bank of operational items and treat them as "Do Not

Score.” This reduced the number of maximum raw points for Science in Grade 8, but did not bear any negative consequences to statistical reliability or content representation.

### 6.3 IRT Models and Calibrations

One parameter Item Response Theory (IRT) model (i.e., *Rasch model*, 1980) was used to calibrate the 2014 Wyoming PAWS assessments for dichotomous items. This measurement model is used regularly to construct test forms, for scaling and equating, and to develop and maintain large item banks. All test analyses, including item model fit analysis, equating, and performance prediction were accomplished within this framework. The statistical software used to calibrate the PAWS operational and field test items that were used in the spring 2014 administration was *WINSTEPS Version 3.68.1* (Linacre, 2007).

The most basic expression of the Rasch model is in the Item Characteristic Curve (ICC). It conceptualizes the probability of a correct response to an item as a function of the student’s ability level and the difficulty of the item. The probability of a correct response is bounded by 1 (certainty of a correct response) and 0 (certainty of an incorrect response). The ability scale is theoretically unbounded. In practice, the ability scale ranges from approximately –4 to +4 logits for heterogeneous ability groups. The probability of an examinee with ability  $\hat{\theta}$  answering item  $i$  with difficulty  $D_i$  is shown in the equation below:

$$P_i(\hat{\theta}) = \frac{\exp(\hat{\theta} - D_i)}{1 + \exp(\hat{\theta} - D_i)}$$

As an example, consider Figure 5.1, in which the response probability curve for a dichotomous item is depicted with a Rasch difficulty ( $D_i$ ) of 0.85. When a person answers a dichotomous item with a difficulty that is at the same level as their ability (ability is represented by  $\theta$  in the equation above), then that person has a 50% chance of answering the item correctly. Another way of expressing this is that if we have a group of 100 people, all of whom have an ability of 0.85, we would expect about 50% of them to answer the item correctly. A person whose ability was above 0.85 would have a higher probability of a correct answer, while a person whose ability is below 0.85 would have a lower probability. This makes intuitive sense and is the basic formulation of Rasch measurement for test items having only two possible scores (i.e., wrong or right).

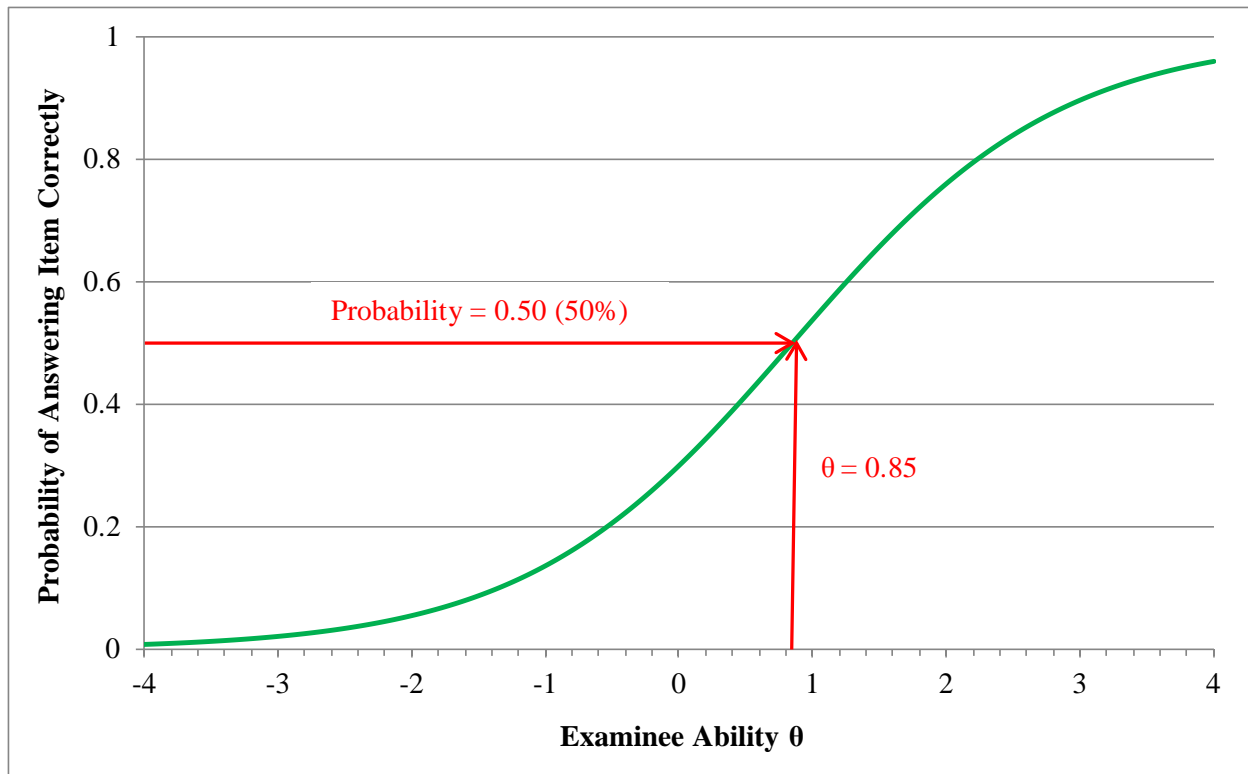


Figure 1. Sample item characteristic curve.

One important property of the Rasch model is its ability to separate the estimation of item/task parameters from the person parameters. With the Rasch model the total raw score is a sufficient statistic for estimating the person's ability (i.e., no additional information is necessary to derive an estimate of the person's level of ability). The total number of responses across examinees in a particular category is a sufficient statistic for estimating the difficulty for that category. Thus with the Rasch model, the same total score will yield the same ability estimate for different examinees, regardless of which *particular* items within the form they answered correctly.

#### 6.4 Fit Statistics for the Rasch Model

Fit statistics are used for evaluating the goodness-of-fit of a model to the data. Fit statistics are calculated by comparing the observed and expected trace lines obtained for an item after parameter estimates are obtained using a particular model. *WINSTEPS* provides two kinds of fit statistics called outfit and infit *mean-squares* that show to what degree the observed data follows the pattern of responses that would be predicted by the model. This indicates how appropriately the model is describing the statistical behavior of the item responses.

Outfit mean-squares are influenced by outliers and are usually easy to diagnose. Infit mean-squares, on the other hand, are influenced by response patterns and are harder to diagnose and remedy. Table 24 presents guidelines for evaluating mean-square fit statistics (Linacre, 2007).

Table 24. Criteria to Evaluate Mean-Square Fit Statistics

Mean-Square	Interpretation
> 2.0	Distorts or degrades the measurement system
1.5 – 2.0	Unproductive for construction of measurement, but not degrading
0.5 – 1.5	Productive for measurement
< 0.5	Unproductive for measurement, but not degrading. May produce misleadingly good reliabilities and separations

In general, mean-squares near 1.0 indicate little distortion of the measurement system, while values less than 1.0 indicate observations that are too predictable (redundancy, model overfit). Values greater than 1.0 indicate unpredictability (unmodeled noise, model underfit).

Appendix M provides Rasch difficulty estimates, standard errors, and infit and outfit statistics for 2014 PAWS operational items. Fit statistics for all but one of the Science items were within the range of 0.5 to 1.5. No operational items exceeded the 2.0 threshold. These results confirm that the Rasch model was appropriate for scaling the 2014 PAWS operational Mathematics and Reading tests, and equating for Science test. Operational classical item statistics are presented in Appendix K.

Appendix L provides IRT statistics and *N*-counts for items field-tested in 2014. Item fit is a factor that is considered during test construction, and items with less than optimal fit statistics that survive data review are not likely to be used on future PAWS forms.

### 6.5 Reading and Mathematics Vertical Scales

According to Young (2006), vertical scales have several important aspects. These include:

- The monitoring of student progress over time within a content area;
- Analyzing the growth patterns for individual students or groups of students in terms of changes in performance and variability from grade to grade;
- Checking on the consistency of achievement-level expectations across grade levels.

It is important to note that vertical scaling produces scales that are *linked* across adjacent grades as opposed to scales that are *equated*. Linked scales are comparable, but have a weaker relationship than equated scales. This relationship is strongest across adjacent grades and weakens as the gap between the grades being compared widens. This is due to the fact that the

tests from adjacent grades cover different subject matter that is specific to their targeted grades. For an equating relationship to exist, the test forms that are being equated should cover the same subject matter. Thus, test forms from the same grade and subject are equated from year to year while test forms from adjacent grades (within grades 3 to 8) and the same subject are linked via the vertical scale.

New scales measuring students' academic performance on Wyoming Content and Performance Standards were established for 2014 PAWS Reading and Mathematics assessments. The vertical scales for Reading and Mathematics were developed in 2014 using an operational and embedded vertical anchor test design. The vertical linking items were embedded within the PAWS 2014 assessment test booklets in the field test positions. The vertical linking items did not count toward a student's scale score.

The Reading scale allows for direct comparison of student test scores across grade levels within a content area. The Mathematics vertical span scale, performing the same function, was designed to address the mathematics blueprint having three content continuums across grades. This scale divides grades 3–8 Mathematics tests into 3 spans (Span I: grades 3–5; Span II: grades 6–7; and Span III: grade 8).

A Rasch model was used for calibration and vertical scaling of 2014 PAWS reading and math tests. A common-item nonequivalent groups design in which students in adjacent grade levels respond to the same items was used to collect data to build a vertical scale for PAWS Reading and Mathematics assessments. The linkages between adjacent grades were established by fixing the item difficulty parameters of the vertical anchor sets in the upper grade to the values obtained from the calibration of these items in the lower grade. Following the decision from Technical Advisory Committee (TAC) meeting, the grade 3 reading scale was chosen to be the base scale for Reading assessments. The grade 4 scale was linked to the grade 3, the grade 5 to the grade 4, etc. As a result, Reading and Mathematics test scores in grades 3–8 are directly comparable across adjacent grades. More information regarding the Reading and Mathematics vertical scales can be found in the PAWS 2014 Calibration and Vertical Scaling Reports for Reading and Mathematics.

### 6.6 Science Equating Analyses

As was previously mentioned, the PAWS Science assessments for 2014 were post-equated, meaning that the item and test statistics used to generate the scoring tables (i.e., tables displaying the relationship between specific raw scores and scale scores for a particular grade level and subject) came from the present (spring 2014) administration. All tests were equated to the pre-existing scale, and so scale scores on the 2014 administration use the same metric as scale scores for the same grade level and subject from previous administrations of the PAWS. The Science tests are only given at grades 4 and 8 and therefore were not vertically scaled.

#### *6.6.1. Calibration and Equating Process for the 2014 Science Administration*

The procedures for equating the 2014 forms of the Science forms to the preexisting scales were similar to those used in 2013. To establish a strong relationship between the 2013 and 2014 forms, each 2014 form had approximately 30% of its items drawn from the set of 2013 operational items. Other items were drawn from the item bank, which was composed of all items used operationally from 2008–2013 (with the exception of those items released publicly as sample PAWS items), and items field tested and accepted at data review from 2008 through 2013.

The tests were equated via common item equating to a calibrated item pool (Kolen and Brennan, 2014). All items were drawn from previous years' administrations and can potentially function as anchor items with their parameters being drawn from their most recent operational use<sup>11</sup>. Nevertheless, in a departure from the practice used in previous administrations, only a core of items that had been operationally used in a previous form and deemed to have reliable bank parameter values were selected for the anchor set. Other items that were only field tested in previous administrations were excluded from the anchor set.

There were some items that were identified as possibly having unreliable statistics from their most recent use and such items were removed from being anchor items. These items included items that were most recently used in the initial 2007 standalone field tests for PAWS. Since the students taking the standalone field tests knew that there were no consequences tied to performance on these tests and that they would not receive any scores from the standalone field test administrations, they likely had less motivation to perform well than did students taking operational tests. Thus, items with statistics derived from these administrations were not used as linking items. Finally, some items had been modified since their most recent use (primarily older items modified to bring them in line with current PAWS item style guidelines). Their previous statistics in the item bank might not be comparable to the statistics of the new modified version of the item. These items were removed from the anchor set. All items that were not used as linking items had their parameters freely estimated while holding the parameters of the remaining anchor items fixed.

Though Rasch (and, in general, IRT) parameters are theoretically invariant across different samples of students, in practice it could be possible for the occurrence of parameter drift. Such

---

<sup>11</sup> ETS used only operationally administered items as anchors. Pearson previously used both field test and operational items as anchors. The change was approved by the WDE.

drift can be the result of shifting emphases in instruction over time, changes in item position from the previous use of the item, contextual effects, or simply random measurement error. Therefore anchor stability should be checked carefully prior to the final calibration analysis to identify any items whose parameters had drifted (i.e., items whose Rasch difficulties estimated from the 2014 administration data differed significantly from their known values used for equating).

Anchor stability analysis was conducted to identify items that were not suitable for use as anchor items. The Robust-Z statistic (Huynh and Meyer, 2010) was utilized to identify items that exhibited instability in their difficulty parameter estimates (multiple choice items) for the 2014 calibration as compared to their difficulty estimates from their most recent previous use.

The first step in computing Robust-Z is to run a WINSTEPS calibration with all items (including those in the anchor set) unanchored (freely calibrated). The Rasch parameter estimates of anchor items in this run and their previous estimates in the item bank were used to calculate the Robust-Z statistics.

Robust-Z is defined as

$$Z = \frac{d - MDN(d)}{0.74 \times IQR(d)}$$

where  $d$  is the difference between the Rasch parameter estimate of an anchor item estimated from the free calibration and its bank parameter estimate,  $MDN(d)$  is the median of  $d$ , and  $IQR(d)$  is the interquartile range of  $d$ . Huynh and Meyer (2010) describe the use of the median and interquartile range as a robustification of the traditional  $z$ -statistic and  $z$ -test. In the above formula, Rasch parameter estimates are Rasch difficulties for multiple choice items (one parameter per item).

Items with a Robust-Z that exceeded 1.645 were deemed to have drifted in difficulty and were considered for being eliminated from the anchor set in the previous protocol. However, ETS retained all flagged item(s) in the anchor set if the items were not identified to be flawed by content experts. This departure from previous protocol was accepted by the WDE based on Yen's (2007) white paper. In the second round of anchor stability checks, anchored items' displacement values were also examined. Linacre (2007, p. 362) describes displacement statistic as:

*... the size of the change in the parameter estimate that would be observed in the next estimation iteration if this parameter was free (unanchored) and all other parameter estimates were anchored at their current values. For a parameter (item or person) that is*



*anchored in the main estimation, (the displacement value) indicates the size of disagreement between an estimate based on the current data and the anchor value.*

If the absolute value of displacement was greater than or equal to 0.5, the item was flagged as having difficulty drift across administrations. The third round of analyses involved examining groups of items for displacements in the same direction, even if those displacements did not individually exceed the threshold value of 0.5 in the second round. This mainly applied to the Reading and Science tests that had sets of items tied to passages, but Mathematics tests were examined as well. If a group of items with something in common (such as a common passage or content area) were all influenced in some way that affected their overall group difficulty in the same way, the cumulative effect of that group on the overall test (specifically, the relationship between raw and scale scores) can be large enough to introduce a significant amount of systematic error into the equating.

The second calibration run of the WINSTEPS software fixed parameters of items in the anchor set to their bank values and freely estimated the parameters of the rest of the items. This procedure enables equating operational test scores from year to year to the baseline scale. The output files that showed the correspondence between raw scores on the test and theta scores (a measure of student ability; see section 5.2) were later used to develop the raw score to scaled score conversion tables (see section 5.6). The theta equivalents for each raw score point were determined iteratively by solving the following equation.

$$TrueScore = \sum_{i=1}^I \sum_{j=0}^{m_i} j \cdot P_{ij}(\hat{\theta})$$

And *True Score* is set to each achievable raw score point to find its theta equivalent.

These theta score estimates were then scaled via constants to the reporting metric. According to Lord and Wingersky (1984), the procedure applied to true scores can be transferred to observed raw scores without any major anomalies in the resulting outcomes.

### 6.7 Translating Raw Scores to Scaled Scores and Performance Levels

Scaled scores on the PAWS Reading, Mathematics, and Science tests ranged generally from 300 to 975 for grades 3–8; the specific minimum and maximum possible scale scores varied by grade and subject. Appendix P provides scale score descriptive statistics for the 2014 PAWS operational reading, mathematics, and science tests.

The following formulae were used to convert the underlying PAWS IRT reading, mathematics, and science scales to the PAWS reporting scale:

$$PAWS \text{ Scaled Score} = \hat{\Theta} \times Slope + Intercept$$

$$PAWS \text{ Scaled SEM} = SEM(\hat{\Theta}) \times Slope$$

where  $\hat{\Theta}$  was the *IRT* ability estimate, and  $SEM(\hat{\Theta})$  was the estimated conditional standard error of measurement (*SEM*) of the ability estimate  $\hat{\Theta}$ . Table 25 also contains the slope, intercept, and LOSS (lowest obtainable scale score) and HOSS (highest obtainable scale score) for the PAWS Reading, Mathematics, and Science scales.

The raw score to scale score conversion tables for the 2014 PAWS Reading, Mathematics, and Science tests can be found in Appendix Q. Conditional standard error estimates and performance levels for the scale scores are also included in these tables. SAWS raw score to performance level conversion tables can also be found in Appendix Q.

Table 25. PAWS Reading, Mathematics, and Science Scaling Constants, Lowest Obtainable Scale Scores, and Highest Obtainable Scale Scores

Grade	Scaling constant	LOSS	HOSS
Reading			
Grade 3	Scaled = $\hat{\Theta} * 43.89281 + 553.1639$	375	800
Grade 4	Scaled = $\hat{\Theta} * 43.89281 + 553.1639$	400	825
Grade 5	Scaled = $\hat{\Theta} * 43.89281 + 553.1639$	425	850
Grade 6	Scaled = $\hat{\Theta} * 43.89281 + 553.1639$	450	875
Grade 7	Scaled = $\hat{\Theta} * 43.89281 + 553.1639$	475	900
Grade 8	Scaled = $\hat{\Theta} * 43.89281 + 553.1639$	500	925
Mathematics			
Grade 3	Scaled = $\hat{\Theta} * 43.4074 + 570.41$	375	850
Grade 4	Scaled = $\hat{\Theta} * 43.4074 + 570.41$	400	875
Grade 5	Scaled = $\hat{\Theta} * 43.4074 + 570.41$	425	900
Grade 6	Scaled = $\hat{\Theta} * 43.4074 + 570.41$	450	925
Grade 7	Scaled = $\hat{\Theta} * 43.4074 + 570.41$	475	950
Grade 8	Scaled = $\hat{\Theta} * 43.4074 + 570.41$	500	975
Science			
Grades 4 and 8	Scaled Score = $\hat{\Theta} * 48.21 + 637.5$	300	900

## 7. PAWS AND SAWS REPORTING

### 7.1 Overview

A thorough understanding of the results of the PAWS and SAWS assessments is essential for all members of the school community (parents, teachers, administrators, and students) to be able to hold students accountable for individual learning progress and delivering targeted intervention as needed to help all students to meet grade level expectations. This level of assessment literacy is only possible if professional educators are well versed in assessment practice and assessment results are presented clearly. Sample student reports are located in Appendices B, C, and D for PAWS and SAWS. Appendix B shows the report for PAWS reading, mathematics, and SAWS students in grades 3, 5, and 7. Appendix C demonstrates the PAWS reading, mathematics, and science reports for grades 4 and 8. Appendix D provides the grade 6 PAWS reading and mathematics reports.

The following reporting information is provided:

- Performance Levels
- Raw and Scaled Scores
- Skill-Reporting Categories
- Production of PAWS and SAWS Individual Student Score Reports

### 7.2 Scaled Scores

The PAWS Reading and Mathematics tests were designed to be comparable across grade levels (vertically) for grades 3–8. The vertical scale scores generally range from 375 to 975 for Reading and Mathematics. Care was taken in crafting the assessment system so that the skills and abilities captured by each grade level assessment (within subject) reflected the same fundamental set of skills. This is the intent of a vertical scaling system. In essence, each PAWS vertical scale reflects a single general underlying construct (e.g., Mathematics ability).

While this is common practice in educational assessment, there are limits to the interpretations based on such scales (Kolen and Brennan, 2014). Where each grade level test is based on a common blueprint design, the grade-level specifics as reflected in the test questions differ from grade to grade. These differences are naturally greater as one compares over wider grade spans. It is thus important to take these underlying factors into consideration when interpreting student performance across grade levels, remembering that the scales for adjacent grades are linked rather than equated. Comparisons across adjacent grades are the most valid.

### 7.3 Performance Levels

Performance classifications are determined by applying the appropriate scale score cuts established from the PAWS standard setting activities described in the 2014 Standard Setting Summary Proficiency Assessments for Wyoming Students (PAWS for Reading, Mathematics) and Student Assessment of Writing Skills (SAWS) (Baron, 2014) and 2008 (Science) PAWS standard setting reports (Pearson, 2008). Tables 26–29 provide the scaled score ranges for the PAWS Reading, Mathematics, Science, and SAWS tests.

Table 26. Proficiency Level Ranges for Grades 3–8 Reading

Grade	Below Basic	Basic	Proficient	Advanced
3	375–552	553–589	590–640	641–800
4	400–565	566–605	606–659	660–825
5	425–577	578–619	620–667	668–850
6	450–588	589–629	630–680	681–875
7	475–605	606–641	642–692	693–900
8	500–615	616–655	656–710	711–925

Table 27. Proficiency Level Ranges for Grades 3–8 Mathematics

Grade	Below Basic	Basic	Proficient	Advanced
3	375–549	550–598	599–659	660–850
4	400–583	584–636	637–696	697–875
5	425–608	609–651	652–726	727–900
6	450–628	629–676	677–742	743–925
7	475–652	653–696	697–752	753–950
8	500–663	664–706	707–762	763–975

Table 28. Proficiency Level Ranges for Grades 4 and 8 Science

Grade	Below Basic	Basic	Proficient	Advanced
4	300–611	612–665	666–725	726–975
8	300–605	606–653	654–713	714–975

Table 29. Proficiency Level Ranges for Grades 3, 5, and 7 SAWS

Grade	Below Basic	Basic	Proficient	Advanced
3	0-8	9-13	14-20	21-36
5	0-8	9-13	14-20	21-36
7	0-8	9-13	14-20	21-36

Descriptions of each performance level provide specific information about the skills and abilities that students at that performance level are typically capable of demonstrating. The performance-level descriptions for Reading, Mathematics, and Science are included on the Student Score Report.

Percentages for all Wyoming students as well as for selected demographic subgroups in each of the four performance levels can be found in Appendix R.

#### 7.4 Content Standard-Level Raw and Scale Scores

Content standard-level scores (by text type for Reading and skill for Science) are provided in the form of scaled and raw scores. The content standard-level scores were produced in the same way as the overall test Scaled scores—a raw score to Scaled score table for each content standard within a particular form/grade/content area combination was derived using the Rasch IRT parameters of the items that mapped to that standard.

The standard-level scaled scores and associated error ranges (student scaled score  $\pm$  one CSEM) are graphically presented on the Student Score Report. Assuming a normal distribution of scaled scores, the probability that a student's true score will be in the range indicated by the error bar is approximately 68%. For Mathematics, scaled scores are provided for Numbers, Operations, and Concepts; Algebra; Geometry; Measurement; and Data Analysis and Probability. For Reading, scale scores are provided by passage type: Functional Texts, Expository Texts, and Narrative Texts. For Science, scaled scores are given by skill type: Observe and Question, Design and Conduct a Scientific Investigation, Organize and Represent Data, and Draw Conclusions and Make Connections. Since measurement error is related to the number of reliable items making up the measure, the error ranges for the standard level scores (i.e., subscale scores) will generally be larger than those for the overall subject-level scale score.

When comparing subscale scores, users should remember that the comparison is affected by measurement error present in both subscales. Generally, the difference between any two subscale scores has a lower level of reliability and a larger SEM than those of the subscales that are being compared. Any decisions based on the comparison between two or more subscale scores should be made with an appropriate degree of caution.

Raw score points earned for each skill-reporting category are also provided relative to total points possible. Skill-reporting categories for Mathematics, Reading, and Science can be found in the blueprints in Appendix A.

#### 7.5 SAWS Raw Scores

The test design for the Spring 2014 SAWS assessment was composed solely of a census with multiple writing prompts. Prompts were aligned to the current (2012) Wyoming writing content

standards. SAWS grade 3 is still in transition. In 2014, grade 3 students responded to two 12-point prompts, one Narrative and the other Informative.

2014 is the base year for grades 5 and 7 with cut points established in July 2014. The students participating in grades 5 and 7 had a single 12-point Narrative writing prompt and a 4 + 8 point set. The 4 + 8 point set requires the student to read a single passage and then answer a 4-point response followed by an 8-point response. Table 30 provides the state level normative percentage distributions for SAWS for 2014.

Table 30. SAWS State Level Normative Percentage Distributions

Grade	State <i>N</i> count	0–6 points	7–12 points	13–18 points	19–24 points
3	7315	5	36	42	17
5	7033	1	27	51	20
7	6763	1	23	49	26

### 7.6 Production of Printed Score Reports for PAWS and SAWS

In final preparation for the production and printing of the PAWS and SAWS combined score reports, the following steps took place at ETS. In the job submission workstation, district orders were submitted in batches for reporting. Upon completion of these jobs, the next step in the process was the production of pilot reports.

The pilot reports allowed the testing and verification of all reporting processes against program reporting requirements. These pilot reports were carefully reviewed by representatives from the following ETS departments: Scoring Operations, Quality Assurance, IT Requirements, IT Scoring Programming, and Contract Testing Program Management. Extensive data checks were performed to verify the validity of reported scores. After verification and sign off by all concerned parties, production reporting commenced.

Individual student reports in paper format were generated for distribution to WDE districts. In addition, a student data file containing student demographic information, item response data, and domain scores was provided to the WDE for Adequate Yearly Progress (AYP) reporting via a secure FTP site. For security purposes, ETS posts to a secure Tumbleweed site. ETS provided secure user IDs and passwords to access the site.

Printed student reports were assembled and packed. Packers visually checked print and form quality during assembly. The reports then moved to pre-ship quality control, where the order received a final quality check prior to shipping. Results were compared against the reporting requirements to verify correct application of the scoring tables and to ensure that all deliverables were present. Each order was then released to shipping. An example of the PAWS/SAWS individual student report is provided in each of Appendices B, C, and D.

### 7.7 Assessment Score Reports: Supplement Guide for Districts and Schools for PAWS and SAWS

The 2014 Wyoming State Assessment Program Score Reports: Interpretation Guide for Teachers for PAWS & SAWS was an online-only version that could be printed by users if desired. It contained explanations of the features and data contained in the PAWS and SAWS reports. It was available on the WDE websites, and was intended for use by all users of the data from the PAWS and SAWS assessment.

## 8. RELIABILITY

### 8.1 Overview

Reliability is the degree to which scores remain consistent over an assessment procedure (Nitko, 2004). Further defined, reliability is the degree to which students' assessment results are consistent when:

- They complete the same task on one, two, or more occasions;
- Two or more raters evaluate their performance on the same task; or
- They complete two or more parallel tasks on one or more occasions.

Consistency of scores over repeated assessment and/or with different raters is the underlying concern of reliability.

This chapter describes the reliability analyses of the 2014 PAWS operational assessments. Internal consistency and interrater reliabilities, classical and conditional standard errors of measurement, and accuracy and consistency results are included.

### 8.2 Internal Consistency Reliability

As a means of gauging score stability, internal consistency reliabilities were computed. Several methods can be used to estimate the internal consistency of a test.

The internal consistency of a test estimates the stability of scores from one sample of content to another. One approach is to split all test questions into two groups and then correlate student scores on the two half-tests. This is known as a split-half estimate of reliability. This method avoids the implications of any changes in the individual by administering only a single test. If scores have a high rate of correlation on the two half-tests, it can be concluded that the test questions complement one another, function well as a group, and measure similar concepts. This also suggests that measurement error is minimal. The split-half method's decision about which questions contribute to each half-test's score can have an impact on the resulting correlation.

As one index of internal consistency, ETS uses Cronbach's coefficient alpha statistic (Cronbach, 1951). The coefficient alpha is the average split-half correlation based on all possible divisions of a test into two parts. Coefficient Alpha is computed using the following formula:



$$r_{xx'} = \left( \frac{N}{N-1} \right) \left( 1 - \frac{\sum s_i^2}{s_x^2} \right)$$

where  $\sum s_i^2$  = sum of all of the item variances,  $s_x^2$  = observed score variance, and

$N$  = the number of items on the test.

Based on the total test, overall alpha statistics suggest reasonable internal consistency reliability for PAWS assessments at all grades and subjects. Alphas were mostly above 0.90 and never lower than 0.87 for any grade/subject combination. These observed reliabilities meet generally accepted industry levels and benchmarks for large-scale assessments. Complete results for PAWS are given in Table 31, including coefficient alpha and the standard error of measurement for each grade and content area. Tables 32–34 provide coefficient alpha and the standard error of measurement for each domain within a grade and content area.

Table 31. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade

Grade	N Counts	Possible Points	Cronbach's Alpha	SEM
Reading				
3	7365	50	0.91	2.88
4	7022	50	0.89	2.89
5	7075	54	0.90	2.85
6	6756	56	0.92	3.05
7	6463	56	0.90	3.19
8	6467	56	0.90	3.00
Mathematics				
3	7369	50	0.90	2.94
4	7026	59	0.89	3.12
5	7076	59	0.93	3.28
6	6759	59	0.92	3.34
7	6467	59	0.91	3.27
8	6470	65	0.92	3.59
Science				
4	7022	50	0.88	3.05
8	6455	46	0.87	3.06

Table 32. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade and Reading Domain

Grade	Domain	N Counts	Possible Points	Cronbach's Alpha	SEM
3	LTKY	7365	20	0.81	1.79
	LTCR	7365	6	0.56	0.93
	INKY	7365	10	0.67	1.32
	INCR	7365	7	0.57	1.18
	LANG	7365	7	0.62	0.99
4	LTKY	7022	15	0.68	1.58
	LTCR	7022	6	0.55	1.01
	INKY	7022	15	0.73	1.56
	INCR	7022	8	0.65	1.06
	LANG	7022	6	0.47	0.95
5	LTKY	7075	14	0.7	1.42
	LTCR	7075	7	0.56	0.86
	INKY	7075	17	0.76	1.71
	INCR	7075	8	0.59	1.15
	LANG	7075	8	0.47	1.16
6	LTKY	6756	15	0.79	1.56
	LTCR	6756	9	0.63	1.16
	INKY	6756	15	0.75	1.70
	INCR	6756	9	0.58	1.36
	LANG	6756	8	0.65	1.12
7	LTKY	6463	13	0.64	1.5
	LTCR	6463	9	0.6	1.2
	INKY	6463	19	0.77	1.87
	INCR	6463	8	0.59	1.15
	LANG	6463	7	0.6	1.14
8	LTKY	6467	12	0.68	1.19
	LTCR	6467	7	0.57	1.05
	INKY	6467	20	0.75	1.85
	INCR	6467	9	0.67	1.21
	LANG	6467	8	0.61	1.00

Table 33. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade and Mathematics Domain

Grade	Domain	N Counts	Possible Points	Cronbach's Alpha	SEM
3	GEOM	7369	6	0.43	0.91
	MEAS	7369	12	0.67	1.55
	ALGE	7369	20	0.82	1.78
	BTEN	7369	6	0.57	0.92
	FRCT	7369	6	0.58	1.10
4	GEOM	7026	6	0.4	1.01
	MEAS	7026	10	0.59	1.34
	ALGE	7026	13	0.73	1.35
	BTEN	7026	10	0.53	1.17
	FRCT	7026	20	0.76	1.86
5	GEOM	7076	6	0.53	1.10
	MEAS	7076	12	0.78	1.45
	ALGE	7076	6	0.54	1.02
	BTEN	7076	16	0.76	1.71
	FRCT	7076	19	0.83	1.90
6	GEOM	6759	6	0.65	1.01
	RELT	6759	10	0.69	1.28
	NMBR	6759	15	0.71	1.72
	EQTN	6759	20	0.81	1.92
	STPR	6759	8	0.68	1.24
7	GEOM	6467	9	0.53	1.37
	RELT	6467	13	0.69	1.50
	NMBR	6467	10	0.69	1.34
	EQTN	6467	18	0.78	1.83
	STPR	6467	9	0.57	1.31
8	GEOM	6470	16	0.75	1.8
	FNCT	6470	14	0.71	1.67
	NMBR	6470	6	0.62	1.05
	EQTN	6470	23	0.83	2.06
	STPR	6470	6	0.52	1.04

Table 34. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade and Science Domain

Grade	Domain	N Counts	Possible Points	Cronbach's Alpha	SEM
4	LIFE	7022	16	0.67	1.72
	PHYS	7022	18	0.74	1.84
	ESCI	7022	16	0.71	1.72
8	LIFE	6455	16	0.71	1.78
	PHYS	6455	18	0.71	1.94
	ESCI	6455	12	0.68	1.47

### 8.3 Interrater Reliability

Rater agreement or consistency is critical for valid test score interpretation for assessments composed of constructed response items requiring human raters to score the performance of students. Interrater agreement provides evidence of the degree to which raters agree in their observations about the qualities evident in students' responses. In order to monitor and evaluate the accuracy of rating, 25% of the responses to SAWS writing prompts were scored twice.

Percentage agreement between two raters is frequently defined as the percentage of exact score and adjacent score agreement. In general, the agreement rates for the traits indicate well above 90% agreement. Tables 35–37 provide the agreement rates for the SAWS operational prompts. Appendix H provides the agreement rates for all SAWS field test prompts. Appendix S provides the SAWS field test demographic performance.

Table 35. SAWS 2014 Overall Interrater Reliability for 12-point Prompt and Trait Rater Agreement

Grade	Trait	N	Rating 1		Rating 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
3	Prompt 1									
	Prompt Total	1607	7.32	2.40	7.18	2.44	0.64	23.65	55.13	
	Idea Development	1607	1.90	0.70	1.88	0.69	0.49	55.38	98.38	0.49
	Organization	1607	1.89	0.73	1.83	0.78	0.58	57.62	98.13	0.58
	Voice	1607	1.78	0.70	1.75	0.71	0.53	58.74	98.13	0.53
Conventions	1607	1.76	0.69	1.73	0.72	0.53	59.43	97.95	0.53	
3	Prompt 2									
	Prompt Total	1739	7.24	2.44	7.11	2.47	0.63	28.87	56.24	
	Idea Development	1739	1.91	0.68	1.88	0.67	0.53	60.09	98.85	0.53
	Organization	1739	1.83	0.68	1.79	0.70	0.51	58.94	98.16	0.51
	Voice	1739	1.79	0.71	1.75	0.71	0.53	58.77	97.93	0.53
Conventions	1739	1.72	0.69	1.68	0.71	0.55	59.34	98.73	0.55	
5	Prompt Total	1706	7.29	2.43	7.22	2.53	0.58	27.20	54.16	
	Idea Development	1706	1.89	0.68	1.87	0.69	0.50	58.91	97.95	0.50
	Organization	1706	1.82	0.68	1.80	0.73	0.52	58.56	98.07	0.52
	Voice	1706	1.80	0.69	1.81	0.72	0.46	54.04	97.48	0.46
	Conventions	1706	1.78	0.69	1.74	0.72	0.49	57.15	97.54	0.49
7	Prompt Total	1661	7.91	2.43	7.87	2.44	0.65	29.44	59.48	
	Idea Development	1661	2.02	0.67	2.01	0.68	0.55	62.67	98.68	0.55
	Organization	1661	2.01	0.70	1.99	0.70	0.59	63.15	98.74	0.59
	Voice	1661	1.95	0.68	1.94	0.69	0.53	59.90	98.62	0.53
	Conventions	1661	1.94	0.69	1.93	0.71	0.51	57.80	98.19	0.51

Table 36. SAWS 2014 Overall Interrater Reliability for 4-point Prompt and Trait Rater Agreement

Grade	Trait	N counts	Rating 1		Rating 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
5	Prompt Total	1726	3.09	0.99	3.05	1.00	0.77	64.89	96.58	
	Response to Text	1726	1.54	0.66	1.53	0.65	0.77	81.81	99.42	0.77
	Holistic	1726	1.55	0.52	1.52	0.52	0.58	77.23	99.94	0.57
7	Prompt Total	1676	2.91	1.04	2.90	1.05	0.75	62.05	94.63	
	Response to Text	1676	1.38	0.70	1.36	0.71	0.75	77.57	99.11	0.75
	Holistic	1676	1.53	0.51	1.53	0.51	0.50	73.93	99.94	0.50

Table 37. SAWS 2014 Overall Interrater Reliability for 8-point Prompt and Trait Rater Agreement

Grade	Trait	N counts	Rating 1		Rating 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
5	Prompt Total	1740	4.90	1.58	4.88	1.57	0.71	35.29	82.07	
	Response to Text	1740	1.34	0.75	1.36	0.73	0.63	67.07	97.47	0.63
	Holistic	1740	3.55	1.13	3.52	1.12	0.66	46.21	91.09	0.66
7	Prompt Total	1683	5.15	1.65	5.10	1.64	0.74	36.01	82.23	
	Response to Text	1683	1.48	0.67	1.48	0.67	0.67	72.73	99.29	0.67
	Holistic	1683	3.67	1.25	3.63	1.25	0.70	43.79	89.19	0.70

#### 8.4 Weighted Kappa

Also provided in Tables 34–37 for operational prompts and Appendix H for field test prompts are the weighted kappas, an index of interrater reliability incorporating a correction for the rate of chance agreement. Weighted kappa was selected since kappa does not take into account the degree of disagreement between observers. It is a generalization of the simple kappa coefficient using weights to quantify the relative difference between categories.

For a writing prompt with  $m$  categories, one can construct an  $m$ -by- $m$  rating table with scores provided by two raters A and B. Define  $P_{ij}$  as the proportion of occurrence of reader A assigning score  $i$  and reader B score  $j$ , respectively, for a given prompt, where  $i = 1, 2, \dots, N; j = 1, 2, \dots, N; i \neq j$ .

The weighted kappa coefficient  $k_{ij}$  is defined as

$$k_w = \frac{P_{o(w)} - P_{e(w)}}{1 - P_{e(w)}}$$

where

$$P_{o(w)} = \sum_i \sum_j w_{ij} P_{ij},$$

$$P_{e(w)} = \sum_i \sum_j w_{ij} P_{i.} P_{.j},$$

$$P_{i.} = \sum_{j=1}^N P_{ij}$$

$$P_{.j} = \sum_{i=1}^N P_{ij}$$

$m$  is the number of categories,  $w_{ij}$  is the Fleiss-Cohen Kappa Coefficient weight (Fleiss and Cohen, 1973), which is defined as:

$$w_{ij} = 1 - \frac{(C_i - C_j)^2}{(C_m - C_1)^2}$$

where  $w_{ij}$  is  $0 \leq w_{ij} \leq 1$ , and  $w_{ij} = w_{ji}$ .  $C_i$  is the score given to the  $i^{th}$  category.

### 8.5 Classical and Conditional Standard Errors of Measurement

Because no assessment measures ability with perfect consistency, it is useful to take into account the likely size of measurement errors. One way to describe the inconsistency of assessment results is to administer the same assessment to a student on multiple occasions and note how much the resulting scores vary. If a student could be assessed on multiple occasions without practice effects, a collection of the student's obtained scores could be compiled. These scores would cluster around an average value. The standard deviation, or spread, of these scores is an estimate of the standard error of measurement (SEM).

The SEM is another index of reliability and provides an estimate of the amount of error in an individual's observed test score. The individual's observed total score is considered an estimate of that individual's true score. Because the standard error of measurement is inversely related to the reliability of a test, the higher the reliability, the lower the standard error of measurement and the more confidence one may have in the accuracy, or precision, of the observed test score. The measurement error is commonly expressed in terms of standard deviation units; that is, the

standard error of measurement is the standard deviation of the measurement error distribution. Under Classical Test Theory and traditional item analysis, we estimate the SEM from:

$$SEM = s_x \sqrt{1 - r_{xx'}}$$

where:  $s_x$  is the observed score standard deviation, and  $r_{xx'}$  is the reliability estimate (coefficient alpha).

In the item response theory (IRT) framework, SEM is estimated as a function of measured ability, and thus is often referred to as a conditional standard error of measurement (CSEM). CSEMs typically are smaller in scaled score units towards the center of the scale where there are more items and more test information and larger at the extremes where there are fewer items and less test information.

Note that the standard error for item difficulty is smallest when the probability of passing is close to the probability of failing. That is, when an item is near the difficulty level for many persons in the sample, the standard error is small (Embretson and Reise, 2000).

Overall Coefficient Alpha and SEM results for PAWS assessments are presented in Table 31. Conditional SEMs for all achievable scores on the assessment are included with the raw score to scaled score tables in Appendix Q for PAWS.

### 8.6 Accuracy and Consistency of Classifications

Analyses were performed using the computer program RelClass (ETS proprietary software) to estimate the accuracy and consistency of decisions about meeting standards on the PAWS assessments. The methods described by Livingston and Lewis (1995) and Young and Yoon (1998) were applied to complete these analyses.

Every discrete test administration will result in some error in the classification of examinees. When an assessment uses performance classifications as the primary method to report test results, accuracy and consistency of decisions become important indicators about the quality of the assessment. This section includes the estimates of decision consistency and accuracy for the PAWS assessments administered in Spring 2014.

The *accuracy* of decisions is represented by the agreement between the classifications based on students' observed scores on the actual test form and the classifications that would have been made based on students' true scores. True scores are assumed to be errorless but are unknown. They can, however, be estimated based on the expected values of test scores over all possible forms of the test. A false positive decision results when a true score corresponds to a classification below a critical cut score (e.g., "does not meet standard"), but the observed score



corresponds to a “meets standard” classification. A false negative decision results when a true score “meets standard,” but the observed score corresponds to a “does not meet standard” classification. Decision *consistency* is the agreement between two non-overlapping and equally difficult forms of the test. This index is estimated using response data from the actual test form and a hypothetical alternate form, based on the actual test form’s estimated reliability.

For each PAWS assessment, the decision consistency and accuracy table includes the proportion of:

- Overall accurate classifications,
- False positives for accurate classifications,
- False negatives for accurate classifications,
- Overall consistent classifications,
- False positives for consistent classifications,
- False negatives for consistent classifications,
- Accuracy around critical cut point (“meets standard” vs. “does not meet standard”), and
- Consistency around critical cut point (“meets standard” vs. “does not meet standard”)

A classification accuracy table is a cross-tabulation of the true score vs. observed score classifications. A classification consistency table is a cross-tabulation of the observed score vs. hypothetical alternate form score classifications.

The proportion of overall accuracy and consistency classifications is computed as the sum of the diagonal cell entries (agreement between observed and true score decisions for accuracy; agreement between observed and hypothetical alternate form score decisions for consistency).

Accuracy and consistency classifications around a critical cut point (e.g., “meets standard” versus “does not meet standard”) are similarly computed by collapsing all classification decisions into a dichotomized distribution around the critical cut point. For each PAWS test, “below basic” and “basic” performance levels result in a “does not meet standard” classification denoted as A in Figure 2; “proficient” and “advanced” performance levels result in the “meets standard” classification indicated as B.

Figure 2. Accuracy or Consistency around Critical Cut Point

Accuracy or Consistency = A + B					
	Below Basic	Basic	Proficient	Advanced	Total
Below Basic	A				
Basic					
Proficient			B		
Advanced					
Total					

Decision accuracy, based on errorless true score classification, is typically higher than decision consistency, which is based on two types of test scores that both contain measurement error. Tables 38–43 present the results of the decision accuracy and consistency of the PAWS cut scores for Reading, Mathematics, and Science. The following information is presented:

- Accuracy classifications;
- False Positives;
- False Negatives; and
- Consistency classifications.

It should be noted that the sum of values of Accuracy, False Positive, and False Negative is equal to 1, but due to rounding errors the sum of the table values may not be equal to 1. False Positive and False Negative classifications refer to the mismatch between student true scores and observed scores. The False Positive value is the proportion of student scores misclassified to the category “*Achieves Proficiency*” when student scores do not meet proficiency. The False Negative value is the proportion of student scores misclassified to the category “*Does Not Achieve Proficiency*” when student scores actually do meet proficiency.

Overall accuracy and consistency ratings range from 0.85 to 0.92, with most results above 0.90. All false negative and false positive results are at or below 0.17. These results suggest acceptable levels of reliability at the cut points for all PAWS assessments.

Table 38. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 3

Subject	N	Accuracy			Consistency			Cut Point Accuracy	Cut Point Consistency
		Overall	False Positive	False Negative	Overall	False Positive	False Negative		
<b>Reading</b>	7365	0.77	0.12	0.11	0.69	0.16	0.15	0.90	0.87
<b>Mathematics</b>	7369	0.79	0.11	0.10	0.71	0.15	0.15	0.91	0.87

Table 39. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 4

Subject	N	Accuracy			Consistency			Cut Point Accuracy	Cut Point Consistency
		Overall	False Positive	False Negative	Overall	False Positive	False Negative		
<b>Reading</b>	7022	0.76	0.13	0.11	0.67	0.17	0.16	0.89	0.85
<b>Mathematics</b>	7026	0.80	0.10	0.10	0.72	0.14	0.14	0.90	0.86
<b>Science</b>	7022	0.78	0.11	0.11	0.69	0.15	0.15	0.90	0.86

Table 40. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 5

Subject	N	Accuracy			Consistency			Cut Point Accuracy	Cut Point Consistency
		Overall	False Positive	False Negative	Overall	False Positive	False Negative		
<b>Reading</b>	7075	0.76	0.12	0.12	0.66	0.17	0.17	0.90	0.87
<b>Mathematics</b>	7077	0.83	0.09	0.09	0.75	0.12	0.12	0.92	0.89

Table 41. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 6

Subject	N	Accuracy			Consistency			Cut Point Accuracy	Cut Point Consistency
		Overall	False Positive	False Negative	Overall	False Positive	False Negative		
<b>Reading</b>	6758	0.79	0.10	0.11	0.71	0.14	0.15	0.92	0.89
<b>Mathematics</b>	6760	0.82	0.09	0.09	0.74	0.13	0.13	0.92	0.88

Table 42. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 7

Subject	<i>N</i>	Accuracy			Consistency			Cut Point Accuracy	Cut Point Consistency
		Overall	False Positive	False Negative	Overall	False Positive	False Negative		
<b>Reading</b>	6796	0.77	0.11	0.12	0.68	0.16	0.16	0.91	0.87
<b>Mathematics</b>	6799	0.80	0.10	0.10	0.71	0.14	0.14	0.91	0.88

Table 43. PAWS 2014 Decision Accuracy and Consistency Indices – Grade 8

Subject	<i>N</i>	Accuracy			Consistency			Cut Point Accuracy	Cut Point Consistency
		Overall	False Positive	False Negative	Overall	False Positive	False Negative		
<b>Reading</b>	6781	0.75	0.13	0.12	0.66	0.17	0.17	0.90	0.87
<b>Mathematics</b>	6784	0.81	0.10	0.10	0.73	0.13	0.14	0.92	0.88
<b>Science</b>	6770	0.77	0.12	0.12	0.67	0.16	0.17	0.89	0.85

## 9. QUALITY CONTROL PROCEDURES

ETS implemented rigorous quality control procedures throughout the test development, administration, scoring, and analyses processes. As part of this effort, ETS program staff consulted with the Office of Professional Standards residing in the legal department. The office publishes and maintains the *ETS Standards for Quality and Fairness*, with the purposes of helping design, develop, and deliver technically sound, fair, and useful products and services, and to help the public and auditors evaluate those products and services.

In addition, every department involved in the program designed and implemented an independent set of procedures to ensure the quality of their products. In the next sections, these quality control procedures are outlined.

### 9.1 Quality Control of Item Development

The item development process for the PAWS and SAWS is described in detail in Chapter 2 of this report. This section highlights the elements of the process devoted specifically to the quality control of item development.

#### *9.1.1. Item and Prompt Specifications*

ETS maintains item specifications for the PAWS and SAWS and has developed an item utilization plan to guide the development of the items for each content area. Item writing emphasis is determined in consultation with the WDE. Adherence to the specifications ensures the maintenance of quality and consistency of the item development process.

#### *9.1.2. Item Writers*

The items for the PAWS and SAWS are written by item writers that have a thorough understanding of the Wyoming Content and Performance Standards. The item writers are carefully screened and selected by senior ETS content staff. Only those with strong content and teaching backgrounds who have experience with students who have severe cognitive disabilities are invited to participate in an extensive training program for item writers.

#### *9.1.3. Internal Contractor Reviews*

Once items have been written, ETS assessment specialists make sure that each item goes through an intensive internal review process. Every step of this process is designed to produce items that exceed industry standards for quality. It includes three rounds of content reviews, two rounds of editorial reviews, an internal fairness review, and a high-level review and approval by a content area director. A carefully designed and monitored workflow and detailed checklists help to ensure that all items meet the specifications for the process.

#### *9.1.4. Content Review*

ETS assessment specialists make sure that the items and related materials comply with ETS's written guidelines for clarity, style, accuracy, and appropriateness and with approved item specifications. The artwork and graphics for the items are created during the internal content review period so assessment specialists can evaluate the correctness and appropriateness of the art early in the item development process. ETS selects visuals that are relevant to the item content and that are easily understood so students do not struggle to determine the purpose or meaning of the questions.

#### *9.1.5. Editorial Review*

Another step in the ETS internal review process involves a team of specially trained editors who check questions for clarity, correctness of language, grade-level appropriateness of language, adherence to style guidelines, and conformity to acceptable item-writing practices. The editorial review also includes rounds of copyediting and proofreading. ETS takes pride in the typographical integrity of the items presented to our clients and strives for error-free items beginning with the initial rounds of review.

#### *9.1.6. Fairness Review*

One of the final steps in the ETS internal review process is to have all items and stimuli reviewed for fairness. Only ETS staff members who have participated in the ETS Fairness Training, a rigorous internal training course, conduct this bias and sensitivity review. These staff members have been trained to identify and eliminate test questions that contain content that could be construed as offensive to, or biased against, members of specific ethnic, racial, or gender groups.

#### *9.1.7. Assessment Director Review*

As a final quality control step, the content area's assessment director or another senior-level content reviewer read each item before it is presented to the WDE.

#### *9.1.8. Data Review of Field Tested Items*

ETS field tests newly developed items to obtain statistical information about item performance. This information is used to evaluate items that are candidates for use in operational test forms. The item statistics are examined carefully at data review meetings, where content experts discuss items that have poor statistics and do not meet the psychometric criteria for item quality. The WDE defines the criteria for acceptable or unacceptable item statistics. This ensures that the item has an appropriate level of difficulty for the target population. The content experts make recommendations about whether to accept or reject each item for inclusion in the PAWS and SAWS item banks.

#### *9.1.9. Quality Control of the Item Bank*

After completion of the pilot analyses, the items are placed in the item bank with their statistics. ETS delivers the prompts to the WDE through an electronic item bank. The item bank database is maintained by a staff of application systems programmers, led by the Item Bank Manager. All processes are logged; all change requests, including item bank updates for prompt availability

status, are tracked. All output and Wyoming item bank deliveries underwent quality control for accuracy.

The quality of the item bank and secure transfer of the Wyoming item bank to the WDE is crucial. The ETS internal item bank database resides on a server within the ETS firewall. Access to the SQL, the server database, is strictly controlled by means of system administration. The electronic item banking application includes a login/password system to authorize access to the database or designated portions of the database. In addition, only users authorized to access the specific database are able to use the item bank. Users are authorized by a designated administrator at the WDE and ETS.

## 9.2 Quality Control of Test Materials

ETS followed a meticulous set of internal quality standards to ensure high-quality printed products for all testing related materials.

- **Publishing and Editing Review**—A three-way review of all project materials was performed internally. After this internal review, assessment materials were forwarded to WDE for review and approval.
- **Printing**—All external printing companies hired to print scannable and nonscannable forms guaranteed the highest level of quality.
- **Multiple Checks**—ETS Program Managers conducted quality checks during the printing process to confirm all requirements for printed materials were met.

Accurate packing, shipping, and collection of test materials were critical for districts and schools to successfully administer the tests. Shipping carriers had online, traceable distribution systems to track all materials.

### *9.2.1. Collecting Test Materials*

After administration, schools returned scorable and nonscorable materials within five working days after the last testing day of each test administration period. Schools were provided UPS return labels with bar-coded information identifying the school. Schools applied the appropriate labels and numbered the cartons prior to returning the materials. All materials were returned via two-day UPS shipment.

ETS closely monitored the return of materials through the “SeNT” system tracking each package of materials shipped out to sites and shipped back to ETS. The Wyoming Customer Support Center at ETS contacted schools not returning materials in a timely manner and worked with them to facilitate the return of the test materials.

### *9.2.2. Processing Test Materials*

Upon receipt of the test materials, ETS used precise inventory and test processing systems, in addition to quality assurance procedures, to maintain an up-to-date accounting of all the testing

materials within their facilities. The materials were removed carefully from the shipping cartons and examined for a number of conditions, including physical damage, shipping errors, and omissions. A visual inspection to compare the number of students recorded on the Header sheets with the number of test and answer books or answer documents in the stack was also conducted.

ETS's image scanning process captured security information electronically and compared scorable material quantities reported on the Headers to actual documents scanned. Schools were contacted by phone if there were any missing shipments or if the quantity of materials returned appeared to be more or less than expected.

### 9.3 Quality Control of Scanning

ETS ensured all student test booklets had been accounted for and processed through scanning, pre-editing, and post-editing processes. All student test and answer books or answer documents returned to ETS were scanned and scored.

The intensity levels of each scanner were constantly monitored throughout each administration for quality control purposes. Intensity diagnostic sheets were run before and during each batch to verify the scanner was working properly. In the event a scanner failed to properly pick up data on the diagnostic sheets, the scanner was recalibrated before it resumed processing student documents.

Documents received in poor condition (torn, folded, or water-stained) that could not be fed through the high-speed scanners were keyed into the system manually.

### 9.4 Quality Control of Psychometric Analyses

The psychometric analyses conducted at ETS underwent comprehensive quality checks by a team of psychometricians and data analysts. Detailed checklists were consulted by members of the team for each of the statistical procedures performed.

Any items flagged for questionable statistical attributes were sent to Assessment Development staff for their review; PAWS and SAWS psychometricians reviewed their comments before prompts were approved to be included in operational forms. Additionally, the statistics imported into the item banking system were thoroughly checked by data analysts and psychometricians before and after the import.

### 9.5 Quality Control of Reporting

For the quality control of Wyoming student reports, three general areas are evaluated, including the following:

- Comparing report formats to input sources from the WDE approved samples
- Validating and verifying the report data by querying the appropriate student data



- Proof reading individual student reports at the WDE and ETS prior to any school district mailings

The student report was required to include a single, accurate WISER ID, a school district name, and a school name. After the draft version of the report was validated against the WDE's requirements, a set of student reports for a pilot district were provided to the WDE for review and approval. ETS sent a PDF of the reports. The WDE and ETS reviewed and signed off on the reports after a thorough review. Upon the WDE's approval of the reports, ETS proceeded with production.

#### *9.5.1. Excluding Student Scores from Summary Reports*

ETS provided specifications to the WDE documenting when to exclude student scores from summary reports. This specification included the logic for handling answer documents, for example, "*was absent,*" "*was not tested due to parent/guardian request,*" or "*did not complete the test due to illness.*"

## 10. HISTORICAL COMPARISONS

Historical comparisons of the PAWS test results are routinely performed to identify trends in examinee performance, in terms of percentage of students meeting standards. As this is the first administration of the PAWS Reading and Mathematics under the new standards 2012 WyCPS, only the 2014 performance is included in Tables 44–47. Tables 48–49 provide a comparison of percentages of the students classified as “Proficient + Advanced” from 2008 to 2014 for PAWS Science.

The percentage of Science students in the equating population classified as “Proficient + Advanced” decreased for grades 4 and increased for grade 8 from 2013. Grade 4 had the largest drop in the percentage of students classified as “Proficient + Advanced,” from 57.5% to 52.5%, a decrease of 5.0%. Grade 8 had an increase of 3.8% from 43.7% to 47.5%. For both grades, the percentage of “Proficient + Advanced” students is within previously observed values for the specific grade.

Because this is also the first administration of the SAWS assessment under the new standards 2012 WyCPS, only the 2014 performance is included in Tables 50–51.

Figures 4 through 5 display the PAWS percentages of students in the equating populations classified as “Proficient + Advanced” from the 2008 through 2014 for each Science grade level. The results for 2010 were not provided due to federal exemption for reporting scores. Displays of Reading and Mathematics students in the equating populations classified as “Proficient + Advanced” are not provided since this is a base year.

Table 44. Scaled Scores Descriptive Statistics for the PAWS Reading Tests

Year	<u>Grade 3</u>			<u>Grade 4</u>			<u>Grade 5</u>			<u>Grade 6</u>			<u>Grade 7</u>			<u>Grade 8</u>		
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>
2014	7365	600.0	50.0	7022	616.0	46.8	7075	626.5	47.4	6758	636.3	48.8	6796	649.3	44.6	6781	661.1	47.7

Table 45. Percentage Proficient and Advanced for the PAWS Reading Tests

Grade	2014
3	61.8
4	63.7
5	58.1
6	56.9
7	58.8
8	57.7

Table 46. Scaled Scores Descriptive Statistics for the Mathematics Tests

	<u>Grade 3</u>			<u>Grade 4</u>			<u>Grade 5</u>			<u>Grade 6</u>			<u>Grade 7</u>			<u>Grade 8</u>		
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>
2014	7369	600.0	50.0	7026	634.7	45.1	7077	659.7	50.6	6760	678.1	48.2	6799	691.8	45.5	6784	707.5	45.0

Table 47. Percentage of Proficient + Advanced Students for the Mathematics Tests

Grade	2014
3	50.4
4	46.6
5	54.0
6	48.5
7	42.8
8	49.2

Table 48. Scaled Scores Descriptive Statistics for the Science Tests

Year	<i>N</i>	<u>Grade 4</u>		<i>N</i>	<u>Grade 8</u>	
		Mean	<i>SD</i>		Mean	<i>SD</i>
2008	6508	665.9	46.5	6588	649.8	44.6
2009	6631	668.1	44.3	6339	647.2	41.2
2010	-	-	-	-	-	-
2011	6680	672.4	42.9	6554	656.5	42.9
2012	6771	677.2	41.1	6752	655.9	44.6
2013	7157	673.1	44.6	6754	651.6	45.6
2014	7022	669.7	46.4	6770	650.9	45.5

Table 49. Percentage of Proficient + Advanced Students for the Science Tests

Grade	2008	2009	2010	2011	2012	2013	2014	Min	Max	Median	2014 Difference from Median	2014 Difference from 2013
4	50.9	50.5		54.5	63.3	57.5	52.5	50.5	63.3	53.0	-0.5	-5.0
8	46.4	42.9		50.7	51.2	43.7	46.8	42.9	51.2	46.6	0.2	3.1

Table 50. Raw Scores Descriptive Statistics for the SAWS Tests

Year	<u>Grade 3</u>			<u>Grade 5</u>			<u>Grade 7</u>		
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>
2014	7315	13.7	4.6	7033	15.0	4.1	6763	15.7	4.3

Table 51. Percentage Proficient and Advanced for the SAWS Tests

Grade	2014
3	50.7
5	63.6
7	69.1

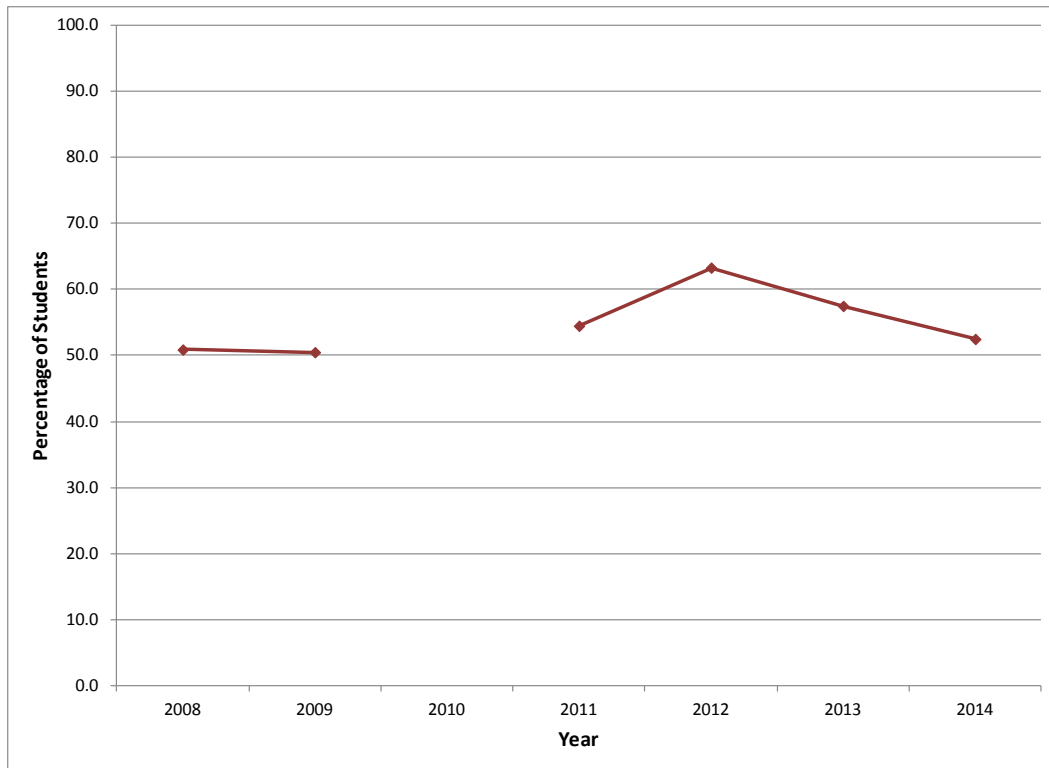


Figure 3. Percentage of Proficient and Advanced Students for Grade 4 Science

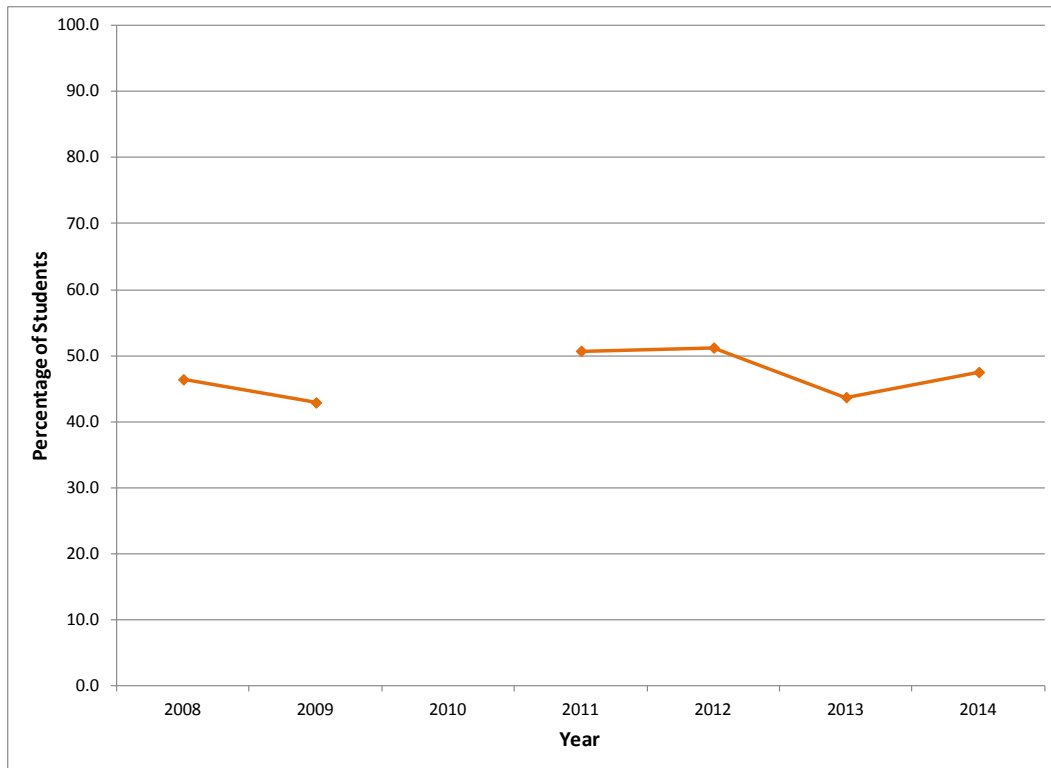


Figure 4. Percentage of Proficient and Advanced Students for Grade 8 Science

## 11. REFERENCES

- American Educational Research Association [AERA], American Psychological Association [APA], and National Council on Measurement in Education [NCME], (2014). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.
- Andrich, A. (1988). *Rasch models for measurement*. Newbury Park, CA: SAGE.
- Andrich, A. (1989). Distinctions between assumptions and requirements in measurement in the social sciences, in J. A. Keats, R. Taft, R. A. Heath, and H. H. Lovibond (Eds.) *Mathematical and theoretical systems* (pp. 7–16). North-Holland: Elsevier Science.
- Baron, P. (2014). 2014 Standard Setting Summary Proficiency Assessments for Wyoming Students (PAWS) and Student Assessment of Writing Skills (SAWS). (Technical Report). Princeton, NJ: Educational Testing Service.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- Dorans, N. J. and Kulick, E. (1983). *Assessing unexpected differential item performance of female candidates on SAT and TSWE forms administered in December 1977: An application of the standardization approach*. Research Report No. 83-9. Princeton, NJ: Educational Testing Service.
- Dorans, N. and Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, 23, 355–368.
- Embretson, S., and Reise, S. (2000). *Item response theory for psychologists*. City, NJ: Erlbaum.
- ETS (2002). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service.
- Hambleton, R. K., Swaminathan, H., and Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: SAGE.
- Holland, P. W., and Thayer, D. T. (1988). Differential item performances and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Eds.), *Test Validity*, (pp. 129–145). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

- Huynh, H. and Meyer, P. (2010). Use of Robust z in detecting unstable items in item response theory models. *Practical Assessment, Research and Evaluation*, 15(2). Retrieved from: <http://pareonline.net/getvn.asp?v=15andN=2>
- IEP definition (2014). Retrieved from <http://www.ncld.org/students-disabilities/iep-504-plan/what-is-iep>).
- Kane, M. (2006). Validation. In R. Brennan (Ed.), *Educational Measurement* (4th ed.). Washington, DC: American Council on Education and National Council on Measurement in Education.
- Karkee, T., Lewis, D. M., and Barton, K. (2005, April). *The effect of including or excluding students with testing accommodations on IRT calibrations*. Paper presented at the meeting of the American Educational Research Association, Montreal, Canada.
- Kim, D., Wang, S., Zhao, Y., and Li, T. (2006, April). *Validation and invariance of factor structure of a statewide reading comprehension test under accommodation and non-accommodation conditions*. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.
- Kolen, M. J. ,and Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). New York, NY: Springer Science + Business Media.
- Linacre, J. M. (2007) *WINSTEPS Rasch measurement computer program and manual (PDF file) v 3.64.2*. Chicago, IL: Winsteps.com
- Livingston, S. A., and Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179–197.
- Lord, F. M., and Wingersky, M. S. (1984). Comparison of IRT true-score and equipercentile observed-score equatings. *Applied Psychological Measurement*, 8, 452–461.
- Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association* 58, 690–700.
- Mantel, N., and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–748.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18, 5–11.
- Nitko, A. J. (2004). *Educational assessments of students*. Englewood Cliffs, NJ: Prentice Hall.
- No Child Left Behind Act of 2001, Pub. L. No. 107–110, 115 Stat. 1425 (2002).



- Pearson. (1988) Technical Report: Science Grades 4, 8, and 11 Standard Setting, 2008. (Technical Report). Iowa City, IA: Pearson.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press.
- Rutherford, J. F. and Ahlgren, A. (1989). *Science for all Americans*. New York, NY; Oxford University Press.
- Ryan, J. P. (1983). Introduction to latent trait analysis and item response theory, in W. E. Hathaway (Ed.), *Testing in the schools. New directions for testing and measurement* (pp. 48–64). San Francisco, CA: Jossey-Bass.
- Webb, N. L. (1999) *Alignment between standards and assessment*, University of Wisconsin Center for Educational Research.
- Wyoming Department of Education (2006). *Wyoming accommodations manual for instruction and assessment: How to select, administer, and evaluate use of accommodations for instruction and assessment of students with disabilities*. Cheyenne, WY: Author.
- Yen, W. (2007). Removing outlier anchor items from IRT equating for educational achievement tests. Unpublished paper.
- Young, M.J. (2006). Vertical scales, in S.M. Downing and T.M. Haladyna (Eds.), *Handbook of test development* (pp.469–485). Mahwah, NJ: Erlbaum.
- Young, M. J., and Yoon, B. (1998, April). *Estimating the consistency and accuracy of classifications in a standards-referenced assessment*. (Technical Report 475). Center for the Study of Evaluation, Standards, and Student Testing. Los Angeles, CA: University of California, Los Angeles.
- U.S. Department of Education, National Center for Education Statistics (NCES), (2001). The Condition of Education 2001. NCES 2001–072, Washington, DC: U.S. Government Printing Office.

## 12. GLOSSARY OF TERMS

The terms below are defined by their application in this document and their common uses in the Wyoming PAWS technical report. Some of the terms refer to complex statistical procedures used in the process of test development. In an effort to avoid the use of excessive technical jargon, definitions have been simplified; however, they should not be considered exhaustive.

**Accommodations** - Changes made in the format or administration of the test to provide options to test takers who are unable to take the original test under standard test conditions.

**Achievement levels** - Descriptions of a test taker's competency in a particular area of knowledge or skill, usually defined as ordered categories on a continuum classified by broad ranges of performance.

**Assessment Descriptions** - These provide skill level descriptions or topics which rely on the structure of the discipline in order to organize instruction. A skill can be defined as somewhere between the breadth of a content standard and the specificity of a benchmark.

**Alternate Assessment** - An assessment that is administered to students for whom the regular assessment with or without an accommodation is inappropriate. It is only used with students who have an individualized education program (IEP) and are unable to respond to accommodated versions of the standard test materials. Wyoming's alternate assessments include Reading, Mathematics, Science and SAWS-ALT administered by the teacher.

**Alignment** - Alignment procedures examine the agreement or match between educational components such as test items and academic standards. To the extent that test items are aligned with academic standards, they are considered to be valid measures of those standards.

**Anchor Sets** - Anchor sets are responses to constructed-response items that best match the criteria on the scoring rubrics. They are selected and assembled during Range Finding. These examples of student work are used to anchor the scoring of the constructed items in the PAWS. The use of anchor sets helps raters assign scores consistently.

**Answer Document** - The form or document on which a student records answers to assessment questions (grades 6–8). These are scannable and have grids for recording student name and demographic information.

**Benchmarks** - These statements specify what students are expected to know and should be able to do at the end of each of the benchmark grade levels in this document, grades 3 through 8. These benchmarks specify the skills and content students must master along the way in order to reach the content standards by the time they graduate.

**Blueprint (Test Blueprint)** - Tests are built to specifications, sometimes called blueprints, in the same way that a house is built to a blueprint. The blueprint specifies such things as reporting categories, number of items for each category, and the number of operational and field test items on the test.

**Common Items** - Test questions that are contained on all test forms and administered to all students in the assessment group.

**Constructed Response Item** - An item for which the student is required to write or draw a response. Such an item must be scored manually. Constructed Response items only appear on the SAWS assessment.

**Content Area** - Subject area; for example, Reading, Mathematics, or Science.

**Content Standards** - These statements define what students are expected to know and should be able to do by the time they graduate. They do not dictate what methodology or instructional materials should be used, nor how the material is delivered.

**Criterion Referenced Test (CRT)** - A customized achievement test that describes student performance in terms of a specific standard. Typically, criterion-referenced testing has been associated with classroom testing where instructional objectives are used. In recent years, standardized testing has moved towards customized criterion-referenced testing in order to provide testing instruments that better align with state and local educational objectives.

**Cut Scores** - A specific point on a score scale, such that scores at or above that point are interpreted or acted upon differently from scores below that point.

**Differential Item Functioning (DIF)** - A statistical procedure for helping detect if an item is differentially difficult for particular groups of test takers with the same ability level. DIF helps determine if members of a particular group have difficulty with an item, not because they know less but because they have different cultural experiences or assumptions. Members of the Item Review panel look at items marked by the DIF procedure and judge whether there was something about the item that was unfair to the group identified.

**Dimensionality** - The extent to which a test item measures more than one ability.

**Embedded Test Model** - Using an operational test to field-test new items or sections. The new items or sections are embedded into the new test and appear to examinees as being indistinguishable from the operational test.

**Equating** - A psychometric process that ensures comparability of scores from one test form to another (e.g., from year to year or from form to form).

**Equivalent forms** - Statistically insignificant differences between forms (i.e., one form is not harder than another).

**Exemplar** - A response to a constructed-response item that is an ideal example of a particular score point of a rubric. Also referred to as an anchor response.

**504 Plan** - An official educational document that may specify a special testing condition (e.g., accommodation) for a student taking an NCLB-related test. In some cases an IEP may specify an alternate assessment or other sources of data related to a student's achievement.

**Field Test** - A field test is a practice run of the items ensuring that test questions are accurate and fair for all students. Statistics produced from field testing will be used in interpreting item behavior/performance and allow for the calibration of item parameters used in equating tests.

**IEP** - Each public school child who receives special education and related services must have an Individualized Education Program (IEP). Each IEP must be designed for one student and must be a truly individualized document. The IEP creates an opportunity for teachers, parents, school administrators, related services personnel, and students (when appropriate) to work together to improve educational results for children with disabilities. (IEP definition, 2014).

**Instructionally Supportive Assessment** - Assessment intended to promote more effective classroom instruction.

**Interrater Reliability** - A method of measuring the agreement among raters scoring the same responses. Compares the scores assigned by one reader to those of another for the same student responses. Reports showing reliability are used to monitor reader performance.

**Item** - A test question. Examples of formats are multiple choice, open-ended (constructed response), and extended response.

**Item Analysis** - Statistical analysis that provides measurement and bias information about items. This information is used for item reviews, test construction, technical reports, and other psychometric documentation. Item analysis may also refer to a quality control step to verify/check answer keys. The item or foil analysis report shows the number and percentage of students responding to each answer choice as well as difficulty values, item-test correlations, for the items.

**Item Bank** - An item bank is a collection of test items, along with associated material (e.g., Reading passages, reviewer's comments) and item statistics. Test items that have passed all reviews are eligible to be put on an operational test.

**Item Calibration** - A process of evaluating item functioning using an Item Response Theory (IRT) model (see description below). The results of item calibration are various item parameter estimates.

**Item Difficulty** - A number that indicates how easy or hard an item is with regard to its intended use. Item difficulty is typically displayed as a  $p$ -value, the proportion of examinees choosing the correct answer. It can also be displayed as a value obtained from an Item Response Theory procedure such as the Rasch logit difficulty or the 3PL theta.

**Item Discrimination** - A number that indicates how well an item differentiates students who know the content measured by the item from those who do not know the content. It is used for indicating how well an item differentiates the more able students from the less able students. Item discrimination is typically displayed as a correlation coefficient with larger positive numbers indicating better discrimination (e.g., 0.42).

**Item Response Theory** - A method of test item analysis that takes into account the ability of the examinee and determines characteristics of the item relative to other items in the test.

**Item Specifications** - Item specifications specify the language and format item writers must follow when constructing items.

**Mantel-Haenszel** - A statistical procedure that examines the differential item functioning (DIF) or the relationship between a score on an item and the different groups answering the item (e.g., gender, race), controlling for ability level. This procedure is used to identify individual items for bias review.

**Operational Test** - Test is administered statewide with standardized procedures and full reporting of scores and stakes for examinees and schools.

**$p$ -value** - Difficulty of an item defined by using the proportion of examinees who answered an item correctly.

**Parallel Forms** - Two or more test forms that are developed for a given exam program, according to the same test blueprint and statistical criteria. The forms should be assembled in such a way that they are as similar to one another as possible.

**Percentile** - The score on a test below which a given percentage of scores fall.

**Performance Level Descriptors** - These statements describe how well students must perform the benchmark standards in order to meet each performance level. The proficient level is required to meet the standards. These descriptors help teachers judge how students are performing in relation to meeting the standards.

**Rangefinding** - The process of selecting responses that exemplify particular score points. The set of responses is used in scoring guides and other training materials that prepare raters for scoring.

**Rasch Model** - A psychometric model from the IRT family of models that permits objective comparisons of individuals, items, etc. Rasch provides both estimates of item difficulty (logit difficulty) and person ability (logit ability) on the same scale. It is used for scaling and equating test forms as well as producing item analyses.

**Raw Score** - The unadjusted score on a test determined by counting the number of correct answers.

**Reliability** - The extent to which test scores are reproducible. If a class of students theoretically took the same test twice in one day and each student's score was the same on the second administration of the test as on the first, the test would be perfectly reliable (1.00). Of course, perfection is not possible and reliabilities in the 0.90s are considered good. In handscoring, reliability (interrater reliability) refers to agreement between raters when assigning scores. Handscoring quality control reports help monitor reader reliability.

**Rollup** - a compilation of individual scores for students into class, school, district, region and/or state level summary reports.

**Rubric** - The criteria used to rate student responses to constructed-response items. Rubrics vary according to the type of item and the goals of the testing program.

**Scaled Score** - A score to which raw scores are converted by numerical transformation. Scale scores allow for comparison of different forms of the test using the same scale.

**Standard Deviation** - A measure of variability, expressed in the same metric as the score. It indicates the spread of test scores around the mean. Assuming a normal distribution, if you know the mean and standard deviation of a distribution, you can determine what proportion of scores falls within one standard deviation of the mean.

**Standard Error of Measurement** - The standard deviation of an individual's observed scores, usually estimated from group data.

**Test and Answer Book** - The form or document on which a student records answers to assessment questions (grades 3–5). These are scannable and have grids for recording student name and demographic information.

**Test Development** - The process of constructing a test. It includes writing the items or test questions and selecting the good items and organizing them into test forms.

**Test Map** - A master document containing a detailed breakdown of a test's specifications by item, objective, cluster, subtest, and all roll-ups involved with each level of reporting category on each testing program. It is considered the master source for information about a test.

**Test Specifications** - Test specifications are the specific rules and characteristics that guide the development of a test. Adherence to test specifications ensures that equivalent test forms are developed annually. Test specifications refer to the overall characteristics of the test content and format that must be followed when constructing tests.

**Validity** - The appropriateness or correctness of inferences, decisions, or descriptions made about individuals, groups, or institutions from test results. There is no such thing as a generically valid test. Validity must be considered in terms of the correctness of a particular inference made from test scores.

# Appendix A: PAWS 2014 Reading, Math, Science, and SAWS Blueprints

## Reading

Table A1. PAWS 2014 Grade 3 Reading Blueprint

### DRAFT Blueprint 2014 - 50 3R OP items aligned to CCSS

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage	
		2013		2014		2015		
Reading Literature			R.03.N skills: 62% (31 items)		Literary portion: approx. 53%		Literary portion: approx. 36%	
Key Ideas and Details								
RL3.1	Ask and answer questions to demonstrate understanding of a text, referring explicitly to the text as the basis for answers.	28						
RL3.2	Recount stories, including fables, folktales, and myths from diverse cultures; determine the central message, lesson, or moral and explain how it is conveyed through key details in the text.							
RL3.3	Describe characters in a story (e.g., their traits, motivations, or feelings) and explain how their actions contribute to the sequence of events.							
Craft and Structure								
RL3.4	Determine the meaning of words and phrases as they are used in a text, distinguishing literal from non-literal language.	3		N.1 = 15 items N.2 = 16 items N.3 = 0 items				
RL3.5	Refer to parts of stories, dramas, and poems when writing or speaking about a text, using terms such as chapter, scene, and stanza; describe how each successive part builds on earlier sections.							
RL3.6	Distinguish their own point of view from that of the narrator or those of the characters.							
Integration of Knowledge and Ideas (see below)								
Range of Reading and Level of Text Complexity								
RL3.10	By the end of the year, read and comprehend literature, including stories, dramas, and poetry, at the high end of the grades 2–3 text complexity band independently and proficiently.							



CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Reading Informational Text			R.03.E skills: 18% (9 items)  E.1 = 4 items E.2 = 0 items E.3 = 5 items  R.03.F skills: 20% (10 items)  F.1 = 10 items F.2 = 0 items		Informational portion: approx. 32%		Informational portion: approx. 32%
Key Ideas and Details							
RI.3.1	Ask and answer questions to demonstrate understanding of a text, referring explicitly to the text as the basis for the answers.	15					
RI.3.2	Determine the main idea of a text; recount the key details and explain how they support the main idea.						
RI.3.3	Describe the relationship between a series of historical events, scientific ideas or concepts, or steps in technical procedures in a text, using language that pertains to time, sequence, and cause/effect.			8-10		7-9	
Craft and Structure							
RI.3.4	Determine the meaning of general academic and domain-specific words and phrases in a text relevant to a grade 3 topic or subject area.	2					
RI.3.5	Use text features and search tools (e.g., key words, sidebars, hyperlinks) to locate information relevant to a given topic efficiently.						
RI.3.6	Distinguish their own point of view from that of the author of a text.			6-8		7-9	
Integration of Knowledge and Ideas (see below)							
Range of Reading and Level of Text Complexity							
RI.3.10	By the end of the year, read and comprehend informational texts, including history/social studies, science, and technical texts, at the high end of the grades 2–3 text complexity band independently and proficiently.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Integration of Knowledge and Ideas*							
			Captured in totals above		*Integration portion: approx. 2%		*Integration portion: approx. 16%
RL3.7	Explain how specific aspects of a text's illustrations contribute to what is conveyed by the words in a story (e.g., create mood, emphasize aspects of a character or setting).						
RL3.8	(Not applicable to literature)						
RL3.9	Compare and contrast the themes, settings, and plots of stories written by the same author about the same or similar characters (e.g., in books from a series).						
RI3.7	Use information gained from illustrations (e.g., maps, photographs) and the words in a text to demonstrate understanding of the text (e.g., where, when, why, and how key events occur).	2		0-2†		7-9	
RI3.8	Describe the logical connection between particular sentences and paragraphs in a text (e.g., comparison, cause/effect, first/second/third in a sequence).						
RI3.9	Compare and contrast the most important points and key details presented in two texts on the same topic.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Language*							
L3.4	Determine or clarify the meaning of unknown and multiple-meaning word and phrases based on grade 3 reading and content, choosing flexibly from a range of strategies.	0	no direct alignment to Wyoming skills: 0%	6-8	*Language portion: approx. 13%	7-9	*Language portion: approx. 16%
L3.4.a	Use sentence-level context as a clue to the meaning of a word or phrase.						
L3.4.b	Determine the meaning of the new word formed when a known affix is added to a known word (e.g., agreeable/disagreeable, comfortable/uncomfortable, care/careless, heat/preheat).						
L3.4.c	Use a known root word as a clue to the meaning of an unknown word with the same root (e.g., company, companion).						
L3.4.d	Use glossaries or beginning dictionaries, both print and digital, to determine or clarify the precise meaning of key words and phrases.						
L3.5	Demonstrate understanding of word relationships and nuances in word meanings.						
L3.5.a	Distinguish the literal and nonliteral meanings of words and phrases in context (e.g., take steps).						
L3.5.b	Identify real-life connections between words and their use (e.g., describe people who are friendly or helpful).						
L3.5.c	Distinguish shades of meaning among related words that describe states of mind or degrees of certainty (e.g., knew, believed, suspected, heard, wondered).						
L3.6	Acquire and use accurately grade-appropriate conversational, general academic, and domain-specific words and phrases, including those that signal spatial and temporal relationships (e.g., After dinner that night we went looking for them).						
suggested total OP items on form:		50		50		50	

\*Integration and Language reporting categories have items associated with both literary and informational passages. The goal will be to strike an overall balance of approx. 50% per genre on the test form.

†Integration of Knowledge and Ideas will not report out in 2014.

Table A2. PAWS 2014 Grade 4 Reading Blueprint

**DRAFT Blueprint 2014 - 50 4R OP items aligned to CCSS**

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Reading Literature			R.04.N skills: 32% (16 items)  N.1 = 7 items N.2 = 9 items N.3 = 0 items				
Key Ideas and Details							
RL.4.1	Refer to details and examples in a text when explaining what the text says explicitly and when drawing inferences from the text.	16					
RL.4.2	Determine a theme of a story, drama, or poem from details in the text; summarize the text.						
RL.4.3	Describe in depth a character, setting, or event in a story or drama, drawing on specific details in the text (e.g., a character's thoughts, words, or actions).						
Craft and Structure							
RL.4.4	Determine the meaning of words and phrases as they are used in a text, including those that allude to significant characters found in mythology (e.g., Herculean).	0			Literary portion: approx. 42%		Literary portion: approx. 35%
RL.4.5	Explain major differences between poems, drama, and prose, and refer to the structural elements of poems (e.g., verse, rhythm, meter) and drama (e.g., casts of characters, settings, descriptions, dialogue, stage directions) when writing or speaking about a text.						
RL.4.6	Compare and contrast the point of view from which different stories are narrated, including the difference between first- and third-person narrations.						
Integration of Knowledge and Ideas (see below)							
Range of Reading and Level of Text Complexity							
RL.4.10	By the end of the year, read and comprehend literature, including stories, dramas, and poetry, in the grades 4–5 text complexity band proficiently, with scaffolding as needed at the high end of the range.						



CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage	
		2013		2014		2015		
Reading Informational Text								
Key Ideas and Details								
RI.4.1	Refer to details and examples in a text when explaining what the text says explicitly and when drawing inferences from the text.	34		R.04.E skills: 40% (20 items)		13-15		10-12
RI.4.2	Determine the main idea of a text and explain how it is supported by key details; summarize the text.							
RI.4.3	Explain events, procedures, ideas, or concepts in a historical, scientific, or technical text, including what happened and why, based on specific information in the text.							
				E.1 = 11 items E.2 = 0 items E.3 = 9 items				
Craft and Structure				R.04.F skills: 28% (14 items)  F.1 = 14 items F.2 = 0 items		Informational portion: approx. 42%		Informational portion: approx. 35%
RI.4.4	Determine the meaning of general academic and domain-specific words or phrases in a text relevant to a grade 4 topic or subject area.	0						
RI.4.5	Describe the overall structure (e.g., chronology, comparison, cause/effect, problem/solution) of events, ideas, concepts, or information in a text or part of a text.							
RI.4.6	Compare and contrast a firsthand and secondhand account of the same event or topic; describe the differences in focus and the information provided.							
Integration of Knowledge and Ideas (see below)								
Range of Reading and Level of Text Complexity								
RI.4.10	By the end of year, read and comprehend informational texts, including history/social studies, science, and technical texts, in the grades 4–5 text complexity band proficiently, with scaffolding as needed at the high end of the range.							

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Integration of Knowledge and Ideas*							
RL4.7	Make connections between the text of a story or drama and a visual or oral presentation of the text, identifying where each version reflects specific descriptions and directions in the text.		Captured in totals above		*Integration portion: approx. 2%		*Integration portion: approx. 14%
RL4.8	(Not applicable to literature)						
RL4.9	Compare and contrast the treatment of similar themes and topics (e.g., opposition of good and evil) and patterns of events (e.g., the quest) in stories, myths, and traditional literature from different cultures.						
RI4.7	Interpret information presented visually, orally, or quantitatively (e.g., in charts, graphs, diagrams, time lines, animations, or interactive elements on Web pages) and explain how the information contributes to an understanding of the text in which it appears.	0		0-2†		6-8	
RI4.8	Explain how an author uses reasons and evidence to support particular points in a text.						
RI4.9	Integrate information from two texts on the same topic in order to write or speak about the subject knowledgeably.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Language*							
L4.4	Determine or clarify the meaning of unknown and multiple-meaning words and phrases based on grade 4 reading and content, choosing flexibly from a range of strategies.	0	no direct alignment to Wyoming skills: 0%	6-8	*Language portion: approx. 14%	7-9	*Language portion: approx. 16%
L4.4.a	Use context (e.g., definitions, examples, or restatements in text) as a clue to the meaning of a word or phrase.						
L4.4.b	Use common, grade-appropriate Greek and Latin affixes and roots as clues to the meaning of a word (e.g., telegraph, photograph, autograph).						
L4.4.c	Consult reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation and determine or clarify the precise meaning of key words and phrases.						
L4.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.						
L4.5.a	Explain the meaning of simple similes and metaphors (e.g., as pretty as a picture) in context.						
L4.5.b	Recognize and explain the meaning of common idioms, adages, and proverbs.						
L4.5.c	Demonstrate understanding of words by relating them to their opposites (antonyms) and to words with similar but not identical meanings (synonyms).						
L4.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases, including those that signal precise actions, emotions, or states of being (e.g., quizzed, whined, stammered) and that are basic to a particular topic (e.g., wildlife, conservation, and endangered when discussing animal preservation).						
suggested total OP items on form:		50		50		50	

\*Integration and Language reporting categories have items associated with both literary and informational passages. The goal will be to strike an overall balance of approx. 50% per genre on the test form.

†Integration of Knowledge and Ideas will not report out in 2014.

Table A3. PAWS 2014 Grade 5 Reading Blueprint

**DRAFT Blueprint 2014 - 54 5R OP items aligned to CCSS**

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Reading Literature							
Key Ideas 0							
RL.5.1	Quote accurately from a text when explaining what the text says explicitly and when drawing inferences from the text.	19	R.05.N skills: 45% (24 items)	12-14	Literary portion: approx. 38%	11-13	Literary portion: approx. 35%
RL.5.2	Determine a theme of a story, drama, or poem from details in the text, including how characters in a story or drama respond to challenges or how the speaker in a poem reflects upon a topic; summarize the text.						
RL.5.3	Compare and contrast two or more characters, settings, or events in a story or drama, drawing on specific details in the text (e.g., how characters interact).						
Craft and Structure							
RL.5.4	Determine the meaning of words and phrases as they are used in a text, including figurative language such as metaphors and similes.	5	N.1 = 14 items N.2 = 10 items N.3 = 0 items	6-8		6-8	
RL.5.5	Explain how a series of chapters, scenes, or stanzas fits together to provide the overall structure of a particular story, drama, or poem.						
RL.5.6	Describe how a narrator's or speaker's point of view influences how events are described.						
Integration of Knowledge and Ideas (see below)							
Range of Reading and Level of Text Complexity							
RL.5.10	By the end of the year, read and comprehend literature, including stories, dramas, and poetry, in the grades 4–5 text complexity band proficiently, with scaffolding as needed at the high end of the range.						



CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Reading Informational Text							
Key Ideas and Details							
RI.5.1	Quote accurately from a text when explaining what the text says explicitly and when drawing inferences from the text.	30	R.05.E skills: 22% (12 items)	15-17	Informational portion: approx. 47%	12-14	Informational portion: approx. 36%
RI.5.2	Determine two or more main ideas of a text and explain how they are supported by key details; summarize the text.						
RI.5.3	Explain the relationships or interactions between two or more individuals, events, ideas, or concepts in a historical, scientific, or technical text based on specific information in the text.						
			E.1 = 7 items E.2 = 0 items E.3 = 5 items				
Craft and Structure							
RI.5.4	Determine the meaning of general academic and domain-specific words and phrases in a text relevant to a grade 5 topic or subject area.	0	R.05.F skills: 33% (18 items)  F.1 = 10 items F.2 = 8 item	7-9	Informational portion: approx. 47%	6-8	Informational portion: approx. 36%
RI.5.5	Compare and contrast the overall structure (e.g., chronology, comparison, cause/effect, problem/solution) of events, ideas, concepts, or information in two or more texts.						
RI.5.6	Analyze multiple accounts of the same event or topic, noting important similarities and differences in the point of view they represent.						
Integration of Knowledge and Ideas (see below)							
Range of Reading and Level of Text Complexity							
RI.5.10	By the end of year, read and comprehend informational texts, including history/social studies, science, and technical texts, in the grades 4–5 text complexity band proficiently, with scaffolding as needed at the high end of the range.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage	
		2013		2014		2015		
Integration of Knowledge and Ideas*								
RL5.7	Analyze how visual and multimedia elements contribute to the meaning, tone, or beauty of a text (e.g., graphic novel; multimedia presentation of fiction, folktale, myth, poem)		Captured in totals above		*Integration portion: approx. 2%		*Integration portion: approx. 13%	
RL5.8	(Not applicable to literature)							
RL5.9	Compare and contrast stories in the same genre (e.g., mysteries and adventure stories) on their approaches to similar themes and topics.							
RI5.7	Draw on information from multiple print or digital sources, demonstrating the ability to locate an answer to a question quickly or to solve a problem efficiently.	0				0-2†		6-8
RI5.8	Explain how an author uses reasons and evidence to support particular points in a text, identifying which reasons and evidence support which point(s).							
RI5.9	Integrate information from several texts on the same topic in order to write or speak about the subject knowledgeably.							

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Language*							
L5.4	Determine or clarify the meaning of unknown and multiple-meaning words and phrases based on grade 5 reading and content, choosing flexibly from a range of strategies.	0	no direct alignment to Wyoming skills: 0%	6-8	*Language portion: approx. 13%	8-10	*Language portion: approx. 16%
L5.4.a	Use context (e.g., cause/effect relationships and comparisons in text) as a clue to the meaning of a word or phrase.						
L5.4.b	Use common, grade-appropriate Greek and Latin affixes and roots as clues to the meaning of a word (e.g., photograph, photosynthesis).						
L5.4.c	Consult reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation and determine or clarify the precise meaning of key words and phrases.						
L5.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.						
L5.5.a	Interpret figurative language, including similes and metaphors, in context.						
L5.5.b	Recognize and explain the meaning of common idioms, adages, and proverbs.						
L5.5.c	Use the relationship between particular words (e.g., synonyms, antonyms, homographs) to better understand each of the words.						
L5.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases, including those that signal contrast, addition, and other logical relationships (e.g., however, although, nevertheless, similarly, moreover, in addition).						
suggested total OP items on form:		54		54		54	

\*Integration and Language reporting categories have items associated with both literary and informational passages. The goal will be to strike an overall balance of approx. 50% per genre on the test form.

†Integration of Knowledge and Ideas will not report out in 2014.



Table A4. PAWS 2014 Grade 6 Reading Blueprint

**DRAFT Blueprint 2014 - 56 6R OP items aligned to CCSS**

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Reading Literature							
Key Ideas and Details							
RL.6.1	Cite textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.	8	R.06.N skills: 41% (23 items)	13-15	Literary portion: approx. 43%	12-14	Literary portion: approx. 36%
RL.6.2	Determine a theme or central idea of a text and how it is conveyed through particular details; provide a summary of the text distinct from personal opinions or judgments.						
RL.6.3	Describe how a particular story's or drama's plot unfolds in a series of episodes as well as how the characters respond or change as the plot moves toward a resolution.						
Craft and Structure							
RL.6.4	Determine the meaning of words and phrases as they are used in a text, including figurative and connotative meanings; analyze the impact of a specific word choice on meaning and tone.	13	N.1 = 9 items N.2 = 7 items N.3 = 7 items	8-10		6-8	
RL.6.5	Analyze how a particular sentence, chapter, scene, or stanza fits into the overall structure of a text and contributes to the development of the theme, setting, or plot.						
RL.6.6	Explain how an author develops the point of view of the narrator or speaker in a text.						
Integration of Knowledge and Ideas (see below)							
Range of Reading and Level of Text Complexity							
RL.6.10	By the end of the year, read and comprehend literature, including stories, dramas, and poems, in the grades 6-8 text complexity band proficiently, with scaffolding as needed at the high end of the range.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Reading Informational Text			R.06.E skills: 25% (14 items)  E.1 = 6 items E.2 = 0 items E.3 = 8 items		Informational portion: approx. 43%		Informational portion: approx. 36%
Key Ideas and Details							
RI.6.1	Cite textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.	11					
RI.6.2	Determine a central idea of a text and how it is conveyed through particular details; provide a summary of the text distinct from personal opinions or judgments.						
RI.6.3	Analyze in detail how a key individual, event, or idea is introduced, illustrated, and elaborated in a text (e.g., through examples or anecdotes).						
Craft and Structure							
RI.6.4	Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings.	2	R.06.F skills: 34% (19 items)  F.1 = 8 items F.2 = 11 items	7-9		6-8	
RI.6.5	Analyze how a particular sentence, paragraph, chapter, or section fits into the overall structure of a text and contributes to the development of the ideas.						
RI.6.6	Determine an author's point of view or purpose in a text and explain how it is conveyed in the text.						
Integration of Knowledge and Ideas (see below)							
Range of Reading and Level of Text Complexity							
RI.6.10	By the end of the year, read and comprehend literary nonfiction in the grades 6–8 text complexity band proficiently, with scaffolding as needed at the high end of the range.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Integration of Knowledge and Ideas*							
RL6.7	Compare and contrast the experience of reading a story, drama, or poem to listening to or viewing an audio, video, or live version of the text, including contrasting what they "see" and "hear" when reading the text to what they perceive when they listen or watch.		Captured in totals above		*Integration portion: approx. 2%		*Integration portion: approx. 14%
RL6.8	(Not applicable to literature)						
RL6.9	Compare and contrast texts in different forms or genres (e.g., stories and poems; historical novels and fantasy stories) in terms of their approaches to similar themes and topics.						
RI6.7	Integrate information presented in different media or formats (e.g., visually, quantitatively) as well as in words to develop a coherent understanding of a topic or issue.	22		0-2†		7-9	
RI6.8	Trace and evaluate the argument and specific claims in a text, distinguishing claims that are supported by reasons and evidence from claims that are not.						
RI6.9	Compare and contrast one author's presentation of events with that of another (e.g., a memoir written by and a biography on the same person).						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Language*							
L6.4	Determine or clarify the meaning of unknown and multiple-meaning words and phrases based on grade 6 reading and content, choosing flexibly from a range of strategies.	0	no direct alignment to Wyoming skills: 0%	6-8	*Language portion: approx. 13%	7-9	*Language portion: approx. 14%
L6.4.a	Use context (e.g., the overall meaning of a sentence or paragraph; a word's position or function in a sentence) as a clue to the meaning of a word or phrase.						
L6.4.b	Use common, grade-appropriate Greek or Latin affixes and roots as clues to the meaning of a word (e.g., audience, auditory, audible).						
L6.4.c	Consult reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation of a word or determine or clarify its precise meaning or its part of speech.						
L6.4.d	Verify the preliminary determination of the meaning of a word or phrase (e.g., by checking the inferred meaning in context or in a dictionary).						
L6.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.						
L6.5.a	Interpret figures of speech (e.g., personification) in context.						
L6.5.b	Use the relationship between particular words (e.g., cause/effect, part/whole, item/category) to better understand each of the words.						
L6.5.c	Distinguish among the connotations (associations) of words with similar denotations (definitions) (e.g., stingy, scrimping, economical, unwhasteful, thrifty).						
L6.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases; gather vocabulary knowledge when considering a word or phrase important to comprehension or expression.						
	suggested total OP items on form:	56		56		56	

\*Integration and Language reporting categories have items associated with both literary and informational passages.

The goal will be to strike an overall balance of approx. 50% per genre on the test form.

†Integration of Knowledge and Ideas will not report out in 2014.



Table A5. PAWS 2014 Grade 7 Reading Blueprint

**DRAFT Blueprint 2014 - 56 7R OP items aligned to CCSS**

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage	
		2013		2014		2015		
Reading Literature								
Key Ideas and Details								
RL.7.1	Cite several pieces of textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.	13	R.07.N skills: 36% (20 items)  N.1 = 7 items N.2 = 8 items N.3 = 5 items	11-13	Literary portion: approx. 37%	9-11	Literary portion: approx. 32%	
RL.7.2	Determine a theme or central idea of a text and analyze its development over the course of the text; provide an objective summary of the text.							
RL.7.3	Analyze how particular elements of a story or drama interact (e.g., how setting shapes the characters or plot).							
Craft and Structure								
RL.7.4	Determine the meaning of words and phrases as they are used in a text, including figurative and connotative meanings; analyze the impact of rhymes and other repetitions of sounds (e.g., alliteration) on a specific verse or stanza of a poem or section of a story or drama.	7		7-9		7-9		
RL.7.5	Analyze how a drama's or poem's form or structure (e.g., soliloquy, sonnet) contributes to its meaning.							
RL.7.6	Analyze how an author develops and contrasts the points of view of different characters or narrators in a text.							
Integration of Knowledge and Ideas (see below)								
Range of Reading and Level of Text Complexity								
RL.7.10	By the end of the year, read and comprehend literature, including stories, dramas, and poems, in the grades 6–8 text complexity band proficiently, with scaffolding as needed at the high end of the range.							

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
			2013	2014		2015	
Reading Informational Text							
Key Ideas and Details							
RI.7.1	Cite several pieces of textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.	9	R.07.E skills: 39% (22 items)	17-19	Informational portion: approx. 49%	14-16	Informational portion: approx. 39%
RI.7.2	Determine two or more central ideas in a text and analyze their development over the course of the text; provide an objective summary of the text.						
RI.7.3	Analyze the interactions between individuals, events, and ideas in a text (e.g., how ideas influence individuals or events, or how individuals influence ideas or events).						
			E.1 = 9 items E.2 = 7 items E.3 = 6 items				
Craft and Structure							
RI.7.4	Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings; analyze the impact of a specific word choice on meaning and tone.	18	R.07.F skills: 25% (14 items)  F.1 = 6 items F.2 = 8 items	7-9	Informational portion: approx. 49%	6-8	Informational portion: approx. 39%
RI.7.5	Analyze the structure an author uses to organize a text, including how the major sections contribute to the whole and to the development of the ideas.						
RI.7.6	Determine an author's point of view or purpose in a text and analyze how the author distinguishes his or her position from that of others.						
Integration of Knowledge and Ideas (see below)							
Range of Reading and Level of Text Complexity							
RI.7.10	By the end of the year, read and comprehend literary nonfiction in the grades 6-8 text complexity band proficiently, with scaffolding as needed at the high end of the range.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Integration of Knowledge and Ideas*							
RL7.7	Compare and contrast a written story, drama, or poem to its audio, filmed, staged, or multimedia version, analyzing the effects of techniques unique to each medium (e.g., lighting, sound, color, or camera focus and angles in a film).		Captured in totals above		*Integration portion: approx. 2%		*Integration portion: approx. 13%
RL7.8	[Not applicable to literature]						
RL7.9	Compare and contrast a fictional portrayal of a time, place, or character and a historical account of the same period as a means of understanding how authors of fiction use or alter history.						
RI7.7	Compare and contrast a text to an audio, video, or multimedia version of the text, analyzing each medium's portrayal of the subject (e.g., how the delivery of a speech affects the impact of the words).	9		0-2†		6-8	
RI7.8	Trace and evaluate the argument and specific claims in a text, assessing whether the reasoning is sound and the evidence is relevant and sufficient to support the claims.						
RI7.9	Analyze how two or more authors writing about the same topic shape their presentations of key information by emphasizing different evidence or advancing different interpretations of facts.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Language*							
L7.4.a	Use context (e.g., the overall meaning of a sentence or paragraph; a word's position or function in a sentence) as a clue to the meaning of a word or phrase.	0	no direct alignment to Wyoming skills: 0%	6-8	*Language portion: approx. 12%	8-10	*Language portion: approx. 16%
L7.4.b	Use common, grade-appropriate Greek or Latin affixes and roots as clues to the meaning of a word (e.g., belligerent, bellicose, rebel).						
L7.4.c	Consult general and specialized reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation of a word or determine or clarify its precise meaning or its part of speech.						
L7.4.d	Verify the preliminary determination of the meaning of a word or phrase (e.g., by checking the inferred meaning in context or in a dictionary).						
L7.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.						
L7.5.a	Interpret figures of speech (e.g., literary, biblical, and mythological allusions) in context.						
L7.5.b	Use the relationship between particular words (e.g., synonym/antonym, analogy) to better understand each of the words.						
L7.5.c	Distinguish among the connotations (associations) of words with similar denotations (definitions) (e.g., refined, respectful, polite, diplomatic, condescending).						
L7.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases; gather vocabulary knowledge when considering a word or phrase important to comprehension or expression.						
suggested total OP items on form:		56		56		56	

\*Integration and Language reporting categories have items associated with both literary and informational passages.

The goal will be to strike an overall balance of approx. 50% per genre on the test form.

\*Integration of Knowledge and Ideas will not report out in 2014.



Table A6. PAWS 2014 Grade 8 Reading Blueprint

**DRAFT Blueprint 2014 - 56 8R OP items aligned to CCSS**

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage			
		2013		2014		2015				
Reading Literature										
Key Ideas and Details										
RL.1	Cite the textual evidence that most strongly supports an analysis of what the text says explicitly as well as inferences drawn from the text.	10	R.08.N skills: 29% (16 items)  N.1 = 4 items N.2 = 6 items N.3 = 6 items	10-12	Literary portion: approx. 35%	9-11	Literary portion: approx. 32%			
RL.2	Determine a theme or central idea of a text and analyze its development over the course of the text, including its relationship to the characters, setting, and plot; provide an objective summary of the text.									
RL.3	Analyze how particular lines of dialogue or incidents in a story or drama propel the action, reveal aspects of a character, or provoke a decision.									
Craft and Structure										
RL.4	Determine the meaning of words and phrases as they are used in a text, including figurative and connotative meanings; analyze the impact of specific word choices on meaning and tone, including analogies or allusions to other texts.	6		7-9		7-9				
RL.5	Compare and contrast the structure of two or more texts and analyze how the differing structure of each text contributes to its meaning and style.									
RL.6	Analyze how differences in the points of view of the characters and the audience or reader (e.g., created through the use of dramatic irony) create such effects as suspense or humor.									
Integration of Knowledge and Ideas (see below)										
Range of Reading and Level of Text Complexity										
RL.10	By the end of the year, read and comprehend literature, including stories, dramas, and poems, at the high end of grades 6–8 text complexity band independently and proficiently.									

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Reading Informational Text							
Key Ideas and Details							
R.18.1	Cite the textual evidence that most strongly supports an analysis of what the text says explicitly as well as inferences drawn from the text.	29	R.04.E skills: 46% (26 items)  E.1 = 12 items E.2 = 9 items E.3 = 5 items  R.04.F skills: 25% (14 items)  F.1 = 7 items F.2 = 7 items	18-20	Informational portion: approx. 50%	12-14	Informational portion: approx. 39%
R.18.2	Determine a central idea of a text and analyze its development over the course of the text, including its relationship to supporting ideas; provide an objective summary of the text.						
R.18.3	Analyze how a text makes connections among and distinctions between individuals, ideas, or events (e.g., through comparisons, analogies, or categories).						
Craft and Structure							
R.18.4	Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings; analyze the impact of specific word choices on meaning and tone, including analogies or allusions to other texts.	5		7-9		8-10	
R.18.5	Analyze in detail the structure of a specific paragraph in a text, including the role of particular sentences in developing and refining a key concept.						
R.18.6	Determine an author's point of view or purpose in a text and analyze how the author acknowledges and responds to conflicting evidence or viewpoints.						
Integration of Knowledge and Ideas (see below)							
Range of Reading and Level of Text Complexity							
R.18.10	By the end of the year, read and comprehend literary nonfiction at the high end of the grades 6–8 text complexity band independently and proficiently.						

CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Integration of Knowledge and Ideas*							
RL.8.7	Analyze the extent to which a filmed or live production of a story or drama stays faithful to or departs from the text or script, evaluating the choices made by the director or actors.		Captured in totals above		*Integration portion: approx. 2%		*Integration portion: approx. 13%
RL.8.8	(Not applicable to literature)						
RL.8.9	Analyze how a modern work of fiction draws on themes, patterns of events, or character types from myths, traditional stories, or religious works such as the Bible, including describing how the material is rendered new.						
RI.8.7	Evaluate the advantages and disadvantages of using different mediums (e.g., print or digital text, video, multimedia) to present a particular topic or idea.						
RI.8.8	Delineate and evaluate the argument and specific claims in a text, assessing whether the reasoning is sound and the evidence is relevant and sufficient; recognize when irrelevant evidence is introduced.	6		0-2†			
RI.8.9	Analyze a case in which two or more texts provide conflicting information on the same topic and identify where the texts disagree on matters of fact or interpretation.						



CCSS code	CCSS text	Current CCSS alignment	Current WY skills alignment	Ideal blueprint	Content coverage	Ideal blueprint	Content coverage
		2013		2014		2015	
Language*							
L8.4.a	Use context (e.g., the overall meaning of a sentence or paragraph; a word's position or function in a sentence) as a clue to the meaning of a word or phrase.	0	no direct alignment to Wyoming skills: 0%	6-8	*Language portion: approx. 13%	8-10	*Language portion: approx. 16%
L8.4.b	Use common, grade-appropriate Greek or Latin affixes and roots as clues to the meaning of a word (e.g., precede, recede, secede).						
L8.4.c	Consult general and specialized reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation of a word or determine or clarify its precise meaning or its part of speech.						
L8.4.d	Verify the preliminary determination of the meaning of a word or phrase (e.g., by checking the inferred meaning in context or in a dictionary).						
L8.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.						
L8.5.a	Interpret figures of speech (e.g. verbal irony, puns) in context.						
L8.5.b	Use the relationship between particular words to better understand each of the words.						
L8.5.c	Distinguish among the connotations (associations) of words with similar denotations (definitions) (e.g., bullheaded, willful, firm, persistent, resolute).						
L8.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases; gather vocabulary knowledge when considering a word or phrase important to comprehension or expression.						
suggested total OP items on form:		56		56		56	

\*Integration and Language reporting categories have items associated with both literary and informational passages. The goal will be to strike an overall balance of approx. 50% per genre on the test form.

\*Integration of Knowledge and Ideas will not report out in 2014.

Mathematics

Table A7. PAWS 2014 Grade 3 Math Blueprint

Domain / Standard	2014 PAWS - 3rd Grade Mathematics Cluster Heading	Focus m = major s = supporting a = additional	Items Per Domain	# of Items / Cluster Heading	PAWS Emphasis
Operations and Algebraic Thinking			20		
3.OA.1	Represent and solve problems involving multiplication and division.	m		6-8	40%
3.OA.2					
3.OA.3					
3.OA.4					
3.OA.5	Understand properties of multiplication and the relationship between multiplication and division.	m		3-5	
3.OA.6					
3.OA.7	Multiply and divide within 100.	m		3-5	
3.OA.8	Solve problems involving the four operations, and identify and explain patterns in arithmetic.	m		4-6	
3.OA.9					
Number and Operations - Base Ten			6		
3.NBT.1	Use place value understanding and properties of operations to perform multi-digit arithmetic.	a		6	12%
3.NBT.2					
3.NBT.3					
Number and Operations - Fractions			6		
3.NF.1	Develop understanding of fractions as numbers.	m		6	12%
3.NF.2					
3.NF.3					
Measurement and Data			12		
3.MD.1	Solve problems involving measurement and estimation of intervals of time, liquid volumes, and masses of objects.	m		3-5	24%
3.MD.2					
3.MD.3	Represent and interpret data.	s		1-3	
3.MD.4					
3.MD.5	Geometric measurement: understand concepts of area and relate area to multiplication and to addition.	m		3-5	
3.MD.6					
3.MD.7					
3.MD.8	Geometric measurement: recognize perimeter as an attribute of plane figures and distinguish between linear and area measures.	a		1-3	
Geometry			6		
3.G.1	Reason with shapes and their attributes.	s		6	12%
3.G.2					
			50	50	100%

Table A8. PAWS 2014 Grade 4 Math Blueprint

Domain / Standard	2014 PAWS - 4th Grade Mathematics Cluster Heading	Focus m = major s = supporting a = additional	Items Per Domain	# of Items / Cluster Heading	PAWS Emphasis
Operations and Algebraic Thinking			13		
4.OA.1	Use the four operations with whole numbers to solve problems.  Gain familiarity with factors and multiples. Generate and analyze patterns.	m		6-8	22%
4.OA.2					
4.OA.3					
4.OA.4		s	2-4		
4.OA.5		a	2-4		
Number and Operations - Base Ten			10		
4.NBT.1	Generalize place value understanding for multi-digit whole numbers.	m		3-5	17%
4.NBT.2					
4.NBT.3					
4.NBT.4	Use place value understanding and properties of operations to perform multi-digit arithmetic.	m		5-7	
4.NBT.5					
4.NBT.6					
Number and Operations - Fractions			20		
4.NF.1	Extend understanding of fraction equivalence and ordering.	m		3-5	34%
4.NF.2					
4.NF.3	Build fractions from unit fractions by applying and extending previous understandings of operations on whole numbers.	m		9-11	
4.NF.4					
4.NF.5	Understand decimal notation for fractions, and compare decimal fractions.	m		5-7	
4.NF.6					
4.NF.7					
Measurement and Data			10		
4.MD.1	Solve problems involving measurement and conversion of measurements from a larger unit to a smaller unit.	s		5-7	17%
4.MD.2					
4.MD.3					
4.MD.4	Represent and interpret data.	s		1-3	
4.MD.5	Geometric measurement: understand concepts of angle and measure angles.	a		1-3	
4.MD.6					
4.MD.7					
Geometry			6		
4.G.1	Draw and identify lines and angles, and classify shapes by properties of their lines and angles.	a		6	10%
4.G.2					
4.G.3					
			59	59	100%

Released May 2013

PAWS Mathematics Blueprint

Wyoming Department of Education



Table A9. PAWS 2014 Grade 5 Math Blueprint







Domain / Standard	2014 PAWS - 5th Grade Mathematics Cluster Heading	Focus m = major s = supporting a = additional	Items Per Domain	# of Items / Cluster Heading	PAWS Emphasis
Operations and Algebraic Thinking			6		
5.OA.1	Write and interpret numerical expressions.	a		2-4	10%
5.OA.2					
5.OA.3	Analyze patterns and relationships.	a		2-4	
Number and Operations - Base Ten			16		
5.NBT.1	Understand the place value system.	m		5-7	27%
5.NBT.2					
5.NBT.3					
5.NBT.4					
5.NBT.5	Perform operations with multi-digit whole numbers and with decimals to hundredths.	m		9-11	
5.NBT.6					
5.NBT.7					
Number and Operations - Fractions			19		
5.NF.1	Use equivalent fractions as a strategy to add and subtract fractions.	m		7-9	32%
5.NF.2					
5.NF.3	Apply and extend previous understandings of multiplication and division to multiply and divide fractions.	m		10-12	
5.NF.4					
5.NF.5					
5.NF.6					
5.NF.7					
Measurement and Data			12		
5.MD.1	Convert like measurement units within a given measurement system.	s		2-4	20%
5.MD.2	Represent and interpret data.	s		1-3	
5.MD.3	Geometric measurement: understand concepts of volume and relate volume to multiplication and to addition.	m		6-8	
5.MD.4					
5.MD.5					
Geometry			6		
5.G.1	Graph points on the coordinate plane to solve real-world and mathematical problems.	a		2-4	10%
5.G.2					
5.G.3	Classify two-dimensional figures into categories based on their properties.	a		2-4	
5.G.4					
			59	59	100%

Released May 2013

PAWS Mathematics Blueprint

Wyoming Department of Education

Table A10. PAWS 2014 Grade 6 Math Blueprint

Domain / Standard	2014 PAWS - 6th Grade Mathematics Cluster Heading	Focus	Items Per Domain	# of Items / Cluster Heading	PAWS Emphasis
Ratios and Proportional Relationships			10		
6.RP.1	Understand ratio concepts and use ratio reasoning to solve problems.	major		10	17%
6.RP.2					
6.RP.3					
The Number System			15		
6.NS.1	Apply and extend previous understandings of multiplication and division to divide fractions by fractions.	major		4-6	25%
6.NS.2	Compute fluently with multi-digit numbers and find common factors.	additional		3-5	
6.NS.3					
6.NS.4					
6.NS.5	Apply and extend previous understandings of numbers to the system of rational numbers.	major		5-7	
6.NS.6					
6.NS.7					
6.NS.8					
Expressions and Equations			20		
6.EE.1	Apply and extend previous understandings of arithmetic to algebraic expressions.	major		7-9	34%
6.EE.2					
6.EE.3					
6.EE.4					
6.EE.5	Reason about and solve one-variable equations and inequalities.	major		6-8	
6.EE.6					
6.EE.7					
6.EE.8					
6.EE.9	Represent and analyze quantitative relationships between dependent and independent variables.	major		4-6	
Geometry			6		
6.G.1	Solve real-world and mathematical problems involving area, surface area, and volume.	supporting		6	10%
6.G.2					
6.G.3					
6.G.4					
Statistics and Probability			8		
6.SP.1	Develop understanding of statistical variability.	additional		2-4	14%
6.SP.2					
6.SP.3					
6.SP.4	Summarize and describe distributions.	additional		4-6	
6.SP.5					
			59	59	100%

Released May 2013

PAWS Mathematics Blueprint








Wyoming Department of Education

Released May 2013

PAWS Mathematics Blueprint

Wyoming Department of Education

Table A11. PAWS 2014 Grade 7 Math Blueprint

Domain / Standard	2014 PAWS - 7th Grade Mathematics Cluster Heading	Focus m = major s = supporting a = additional	Items Per Domain	# of Items / Cluster Heading	PAWS Emphasis
Ratios and Proportional Relationships			13		
7.RP.1	Analyze proportional relationships and use them to solve real-world and mathematical problems.	m		13	22%
7.RP.2					
7.RP.3					
The Number System			10		
7.NS.1	Apply and extend previous understandings of operations with fractions to add, subtract, multiply, and divide rational numbers.	m		10	17%
7.NS.2					
7.NS.3					
Expressions and Equations			18		
7.EE.1	Use properties of operations to generate equivalent expressions.	m		5-7	31%
7.EE.2					
7.EE.3	Solve real-life and mathematical problems using numerical and algebraic expressions and equations.	m		11-13	
7.EE.4					
Geometry			9		
7.G.1	Draw, construct, and describe geometrical figures and describe the relationships between them.	a		1-3	15%
7.G.2					
7.G.3					
7.G.4	Solve real-life and mathematical problems involving angle measure, area, surface area, and volume.	a		6-8	
7.G.5					
7.G.6					
Statistics and Probability			9		
7.SP.1	Use random sampling to draw inferences about a population.	s		2-4	15%
7.SP.2					
7.SP.3	Draw informal comparative inferences about two populations.	a		1-3	
7.SP.4					
7.SP.5	Investigate chance processes and develop, use, and evaluate probability models.	s		3-5	
7.SP.6					
7.SP.7					
7.SP.8					
			59	59	100.00%








Released May 2013

PAWS Mathematics Blueprint

Wyoming Department of Education



Table A12. PAWS 2014 Grade 8 Math Blueprint

Domain / Standard	2014 PAWS - 8th Grade Mathematics Cluster Heading	Focus m = major s = supporting a = additional	Items Per Domain	# of Items / Cluster Heading	PAWS Emphasis
The Number System			6		
8.NS.1	Know that there are numbers that are not rational, and approximate them by rational numbers.	supporting		6	9%
8.NS.2					
Expressions and Equations			23		
8.EE.1	Work with radicals and integer exponents.	major		6-8	35%
8.EE.2					
8.EE.3					
8.EE.4					
8.EE.5	Understand the connections between proportional relationships, lines, and linear equations.	major		5-7	
8.EE.6	Analyze and solve linear equations and pairs of simultaneous linear equations.	major		9-11	
8.EE.7					
8.EE.8					
Functions			14		
8.F.1	Define, evaluate, and compare functions.	major		7-9	22%
8.F.2					
8.F.3					
8.F.4	Use functions to model relationships between quantities.	supporting		5-7	
8.F.5					
Geometry			16		
8.G.1	Understand congruence and similarity using physical models, transparencies, or geometry software.	major		6-8	25%
8.G.2					
8.G.3					
8.G.4					
8.G.5					
8.G.6	Understand and apply the Pythagorean Theorem.	major		5-7	
8.G.7					
8.G.8					
8.G.9	Solve real-world and mathematical problems involving volume of cylinders, cones, and spheres.	additional		2-4	
Statistics and Probability			6		
8.SP.1	Investigate patterns of association in bivariate data.	supporting		6	9%
8.SP.2					
8.SP.3					
8.SP.4					
			65	65	100.00%

Released May 2013

PAWS Mathematics Blueprint

Wyoming Department of Education

Released May 2013

PAWS Mathematics Blueprint

Wyoming Department of Education

# Science

Table A13. PAWS 2014 Science Grade 4 Blueprint

Standards	Total Points per Standard	Skills	Branches	Benchmarks	Total # of Items per Benchmark (MC = 1pt. each)	Items Aligning Skill & Benchmark	of Items per Branches (MC Items = 1)	Age of Test Items per Branch			
Concepts & Processes	50	I.1 Observe and Question	Life Science	4.1.1 Characteristics of Organisms: Students describe observable characteristics of living things, including structures that serve specific functions and everyday behaviors.	5-6	I.1 - 0-1 I.2 - 1-2 I.3 - 0-1 I.4 - 0-1	16	32.00%			
				4.1.2 Life Cycles of Organisms: Students sequence life cycles of living things, and recognize that plants and animals resemble their parents.	5-6	I.1 - 0-1 I.2 - 1-2 I.3 - 0-1 I.4 - 1-2					
				4.1.3 Organisms and Their Environments: Students show connections between living things, their basic needs, and the environments.	4-5	I.1 - 0-1 I.2 - 1-2 I.3 - 0-1 I.4 - 0-1					
		I.2 Design and Conduct a Scientific Investigation		Earth and Space Science	4.1.4 Properties of Earth Materials: Students investigate water, air, rocks, and soils to compare basic properties of earth materials.	4-5			I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 1-2	16	32.00%
					4.1.5 Objects in the Sky: Students describe observable objects in the sky and their patterns of movement.	5-6			I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 1-2		
					4.1.6 Changes in Earth and Sky: Students describe observable changes in earth and sky, including rapid and gradual changes to the earth's surface, and daily and seasonal changes in the weather.	5-6			I.1 - 0-1 I.2 - 1-2 I.3 - 1-2 I.4 - 1-2		
		I.3 Organize and Represent Data	Physical Science		4.1.7 Properties of Objects: Students classify objects by properties that can be observed, measured, and recorded, including color, shape, size, weight, volume, texture, and temperature.	4-5	I.1 - 0-1 I.2 - 1-2 I.3 - 0-1 I.4 - 0-1	18	36.00%		
					4.1.8 Changes in States of Matter: Students demonstrate that the processes of heating and cooling can change matter from one state to another.	4-5	I.1 - 1-2 I.2 - 0-1 I.3 - 0-1 I.4 - 1-2				
					4.1.9 Physical Phenomena: Students investigate physical phenomena commonly encountered in daily life, including light, heat, electricity, sound, and magnetism.	4-5	I.1 - 0-1 I.2 - 1-2 I.3 - 0-1 I.4 - 1-2				
		I.4 Draw Conclusions and Make Connections		Physical Science	4.1.10 Position and Motion of Objects: Students demonstrate that pushing and pulling can change the position and motion of objects.	4-5	I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 0-1				



<b>Science as Inquiry</b>			4.2.1 Students research answers to science questions and present findings through appropriate means.	<b>Not Assessed</b>
			4.2.2 Students use the inquiry process to conduct simple scientific investigations: 1) Collect and organize data; 2) Use data to construct simple graphs, charts, diagrams, and/or model; 3) Draw conclusions and accurately communicate results, making connections to daily life; 4) Pose or identify questions and make predictions; and 5) Conduct investigations to answer questions and check predictions	<b>Assessed with Concepts &amp; Processes</b>
			4.2.3 Students identify and use appropriate scientific equipment.	
			4.2.4 Students properly use safety equipment and recognize hazards and safety symbols while practicing standard safety procedures.	

Table A14. PAWS 2014 Science Grade 8 Blueprint

Standards	Total Points per Standard	Skills	Branches	Benchmarks	Total # of Items per Benchmark (MC = 1pt. each)	# of Items Aligning Skill & Benchmark	Total # of Items per Branch (MC Items = 1 pt. each)	Percentage of Test Items per Branch (%)
Concepts and Processes	50	I.1 Observe and Question  I.2 Design and Conduct a Scientific Investigation  I.3 Organize and Represent Data  I.4 Draw Conclusions and Make Connections	Life Science	<u>8.1.1 Levels of Organization in Living Systems:</u> Students model the cell as the basic unit of a living system. They realize that all functions that sustain life act within a single cell and cells differentiate into specialized cells, tissues, organs, and organ systems.	2-3	I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 0-1	16	32%
				<u>8.1.2 Reproduction and Heredity:</u> Students describe reproduction as a characteristic of all living systems, which is essential to the continuation of species, and identify and interpret traits, patterns of inheritance, and the interaction between genetics and environment.		I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 0-1		
				<u>8.1.3 Evolution as a Theory:</u> Students explain evolution as a theory and apply the theory to the diversity of species, which results from natural selection and the acquisition of unique characteristics through biological adaptation.		I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 0-1		
				<u>8.1.4 Diversity of Organisms:</u> Students investigate the interconnectedness of organisms, identifying similarity and diversity of organisms through a classification system of hierarchical relationships and structural homologies.		I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 0-1		
				<u>8.1.5 Behavior and Adaptation:</u> Students recognize behavior as a response of an organism to an internal or environmental stimulus and connect the characteristics and behaviors of an organism to biological adaptation.	2-3	I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 0-1		
Concepts and Processes	50	I.1 Observe and Question  I.2 Design and Conduct a Scientific Investigation  I.3 Organize and Represent Data  I.4 Draw Conclusions and Make Connections	Earth and Space Science	<u>8.1.6 Interrelationships of Populations and Ecosystems:</u> Students illustrate populations of organisms and their interconnection within an ecosystem, identifying relationships among producers, consumers, and	2-3	I.1 - 0-1 I.2 - 0-1 I.3 - 0-1 I.4 - 0-1	16	32.00%
				<u>8.1.7 The Earth in the Solar System:</u> Students describe Earth as the third planet in the Solar System and understand the effects of the sun as a major source of energy, gravitational forces, and motions of objects in the Solar System.		I.1 - 0-1 I.2 - 1-2 I.3 - 0-1 I.4 - 1-2		
				<u>8.1.8 The Structure of the Earth System:</u> Students examine the structure of the Earth, identifying layers of the Earth, considering plate movement and its effect, and recognizing landforms resulting from constructive and destructive forces.		I.1 - 1-2 I.2 - 1-2 I.3 - 0-1 I.4 - 1-2		
				<u>8.1.9 The Earth's History:</u> Students systematize the Earth's history in terms of geologic evidence, comparing past and present Earth processes and identifying catastrophic events and fossil evidence.		I.1 - 0-1 I.2 - 0-1 I.3 - 1-2 I.4 - 1-2		
					5-6			
					5-6			
					5-6			

		I.1 Observe and Question	I.2 Design and Conduct a Scientific Investigation	I.3 Organize and Represent Data	I.4 Draw Conclusions and Make Connections	8.1.10 The Structure and Properties of Matter: Students identify characteristic properties of matter such as density, solubility, and boiling point and understand that elements are the basic components of matter.	3-4	1.1 - 0-1 1.2 - 0-1 1.3 - 0-1 1.4 - 0-1	18	36.00%

Science as Inquiry						8.2.1 Students research answers to science questions and present findings through appropriate means.	4	Assessed with Concepts & Processes
						8.2.2 Students use the inquiry to conduct scientific investigations: 1) Ask questions that lead to conducting an investigation; 2) Collect, organize, and analyze and appropriately represent data; 3) Draw conclusions based on evidence and make connections to applied scientific concepts; 4) Clearly and accurately communicate the result of the investigations.		
						8.2.3 Students clearly and accurately communicate the result of their own work, as well as information obtained from other sources.		
						8.2.4 Students recognize the relationship between science and technology in meeting human needs.		
						8.2.5 Students properly use appropriate scientific and safety equipment, recognize hazards and safety symbols, and observe standard safety procedures.		

# SAWS

Table A15. SAWS 2014 Writing Grade 3 Blueprint

3W 2014									
Operational Items									
Standard	Item Type		Trait	Total Points per Standard	Total Points per Trait	Percentage of Test Items per Mode	Point Distribution		
	Text-based	Stand-alone					Short Response (4 pt)	Extended Response (8 pt)	Writing Prompt (12 pts)
W3.2 Informative/ Explanatory Writing		✓	Item Development	12	3	50%			✓
			Organization		3				
			Voice		3				
			Conventions		3				
W3.3 Narrative Writing*		✓	Item Development	12	3	50%			✓
			Organization		3				
			Voice		3				
			Conventions		3				
Total Points:				24					

\* The W3.3 Narrative Writing OP prompt from the 2013 admin will be re-seeded in the 2014 OP form.

Table A16. SAWS 2014 Writing Grade 5 Blueprint

5W 2014									
Operational Items									
Standard	Item Type		Trait	Total Points per Standard	Total Points per Trait	Percentage of Test Items per Mode	Point Distribution		
	Passage-based	Stand-alone					Short Response (4 pt)	Extended Response (8 pt)	Writing Prompt (12 pts)
W5.2 Informative/ Explanatory Writing		✓	Item Development	12	3	50%			✓
			Organization		3				
			Voice		3				
			Conventions		3				
W5.8 Response to Literary or Informational Text Writing	✓		holistically scored	4 + 8	holistically scored	50%		✓	
Total Points:				24					

Table A17. SAWS 2014 Writing Grade 7 Blueprint

7W 2014									
Operational Items									
Standard	Item Type		Trait	Total Points per Standard	Total Points per Trait	Percentage of Test Items per Mode	Point Distribution		
	Passage-based	Stand-alone					Short Response (4 pt)	Extended Response (8 pt)	Writing Prompt (12 pts)
W7.2 Informative/ Explanatory Writing		✓	Item Development	12	3	50%			✓
			Organization		3				
			Voice		3				
			Conventions		3				
W7.9 Response to Literary or Informational Text Writing	✓		holistically scored	4 + 8	holistically scored	50%	✓	✓	
							✓		
Total Points:				24					

Appendix B: Sample PAWS/SAWS Student Reports: Grade 3 Reading,  
Mathematics, and SAWS (exemplar for Grades 5 and 7)





## DEPARTMENT OF EDUCATION

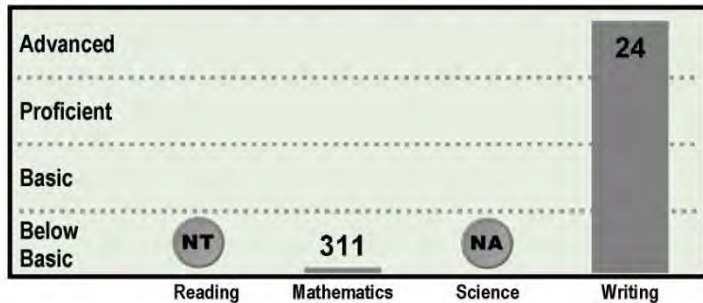
# Student Report

**First Name:** CHECK  
**Middle Initial:** J  
**Last Name:** AUSTIN

**Grade:** 3  
**Birthdate:** 01/19/2005  
**Student ID:** 23931014  
**Test Window:** 10/21/13 - 06/30/14  
**School:** Laura Irwin Elementary

**District:** Big Horn CSD # 4

	Score Ranges				Your Student	State Percentile Rank
	Below Basic	Basic	Proficient	Advanced		
Reading	300-519	520-583	584-660	661-975	NT	NT
Mathematics	300-557	558-599	600-679	680-975	311	11%
Science	NA	NA	NA	NA	NA	NA
Writing	0-7	8-15	16-21	22-24	24	99%



Lexile® Measure	---	Quantile® Measure	310Q
-----------------	-----	-------------------	------

### Purpose of Report

This report provides information about your child's achievement on the Proficiency Assessment for Wyoming Students (PAWS) and the Student Assessment of Writing Skills (SAWS). This report will help you understand your child's performance in reading and math (grades 3-8), science (grades 4 and 8), and writing (grades 3, 5, and 7) for the 2013-2014 school year.

### Glossary of Terms

**Scale Score:** Your child's raw score (total number of points earned) transformed into a score on a scale.

**Score Ranges:** The range of scale score points assigned to each proficiency level. Each subject area has a unique range of points.

**Domain Performance:** Describes your child's performance in sub-categories (domains) of each content area. The Scale Score column indicates your child's performance in relation to the overall reading scale (above). The Domain Performance column shows your child's percent correct in each of the measured domains.

**State Percentile Rank:** State Percentile Rank indicates your child's performance in relation to other Wyoming students in the same grade. The percentile shows the percentage of Wyoming students in the same grade who obtained scores equal to or less than your child's score.

**Lexile Measure:** Helps readers select materials at their reading level. This can serve as a guide in selecting books for your child.

**Quantile Measure:** Similar to the Lexile and can help you identify math activities to do at home. These activities will help your child practice mathematical skills leading to increased mathematical understanding.

**Not Tested (NT):** Your child did not take this part of the assessment.

**Not Applicable (NA):** This content area is not tested at this grade level.

**For Additional Resources and Information-**  
**Visit the Wyoming Department of Education online**

Go to the Wyoming Department of Education's website at [edu.wyoming.gov](http://edu.wyoming.gov) for more information about the PAWS and SAWS Assessments.

### CONTENT PERFORMANCE BY DOMAIN

PAWS		Scale Score	Domain Performance (% Correct)
Reading	Literature: Craft and Structure	NT	NT
	Literature: Key Ideas and Details		
	Informational Text: Craft and Structure		
	Informational Text: Key Ideas and Details		
Math	Language	401	17%
	Number Operation – Base 10		
	Geometry		
	Number Operations - Fractions		
	Operations and Algebraic Thinking		
Science	Measurement & Data	702	17%
	NA		
SAWS		Total Possible	Student Score
Writing	<b>Narrative Total</b>	<b>12</b>	<b>12</b>
	Idea Development	3	3
	Organization	3	3
	Voice	3	3
	Conventions	3	3
	<b>Informative/Explanatory Total</b>	<b>12</b>	<b>12</b>
	Idea Development	3	3
	Organization	3	3
	Voice	3	3
	Conventions	3	3

For more information, and to search for books by Lexile measure, visit [www.Lexile.com](http://www.Lexile.com).  
For more information about Quantile measures, visit the Math @ Home section at [www.Quantiles.com](http://www.Quantiles.com)

## A Guide to the Score Report

Several types of information are provided in this report. One is a description of your child's performance level (advanced, proficient, basic, or below basic) in each content area. Another is the scale score your child earned for each content area. Within each content area you can see your child's performance in each domain of that content area. These results provide your child's school with information about how well your child is learning the Wyoming Content Standards.

The box chart at the top right of the first page indicates the ranges of the scale scores in each content area. The bar graph below this chart depicts your child's scale score in each of the content areas. Different scales are used for each content area, so the numerical scores from each content area are not comparable numbers. The placement of the score in each proficiency band is the important information for you and your child's teacher to know. Below this chart you will see the Lexile and Quantile scores for your child. Entering your child's scores on these websites will provide you access to valuable tools and resources which can support your child's academic growth. The Content Performance by Domain box on the lower right side of the front page contains more detailed information about your child's performance in each content area. This information helps you and your child's teacher identify specific areas of strength and areas in need of improvement. The percent correct column indicates the percentage of items in that domain that your child answered correctly on the test. The SAWS writing information (bottom right) shows the raw score (total earned and total possible) for your child on the writing test.

### Performance Level Descriptors

	Reading	Mathematics	Science	Writing
Advanced	Students demonstrate thorough ability to comprehend implied main ideas, make subtle connections within/across texts, answer questions, and use text features to locate information. They understand complex words and phrases.	Students demonstrate an in-depth understanding of multiplication and division using strategies; fractional concepts; area and perimeter; computing the sum or difference of whole numbers, time intervals, and measurements; analyzing data and 2-D shapes.	Not Applicable.	Students develop a clear and focused main idea in response to the topic; effectively group similar ideas together with a topic sentence; use a variety of descriptive words/sentences; show consistent use of writing mechanics with few errors.
Proficient	Students demonstrate adequate ability to comprehend main ideas, make connections within/across texts, answer questions, and use text features to locate information. They understand words and phrases.	Students demonstrate a solid understanding of multiplication and division using strategies; fractional concepts; area and perimeter; computing the sum or difference of whole numbers, time intervals, and measurements; describing data and 2-D shapes.	Not Applicable.	Students present a main idea in response to the topic; group similar ideas together with a topic sentence; use some descriptive words/sentences; show adequate use of writing mechanics with some errors.
Basic	Students demonstrate partial ability to comprehend main ideas, make simple connections within/across texts, answer simple questions, and use text features to locate information. They understand simple words and phrases.	Students demonstrate a partial understanding of multiplication and division using strategies; fractional concepts; area and perimeter; computing the sum or difference of whole numbers, time intervals, and measurements; describing data and 2-D shapes.	Not Applicable.	Students attempt to present a main idea in response to the topic; group ideas together without a topic sentence; use basic word choice with limited descriptive words/simple sentences; show basic use of writing mechanics with multiple errors.
Below Basic	Students require extensive support or provide little or no evidence in meeting the standard.	Students require extensive support or provide little or no evidence in meeting the standard.	Not Applicable.	Students require extensive support or provide little or no evidence in meeting the standard.

### What Can You Do at Home?

Reading	Mathematics	Science	Writing
<ul style="list-style-type: none"> <li>• Read for at least thirty minutes per night.</li> <li>• Read aloud to your child.</li> <li>• Model being a reader.</li> <li>• Ask who, what, where, when, why and how, questions about the book your child is reading.</li> <li>• Ask your child to jot notes about his or her reading.</li> </ul>	<ul style="list-style-type: none"> <li>• Be familiar with what your child is learning at school and work on those concepts in different contexts.</li> <li>• Have a specific place for homework at home where you can check your child's progress and concepts being taught at any time.</li> <li>• Practice fluency facts.</li> </ul>	<ul style="list-style-type: none"> <li>• Promote investigative activities that happen outside.</li> <li>• Turn everyday household activities into experiments (baking, cleaning, etc.).</li> <li>• Form hypotheses when trying to work out everyday problems.</li> </ul>	<ul style="list-style-type: none"> <li>• Keep a journal or diary at home.</li> <li>• Find a pen-pal and write often.</li> <li>• Find creative writing topics and work on stories together with your child.</li> <li>• Ask your child's teacher for a writing rubric to know what exactly to work on at home.</li> <li>• Model your own writing.</li> </ul>



Appendix C: Sample PAWS/SAWS Student Reports: Grade 4 Reading,  
Mathematics, and SAWS (exemplar for Grade 8)



## DEPARTMENT OF EDUCATION

# Student Report

**First Name:** AIDAN  
**Middle Initial:**  
**Last Name:** AUSTIN

**Grade:** 4  
**Birthdate:** 03/11/2004  
**Student ID:** 23930030  
**Test Window:** 10/21/13 - 06/30/14  
**School:** Laura Irwin Elementary

**District:** Big Horn CSD # 4

	Score Ranges				Your Student	State Percentile Rank
	Below Basic	Basic	Proficient	Advanced		
Reading	300-569	570-633	634-699	700-975	NT	NT
Mathematics	300-583	584-619	620-697	698-975	NT	NT
Science	300-611	612-665	666-725	726-975	309	12%
Writing	NA	NA	NA	NA	NA	NA

Advanced

Proficient

Basic

Below Basic

NT

NT

309

NA

Reading

Mathematics

Science

Writing

Lexile® Measure

---

Quantile® Measure

---

### Purpose of Report

This report provides information about your child's achievement on the Proficiency Assessment for Wyoming Students (PAWS) and the Student Assessment of Writing Skills (SAWS). This report will help you understand your child's performance in reading and math (grades 3-8), science (grades 4 and 8), and writing (grades 3, 5, and 7) for the 2013-2014 school year.

### Glossary of Terms

**Scale Score:** Your child's raw score (total number of points earned) transformed into a score on a scale.

**Score Ranges:** The range of scale score points assigned to each proficiency level. Each subject area has a unique range of points.

**Domain Performance:** Describes your child's performance in sub-categories (domains) of each content area. The Scale Score column indicates your child's performance in relation to the overall reading scale (above). The Domain Performance column shows your child's percent correct in each of the measured domains.

**State Percentile Rank:** State Percentile Rank indicates your child's performance in relation to other Wyoming students in the same grade. The percentile shows the percentage of Wyoming students in the same grade who obtained scores equal to or less than your child's score.

**Lexile Measure:** Helps readers select materials at their reading level. This can serve as a guide in selecting books for your child.

**Quantile Measure:** Similar to the Lexile and can help you identify math activities to do at home. These activities will help your child practice mathematical skills leading to increased mathematical understanding.

**Not Tested (NT):** Your child did not take this part of the assessment.

**Not Applicable (NA):** This content area is not tested at this grade level.

**For Additional Resources and Information-**  
**Visit the Wyoming Department of Education online**

Go to the Wyoming Department of Education's website at [edu.wyoming.gov](http://edu.wyoming.gov) for more information about the PAWS and SAWS Assessments.

### CONTENT PERFORMANCE BY DOMAIN

PAWS		Scale Score	Domain Performance (% Correct)
Reading	Literature: Key Ideas and Details	NT	NT
	Literature: Craft and Structure		
	Informational Text: Key Ideas and Details		
	Informational Text: Craft and Structure		
Math	Language	NT	NT
	Operation & Algebraic Thinking		
	Number Operations – Base 10		
	Number Operations - Fractions		
Science	Measurement & Data	501	6%
	Geometry		
	Life Science		
	Physical Science		
Writing	Earth & Space Science	804	22%
SAWS		Total Possible	Student Score
Writing	NA		
	NA		

For more information, and to search for books by Lexile measure, visit [www.Lexile.com](http://www.Lexile.com).  
For more information about Quantile measures, visit the Math @ Home section at [www.Quantiles.com](http://www.Quantiles.com)

## A Guide to the Score Report

Several types of information are provided in this report. One is a description of your child's performance level (advanced, proficient, basic, or below basic) in each content area. Another is the scale score your child earned for each content area. Within each content area you can see your child's performance in each domain of that content area. These results provide your child's school with information about how well your child is learning the Wyoming Content Standards.

The box chart at the top right of the first page indicates the ranges of the scale scores in each content area. The bar graph below this chart depicts your child's scale score in each of the content areas. Different scales are used for each content area, so the numerical scores from each content area are not comparable numbers. The placement of the score in each proficiency band is the important information for you and your child's teacher to know. Below this chart you will see the Lexile and Quantile scores for your child. Entering your child's scores on these websites will provide you access to valuable tools and resources which can support your child's academic growth. The Content Performance by Domain box on the lower right side of the front page contains more detailed information about your child's performance in each content area. This information helps you and your child's teacher identify specific areas of strength and areas in need of improvement. The percent correct column indicates the percentage of items in that domain that your child answered correctly on the test. The SAWS writing information (bottom right) shows the raw score (total earned and total possible) for your child on the writing test.

### Performance Level Descriptors

	Reading	Mathematics	Science	Writing
Advanced	Students demonstrate thorough ability to comprehend implied main ideas, answer questions, and explain structural comparisons between one or more texts. They understand complex words and phrases.	Students demonstrate an in-depth understanding of whole numbers and measurements to solve problems using the four operations; whole number place value; equivalent fractions or decimals; fraction operations; classifying and analyzing geometric shapes.	Students demonstrate in-depth science knowledge in complex situations to analyze, construct, and use information and ideas. They evaluate conclusions based on evidence to explain phenomena and generalize cause and effect relationships.	Not Applicable.
Proficient	Students demonstrate adequate ability to comprehend main ideas, answer questions about text, and explain structural comparisons between one or more texts. They understand words and phrases.	Students demonstrate a solid understanding of whole numbers and measurements to solve problems using the four operations; whole number place value; equivalent fractions or decimals; fraction operations; identifying or classifying geometric shapes.	Students demonstrate solid understanding of concepts and skills to analyze information and use ideas. They can describe and recall facts, perform investigations, record results, predict outcomes and provide solutions based on evidence.	Not Applicable.
Basic	Students demonstrate partial ability to comprehend main ideas, answer questions, and explain structural comparisons between one or more texts. They understand simple words and phrases.	Students demonstrate a partial understanding of whole numbers and measurements to solve problems using the four operations; whole number place value; equivalent fractions or decimals; fraction operations; identifying or classifying geometric shapes.	Students demonstrate partial understanding or require assistance to use concepts and skills to analyze facts and ideas. They demonstrate or explain basic scientific principles and observations with support.	Not Applicable.
Below Basic	Students require extensive support or provide little or no evidence in meeting the standard.	Students require extensive support or provide little or no evidence in meeting the standard.	Students require extensive support or provide little or no evidence in meeting the standard.	Not Applicable.

### What Can You Do at Home?

Reading	Mathematics	Science	Writing
<ul style="list-style-type: none"> <li>• Read for at least thirty minutes per night.</li> <li>• Read aloud to your child.</li> <li>• Model being a reader.</li> <li>• Ask who, what, where, when, why and how, questions about the book your child is reading.</li> <li>• Ask your child to jot notes about his or her reading.</li> </ul>	<ul style="list-style-type: none"> <li>• Be familiar with what your child is learning at school and work on those concepts in different contexts.</li> <li>• Have a specific place for homework at home where you can check your child's progress and concepts being taught at any time.</li> <li>• Practice fluency facts.</li> </ul>	<ul style="list-style-type: none"> <li>• Promote investigative activities that happen outside.</li> <li>• Turn everyday household activities into experiments (baking, cleaning, etc.).</li> <li>• Form hypotheses when trying to work out everyday problems.</li> </ul>	<ul style="list-style-type: none"> <li>• Keep a journal or diary at home.</li> <li>• Find a pen-pal and write often.</li> <li>• Find creative writing topics and work on stories together with your child.</li> <li>• Ask your child's teacher for a writing rubric to know what exactly to work on at home.</li> <li>• Model your own writing.</li> </ul>

## Appendix D: Sample PAWS/SAWS Student Reports: Grade 6 Reading and Mathematics





## DEPARTMENT OF EDUCATION

# Student Report

**First Name:** AMY  
**Middle Initial:** A  
**Last Name:** COOPER  
  
**Grade:** 6  
**Birthdate:** 11/05/2002  
**Student ID:** 94730005  
**Test Window:** 10/21/13 - 06/30/14  
**School:** Dubois Middle School  
**District:** Dubois School District

	Score Ranges				Your Student	State Percentile Rank
	Below Basic	Basic	Proficient	Advanced		
Reading	300-593	594-649	650-717	718-975	515	21%
Mathematics	300-631	632-662	663-740	741-975	614	16%
Science	NA	NA	NA	NA	NA	NA
Writing	NA	NA	NA	NA	NA	NA

Advanced

Proficient

Basic

Below Basic

515

614

NA

NA

Reading

Mathematics

Science

Writing

Lexile® Measure

615L

Quantile® Measure

614Q

### Purpose of Report

This report provides information about your child's achievement on the Proficiency Assessment for Wyoming Students (PAWS) and the Student Assessment of Writing Skills (SAWS). This report will help you understand your child's performance in reading and math (grades 3-8), science (grades 4 and 8), and writing (grades 3, 5, and 7) for the 2013-2014 school year.

### Glossary of Terms

**Scale Score:** Your child's raw score (total number of points earned) transformed into a score on a scale.

**Score Ranges:** The range of scale score points assigned to each proficiency level. Each subject area has a unique range of points.

**Domain Performance:** Describes your child's performance in sub-categories (domains) of each content area. The Scale Score column indicates your child's performance in relation to the overall reading scale (above). The Domain Performance column shows your child's percent correct in each of the measured domains.

**State Percentile Rank:** State Percentile Rank indicates your child's performance in relation to other Wyoming students in the same grade. The percentile shows the percentage of Wyoming students in the same grade who obtained scores equal to or less than your child's score.

**Lexile Measure:** Helps readers select materials at their reading level. This can serve as a guide in selecting books for your child.

**Quantile Measure:** Similar to the Lexile and can help you identify math activities to do at home. These activities will help your child practice mathematical skills leading to increased mathematical understanding.

**Not Tested (NT):** Your child did not take this part of the assessment.

**Not Applicable (NA):** This content area is not tested at this grade level.

For Additional Resources and Information- Visit the Wyoming Department of Education online

Go to the Wyoming Department of Education's website at [edu.wyoming.gov](http://edu.wyoming.gov) for more information about the PAWS and SAWS Assessments.

### CONTENT PERFORMANCE BY DOMAIN

PAWS		Scale Score	Domain Performance (% Correct)
Reading	Literature: Key Ideas and Details	306	40%
	Literature: Craft and Structure	303	33%
	Informational Text: Key Ideas and Details	402	13%
	Informational Text: Craft and Structure	402	22%
	Language	802	25%
Math	Ratio & Proportional Relationship	502	20%
	The Number System	304	27%
	Expressions & Equations	302	10%
	Geometry	303	50%
	Statistics & Probability	803	38%
Science	NA		
Writing	SAWS	Total Possible	Student Score
	NA		
	NA		

For more information, and to search for books by Lexile measure, visit [www.Lexile.com](http://www.Lexile.com). For more information about Quantile measures, visit the Math @ Home section at [www.Quantiles.com](http://www.Quantiles.com)

## A Guide to the Score Report

Several types of information are provided in this report. One is a description of your child's performance level (advanced, proficient, basic, or below basic) in each content area. Another is the scale score your child earned for each content area. Within each content area you can see your child's performance in each domain of that content area. These results provide your child's school with information about how well your child is learning the Wyoming Content Standards.

The box chart at the top right of the first page indicates the ranges of the scale scores in each content area. The bar graph below this chart depicts your child's scale score in each of the content areas. Different scales are used for each content area, so the numerical scores from each content area are not comparable numbers. The placement of the score in each proficiency band is the important information for you and your child's teacher to know. Below this chart you will see the Lexile and Quantile scores for your child. Entering your child's scores on these websites will provide you access to valuable tools and resources which can support your child's academic growth. The Content Performance by Domain box on the lower right side of the front page contains more detailed information about your child's performance in each content area. This information helps you and your child's teacher identify specific areas of strength and areas in need of improvement. The percent correct column indicates the percentage of items in that domain that your child answered correctly on the test. The SAWS writing information (bottom right) shows the raw score (total earned and total possible) for your child on the writing test.

### Performance Level Descriptors

	Reading	Mathematics	Science	Writing
Advanced	Students demonstrate thorough comprehension by supporting inferences, analyzing points of view, and making comparisons between themes, topics, and text structures within/across texts. They understand complex words and phrases.	Students demonstrate an in-depth understanding of ratios, positive rational numbers, and geometric properties to solve problems; negative numbers; writing, interpreting, and using expressions, equations, and inequalities; summarizing data.	Not Applicable.	Not Applicable.
Proficient	Students demonstrate adequate comprehension by supporting inferences, explaining points of view, and making comparisons between themes, topics, and text structures within/across texts. They understand words and phrases.	Students demonstrate a solid understanding of ratios, positive rational numbers, and geometric properties to solve problems; negative numbers; writing, interpreting, and using expressions, equations, and inequalities; summarizing data.	Not Applicable.	Not Applicable.
Basic	Students demonstrate partial comprehension by supporting inferences, identifying points of view, and making comparisons between themes, topics, and text structures within/across texts. They understand simple words and phrases.	Students demonstrate a partial understanding of ratios, positive rational numbers, and geometric properties to solve problems; negative numbers; writing, interpreting, and using expressions, equations, and inequalities; summarizing data.	Not Applicable.	Not Applicable.
Below Basic	Students require extensive support or provide little or no evidence in meeting the standard.	Students require extensive support or provide little or no evidence in meeting the standard.	Not Applicable.	Not Applicable.

### What Can You Do at Home?

Reading	Mathematics	Science	Writing
<ul style="list-style-type: none"> <li>• Read for at least thirty minutes per night.</li> <li>• Read aloud to your child.</li> <li>• Model being a reader.</li> <li>• Ask who, what, where, when, why and how, questions about the book your child is reading.</li> <li>• Ask your child to jot notes about his or her reading.</li> </ul>	<ul style="list-style-type: none"> <li>• Be familiar with what your child is learning at school and work on those concepts in different contexts.</li> <li>• Have a specific place for homework at home where you can check your child's progress and concepts being taught at any time.</li> <li>• Practice fluency facts.</li> </ul>	<ul style="list-style-type: none"> <li>• Promote investigative activities that happen outside.</li> <li>• Turn everyday household activities into experiments (baking, cleaning, etc.).</li> <li>• Form hypotheses when trying to work out everyday problems.</li> </ul>	<ul style="list-style-type: none"> <li>• Keep a journal or diary at home.</li> <li>• Find a pen-pal and write often.</li> <li>• Find creative writing topics and work on stories together with your child.</li> <li>• Ask your child's teacher for a writing rubric to know what exactly to work on at home.</li> <li>• Model your own writing.</li> </ul>

## Appendix E: SAWS Field Test Subscale Correlations

Table E1. Grade 5 SAWS 2014 Field Test Subscale Correlations – 12-point prompt

Form	Score	Total	Idea Development	Organization	Voice	Conventions
1	Total	1.00	0.91	0.90	0.91	0.86
	Idea Development		1.00	0.77	0.79	0.69
	Organization			1.00	0.76	0.69
	Voice				1.00	0.71
	Conventions					1.00
2	Total	1.00	0.90	0.89	0.92	0.86
	Idea Development		1.00	0.77	0.79	0.67
	Organization			1.00	0.74	0.66
	Voice				1.00	0.74
	Conventions					1.00
3	Total	1.00	0.89	0.89	0.92	0.86
	Idea Development		1.00	0.74	0.79	0.65
	Organization			1.00	0.75	0.69
	Voice				1.00	0.74
	Conventions					1.00

Table E2. Grade 5 SAWS 2014 Field Test Subscale Correlations – 4-point prompt

Form	Score	Total	Response-to-Text	Holistic
4	Total	1.00	0.86	0.80
	Response-to-Text		1.00	0.39
	Holistic			1.00
5	Total	1.00	0.84	0.82
	Response-to-Text		1.00	0.38
	Holistic			1.00
6	Total	1.00	0.87	0.83
	Response-to-Text		1.00	0.44
	Holistic			1.00
7	Total	1.00	0.89	0.78
	Response-to-Text		1.00	0.41
	Holistic			1.00
8	Total	1.00	0.88	0.81
	Response-to-Text		1.00	0.43
	Holistic			1.00

Table E3. Grade 5 SAWS 2014 Field Test Subscale Correlations – 8-point prompt

Form	Score	Total	Response-to-Text	Holistic
4	Total	1.00	0.73	0.95
	Response-to-Text		1.00	0.48
	Holistic			1.00
5	Total	1.00	0.76	0.96
	Response-to-Text		1.00	0.56
	Holistic			1.00
6	Total	1.00	0.77	0.92
	Response-to-Text		1.00	0.45
	Holistic			1.00
7	Total	1.00	0.80	0.93
	Response-to-Text		1.00	0.53
	Holistic			1.00
8	Total	1.00	0.76	0.94
	Response-to-Text		1.00	0.49
	Holistic			1.00

Table E4. Grade 7 SAWS 2014 Field Test Subscale Correlations – 12-point prompt

Form	Score	Total	Idea Development	Organization	Voice	Conventions
1	Total	1.00	0.91	0.92	0.93	0.87
	Idea Development		1.00	0.82	0.82	0.69
	Organization			1.00	0.80	0.71
	Voice				1.00	0.76
	Conventions					1.00
2	Total	1.00	0.89	0.89	0.90	0.84
	Idea Development		1.00	0.76	0.76	0.63
	Organization			1.00	0.71	0.63
	Voice				1.00	0.69
	Conventions					1.00
3	Total	1.00	0.92	0.92	0.92	0.85
	Idea Development		1.00	0.82	0.81	0.67
	Organization			1.00	0.79	0.71
	Voice				1.00	0.70
	Conventions					1.00



Table E5. Grade 7 SAWS 2014 Field Test Subscale Correlations – 4-point prompt

Form	Score	Total	Response-to-Text	Holistic
4	Total	1.00	0.90	0.83
	Response-to-Text		1.00	0.50
	Holistic			1.00
5	Total	1.00	0.86	0.79
	Response-to-Text		1.00	0.37
	Holistic			1.00
6	Total	1.00	0.88	0.82
	Response-to-Text		1.00	0.45
	Holistic			1.00
7	Total	1.00	0.90	0.85
	Response-to-Text		1.00	0.53
	Holistic			1.00
8	Total	1.00	0.89	0.82
	Response-to-Text		1.00	0.48
	Holistic			1.00

Table E6. Grade 7 SAWS 2014 Field Test Subscale Correlations – 8-point prompt

Form	Score	Total	Response-to-Text	Holistic
4	Total	1.00	0.75	0.96
	Response-to-Text		1.00	0.53
	Holistic			1.00
5	Total	1.00	0.77	0.96
	Response-to-Text		1.00	0.55
	Holistic			1.00
6	Total	1.00	0.79	0.94
	Response-to-Text		1.00	0.53
	Holistic			1.00
7	Total	1.00	0.76	0.96
	Response-to-Text		1.00	0.56
	Holistic			1.00
8	Total	1.00	0.73	0.95
	Response-to-Text		1.00	0.48
	Holistic			1.00

## Appendix F: PAWS and SAWS Operational Subscale Correlations

### PAWS

Table F1. Grade 3 PAWS Total Test and Subscale Correlations

	Reading Total	LitKey	LitCrft	InfKey	InfCrft	Language	Math Total	Geometry	Measure	Algebra	Base Ten	Fraction
Reading Total	1.00	0.93	0.77	0.85	0.79	0.81	0.70	0.49	0.60	0.66	0.54	0.46
LitKey		1.00	0.66	0.69	0.63	0.68	0.63	0.43	0.55	0.59	0.49	0.41
LitCrft			1.00	0.58	0.55	0.61	0.51	0.37	0.43	0.48	0.40	0.32
InfKey				1.00	0.61	0.63	0.62	0.42	0.54	0.58	0.47	0.43
InfCrft					1.00	0.59	0.57	0.39	0.50	0.53	0.43	0.40
Language						1.00	0.58	0.41	0.49	0.54	0.45	0.37
Math Total							1.00	0.63	0.86	0.93	0.77	0.71
Geometry								1.00	0.47	0.51	0.41	0.38
Measure									1.00	0.71	0.59	0.53
Algebra										1.00	0.70	0.57
Base Ten											1.00	0.45
Fraction												1.00

Table F2. Grade 4 PAWS Total Test and Subscale Correlations

	Reading Total	LitKey	LitCrft	InfKey	InfCrft	Language	Math Total	Geometry	Measure	Algebra	Base Ten	Fraction	Science Total	LifeSci	PhysSci	EarthSci
Reading Total	1.00	0.87	0.78	0.89	0.83	0.72	0.69	0.39	0.53	0.65	0.53	0.59	0.76	0.67	0.69	0.68
LitKey		1.00	0.60	0.68	0.64	0.55	0.58	0.32	0.44	0.55	0.45	0.49	0.65	0.57	0.58	0.57
LitCrft			1.00	0.62	0.60	0.51	0.54	0.32	0.42	0.52	0.40	0.46	0.59	0.52	0.54	0.53
InfKey				1.00	0.68	0.56	0.63	0.36	0.50	0.58	0.48	0.54	0.71	0.62	0.64	0.63
InfCrft					1.00	0.56	0.58	0.32	0.44	0.55	0.44	0.49	0.63	0.55	0.56	0.57
Language						1.00	0.49	0.28	0.38	0.46	0.38	0.41	0.52	0.46	0.48	0.47
Math Total							1.00	0.58	0.80	0.84	0.75	0.90	0.73	0.65	0.67	0.65
Geometry								1.00	0.41	0.40	0.36	0.43	0.44	0.38	0.40	0.39
Measure									1.00	0.60	0.50	0.63	0.60	0.53	0.54	0.53
Algebra										1.00	0.60	0.65	0.65	0.58	0.59	0.57
Base Ten											1.00	0.59	0.53	0.48	0.48	0.47
Fraction												1.00	0.64	0.57	0.58	0.56
Science Total													1.00	0.88	0.91	0.89
LifeSci														1.00	0.70	0.68
PhysSci															1.00	0.71
EarthSci																1.00

Table F3. Grade 5 PAWS Total Test and Subscale Correlations

	Reading Total	LitKey	LitCrft	InfKey	InfCrft	Language	Math Total	Geometry	Measure	Algebra	Base Ten	Fraction
Reading Total	1.00	0.86	0.75	0.91	0.80	0.73	0.69	0.52	0.60	0.57	0.64	0.61
LitKey		1.00	0.63	0.68	0.58	0.53	0.57	0.41	0.49	0.48	0.53	0.49
LitCrft			1.00	0.59	0.53	0.50	0.48	0.36	0.42	0.41	0.45	0.41
InfKey				1.00	0.68	0.57	0.66	0.49	0.58	0.53	0.60	0.59
InfCrft					1.00	0.51	0.58	0.44	0.50	0.48	0.53	0.51
Language						1.00	0.48	0.36	0.42	0.39	0.45	0.43
Math Total							1.00	0.72	0.88	0.74	0.90	0.92
Geometry								1.00	0.57	0.51	0.59	0.59
Measure									1.00	0.60	0.71	0.74
Algebra										1.00	0.63	0.61
Base Ten											1.00	0.75
Fraction												1.00

Table F4. Grade 6 PAWS Total Test and Subscale Correlations

	Reading Total	LitKey	LitCrft	InfKey	InfCrft	Language	Math Total	Geometry	Measure	Algebra	Base Ten	Fraction
Reading Total	1.00	0.89	0.83	0.88	0.80	0.82	0.73	0.52	0.64	0.62	0.67	0.59
LitKey		1.00	0.70	0.69	0.63	0.67	0.62	0.42	0.55	0.54	0.57	0.50
LitCrft			1.00	0.65	0.59	0.62	0.60	0.41	0.51	0.50	0.55	0.49
InfKey				1.00	0.64	0.65	0.69	0.50	0.60	0.58	0.63	0.55
InfCrft					1.00	0.61	0.57	0.40	0.50	0.48	0.53	0.46
Language						1.00	0.60	0.44	0.54	0.50	0.54	0.48
Math Total							1.00	0.74	0.81	0.87	0.92	0.79
Geometry								1.00	0.54	0.57	0.63	0.52
Measure									1.00	0.63	0.68	0.59
Algebra										1.00	0.73	0.61
Base Ten											1.00	0.66
Fraction												1.00

Table F5. Grade 7 PAWS Total Test and Subscale Correlations

	Reading Total	LitKey	LitCrft	InfKey	InfCrft	Language	Math Total	Geometry	Measure	Algebra	Base Ten	Fraction
Reading Total	1.00	0.83	0.80	0.91	0.81	0.79	0.73	0.49	0.65	0.63	0.66	0.61
LitKey		1.00	0.62	0.67	0.60	0.57	0.56	0.37	0.51	0.49	0.50	0.48
LitCrft			1.00	0.62	0.58	0.58	0.54	0.36	0.49	0.47	0.49	0.45
InfKey				1.00	0.66	0.65	0.68	0.46	0.61	0.58	0.62	0.57
InfCrft					1.00	0.60	0.60	0.41	0.54	0.52	0.55	0.49
Language						1.00	0.62	0.42	0.55	0.53	0.56	0.51
Math Total							1.00	0.73	0.87	0.84	0.92	0.80
Geometry								1.00	0.54	0.52	0.59	0.50
Measure									1.00	0.67	0.74	0.64
Algebra										1.00	0.72	0.62
Base Ten											1.00	0.65
Fraction												1.00

Table F6. Grade 8 PAWS Total Test and Subscale Correlations

	Reading Total	LitKey	LitCrft	InfKey	InfCrft	Language	Math Total	Geometry	Measure	Algebra	Base Ten	Fraction	Science Total	LifeSci	PhysSci	EarthSci
Reading Total	1.00	0.83	0.78	0.91	0.84	0.79	0.70	0.57	0.61	0.49	0.65	0.56	0.76	0.68	0.67	0.67
LitKey		1.00	0.60	0.66	0.64	0.61	0.55	0.45	0.48	0.37	0.51	0.44	0.60	0.54	0.53	0.53
LitCrft			1.00	0.62	0.60	0.55	0.52	0.42	0.46	0.36	0.49	0.42	0.58	0.53	0.52	0.51
InfKey				1.00	0.69	0.64	0.66	0.54	0.58	0.46	0.61	0.52	0.72	0.64	0.65	0.63
InfCrft					1.00	0.62	0.60	0.49	0.52	0.43	0.56	0.49	0.64	0.58	0.57	0.57
Language						1.00	0.55	0.44	0.48	0.39	0.51	0.45	0.59	0.53	0.52	0.52
Math Total							1.00	0.85	0.88	0.74	0.93	0.67	0.75	0.67	0.69	0.64
Geometry								1.00	0.66	0.58	0.69	0.49	0.64	0.57	0.58	0.55
Measure									1.00	0.58	0.77	0.54	0.66	0.59	0.60	0.56
Algebra										1.00	0.63	0.43	0.54	0.48	0.50	0.45
Base Ten											1.00	0.56	0.68	0.61	0.62	0.58
Fraction												1.00	0.58	0.52	0.51	0.51
Science Total													1.00	0.90	0.91	0.86
LifeSci														1.00	0.71	0.67
PhysSci															1.00	0.67
EarthSci																1.00

SAWS

Table F7. SAWS 2014 Operational Test Subscale Correlations – 12-point Prompt

Grade	Score	Total	Idea Development	Organization	Voice	Conventions
Prompt 1						
3	Total	1.00	0.87	0.84	0.88	0.84
	Idea Development		1.00	0.63	0.75	0.63
	Organization			1.00	0.61	0.59
	Voice				1.00	0.68
	Conventions					1.00
Prompt 2						
3	Total	1.00	0.89	0.90	0.90	0.85
	Idea Development		1.00	0.78	0.76	0.62
	Organization			1.00	0.74	0.67
	Voice				1.00	0.69
	Conventions					1.00
5	Total	1.00	0.89	0.89	0.90	0.86
	Idea Development		1.00	0.75	0.76	0.66
	Organization			1.00	0.73	0.68
	Voice				1.00	0.71
	Conventions					1.00
7	Total	1.00	0.90	0.90	0.90	0.86
	Idea Development		1.00	0.77	0.77	0.65
	Organization			1.00	0.74	0.67
	Voice				1.00	0.71
	Conventions					1.00

Table F8. SAWS 2014 Operational Test Subscale Correlations – 4-point Prompt

Grade	Score	Total	Response-to-Text	Holistic
5	Total	1.00	0.88	0.81
	Response-to-Text		1.00	0.44
	Holistic			1.00
7	Total	1.00	0.90	0.80
	Response-to-Text		1.00	0.47
	Holistic			1.00



Table F9. SAWS 2014 Operational Test Subscale Correlations –8-point Prompt

Grade	Score	Total	Response-to-Text	Holistic
5	Total	1.00	0.76	0.90
	Response-to-Text		1.00	0.40
	Holistic			1.00
7	Total	1.00	0.72	0.93
	Response-to-Text		1.00	0.41
	Holistic			1.00

## Appendix G: DIF Results for Field Test 2014 Items<sup>12</sup>

### Reading

Table G1. Grade 3 Reading DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs.</u> <u>Female</u>		<u>White vs.</u> <u>Asian</u>		<u>White vs.</u> <u>African American</u>		<u>White vs.</u> <u>Hispanic/Latino</u>		<u>White vs.</u> <u>Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	0	0	0	0
<b>B-</b>	8	6.7	0	0	0	0	4	3.4	0	0
<b>A</b>	106	89.1	0	0	0	0	73	61.3	0	0
<b>B+</b>	5	4.2	0	0	0	0	2	1.7	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL N</b>	0	0	119	100	119	100	40	33.6	119	100
<b>TOTAL</b>	119	100	119	100	119	100	119	100	119	100

<sup>12</sup> Not all percentages will sum to a total of 100 due to rounding.

Table G2. Grade 4 Reading DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	0	0	0	0
<b>B-</b>	3	2.7	0	0	0	0	2	1.8	0	0
<b>A</b>	102	90.3	0	0	0	0	28	24.8	0	0
<b>B+</b>	8	7.1	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	113	100	113	100	83	73.5	113	100
<b>TOTAL</b>	113	100	113	100	113	100	113	100	113	100

Table G3. Grade 5 Reading DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	0	0	0	0
<b>B-</b>	1	0.9	0	0	0	0	0	0	0	0
<b>A</b>	102	91.1	0	0	0	0	26	23.2	0	0
<b>B+</b>	9	8.0	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	112	100	112	100	86	76.8	112	100
<b>TOTAL</b>	112	100	112	100	112	100	112	100	112	100

Table G4. Grade 6 Reading DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	1	0.9	0	0	0	0	0	0	0	0
<b>B-</b>	4	3.6	0	0	0	0	0	0	0	0
<b>A</b>	96	85.7	0	0	0	0	14	12.5	0	0
<b>B+</b>	11	9.8	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	112	100	112	100	98	87.5	112	100
<b>TOTAL</b>	112	100	112	100	112	100	112	100	112	100

Table G5. Grade 7 Reading DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	0	0	0	0
<b>B-</b>	4	3.6	0	0	0	0	0	0	0	0
<b>A</b>	97	86.6	0	0	0	0	15	13.4	0	0
<b>B+</b>	10	8.9	0	0	0	0	0	0	0	0
<b>C+</b>	1	0.9	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	112	100	112	100	97	86.6	112	100
<b>TOTAL</b>	112	100	112	100	112	100	112	100	112	100

Table G6. Grade 8 Reading DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs.</u> <u>Female</u>		<u>White vs.</u> <u>Asian</u>		<u>White vs.</u> <u>African American</u>		<u>White vs.</u> <u>Hispanic/Latino</u>		<u>White vs.</u> <u>Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	1	0.8	0	0	0	0	0	0	0	0
<b>B-</b>	5	4.2	0	0	0	0	1	0.8	0	0
<b>A</b>	106	88.3	0	0	0	0	12	10.0	0	0
<b>B+</b>	7	5.8	0	0	0	0	0	0	0	0
<b>C+</b>	1	0.8	0	0	0	0	0	0	0	0
<b>SMALL N</b>	0	0	120	100	120	100	107	89.2	120	100
<b>TOTAL</b>	120	100	120	100	120	100	120	100	120	100

Mathematics

Table G7. Grade 3 Mathematics DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	1	0.9	0	0
<b>B-</b>	9	7.8	0	0	0	0	3	2.6	0	0
<b>A</b>	99	86.1	0	0	0	0	91	79.1	0	0
<b>B+</b>	5	4.3	0	0	0	0	4	3.5	0	0
<b>C+</b>	2	1.7	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	115	100	115	100	16	13.9	115	100
<b>TOTAL</b>	115	100	115	100	115	100	115	100	115	100

Table G8. Grade 4 Mathematics DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	1	0.9	0	0
<b>B-</b>	4	3.6	0	0	0	0	2	1.8	0	0
<b>A</b>	104	94.5	0	0	0	0	28	25.5	0	0
<b>B+</b>	2	1.8	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	1	0.9	0	0
<b>SMALL <i>N</i></b>	0	0	110	100	110	100	78	70.9	110	100
<b>TOTAL</b>	110	100	110	100	110	100	110	100	110	100

Table G9. Grade 5 Mathematics DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	0	0	0	0
<b>B-</b>	1	0.9	0	0	0	0	0	0	0	0
<b>A</b>	104	94.5	0	0	0	0	27	24.5	0	0
<b>B+</b>	5	4.5	0	0	0	0	1	0.9	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	110	100	110	100	82	74.5	110	100
<b>TOTAL</b>	110	100	110	100	110	100	110	100	110	100

Table G10. Grade 6 Mathematics DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	0	0	0	0
<b>B-</b>	5	4.5	0	0	0	0	2	1.8	0	0
<b>A</b>	103	93.6	0	0	0	0	25	22.7	0	0
<b>B+</b>	2	1.8	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	1	0.9	0	0
<b>SMALL <i>N</i></b>	0	0	110	100	110	100	82	74.5	110	100
<b>TOTAL</b>	110	100	110	100	110	100	110	100	110	100

Table G11. Grade 7 Mathematics DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	1	0.9	0	0	0	0	0	0	0	0
<b>B-</b>	4	3.6	0	0	0	0	0	0	0	0
<b>A</b>	96	85.7	0	0	0	0	14	12.5	0	0
<b>B+</b>	11	9.8	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	112	100	112	100	98	87.5	112	100
<b>TOTAL</b>	112	100	112	100	112	100	112	100	112	100

Table G12. Grade 8 Mathematics DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	1	0.8	0	0	0	0	0	0	0	0
<b>B-</b>	5	4.2	0	0	0	0	1	0.8	0	0
<b>A</b>	106	88.3	0	0	0	0	12	10.0	0	0
<b>B+</b>	7	5.8	0	0	0	0	0	0	0	0
<b>C+</b>	1	0.8	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	120	100	120	100	107	89.2	120	100
<b>TOTAL</b>	120	100	120	100	120	100	120	100	120	100



Science

Table G13. Grade 4 Science DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	1	0.8	0	0	0	0	0	0	0	0
<b>B-</b>	4	3.4	0	0	0	0	2	1.7	0	0
<b>A</b>	113	95.0	0	0	0	0	57	47.9	0	0
<b>B+</b>	1	0.8	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	119	100	119	100	60	50.4	119	100
<b>TOTAL</b>	119	100	119	100	119	100	119	100	119	100

Table G14. Grade 8 Science DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	0	0	0	0
<b>B-</b>	9	7.6	0	0	0	0	1	0.8	0	0
<b>A</b>	109	91.6	0	0	0	0	23	19.3	0	0
<b>B+</b>	1	0.8	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL <i>N</i></b>	0	0	119	100	119	100	95	79.8	119	100
<b>TOTAL</b>	119	100	119	100	119	100	119	100	119	100

SAWS

Table G15. Grade 3 SAWS DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
C-	00	0	0	0	0	0	0	0	0	0
B-	00	00	0	0	0	0	1	16.7	0	0
A	5	83.3	0	0	0	0	5	83.3	0	0
B+	1	16.7	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL <i>N</i>	0	0	6	100	6	100	0	0	6	100
TOTAL	6	100	6	100	6	100	6	100	6	100

Table G16. Grade 5 SAWS DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs. Female</u>		<u>White vs. Asian</u>		<u>White vs. African American</u>		<u>White vs. Hispanic/Latino</u>		<u>White vs. Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
C-	0	0	0	0	0	0	0	0	0	0
B-	0	0	0	0	0	0	1	7.7	0	0
A	5	38.5	0	0	0	0	11	84.6	0	0
B+	7	53.8	0	0	0	0	1	7.7	0	0
C+	1	7.7	0	0	0	0	0	0	0	0
SMALL <i>N</i>	0	0	13	100	13	100	0	0	13	100
TOTAL	13	100	13	100	13	100	13	100	13	100

Table G17. Grade 7 SAWS DIF Summary Statistics for Embedded Field Test Items

DIF Category	<u>Male vs.</u> <u>Female</u>		<u>White vs.</u> <u>Asian</u>		<u>White vs.</u> <u>African American</u>		<u>White vs.</u> <u>Hispanic/Latino</u>		<u>White vs.</u> <u>Native American</u>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>C-</b>	0	0	0	0	0	0	0	0	0	0
<b>B-</b>	0	0	0	0	0	0	0	0	0	0
<b>A</b>	13	100	0	0	0	0	13	100	0	0
<b>B+</b>	0	0	0	0	0	0	0	0	0	0
<b>C+</b>	0	0	0	0	0	0	0	0	0	0
<b>SMALL N</b>	0	0	13	100	13	100	0	0	13	100
<b>TOTAL</b>	13	100	13	100	13	100	13	100	13	100

## Appendix H: SAWS Field Test Rater Reliability

Table H1. Grade 3 Field Test Trait Rater Agreement and Weighted Kappa – 8-point Prompt

Form	N	<u>Rater 1</u>		<u>Rater 2</u>		CORR	<u>Percentages of Agreement</u>		Weighted Kappa
		Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
1	312	5.68	1.76	5.96	1.72	0.57	51.92	51.92	0.57
2	302	5.72	1.80	5.59	1.80	0.68	60.93	60.93	0.68
3	303	5.41	1.90	5.27	1.75	0.72	58.75	58.75	0.72
4	307	5.30	1.73	5.36	1.87	0.64	51.79	51.79	0.64
5	305	5.39	1.98	5.38	1.93	0.69	56.72	56.72	0.69
6	307	5.35	1.75	5.36	1.77	0.61	54.07	54.07	0.61

Table H2. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 12-point prompts

Form	Trait	N	<u>Rater 1</u>		<u>Rater 2</u>		CORR	<u>Percentages of Agreement</u>		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
1	Total	218	7.89	2.59	7.76	2.50	0.64	31.19	53.21	
	Idea Development	218	2.02	0.72	1.99	0.70	0.50	59.17	96.79	0.50
	Organization	218	1.96	0.74	1.92	0.73	0.55	57.80	97.71	0.55
	Voice	218	1.97	0.71	1.95	0.71	0.56	58.72	99.08	0.56
	Conventions	218	1.94	0.73	1.91	0.71	0.57	61.47	98.17	0.57
2	Total	212	7.28	2.46	7.28	2.58	0.71	30.66	57.55	
	Idea Development	212	1.92	0.69	1.96	0.71	0.60	63.68	99.06	0.60
	Organization	212	1.84	0.71	1.81	0.73	0.61	61.79	99.06	0.60
	Voice	212	1.81	0.69	1.77	0.73	0.58	60.85	99.06	0.58
	Conventions	212	1.72	0.70	1.75	0.70	0.57	58.02	100.00	0.57
3	Total	220	7.50	2.47	7.70	2.41	0.65	30.91	56.82	
	Idea Development	220	1.96	0.67	2.06	0.69	0.54	63.64	97.73	0.53
	Organization	220	1.86	0.68	1.92	0.68	0.54	60.00	99.09	0.54
	Voice	220	1.91	0.72	1.90	0.68	0.58	60.00	99.55	0.58
	Conventions	220	1.77	0.69	1.82	0.70	0.52	58.64	98.18	0.52

Table H3. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 4-point prompts

Form	Trait	N	Rater 1		Rater 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
4	Total	217	3.06	0.95	3.09	0.95	0.66	56.68	94.93	
	Response-to-Text	217	1.57	0.62	1.57	0.61	0.57	71.89	98.62	0.57
	Holistic	217	1.49	0.53	1.52	0.53	0.59	76.96	100.00	0.59
5	Total	225	3.28	0.86	3.20	0.89	0.65	60.44	95.11	
	Response-to-Text	225	1.71	0.50	1.68	0.54	0.56	76.00	100.00	0.56
	Holistic	225	1.57	0.52	1.52	0.53	0.55	75.11	100.00	0.55
6	Total	216	3.07	0.94	3.20	0.87	0.58	59.72	89.81	
	Response-to-Text	216	1.60	0.55	1.68	0.52	0.56	74.07	100.00	0.56
	Holistic	216	1.47	0.52	1.52	0.52	0.49	72.69	100.00	0.49
7	Total	220	2.76	1.05	2.71	1.07	0.68	52.73	92.27	
	Response-to-Text	220	1.34	0.71	1.31	0.72	0.63	66.36	98.64	0.63
	Holistic	220	1.43	0.50	1.40	0.50	0.58	78.64	100.00	0.58
8	Total	216	3.02	1.04	2.96	1.14	0.70	57.87	92.13	
	Response-to-Text	216	1.51	0.66	1.44	0.71	0.64	70.83	98.15	0.63
	Holistic	216	1.50	0.54	1.52	0.55	0.57	76.39	99.54	0.57

Table H4. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 8-point prompts

Form	Trait	N	Rater 1		Rater 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
4	Total	215	5.47	1.63	5.42	1.64	0.70	34.42	79.07	
	Response-to-Text	215	1.68	0.56	1.68	0.56	0.59	75.81	99.53	0.59
	Holistic	215	3.78	1.34	3.74	1.33	0.63	42.79	82.79	0.63
5	Total	214	5.43	1.74	5.45	1.75	0.71	33.18	72.90	
	Response-to-Text	214	1.65	0.59	1.69	0.57	0.63	77.57	99.07	0.63
	Holistic	214	3.77	1.35	3.76	1.36	0.66	36.92	82.24	0.66
6	Total	216	4.69	1.85	4.70	1.70	0.68	30.09	74.07	
	Response-to-Text	216	1.06	0.84	1.07	0.79	0.54	56.02	94.44	0.54
	Holistic	216	3.63	1.27	3.63	1.24	0.67	43.06	87.50	0.67
7	Total	217	4.87	1.63	5.04	1.60	0.74	38.71	82.03	
	Response-to-Text	217	1.38	0.69	1.42	0.65	0.52	66.36	96.77	0.52
	Holistic	217	3.49	1.17	3.62	1.18	0.73	48.85	92.17	0.73
8	Total	222	5.38	1.47	5.36	1.59	0.70	34.23	82.43	
	Response-to-Text	222	1.73	0.52	1.70	0.60	0.55	76.58	98.20	0.54
	Holistic	222	3.66	1.12	3.67	1.19	0.65	44.14	89.64	0.65

Table H5. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 12-point prompts

Form	Trait	N	Rater 1		Rater 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
1	Total	199	7.99	2.76	7.90	2.80	0.71	33.67	60.80	
	Idea Development	199	2.06	0.76	2.07	0.79	0.58	62.31	95.98	0.58
	Organization	199	2.03	0.76	1.96	0.81	0.67	63.82	98.49	0.67
	Voice	199	1.95	0.75	1.94	0.77	0.64	62.81	98.49	0.64
	Conventions	199	1.94	0.71	1.93	0.74	0.61	61.81	99.00	0.61
2	Total	205	8.01	2.35	8.07	2.19	0.71	31.71	66.34	
	Idea Development	205	2.11	0.65	2.08	0.58	0.59	71.71	99.02	0.59
	Organization	205	2.00	0.71	2.00	0.68	0.59	62.44	99.51	0.59
	Voice	205	1.99	0.67	2.01	0.63	0.60	66.34	100.00	0.60
	Conventions	205	1.92	0.67	1.98	0.60	0.57	68.29	99.02	0.57
3	Total	204	7.59	2.66	7.72	2.68	0.69	32.35	58.82	
	Idea Development	204	1.95	0.74	2.01	0.75	0.60	60.78	98.04	0.59
	Organization	204	1.90	0.76	1.96	0.74	0.58	57.84	98.04	0.57
	Voice	204	1.89	0.75	1.87	0.75	0.64	64.22	98.53	0.64
	Conventions	204	1.85	0.69	1.87	0.77	0.55	60.78	97.06	0.55

Table H6. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 4-point prompts

Form	Trait	N	Rater 1		Rater 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
4	Total	211	2.92	1.05	2.97	1.05	0.69	55.92	93.36	
	Response-to-Text	211	1.38	0.65	1.38	0.66	0.51	63.51	98.10	0.51
	Holistic	211	1.54	0.54	1.59	0.53	0.59	76.30	100.00	0.58
5	Total	217	3.17	0.94	3.20	0.94	0.60	54.84	91.71	
	Response-to-Text	217	1.53	0.64	1.54	0.62	0.54	65.44	99.54	0.54
	Holistic	217	1.64	0.51	1.66	0.50	0.51	75.12	100.00	0.51
6	Total	218	3.14	0.94	3.12	0.93	0.61	58.26	91.74	
	Response-to-Text	218	1.56	0.58	1.55	0.60	0.51	70.18	98.62	0.51
	Holistic	218	1.58	0.51	1.57	0.51	0.48	72.48	100.00	0.48
7	Total	212	3.17	0.93	3.27	0.91	0.71	62.74	95.75	
	Response-to-Text	212	1.57	0.61	1.62	0.58	0.57	70.75	99.53	0.57
	Holistic	212	1.60	0.50	1.65	0.49	0.51	75.94	100.00	0.51
8	Total	217	2.98	0.98	2.98	1.00	0.61	51.61	91.71	
	Response-to-Text	217	1.51	0.62	1.47	0.64	0.45	61.29	98.16	0.44
	Holistic	217	1.47	0.51	1.51	0.51	0.61	79.72	100.00	0.61

Table H7. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 8-point prompts

Form	Trait	<i>N</i>	<u>Rater 1</u>		<u>Rater 2</u>		CORR	<u>Percentages of Agreement</u>		Weighted Kappa
			Mean	<i>SD</i>	Mean	<i>SD</i>		Exact	Exact + adjacent agreement only	
4	Total	220	5.29	1.63	5.44	1.61	0.66	32.27	75.91	
	Response-to-Text	220	1.58	0.57	1.63	0.55	0.44	67.73	99.09	0.44
	Holistic	220	3.70	1.25	3.81	1.27	0.63	40.00	84.55	0.63
5	Total	214	5.20	1.66	5.26	1.77	0.60	28.97	68.69	
	Response-to-Text	214	1.51	0.59	1.55	0.59	0.41	65.42	98.13	0.41
	Holistic	214	3.68	1.28	3.71	1.36	0.57	34.11	78.04	0.57
6	Total	215	4.87	1.87	5.05	1.85	0.81	36.28	82.79	
	Response-to-Text	215	1.40	0.78	1.52	0.75	0.71	75.81	96.28	0.70
	Holistic	215	3.46	1.35	3.53	1.36	0.76	45.58	89.77	0.76
7	Total	207	5.47	1.55	5.43	1.63	0.55	31.88	70.53	
	Response-to-Text	207	1.65	0.53	1.65	0.53	0.40	69.57	99.03	0.40
	Holistic	207	3.82	1.21	3.78	1.30	0.56	34.30	81.16	0.56
8	Total	215	5.18	1.62	5.27	1.68	0.71	40.00	79.53	
	Response-to-Text	215	1.67	0.58	1.67	0.60	0.65	78.60	99.07	0.65
	Holistic	215	3.51	1.23	3.59	1.29	0.67	43.26	88.37	0.67

## Appendix I: SAWS Operational Rater Reliability

Table II. SAWS 2014 Overall Interrater Reliability for 12-point Prompt and Trait Rater Agreement

Grade	Trait	<i>N</i>	<u>Rating 1</u>		<u>Rating 2</u>		CORR	<u>Percentages of Agreement</u>		Weighted Kappa
			Mean	<i>SD</i>	Mean	<i>SD</i>		Exact	Exact + adjacent agreement only	
3			Prompt 1							
	Prompt Total	1607	7.32	2.40	7.18	2.44	0.64	23.65	55.13	
	Idea Development	1607	1.90	0.70	1.88	0.69	0.49	55.38	98.38	0.49
	Organization	1607	1.89	0.73	1.83	0.78	0.58	57.62	98.13	0.58
	Voice	1607	1.78	0.70	1.75	0.71	0.53	58.74	98.13	0.53
	Conventions	1607	1.76	0.69	1.73	0.72	0.53	59.43	97.95	0.53
3			Prompt 2							
	Prompt Total	1739	7.24	2.44	7.11	2.47	0.63	28.87	56.24	
	Idea Development	1739	1.91	0.68	1.88	0.67	0.53	60.09	98.85	0.53
	Organization	1739	1.83	0.68	1.79	0.70	0.51	58.94	98.16	0.51
	Voice	1739	1.79	0.71	1.75	0.71	0.53	58.77	97.93	0.53
	Conventions	1739	1.72	0.69	1.68	0.71	0.55	59.34	98.73	0.55
5	Prompt Total	1706	7.29	2.43	7.22	2.53	0.58	27.20	54.16	
	Idea Development	1706	1.89	0.68	1.87	0.69	0.50	58.91	97.95	0.50
	Organization	1706	1.82	0.68	1.80	0.73	0.52	58.56	98.07	0.52
	Voice	1706	1.80	0.69	1.81	0.72	0.46	54.04	97.48	0.46
	Conventions	1706	1.78	0.69	1.74	0.72	0.49	57.15	97.54	0.49
7	Prompt Total	1661	7.91	2.43	7.87	2.44	0.65	29.44	59.48	
	Idea Development	1661	2.02	0.67	2.01	0.68	0.55	62.67	98.68	0.55
	Organization	1661	2.01	0.70	1.99	0.70	0.59	63.15	98.74	0.59
	Voice	1661	1.95	0.68	1.94	0.69	0.53	59.90	98.62	0.53
	Conventions	1661	1.94	0.69	1.93	0.71	0.51	57.80	98.19	0.51



Table I2. SAWS 2014 Overall Interrater Reliability for 4-point Prompt and Trait Rater Agreement

Grade	Trait	<i>N</i>	<u>Rating 1</u>		<u>Rating 2</u>		CORR	<u>Percentages of Agreement</u>		Weighted Kappa
			Mean	<i>SD</i>	Mean	<i>SD</i>		Exact	Exact + adjacent agreement only	
5	Prompt Total	1726	3.09	0.99	3.05	1.00	0.77	64.89	96.58	
	Response-to-Text	1726	1.54	0.66	1.53	0.65	0.77	81.81	99.42	0.77
	Holistic	1726	1.55	0.52	1.52	0.52	0.58	77.23	99.94	0.57
7	Prompt Total	1676	2.91	1.04	2.90	1.05	0.75	62.05	94.63	
	Response-to-Text	1676	1.38	0.70	1.36	0.71	0.75	77.57	99.11	0.75
	Holistic	1676	1.53	0.51	1.53	0.51	0.50	73.93	99.94	0.50

Table I3. SAWS 2014 Overall Interrater Reliability for 8-point Prompt and Trait Rater Agreement

Grade	Trait	<i>N</i>	<u>Rating 1</u>		<u>Rating 2</u>		CORR	<u>Percentages of Agreement</u>		Weighted Kappa
			Mean	<i>SD</i>	Mean	<i>SD</i>		Exact	Exact + adjacent agreement only	
5	Prompt Total	1740	4.90	1.58	4.88	1.57	0.71	35.29	82.07	
	Response-to-Text	1740	1.34	0.75	1.36	0.73	0.63	67.07	97.47	0.63
	Holistic	1740	3.55	1.13	3.52	1.12	0.66	46.21	91.09	0.66
7	Prompt Total	1683	5.15	1.65	5.10	1.64	0.74	36.01	82.23	
	Response-to-Text	1683	1.48	0.67	1.48	0.67	0.67	72.73	99.29	0.67
	Holistic	1683	3.67	1.25	3.63	1.25	0.70	43.79	89.19	0.70

## Appendix J: Classical Item Statistics for 2014 Field Test Items

### Reading

Table J1. Reading Grade 3 Classical Statistics for Field Test Items

Accession Number	N	Average Item	Point Biserial Corr.
		Score	
		Form 1	
VF798298	1509	0.79	0.38
VF815015	777	0.78	0.49
VF815011	777	0.79	0.20
VF815010	777	0.87	0.48
VF815012	777	0.65	0.46
VF815022	777	0.84	0.39
VF815020	777	0.67	0.22
VF798266	777	0.54	0.38
VF798299	777	0.68	0.49
VF798292	777	0.72	0.48
VF798282	777	0.27	0.27
VF798239	777	0.80	0.40
VF885220	777	0.76	0.59
VF885209	777	0.40	0.25
Form 2			
VF885358	731	0.66	0.48
VF885388	731	0.81	0.52
Form 3			
VF815562	734	0.90	0.50
VF815556	734	0.92	0.35
VF815598	734	0.31	0.12
VF815575	734	0.45	0.08
VF815528	734	0.85	0.47
VF815537	734	0.92	0.42
VF885405	734	0.89	0.43
VF885201	734	0.43	0.07
Form 4			

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF815017	732	0.70	0.39
VF815018	732	0.91	0.50
VF815014	732	0.71	0.44
VF815009	732	0.59	0.07
VF815019	732	0.88	0.50
VF815021	732	0.85	0.49
VF798274	732	0.28	0.05
VF798298	1509	0.79	0.38
VF798297	732	0.27	0.02
VF798301	732	0.54	0.37
VF798290	732	0.58	0.44
VF798300	732	0.36	0.23
VF885322	732	0.85	0.48
VF885192	732	0.89	0.53
Form 5			
VF884430	1461	0.25	0.16
VF814997	723	0.84	0.50
VF814980	723	0.82	0.44
VF814974	723	0.78	0.55
VF814982	723	0.82	0.25
VF814989	723	0.69	0.26
VF814966	723	0.91	0.50
VF884228	723	0.39	0.20
VF884215	723	0.49	0.21
VF884415	723	0.78	0.35
VF884522	723	0.72	0.43
VF884498	723	0.57	0.28
VF885399	723	0.70	0.46
VF885379	723	0.55	0.32
Form 6			
VF814976	738	0.75	0.44

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF814978	738	0.85	0.56
VF814983	738	0.72	0.42
VF814992	738	0.86	0.42
VF814971	738	0.46	0.23
VF814994	738	0.93	0.48
VF884237	738	0.64	0.49
VF884239	738	0.60	0.20
VF884430	1461	0.25	0.16
VF884250	738	0.89	0.44
VF884503	738	0.33	0.17
VF884518	738	0.60	0.34
VF885412	738	0.77	0.43
VF885214	738	0.68	0.41

---

Form 7

VF814724	741	0.47	0.23
VF814758	741	0.52	0.39
VF814839	741	0.68	0.27
VF814748	741	0.60	0.23
VF814762	741	0.88	0.44
VF814688	741	0.67	0.19
VF883326	741	0.66	0.22
VF883330	741	0.54	0.35
VF883549	741	0.81	0.49
VF883561	741	0.54	0.27
VF883619	741	0.74	0.42
VF883622	741	0.76	0.29
VF885434	741	0.70	0.56
VF885162	741	0.86	0.40

---



---

Form 8

VF814737	732	0.77	0.52
VF814753	732	0.44	0.28

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF814821	732	0.30	0.10
VF814829	732	0.53	0.31
VF814673	732	0.90	0.40
VF814681	732	0.79	0.48
VF882884	732	0.50	0.43
VF882936	732	0.65	0.39
VF883543	732	0.69	0.31
VF883364	732	0.93	0.45
VF883614	732	0.76	0.34
VF883610	732	0.56	0.41
VF885187	732	0.61	0.40
VF885218	732	0.72	0.60
Form 9			
VF821218	726	0.78	0.36
VF821123	726	0.73	0.32
VF821312	726	0.60	0.35
VF821292	726	0.22	0.12
VF821362	726	0.60	0.43
VF821338	726	0.65	0.45
VF821088	726	0.85	0.31
VF821030	726	0.62	0.32
VF821078	726	0.38	0.33
VF821006	726	0.86	0.44
VF821011	726	0.61	0.33
VF821070	726	0.87	0.52
VF885423	726	0.79	0.48
VF885198	726	0.80	0.45

Form 10			
VF821120	731	0.74	0.45
VF821206	731	0.28	0.27
VF821272	731	0.88	0.56

---

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF821320	731	0.51	0.34
VF821332	731	0.71	0.49
VF821360	731	0.49	0.47
VF821072	731	0.49	0.20
VF821062	731	0.69	0.31
VF821037	731	0.78	0.46
VF821055	731	0.78	0.36
VF821065	731	0.73	0.44
VF821024	731	0.49	0.21
VF885384	731	0.94	0.42
VF885340	731	0.86	0.49

---

Table J2. Reading Grade 4 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF822267	761	0.86	0.45
VF822250	761	0.54	0.29
VF822284	761	0.70	0.32
VF822291	761	0.60	0.37
VF822301	761	0.73	0.32
VF822303	761	0.68	0.52
VF862927	761	0.81	0.44
VF862890	761	0.91	0.38
VF862920	761	0.90	0.46
VF862909	761	0.63	0.39
VF862957	761	0.62	0.52
VF862882	761	0.74	0.43
VF885009	761	0.84	0.52
VF885043	761	0.83	0.42
Form 2			
VF885156	702	0.42	0.17
VF885173	702	0.84	0.46
Form 3			
VF885200	706	0.87	0.52
VF885215	706	0.57	0.16
Form 4			
VF822261	699	0.93	0.48
VF822269	699	0.38	0.01
VF822292	699	0.78	0.51
VF822294	699	0.65	0.21
VF822302	699	0.62	0.33
VF822298	699	0.43	0.03
VF862893	699	0.71	0.29
VF862946	699	0.90	0.38
VF862897	699	0.91	0.39
VF862965	699	0.75	0.35
VF862952	699	0.64	0.41
VF862870	699	0.82	0.37
VF885233	699	0.79	0.45
VF885037	699	0.65	0.39

Form 5

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF884830	700	0.64	0.38
VF884836	700	0.55	0.42
VF884906	700	0.58	0.23
VF884910	700	0.80	0.28
VF884918	1402	0.76	0.34
VF884913	700	0.36	0.26
VF880683	700	0.65	0.39
VF880649	700	0.45	0.19
VF880694	700	0.77	0.17
VF880689	700	0.94	0.35
VF880676	700	0.86	0.42
VF880576	700	0.75	0.42
VF885059	700	0.41	0.35
VF885226	700	0.87	0.46
Form 6			
VF884918	1402	0.76	0.34
VF884828	702	0.73	0.39
VF884843	702	0.87	0.51
VF884900	702	0.85	0.45
VF884896	702	0.80	0.28
VF884925	702	0.75	0.33
VF880629	702	0.85	0.37
VF880664	702	0.78	0.42
VF880611	702	0.83	0.35
VF880672	702	0.51	0.15
VF880678	702	0.71	0.28
VF880686	702	0.91	0.30
VF885205	702	0.71	0.53
VF885195	702	0.79	0.42
VF884587	696	0.84	0.35
VF884603	696	0.70	0.21
VF884592	696	0.84	0.26



Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF884608	696	0.53	0.36
VF884605	696	0.82	0.44
VF884561	696	0.62	0.30
VF864776	696	0.55	0.26
VF864822	696	0.66	0.44
VF864878	696	0.45	0.27
VF864868	696	0.56	0.46
VF864893	696	0.54	0.24
VF864887	696	0.51	0.27
VF885219	696	0.66	0.47
VF885232	696	0.87	0.44
Form 8			
VF884590	682	0.12	0.20
VF884582	682	0.84	0.43
VF884602	682	0.56	0.33
VF884611	682	0.39	0.04
VF884593	682	0.58	0.30
VF884577	682	0.38	0.29
VF864828	682	0.27	0.11
VF864786	682	0.56	0.37
VF864876	682	0.43	0.22
VF864861	682	0.71	0.45
VF864895	682	0.35	0.24
VF864889	682	0.52	0.42
VF885028	682	0.84	0.43
VF885064	682	0.53	0.27

Form 9			
VF884781	690	0.83	0.37
VF884769	690	0.80	0.27
VF884813	690	0.45	0.18
VF884777	690	0.71	0.31

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF884817	690	0.57	0.26
VF884734	690	0.45	0.17
VF880215	690	0.72	0.40
VF880200	690	0.72	0.47
VF880326	690	0.67	0.29
VF880314	690	0.30	0.14
VF880345	690	0.61	0.38
VF880343	690	0.61	0.31
VF885078	690	0.84	0.39
VF885092	690	0.69	0.38
Form 10			
VF884743	684	0.64	0.37
VF884766	684	0.53	0.28
VF884773	684	0.81	0.35
VF884802	684	0.43	0.01
VF884723	684	0.95	0.41
VF884807	684	0.69	0.19
VF880204	684	0.47	0.39
VF880210	684	0.50	0.26
VF880311	684	0.55	0.35
VF880321	684	0.45	0.30
VF880354	684	0.41	0.03
VF880350	684	0.60	0.30
VF885228	684	0.80	0.46
VF885166	684	0.85	0.45

Table J3. Reading Grade 5 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF884413	738	0.93	0.43
VF884409	738	0.62	0.28
VF884354	738	0.88	0.39
VF884360	738	0.32	0.16
VF888390	738	0.76	0.31
VF884312	738	0.59	0.46
VF884476	738	0.53	0.37
VF884481	738	0.76	0.45
VF884509	738	0.58	0.47
VF884517	738	0.77	0.53
VF884556	738	0.54	0.27
VF884535	738	0.64	0.33
VF885335	738	0.86	0.52
VF885180	738	0.56	0.33
Form 2			
VF885197	705	0.80	0.48
VF885191	705	0.83	0.56
Form 3			
VF885224	701	0.73	0.38
VF885329	701	0.58	0.38
Form 4			
VF884420	701	0.21	-0.05
VF884348	701	0.46	0.15
VF884341	701	0.75	0.46
VF884405	701	0.80	0.25
VF884336	701	0.61	0.36
VF884333	701	0.72	0.44
VF909884	701	0.53	0.35
VF884489	701	0.93	0.42
VF884524	701	0.60	0.27
VF884520	701	0.71	0.28
VF884559	701	0.26	0.01
VF884567	701	0.54	0.34
VF885345	701	0.55	0.26
VF885142	701	0.59	0.21

Form 5

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF882694	703	0.14	-0.29
VF880876	703	0.37	0.06
VF882773	703	0.56	0.09
VF882778	703	0.41	0.03
VF909893	703	0.64	0.29
VF882794	703	0.54	0.18
VF822491	703	0.81	0.50
VF822534	703	0.83	0.41
VF822545	703	0.81	0.35
VF822556	703	0.71	0.47
VF822548	703	0.48	0.32
VF822551	703	0.77	0.34
VF885161	703	0.74	0.36
VF885146	703	0.81	0.49
Form 6			
VF880864	711	0.92	0.35
VF881653	711	0.55	0.12
VF882769	711	0.75	0.25
VF882762	711	0.54	0.30
VF882790	711	0.56	0.27
VF882786	711	0.72	0.24
VF822538	711	0.54	0.36
VF822463	711	0.88	0.37
VF822542	711	0.83	0.31
VF822571	711	0.70	0.28
VF822496	711	0.59	0.29
VF822549	711	0.74	0.34
VF885134	711	0.81	0.43
VF885204	711	0.43	0.32

Form 7			
VF822290	705	0.52	0.21

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF822271	705	0.13	0.09
VF822280	705	0.86	0.28
VF822287	705	0.31	0.20
VF822278	705	0.45	0.15
VF822285	705	0.46	0.10
VF814960	705	0.82	0.34
VF814977	705	0.68	0.41
VF814970	705	0.78	0.40
VF814962	705	0.52	0.31
VF814973	705	0.61	0.19
VF814958	705	0.61	0.24
VF885217	705	0.80	0.41
VF885212	705	0.87	0.46
Form 8			
VF822288	697	0.93	0.27
VF822275	697	0.34	0.26
VF822282	697	0.76	0.30
VF822276	697	0.70	0.28
VF822283	697	0.47	0.20
VF822259	697	0.81	0.30
VF814959	697	0.33	0.17
VF814961	697	0.66	0.34
VF814963	697	0.24	0.18
VF814968	697	0.67	0.37
VF814975	697	0.89	0.48
VF814956	697	0.52	0.29
VF885167	697	0.55	0.24
VF885154	697	0.76	0.43

Form 9			
VF822708	706	0.40	0.01
VF822723	706	0.92	0.44

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF822757	706	0.72	0.24
VF822776	706	0.95	0.34
VF822832	706	0.73	0.35
VF822823	706	0.65	0.29
VF884191	706	0.77	0.44
VF884224	706	0.42	0.10
VF884208	706	0.46	0.20
VF884240	706	0.49	0.34
VF884152	706	0.53	0.23
VF884231	706	0.86	0.45
VF885221	706	0.60	0.31
VF885202	706	0.52	0.25
Form 10			
VF822732	708	0.92	0.28
VF822718	708	0.81	0.27
VF822785	708	0.62	0.37
VF822797	708	0.81	0.29
VF822829	708	0.88	0.31
VF822821	708	0.63	0.23
VF884196	708	0.60	0.34
VF884218	708	0.75	0.43
VF884226	708	0.56	0.53
VF884213	708	0.53	0.28
VF884236	708	0.69	0.46
VF884158	708	0.26	0.25
VF885158	708	0.79	0.52
VF885314	708	0.75	0.44

Table J4. Reading Grade 6 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
Form 1			
VF883357	800	0.68	0.33
VF883356	800	0.52	0.33
VF883348	800	0.57	0.37
VF883351	800	0.34	0.06
VF883334	800	0.50	0.30
VF883365	800	0.47	0.28
VF884733	800	0.57	0.28
VF884751	800	0.61	0.19
VF884844	800	0.58	0.36
VF884814	800	0.32	-0.02
VF884886	800	0.39	0.19
VF884880	800	0.67	0.27
VF885006	800	0.41	0.23
VF884659	800	0.66	0.43
Form 2			
VF884676	655	0.74	0.42
VF884630	655	0.63	0.27
Form 3			
VF884677	656	0.35	0.41
VF884693	656	0.77	0.44
Form 4			
VF883345	676	0.49	0.27
VF883367	676	0.74	0.42
VF883338	676	0.44	0.28
VF883354	676	0.75	0.41
VF883361	676	0.90	0.43
VF883331	676	0.88	0.29
VF884740	676	0.52	0.33
VF884772	676	0.43	0.36
VF884853	676	0.18	0.04
VF884808	676	0.40	0.17
VF884876	676	0.36	0.16
VF884857	676	0.66	0.38
VF884624	676	0.80	0.34
VF884631	676	0.49	0.20
Form 5			

---

Average Item			
Accession Number	<i>N</i>	Score	Point Biserial Corr.
VF885203	679	0.49	0.32
VF885189	679	0.61	0.23
VF885141	679	0.52	0.25
VF885148	679	0.72	0.44
VF885178	679	0.65	0.36
VF885098	679	0.62	0.03
VF805047	679	0.76	0.30
VF805061	679	0.66	0.42
VF805054	679	0.59	0.25
VF805825	679	0.47	0.13
VF805822	679	0.58	0.33
VF804276	679	0.46	0.14
VF884658	679	0.46	0.32
VF884665	679	0.84	0.41

---

Form 6			
VF885211	645	0.36	0.23
VF885199	645	0.34	0.29
VF885144	645	0.49	0.32
VF885152	645	0.87	0.31
VF885193	645	0.54	0.11
VF885113	645	0.62	0.34
VF804289	645	0.66	0.34
VF805824	645	0.69	0.22
VF805055	645	0.43	0.31
VF805049	645	0.41	0.35
VF805052	645	0.78	0.29
VF804261	645	0.87	0.33
VF884669	645	0.66	0.39
VF884654	645	0.47	0.29

---



---

Form 7			
VF820258	648	0.88	0.20
VF820467	648	0.63	0.42

---



Average Item			
Accession Number	<i>N</i>	Score	Point Biserial Corr.
VF820394	648	0.91	0.28
VF820332	648	0.75	0.37
VF820193	648	0.30	0.33
VF820442	648	0.60	0.23
VF821684	648	0.25	0.27
VF821664	648	0.74	0.47
VF821580	648	0.70	0.31
VF821704	648	0.85	0.40
VF821619	648	0.62	0.31
VF821542	648	0.90	0.35
VF885013	648	0.18	0.09
VF884657	648	0.47	0.41
Form 8			
VF820463	656	0.50	0.25
VF820281	656	0.79	0.21
VF820354	656	0.69	0.44
VF820218	656	0.65	0.38
VF820310	656	0.74	0.33
VF820457	656	0.77	0.34
VF821558	656	0.82	0.28
VF821572	656	0.81	0.27
VF821721	656	0.75	0.41
VF821673	656	0.53	0.31
VF821709	656	0.71	0.24
VF821429	656	0.73	0.43
VF884974	656	0.67	0.44
VF884689	656	0.34	0.25

Form 9			
VF814337	662	0.57	0.30
VF814311	662	0.83	0.41
VF814388	662	0.68	0.40
VF814382	662	0.76	0.50

Average Item			
Accession Number	<i>N</i>	Score	Point Biserial Corr.
VF814394	662	0.45	0.14
VF814391	662	0.73	0.38
VF883112	662	0.44	0.34
VF883100	662	0.76	0.10
VF883095	662	0.80	0.33
VF883152	662	0.80	0.39
VF883106	662	0.63	0.32
VF883066	662	0.92	0.31
VF885026	662	0.42	0.06
VF884626	662	0.64	0.30
Form 10			
VF814327	680	0.74	0.49
VF814300	680	0.65	0.52
VF814384	680	0.37	0.13
VF814358	680	0.89	0.36
VF814392	680	0.63	0.29
VF814393	680	0.37	0.26
VF883144	680	0.70	0.37
VF883088	680	0.63	0.20
VF883072	680	0.44	0.19
VF883158	680	0.50	0.19
VF883061	680	0.71	0.52
VF883052	680	0.82	0.37
VF884988	680	0.63	0.40
VF884628	680	0.51	0.35

Table J5. Reading Grade 7 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF820419	768	0.60	0.42
VF820422	768	0.65	0.47
VF820444	768	0.89	0.36
VF820435	768	0.35	0.26
VF820404	768	0.42	0.34
VF820464	768	0.59	0.21
VF864796	768	0.73	0.38
VF864756	768	0.41	0.35
VF864677	768	0.67	0.43
VF864684	768	0.70	0.31
VF864681	768	0.46	0.22
VF864667	768	0.77	0.30
VF885647	768	0.64	0.52
VF885607	768	0.55	0.25
Form 2			
VF885786	660	0.86	0.30
VF885813	660	0.49	0.30
Form 3			
VF885820	1319	0.49	0.50
VF885815	668	0.67	0.49
Form 4			
VF820466	671	0.47	0.25
VF820430	671	0.72	0.42
VF820412	671	0.77	0.31
VF820449	671	0.69	0.25
VF820438	671	0.88	0.45
VF820391	671	0.70	0.30
VF864785	671	0.47	0.24
VF864676	671	0.84	0.37
VF864750	671	0.51	0.23
VF864789	671	0.19	-0.02
VF864685	671	0.76	0.43
VF864668	671	0.54	0.28
VF885757	671	0.71	0.23
VF885485	671	0.83	0.31

Form 5

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF814792	679	0.81	0.39
VF814759	679	0.58	0.24
VF814742	679	0.68	0.27
VF814788	679	0.37	0.28
VF814770	679	0.49	0.50
VF814720	679	0.76	0.35
VF865166	679	0.78	0.32
VF865194	679	0.35	0.08
VF865169	679	0.52	0.19
VF865185	679	0.51	0.47
VF865189	679	0.73	0.20
VF865164	679	0.16	0.01
VF885809	679	0.74	0.32
VF885612	679	0.47	0.32
Form 6			
VF814809	651	0.49	0.16
VF814766	651	0.61	0.47
VF814800	651	0.34	0.08
VF814826	651	0.78	0.37
VF814781	651	0.80	0.24
VF814702	651	0.92	0.32
VF865186	651	0.39	0.22
VF865195	651	0.60	0.16
VF865182	651	0.59	0.32
VF865187	651	0.46	0.27
VF865165	651	0.83	0.35
VF865141	651	0.76	0.38
VF885820	1319	0.49	0.50
VF885443	651	0.64	0.52

Form 7			
VF865426	675	0.78	0.36

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF865456	675	0.48	0.06
VF865482	675	0.77	0.20
VF865473	675	0.45	0.41
VF865624	675	0.64	0.31
VF865614	675	0.47	0.32
VF883991	675	0.56	0.27
VF883986	675	0.63	0.22
VF884006	675	0.76	0.36
VF884003	675	0.48	0.32
VF883998	675	0.64	0.28
VF883976	675	0.77	0.49
VF885659	675	0.35	0.28
VF910031	675	0.74	0.49
Form 8			
VF865413	666	0.70	0.34
VF865388	666	0.90	0.44
VF865477	666	0.80	0.32
VF865494	666	0.55	0.27
VF865627	666	0.74	0.39
VF906623	666	0.44	0.23
VF883995	666	0.67	0.13
VF883997	666	0.93	0.35
VF884008	666	0.66	0.45
VF884005	666	0.53	0.32
VF883972	666	0.82	0.35
VF883999	666	0.64	0.36
VF885797	666	0.86	0.46
VF885440	666	0.44	0.21

Form 9			
VF884891	679	0.57	0.22
VF885076	679	0.64	0.35

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF885063	679	0.22	0.05
VF884887	679	0.67	0.38
VF885031	679	0.36	0.16
VF884846	679	0.91	0.38
VF864898	679	0.88	0.25
VF864910	679	0.28	-0.05
VF865078	679	0.44	0.25
VF865063	679	0.53	0.26
VF865100	679	0.46	0.17
VF865094	679	0.40	0.32
VF885398	679	0.60	0.37
VF885385	679	0.79	0.51
Form 10			
VF884878	678	0.65	0.48
VF884859	678	0.53	0.32
VF885046	678	0.62	0.31
VF885072	678	0.56	0.35
VF885060	678	0.47	0.16
VF884855	678	0.91	0.42
VF864902	678	0.74	0.48
VF865004	678	0.41	0.07
VF865057	678	0.69	0.53
VF865072	678	0.76	0.38
VF865104	678	0.56	0.45
VF865088	678	0.62	0.34
VF885375	678	0.77	0.33
VF885333	678	0.77	0.40

Table J6. Reading Grade 8 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF865107	793	0.63	0.27
VF864994	793	0.82	0.42
VF865075	793	0.47	0.17
VF865060	793	0.52	0.32
VF994816	793	0.42	0.18
VF997001	793	0.35	0.19
VF819971	793	0.42	0.21
VF819982	793	0.43	0.10
VF820236	793	0.41	0.27
VF820165	793	0.51	0.32
VF820159	793	0.71	0.26
VF820261	793	0.59	0.35
VF883743	793	0.74	0.34
VF883621	793	0.65	0.41
Form 2			
VF820781	659	0.57	0.37
VF820771	659	0.47	0.31
VF820720	659	0.56	0.31
VF820786	659	0.67	0.16
VF820796	659	0.46	0.20
VF820792	659	0.65	0.24
VF883642	659	0.58	0.28
VF883685	659	0.55	0.17
Form 3			
VF820777	662	0.82	0.44
VF820740	662	0.76	0.39
VF820727	662	0.61	0.48
VF820734	662	0.91	0.41
VF820750	662	0.75	0.32
VF820801	662	0.91	0.38
VF883708	662	0.60	0.30
VF883653	662	0.89	0.43
Form 4			

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF865101	668	0.80	0.27
VF865111	668	0.85	0.34
VF865050	668	0.50	0.36
VF865091	668	0.78	0.36
VF865171	668	0.68	0.28
VF865178	668	0.33	0.13
VF820011	668	0.65	0.21
VF819976	668	0.54	0.45
VF820170	668	0.46	0.34
VF820174	668	0.75	0.47
VF820025	668	0.80	0.42
VF820249	668	0.76	0.23
VF883823	668	0.78	0.40
VF883674	668	0.54	0.33
Form 5			
VF866201	679	0.88	0.34
VF866195	679	0.74	0.45
VF866192	679	0.53	0.12
VF866296	679	0.59	0.31
VF866316	679	0.92	0.28
VF866341	679	0.81	0.44
VF867326	679	0.73	0.46
VF867246	679	0.71	0.45
VF867293	679	0.62	0.45
VF867267	679	0.44	0.12
VF867355	679	0.37	0.28
VF867368	679	0.60	0.28
VF883554	679	0.50	0.24
VF883624	679	0.70	0.45

Form 6			
VF866325	655	0.73	0.35



Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF866173	655	0.82	0.41
VF866186	655	0.63	0.09
VF866307	655	0.93	0.45
VF866228	655	0.73	0.33
VF866331	655	0.87	0.42
VF867197	655	0.76	0.44
VF867239	655	0.66	0.35
VF867333	655	0.70	0.46
VF867305	655	0.62	0.21
VF867338	655	0.68	0.37
VF867274	655	0.73	0.34
VF883655	655	0.58	0.38
VF883680	655	0.67	0.16
Form 7			
VF813900	657	0.74	0.39
VF813664	657	0.69	0.16
VF813654	657	0.94	0.39
VF813646	657	0.26	0.26
VF813874	657	0.81	0.49
VF813924	657	0.62	0.40
VF813668	657	0.89	0.37
VF813648	657	0.84	0.44
VF813641	657	0.87	0.25
VF813649	657	0.81	0.46
VF813671	657	0.59	0.31
VF813673	657	0.64	0.32
VF883695	657	0.35	-0.09
VF883716	657	0.81	0.34

Form 8			
VF813601	667	0.58	0.33
VF813639	667	0.61	0.38

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF813842	667	0.78	0.38
VF813879	667	0.63	0.36
VF864943	667	0.83	0.47
VF813904	667	0.88	0.36
VF813667	667	0.56	0.36
VF813653	667	0.28	0.04
VF813655	667	0.76	0.37
VF813657	667	0.85	0.37
VF813645	667	0.35	0.26
VF813670	667	0.80	0.24
VF883726	667	0.21	0.20
VF883732	667	0.39	0.24

---

Form 9

VF812806	662	0.83	0.28
VF812982	662	0.52	0.20
VF812809	662	0.63	0.39
VF812818	662	0.58	0.39
VF812971	662	0.66	0.25
VF812965	662	0.95	0.29
VF884543	662	0.37	0.16
VF884544	662	0.44	0.16
VF884581	662	0.28	0.07
VF884589	662	0.54	0.15
VF884613	662	0.53	0.37
VF884609	662	0.44	0.25
VF883701	662	0.65	0.27
VF883817	662	0.32	0.25

---



---

Form 10

VF812813	679	0.62	0.16
VF812960	679	0.55	0.20
VF812799	679	0.72	0.38

---

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF812800	679	0.53	0.25
VF812796	679	0.45	0.32
VF812988	679	0.67	0.17
VF884547	679	0.28	-0.01
VF884552	679	0.53	0.21
VF884586	679	0.17	-0.11
VF884583	679	0.80	0.41
VF884614	679	0.15	-0.12
VF884606	679	0.37	0.06
VF883629	679	0.46	0.31
VF883736	679	0.42	0.39

---

Mathematics

Table J7. Mathematics Grade 3 Classical Statistics for Field Test Items

Average Item			
Accession Number	<i>N</i>	Score	Point Biserial Corr.
Form 1			
VF803080	789	0.74	0.41
VF867016	789	0.56	0.48
VF865397	789	0.76	0.38
VF822811	789	0.46	0.43
VF867203	789	0.39	0.31
VF867073	789	0.47	0.31
VF866360	789	0.85	0.48
VF867061	789	0.62	0.44
VF866931	789	0.56	0.33
VF865404	789	0.34	0.28
VF865420	789	0.38	0.56
VF866941	789	0.28	0.19
VF803183	789	0.57	0.43
Form 2			
VF803121	724	0.53	0.28
VF821403	724	0.85	0.32
VF822819	724	0.26	0.14
VF867001	724	0.68	0.43
VF819629	724	0.96	0.12
VF866364	724	0.78	0.36
VF867181	724	0.22	0.10
VF818296	724	0.76	0.31
Form 3			
VF803161	726	0.63	0.35
VF821680	726	0.48	0.39
VF737752	726	0.46	0.56
VF867075	726	0.66	0.33
VF866952	726	0.48	0.42
VF865570	726	0.55	0.50
VF821770	726	0.46	0.26
VF803199	726	0.60	0.45

Form 4

Average Item			
Accession Number	<i>N</i>	Score	Point Biserial Corr.
VF803172	723	0.54	0.43
VF866981	723	0.76	0.46
VF737767	723	0.60	0.51
VF740960	723	0.50	0.42
VF866256	723	0.40	0.43
VF819315	723	0.63	0.47
VF740954	723	0.72	0.49
VF865462	723	0.68	0.45
Form 5			
VF866354	717	0.81	0.45
VF866961	717	0.52	0.48
VF865381	717	0.67	0.53
VF740917	717	0.46	0.49
VF740830	717	0.52	0.26
VF821698	717	0.44	0.29
VF822773	717	0.10	0.04
VF819577	717	0.74	0.45
VF387508	717	0.88	0.31
VF865285	717	0.84	0.40
VF740915	717	0.24	0.39
VF819363	717	0.33	0.28
VF822685	717	0.58	0.49
Form 6			
VF866235	740	0.67	0.42
VF867176	740	0.17	0.32
VF821665	740	0.67	0.28
VF819348	740	0.76	0.24
VF866906	740	0.57	0.44
VF819555	740	0.61	0.48
VF865389	740	0.63	0.46
VF740959	740	0.33	0.09
VF865371	740	0.50	0.44
VF821723	740	0.45	0.34
VF867224	740	0.76	0.43
VF821767	740	0.39	0.29
VF822716	740	0.44	0.32
Form 7			
VF866264	744	0.70	0.40
VF821729	744	0.56	0.40

Accession Number	N	Average Item	
		Score	Point Biserial Corr.
VF865449	744	0.88	0.23
VF819660	744	0.35	0.27
VF803266	744	0.46	0.46
VF819299	744	0.52	0.39
VF865405	744	0.57	0.41
VF865478	744	0.44	0.30
VF737761	744	0.23	-0.10
VF740949	744	0.66	0.21
VF821407	744	0.91	0.31
VF803242	744	0.49	0.46
VF822742	744	0.39	0.09
Form 8			
VF866946	741	0.58	0.56
VF866996	741	0.32	0.28
VF819669	741	0.44	0.23
VF821481	741	0.82	0.37
VF819375	741	0.42	0.52
VF822709	741	0.66	0.38
VF803290	741	0.64	0.34
VF818374	741	0.60	0.55
VF865302	741	0.29	0.22
VF821738	741	0.91	0.33
VF819676	741	0.84	0.18
VF818365	741	0.76	0.28
VF822822	741	0.73	0.44
Form 9			
VF822784	735	0.30	0.16
VF866988	735	0.86	0.43
VF867009	735	0.43	0.37
VF822725	735	0.26	0.22
VF865323	735	0.41	0.31
VF819654	735	0.44	0.32
VF819622	735	0.96	0.22
VF867066	735	0.84	0.46
VF865468	735	0.37	0.34
VF819598	735	0.24	0.13
VF866898	735	0.75	0.34
VF740957	735	0.21	0.18
VF821652	735	0.54	0.34
Form 10			
VF740890	730	0.51	0.40
VF821745	730	0.49	0.33
VF819675	730	0.89	0.29

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF819337	730	0.47	0.49
VF803307	730	0.42	0.47
VF865488	730	0.51	0.22
VF866888	730	0.35	0.41
VF819543	730	0.56	0.46
VF867023	730	0.33	0.19
VF819639	730	0.15	0.35
VF865414	730	0.67	0.38
VF865316	730	0.84	0.45
VF821387	730	0.45	0.46

Table J8. Mathematics Grade 4 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF816041	766	0.61	0.34
VF865554	766	0.64	0.31
VF880261	766	0.32	0.35
VF867083	766	0.56	0.28
VF741942	766	0.31	0.19
VF800697	766	0.67	0.12
VF866686	766	0.42	0.24
VF880294	766	0.57	0.53
VF816026	766	0.46	0.48
VF816159	766	0.58	0.54
VF801227	766	0.36	0.31
VF823081	766	0.47	0.40
VF823371	766	0.71	0.29
VF866870	766	0.34	0.23
Form 2			
VF816048	704	0.72	0.42
VF823138	704	0.70	0.49
VF880252	704	0.74	0.46
VF880336	704	0.30	0.40
Form 3			
VF822848	699	0.77	0.41
VF741944	699	0.84	0.25
VF823036	699	0.23	0.26
VF880325	699	0.58	0.28
Form 4			
VF866662	700	0.62	0.42
VF880413	700	0.53	0.49
VF815936	700	0.41	0.37
VF741929	700	0.32	0.30

Form 5



Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF816151	694	0.85	0.34
VF880305	694	0.35	0.26
VF823330	694	0.30	0.38
VF880421	694	0.47	0.55
VF866416	694	0.37	0.26
VF867088	694	0.42	0.39
VF816057	694	0.66	0.38
VF801810	694	0.44	0.37
VF815880	694	0.70	0.30
VF816028	694	0.66	0.41
VF864158	694	0.13	0.22
VF815975	694	0.98	0.15
VF866672	694	0.36	0.42
VF866699	694	0.55	0.50
Form 6			
VF822854	687	0.91	0.34
VF815849	687	0.26	0.15
VF866677	687	0.72	0.39
VF801214	687	0.42	0.40
VF815888	687	0.57	0.21
VF815962	687	0.69	0.36
VF880334	687	0.29	0.11
VF741950	687	0.75	0.41
VF864145	687	0.46	0.25
VF863975	687	0.83	0.27
VF815942	687	0.79	0.26
VF866847	687	0.32	0.08
VF823410	687	0.38	0.46
VF880341	687	0.24	0.40

Form 7			
VF822864	690	0.33	0.30

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF815875	690	0.66	0.47
VF823304	690	0.68	0.15
VF880328	690	0.34	0.41
VF800875	690	0.73	0.50
VF741919	690	0.56	0.30
VF741945	690	0.37	0.43
VF801835	690	0.69	0.24
VF864141	690	0.24	0.12
VF867091	690	0.32	0.10
VF864104	690	0.38	0.54
VF864153	690	0.32	0.46
VF815948	690	0.74	0.33
VF815909	690	0.87	0.43
Form 8			
VF822870	697	0.47	0.40
VF865651	697	0.52	0.43
VF880443	697	0.37	0.47
VF867078	697	0.31	0.44
VF741948	697	0.37	0.17
VF866702	697	0.05	0.05
VF864149	697	0.49	0.33
VF822874	697	0.26	0.20
VF866381	697	0.61	0.51
VF815957	697	0.65	0.38
VF866410	697	0.22	0.21
VF864100	697	0.71	0.38
VF800889	697	0.50	0.54
VF866368	697	0.20	0.18

Form 9			
VF864111	688	0.41	0.38
VF864078	688	0.83	0.39

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF742706	688	0.64	0.37
VF866714	688	0.83	0.43
VF866709	688	0.59	0.36
VF816162	688	0.52	0.40
VF866402	688	0.76	0.31
VF823145	688	0.81	0.31
VF866392	688	0.21	0.11
VF741947	688	0.35	0.23
VF866857	688	0.55	0.43
VF864035	688	0.18	0.35
VF741936	688	0.13	0.18
VF815303	688	0.13	0.11
Form 10			
VF823141	701	0.37	0.33
VF864051	701	0.77	0.46
VF867086	701	0.65	0.47
VF866691	701	0.84	0.13
VF867084	701	0.37	0.40
VF815900	701	0.53	0.29
VF866696	701	0.42	0.30
VF823000	701	0.69	0.51
VF866830	701	0.61	0.46
VF880274	701	0.55	0.34
VF741924	701	0.74	0.40
VF741949	701	0.34	0.45
VF880269	701	0.82	0.45
VF823196	701	0.52	0.37

Table J9. Mathematics Grade 5 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF802791	757	0.13	0.31
VF801975	757	0.25	0.33
VF798083	757	0.67	0.49
VF819978	757	0.62	0.48
VF866083	757	0.54	0.40
VF736258	757	0.24	0.28
VF740925	757	0.62	0.42
VF823764	757	0.71	0.49
VF880826	757	0.48	0.27
VF741570	757	0.72	0.33
VF816137	757	0.37	0.22
VF823490	757	0.54	0.46
VF864609	757	0.43	0.53
VF819989	757	0.42	0.45
Form 2			
VF802051	705	0.47	0.15
VF741381	705	0.26	0.25
VF864581	705	0.35	0.10
VF865968	705	0.26	0.24
Form 3			
VF823809	704	0.75	-0.12
VF819955	704	0.25	0.27
VF823498	704	0.08	0.08
VF880803	704	0.57	0.48
Form 4			
VF823474	716	0.57	0.48
VF740894	716	0.83	0.39
VF802821	716	0.25	0.35
VF736492	716	0.63	0.49

Form 5

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF864587	711	0.62	0.37
VF741081	711	0.59	0.33
VF741507	711	0.30	0.23
VF880721	711	0.42	0.41
VF802763	711	0.37	0.17
VF864638	711	0.21	0.23
VF823729	711	0.54	0.38
VF823759	711	0.33	0.41
VF801992	711	0.70	0.46
VF736482	711	0.87	0.26
VF816183	711	0.38	0.40
VF736475	711	0.34	0.17
VF816152	711	0.71	0.31
VF741093	711	0.37	0.40
Form 6			
VF802778	708	0.34	0.27
VF741450	708	0.77	0.38
VF816021	708	0.52	0.43
VF866061	708	0.46	0.43
VF864521	708	0.54	0.58
VF741371	708	0.82	0.38
VF864548	708	0.26	0.35
VF816005	708	0.26	0.39
VF823504	708	0.33	0.05
VF823790	708	0.61	0.25
VF741416	708	0.43	0.21
VF864628	708	0.65	0.14
VF736633	708	0.70	0.47
VF864671	708	0.25	-0.03

Form 7			
VF823638	697	0.70	0.42

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF797033	697	0.48	0.11
VF865989	697	0.35	0.58
VF864536	697	0.25	0.39
VF741106	697	0.75	0.31
VF802069	697	0.32	0.13
VF823779	697	0.43	0.11
VF880813	697	0.54	0.49
VF864590	697	0.42	0.25
VF815866	697	0.43	0.15
VF819994	697	0.56	0.37
VF864614	697	0.63	0.35
VF866022	697	0.25	0.30
VF797963	697	0.55	0.41
Form 8			
VF741941	693	0.60	0.39
VF865997	693	0.22	0.31
VF866065	693	0.54	0.32
VF741573	693	0.73	0.29
VF802894	693	0.51	0.27
VF880786	693	0.37	0.46
VF866009	693	0.84	0.37
VF802089	693	0.29	0.16
VF741405	693	0.29	0.13
VF802032	693	0.45	0.07
VF736503	693	0.84	0.40
VF741052	693	0.65	0.42
VF823652	693	0.55	0.54
VF802860	693	0.41	-0.02

Form 9			
VF864604	691	0.49	0.37
VF815846	691	0.51	0.36

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF741193	691	0.76	0.34
VF815902	691	0.40	0.41
VF801897	691	0.53	0.62
VF864641	691	0.10	0.11
VF741382	691	0.66	0.33
VF823838	691	0.85	0.29
VF797110	691	0.51	0.37
VF797938	691	0.63	0.42
VF802014	691	0.64	0.55
VF736495	691	0.74	0.32
VF866037	691	0.37	0.38
VF815982	691	0.35	0.42
Form 10			
VF819900	695	0.84	0.25
VF823819	695	0.45	0.14
VF802870	695	0.56	0.21
VF736438	695	0.26	0.23
VF741539	695	0.73	0.43
VF740936	695	0.73	0.42
VF736524	695	0.78	0.39
VF741389	695	0.49	0.32
VF864618	695	0.29	0.06
VF815953	695	0.65	0.34
VF802847	695	0.09	0.26
VF866034	695	0.44	0.01
VF864557	695	0.30	0.39
VF866103	695	0.57	0.31

Table J10. Mathematics Grade 6 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF862699	806	0.61	0.42
VF741557	806	0.42	0.44
VF741723	806	0.51	0.24
VF810665	806	0.30	0.30
VF883019	806	0.58	0.19
VF741771	806	0.50	0.26
VF797171	806	0.38	0.10
VF810667	806	0.27	0.39
VF797964	806	0.32	0.26
VF865682	806	0.28	0.27
VF822031	806	0.48	0.30
VF803280	806	0.53	0.41
VF741572	806	0.48	0.40
VF797954	806	0.19	-0.18
Form 2			
VF741728	655	0.50	0.52
VF803302	655	0.52	0.47
VF865661	655	0.35	0.16
VF821954	655	0.31	0.25
VF862885	664	0.38	0.12
Form 3			
VF803311	664	0.47	0.15
VF865649	664	0.61	0.41
VF882800	664	0.53	0.27
Form 4			
VF882956	679	0.79	0.32
VF803328	679	0.43	0.31
VF865678	679	0.56	0.38
VF821946	679	0.57	0.30



Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 5			
VF741574	681	0.49	0.21
VF862858	681	0.60	0.29
VF741711	681	0.56	0.37
VF741566	681	0.29	0.09
VF809034	681	0.70	0.37
VF741928	681	0.77	0.39
VF812407	681	0.39	0.33
VF803399	681	0.41	0.33
VF797120	681	0.28	0.07
VF741533	681	0.44	0.32
VF821929	681	0.74	0.31
VF797970	681	0.77	0.48
VF741515	681	0.28	0.29
VF866206	681	0.39	0.09
Form 6			
VF883002	662	0.58	0.03
VF809062	662	0.67	0.41
VF862786	662	0.56	0.50
VF882803	662	0.25	-0.09
VF741578	662	0.53	0.30
VF797163	662	0.45	0.16
VF741576	662	0.29	0.15
VF803386	662	0.31	0.34
VF741934	662	0.44	0.30
VF865621	662	0.25	0.21
VF821920	662	0.65	0.30
VF797977	662	0.73	0.46
VF822004	662	0.71	0.47
VF797944	662	0.14	-0.07

Form 7			
VF741690	637	0.62	0.16

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF810689	637	0.60	0.52
VF862804	637	0.81	0.27
VF812185	637	0.60	0.34
VF882963	637	0.38	0.17
VF821992	637	0.41	0.16
VF865650	637	0.62	0.29
VF741935	637	0.43	0.36
VF866265	637	0.79	0.25
VF865654	637	0.28	0.23
VF882811	637	0.42	0.40
VF821988	637	0.66	0.25
VF821963	637	0.62	0.45
VF866221	637	0.64	0.40
Form 8			
VF811515	654	0.40	0.44
VF741538	654	0.75	0.49
VF882993	654	0.51	0.34
VF809839	654	0.62	0.58
VF741699	654	0.15	0.23
VF883062	654	0.30	0.10
VF882780	654	0.39	0.05
VF865635	654	0.50	0.47
VF882808	654	0.36	0.26
VF797981	654	0.56	0.25
VF883067	654	0.62	0.25
VF797996	654	0.74	0.42
VF866290	654	0.54	0.31
VF821976	654	0.46	0.42

Form 9			
VF741668	664	0.33	0.34
VF741549	664	0.76	0.45

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF741781	664	0.87	0.19
VF810696	664	0.52	0.49
VF809076	664	0.16	0.06
VF821906	664	0.78	0.30
VF882789	664	0.23	0.05
VF803317	664	0.55	0.39
VF883071	664	0.81	0.33
VF865668	664	0.48	0.31
VF803393	664	0.76	0.43
VF821939	664	0.22	0.14
VF822023	664	0.68	0.39
VF866230	664	0.57	0.42
Form 10			
VF862813	657	0.48	0.22
VF810701	657	0.35	0.24
VF741692	657	0.32	0.33
VF741562	657	0.32	-0.06
VF811529	657	0.27	0.10
VF865671	657	0.52	0.39
VF882795	657	0.56	0.26
VF821998	657	0.70	0.40
VF866278	657	0.44	0.32
VF803324	657	0.59	0.41
VF741859	657	0.38	0.10
VF803293	657	0.44	0.38
VF822007	657	0.37	0.33
VF866301	657	0.13	0.00

Table J11. Mathematics Grade 7 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF880308	779	0.59	0.37
VF880331	779	0.27	0.32
VF866890	779	0.60	0.46
VF880323	779	0.50	0.21
VF823091	779	0.28	0.15
VF736963	779	0.35	0.27
VF883244	779	0.53	0.35
VF866547	779	0.20	0.14
VF736931	779	0.43	0.36
VF882920	779	0.30	0.39
VF867243	779	0.47	0.29
VF883150	779	0.57	0.28
VF880171	779	0.23	0.22
VF736947	779	0.61	0.43
Form 2			
VF818173	657	0.24	0.26
VF799825	657	0.79	0.39
VF882559	657	0.33	0.12
VF866826	657	0.34	0.08
VF866499	662	0.43	0.40
VF822880	662	0.44	0.37
Form 3			
VF800078	662	0.47	0.15
VF822986	662	0.59	0.30
Form 4			
VF813483	674	0.32	0.49
VF819351	674	0.28	0.31
VF867315	674	0.41	0.18
VF882910	674	0.53	0.34

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 5			
VF736959	681	0.36	0.33
VF867219	681	0.29	0.20
VF818177	681	0.35	0.42
VF880250	681	0.38	0.26
VF736957	681	0.31	0.05
VF867395	681	0.58	0.32
VF870864	681	0.49	0.26
VF882715	681	0.28	0.24
VF867292	681	0.18	0.04
VF818184	681	0.47	0.30
VF736938	681	0.60	0.49
VF822889	681	0.17	0.32
VF866506	681	0.28	0.24
VF819294	681	0.47	0.42
Form 6			
VF822884	650	0.14	0.18
VF817427	650	0.39	0.19
VF800136	650	0.65	0.33
VF813502	650	0.60	0.49
VF813530	650	0.38	0.27
VF819535	650	0.21	0.22
VF813096	650	0.46	0.42
VF736961	650	0.32	0.30
VF867377	650	0.26	0.00
VF818347	650	0.83	0.10
VF866386	650	0.41	0.28
VF881807	650	0.67	0.51
VF868691	650	0.33	0.24
VF883264	650	0.40	0.44

Form 7			
VF800144	686	0.62	0.41

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF823026	686	0.27	0.21
VF818181	686	0.37	0.40
VF867610	686	0.32	0.30
VF866421	686	0.69	0.28
VF818335	686	0.47	0.17
VF880897	686	0.73	0.44
VF866491	686	0.25	0.46
VF800055	686	0.80	0.40
VF867323	686	0.43	0.04
VF736940	686	0.35	0.39
VF883129	686	0.58	0.42
VF866539	686	0.32	0.24
VF882732	686	0.71	0.37
Form 8			
VF818182	659	0.36	0.35
VF866376	659	0.18	-0.01
VF800133	659	0.30	0.44
VF867183	659	0.24	0.21
VF866963	659	0.32	0.49
VF800103	659	0.44	0.49
VF867307	659	0.66	0.23
VF883138	659	0.38	0.19
VF869623	659	0.56	0.49
VF822997	659	0.76	0.38
VF819696	659	0.23	0.12
VF882691	659	0.49	0.42
VF882746	659	0.34	0.22
VF866531	659	0.40	0.36

Form 9			
VF736954	674	0.53	0.50
VF818180	674	0.39	0.20

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF867060	674	0.26	0.36
VF818174	674	0.68	0.37
VF813104	674	0.35	0.32
VF819358	674	0.39	0.28
VF867260	674	0.66	0.54
VF866401	674	0.35	0.10
VF819694	674	0.38	-0.06
VF882739	674	0.62	0.24
VF813100	674	0.46	0.10
VF883156	674	0.46	0.46
VF819306	674	0.67	0.48
VF867401	674	0.40	0.29
Form 10			
VF823079	677	0.12	0.20
VF880312	677	0.35	0.13
VF867038	677	0.44	0.22
VF818183	677	0.46	0.35
VF813490	677	0.51	0.51
VF866423	677	0.26	0.24
VF736941	677	0.27	0.24
VF882946	677	0.76	0.41
VF867365	677	0.27	-0.03
VF867256	677	0.48	0.43
VF880886	677	0.58	0.20
VF799837	677	0.74	0.43
VF818361	677	0.59	0.20
VF883220	677	0.34	-0.02

Table J12. Mathematics Grade 8 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF802927	828	0.39	0.37
VF885497	828	0.40	0.39
VF802937	828	0.48	0.31
VF812962	828	0.55	0.54
VF810708	828	0.73	0.35
VF810643	828	0.28	0.19
VF823784	828	0.63	0.33
VF880641	828	0.28	0.39
VF865981	828	0.39	0.15
VF883670	828	0.48	0.39
VF812445	828	0.19	0.10
VF809001	828	0.38	0.48
VF883722	828	0.28	0.22
Form 2			
VF803463	651	0.88	0.37
VF885510	651	0.53	0.46
VF865673	651	0.35	0.24
VF809017	651	0.26	0.24
VF865996	651	0.41	0.41
VF883593	651	0.65	0.36
VF823806	651	0.55	0.31
VF885529	651	0.30	0.19
Form 3			
VF823449	661	0.20	0.03
VF885500	661	0.51	0.33
VF880669	661	0.46	0.50
VF885549	661	0.35	0.37
VF863266	661	0.33	0.17
VF803474	661	0.71	0.40
VF866052	661	0.50	0.32
VF880638	661	0.29	0.24
Form 4			
VF812762	657	0.42	0.37
VF880849	657	0.50	0.26
VF866191	657	0.24	0.16
VF880501	657	0.64	0.42
VF863351	657	0.53	0.15
VF883663	657	0.42	0.25
VF883692	657	0.34	0.16
VF822454	657	0.60	0.46



Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 5			
VF812728	664	0.13	0.05
VF880420	664	0.43	0.37
VF802935	664	0.60	0.18
VF802924	664	0.39	0.13
VF883687	664	0.32	0.15
VF883715	664	0.66	0.43
VF823932	664	0.36	0.11
VF804251	664	0.57	0.36
VF866035	664	0.27	0.04
VF804267	664	0.46	0.24
VF822412	664	0.60	0.47
VF811990	664	0.17	0.20
VF812983	664	0.42	0.36
Form 6			
VF802939	637	0.23	0.16
VF880528	637	0.57	0.28
VF823336	637	0.17	-0.08
VF802931	637	0.73	0.49
VF883657	637	0.52	0.16
VF883648	637	0.79	0.35
VF823206	637	0.37	0.44
VF880628	637	0.45	0.42
VF863290	637	0.29	0.25
VF809061	637	0.47	0.24
VF863242	637	0.57	0.34
VF812970	637	0.34	0.19
VF804260	637	0.57	0.44
Form 7			
VF885577	652	0.52	0.45
VF880798	652	0.65	0.40
VF823406	652	0.18	0.02
VF802936	652	0.73	0.44
VF866220	652	0.38	0.24
VF809049	652	0.58	0.41
VF823366	652	0.34	0.34
VF863280	652	0.37	0.44
VF823921	652	0.61	0.30
VF822425	652	0.53	0.36
VF804282	652	0.29	0.08
VF885555	652	0.61	0.29
VF880680	652	0.56	0.39
Form 8			
VF885519	688	0.71	0.31

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF880697	688	0.38	0.18
VF802934	688	0.68	0.39
VF823444	688	0.52	0.44
VF883698	688	0.53	0.34
VF883641	688	0.54	0.25
VF805819	688	0.41	0.34
VF823294	688	0.52	0.47
VF880525	688	0.42	0.21
VF822402	688	0.25	0.28
VF822441	688	0.46	0.45
VF810683	688	0.43	0.16
VF880559	688	0.34	0.05
Form 9			
VF802938	663	0.43	0.33
VF885483	663	0.54	0.32
VF880493	663	0.52	0.30
VF823307	663	0.55	0.45
VF804256	663	0.55	0.29
VF880675	663	0.62	0.51
VF823748	663	0.35	0.11
VF863346	663	0.58	0.46
VF809838	663	0.66	0.40
VF866181	663	0.32	0.22
VF812997	663	0.46	0.36
VF883707	663	0.44	0.42
VF866064	663	0.48	0.38
Form 10			
VF812743	683	0.31	0.21
VF880512	683	0.45	0.41
VF802932	683	0.49	0.06
VF823432	683	0.66	0.45
VF880646	683	0.57	0.43
VF885561	683	0.32	0.13
VF823736	683	0.40	0.31
VF880613	683	0.69	0.44
VF866019	683	0.30	0.25
VF863323	683	0.15	0.12
VF865675	683	0.46	0.34
VF823848	683	0.54	0.32
VF822465	683	0.49	0.34

Science

Table J13. Science Grade 4 Classical Statistics for Field Test Items

Accession Number	N	Average Item Score	Point Biserial Corr.
Form 1			
VF800147	750	0.66	0.22
VF800163	750	0.61	0.26
VF800182	750	0.63	0.34
VF800193	750	0.63	0.46
VF801233	750	0.31	0.30
VF656005	750	0.82	0.35
VF656072	750	0.55	0.33
VF656851	750	0.43	0.03
VF671249	750	0.66	0.44
VF671205	750	0.30	0.13
VF671215	750	0.38	0.03
VF671241	750	0.22	0.06
VF800026	691	0.56	0.33
Form 2			
VF800044	691	0.38	0.19
VF800096	691	0.65	0.26
VF800109	691	0.80	0.38
VF800118	691	0.71	0.11
VF656085	691	0.60	0.33
VF656089	691	0.06	-0.01
VF671330	691	0.80	0.48
VF671338	691	0.45	0.42
VF671340	691	0.33	0.25
VF671344	691	0.38	0.19
VF671357	691	0.62	0.35
Form 3			
VF656139	685	0.47	0.37
VF656143	685	0.52	0.33
VF656150	685	0.79	0.14
VF656157	685	0.73	0.48
VF656175	685	0.64	0.17
VF656106	685	0.54	0.33
VF656109	685	0.55	0.35
VF656178	685	0.68	0.57
VF656181	685	0.64	0.42
VF656190	685	0.46	0.40
VF656195	685	0.43	0.28
VF656227	685	0.63	0.46
Form 4			

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF656180	691	0.66	0.38
VF656183	691	0.72	0.19
VF656202	691	0.51	0.39
VF656217	691	0.58	0.17
VF656228	691	0.49	0.25
VF800157	691	0.89	0.35
VF800137	691	0.97	0.21
VF815566	691	0.66	0.29
VF815606	691	0.24	-0.04
VF815613	691	0.23	0.03
VF815620	691	0.35	0.21
VF815623	691	0.25	0.11
Form 5			
VF814118	683	0.71	0.26
VF814057	683	0.93	0.27
VF814125	683	0.37	0.14
VF814129	683	0.30	0.19
VF814143	683	0.59	0.43
VF800175	683	0.26	0.19
VF801217	683	0.49	0.19
VF800030	683	0.56	0.29
VF799850	683	0.50	0.28
VF800090	683	0.67	0.42
VF800059	683	0.50	0.26
VF800018	683	0.22	0.21
Form 6			
VF656211	704	0.55	0.21
VF656239	704	0.74	0.25
VF656220	704	0.63	0.35
VF656177	704	0.85	0.38
VF656245	704	0.82	0.32
VF671126	704	0.67	0.28
VF671189	704	0.34	0.27
VF656237	704	0.30	0.02
VF656226	704	0.59	0.41
VF656221	704	0.51	0.40
VF656218	704	0.28	0.16
VF656179	704	0.68	0.47
Form 7			
VF671286	705	0.82	0.18

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF671354	705	0.17	0.05
VF671349	705	0.52	0.35
VF671365	705	0.76	0.38
VF671318	705	0.70	0.16
VF815653	705	0.50	0.02
VF815662	705	0.69	0.41
VF814054	705	0.43	0.40
VF814067	705	0.71	0.45
VF814152	705	0.81	0.41
VF814064	705	0.60	0.37
VF814112	705	0.77	0.38
Form 8			
VF815516	703	0.77	0.29
VF814294	703	0.75	0.32
VF814286	703	0.52	0.38
VF814283	703	0.65	0.45
VF814431	703	0.78	0.42
VF801255	703	0.46	0.33
VF801247	703	0.44	0.25
VF815658	703	0.57	0.30
VF815665	703	0.41	0.31
VF815667	703	0.70	0.29
VF815652	703	0.39	0.34
VF815668	703	0.50	0.24
Form 9			
VF656748	704	0.55	0.29
VF656812	704	0.57	0.27
VF656804	704	0.46	0.12
VF656846	704	0.32	0.10
VF656794	704	0.98	0.18
VF814076	704	0.70	0.37
VF814089	704	0.69	0.27
VF814292	704	0.67	0.52
VF815472	704	0.62	0.50
VF814288	704	0.63	0.46
VF814285	704	0.28	0.09
VF814290	704	0.46	0.38
Form 10			
VF656800	706	0.33	0.31
VF656785	706	0.69	0.25

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF656808	706	0.38	0.04
VF656820	706	0.67	0.24
VF656815	706	0.42	0.26
VF815601	706	0.40	0.29
VF815618	706	0.43	0.13
VF815654	706	0.38	0.45
VF815661	706	0.58	0.29
VF815666	706	0.41	0.14
VF815670	706	0.42	0.25
VF815664	706	0.51	0.26

Table J14. Science Grade 8 Classical Statistics for Field Test Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
Form 1			
VF671265	785	0.83	0.22
VF671278	785	0.73	0.23
VF671327	785	0.33	0.10
VF671334	785	0.48	0.43
VF671285	785	0.70	0.35
VF813872	785	0.44	0.32
VF813811	785	0.42	0.15
VF735904	785	0.69	0.43
VF735980	785	0.73	0.14
VF736039	785	0.80	0.29
VF735995	785	0.53	0.24
VF735983	785	0.52	0.31
Form 2			
VF737468	656	0.83	0.21
VF737466	656	0.73	0.24
VF737445	656	0.78	0.34
VF737472	656	0.36	-0.05
VF737473	656	0.62	0.36
VF813803	656	0.84	0.29
VF813860	656	0.59	0.23
VF735043	656	0.67	0.06
VF734993	656	0.50	0.17
VF735020	656	0.34	0.09
VF735039	656	0.56	0.35
VF735045	656	0.39	0.23
Form 3			
VF812720	671	0.72	0.36
VF812729	671	0.39	-0.06
VF812711	671	0.55	0.39
VF812684	671	0.80	0.44
VF812703	671	0.83	0.29
VF813905	671	0.80	0.37
VF813827	671	0.72	0.32
VF671280	671	0.68	0.32
VF671343	671	0.57	0.28
VF671294	671	0.70	0.39
VF671350	671	0.81	0.35
VF671352	671	0.60	0.30

Form 4

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF813963	664	0.74	0.40
VF814074	664	0.31	0.06
VF814058	664	0.53	0.34
VF814043	664	0.58	0.33
VF814084	664	0.56	0.27
VF813865	664	0.59	0.30
VF813887	664	0.82	0.48
VF812690	664	0.47	0.20
VF812745	664	0.57	0.20
VF812741	664	0.70	0.33
VF812733	664	0.51	0.25
VF812726	664	0.59	0.28
Form 5			
VF735041	685	0.60	0.31
VF735007	685	0.51	0.29
VF735827	685	0.18	0.30
VF735791	685	0.81	0.31
VF735035	685	0.42	0.06
VF813878	685	0.47	0.17
VF813848	685	0.45	0.13
VF823979	685	0.58	0.33
VF824038	685	0.74	0.31
VF824046	685	0.72	0.35
VF824033	685	0.35	0.31
VF823985	685	0.71	0.41
Form 6			
VF813970	649	0.43	0.07
VF814052	649	0.66	0.32
VF814047	649	0.53	0.17
VF814096	649	0.69	0.34
VF814068	649	0.20	0.10
VF671317	649	0.92	0.37
VF671342	649	0.39	0.20
VF736114	649	0.95	0.19
VF736075	649	0.56	0.26
VF735912	649	0.65	0.38
VF735917	649	0.69	0.45
VF735960	649	0.31	0.17
Form 7			
VF824043	657	0.36	0.17



Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF823953	657	0.58	0.28
VF823970	657	0.53	0.34
VF824010	657	0.37	0.15
VF824019	657	0.89	0.37
VF736122	657	0.64	0.12
VF736130	657	0.40	0.27
VF815543	657	0.25	-0.17
VF815578	657	0.48	0.41
VF815494	657	0.58	0.37
VF815561	657	0.46	0.31
VF815521	657	0.90	0.36
Form 8			
VF684529	684	0.71	0.16
VF685187	684	0.65	0.39
VF685871	684	0.62	0.24
VF671386	684	0.53	0.33
VF686532	684	0.14	0.09
VF735828	684	0.42	0.25
VF735123	684	0.36	0.15
VF671359	684	0.39	0.21
VF671184	684	0.67	0.40
VF671279	684	0.27	0.08
VF671247	684	0.52	0.28
VF671361	684	0.71	0.15
Form 9			
VF684522	654	0.72	0.34
VF684505	654	0.38	0.12
VF686540	654	0.51	0.11
VF685863	654	0.72	0.28
VF687024	654	0.84	0.39
VF824049	654	0.61	0.23
VF824029	654	0.53	0.40
VF862725	654	0.75	0.31
VF862740	654	0.39	0.30
VF862778	654	0.32	0.24
VF862757	654	0.52	0.39
VF862752	654	0.72	0.39
Form 10			
VF862697	665	0.48	0.38
VF862703	665	0.70	0.20

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF862684	665	0.83	0.34
VF862659	665	0.58	0.33
VF862718	665	0.72	0.22
VF671315	665	0.80	0.34
VF671269	665	0.68	0.38
VF815568	665	0.28	0.30
VF815593	665	0.66	0.35
VF815536	665	0.65	0.29
VF815541	665	0.64	0.40
VF815587	665	0.38	0.28

SAWS

Table J15. Distributions of SAWS Grade 3 Field Test –8-point Prompt

Score	Form											
	<u>1</u>		<u>2</u>		<u>3</u>		<u>4</u>		<u>5</u>		<u>6</u>	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	6	0.5	3	0.2	12	1.0	5	0.4	8	0.7	7	0.6
1	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
2	58	4.7	95	7.8	111	9.2	96	7.9	132	10.7	124	10.2
3	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
4	328	26.7	332	27.3	390	32.2	413	34.1	316	25.7	369	30.4
5	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
6	524	42.7	494	40.6	441	36.4	470	38.8	491	40.0	481	39.6
7	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
8	311	25.4	294	24.1	259	21.4	228	18.8	282	23.0	235	19.3
<i>N</i>	1227		1218		1213		1212		1229		1216	
MEAN	5.75		5.61		5.36		5.35		5.48		5.34	
<i>SD</i>	1.72		1.80		1.88		1.77		1.90		1.84	

Table J16. Distributions of Scores for SAWS Grade 5 – 12-point Prompt

Score	Form					
	<u>1</u>		<u>2</u>		<u>3</u>	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	5	0.6	6	0.7	2	0.2
1	0	0.0	4	0.5	3	0.3
2	4	0.4	5	0.6	1	0.1
3	10	1.1	13	1.5	8	0.9
4	125	14.0	144	16.4	149	16.9
5	65	7.3	72	8.2	74	8.4
6	85	9.5	95	10.8	94	10.7
7	102	11.4	111	12.6	111	12.6
8	226	25.3	198	22.5	227	25.8
9	52	5.8	43	4.9	39	4.4
10	56	6.3	61	6.9	36	4.1
11	54	6.1	41	4.7	44	5.0
12	108	12.1	87	9.9	92	10.5
<i>N</i>	892		880		880	
MEAN	7.64		7.29		7.33	
<i>SD</i>	2.62		2.64		2.55	

Table J17. Distributions of Scores for SAWS Grade 5 by Trait – 12-point Prompt

Form	Score	Idea Development		Organization		Voice		Conventions	
		<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
1	0	6	0.7	11	1.2	5	0.6	17	1.9
	1	225	25.2	285	32.0	259	29.0	248	27.8
	2	451	50.6	396	44.4	437	49.0	456	51.1
	3	210	23.5	200	22.4	191	21.4	171	19.2
	<i>N</i>	892		892		892		892	
	MEAN	1.97		1.88		1.91		1.88	
	<i>SD</i>	0.72		0.76		0.72		0.73	
2	0	7	0.8	25	2.8	15	1.7	16	1.8
	1	239	27.2	283	32.2	301	34.2	328	37.3
	2	445	50.6	407	46.3	400	45.5	405	46.0
	3	189	21.5	165	18.8	164	18.6	131	14.9
	<i>N</i>	880		880		880		880	
	MEAN	1.93		1.81		1.81		1.74	
	<i>SD</i>	0.72		0.77		0.75		0.73	
3	0	4	0.5	7	0.8	6	0.7	10	1.1
	1	232	26.4	302	34.3	288	32.7	347	39.4
	2	464	52.7	426	48.4	406	46.1	394	44.8
	3	180	20.5	145	16.5	180	20.5	129	14.7
	<i>N</i>	880		880		880		880	
	MEAN	1.93		1.81		1.86		1.73	
	<i>SD</i>	0.69		0.71		0.74		0.72	

Table J18. Distributions of Scores for SAWS Grade 5 – 4-point Prompt<sup>13</sup>

Score	Form									
	4		5		6		7		8	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	10	1.1	6	0.7	5	0.6	6	0.7	7	0.8
1	44	5.0	27	3.1	34	3.9	92	10.5	63	7.2
2	183	20.8	146	16.6	179	20.6	277	31.7	176	20.2
3	286	32.5	287	32.6	252	28.9	242	27.7	280	32.1
4	358	40.6	415	47.1	401	46.0	258	29.5	347	39.8
<i>N</i>	881		881		871		875		873	
MEAN	3.06		3.22		3.16		2.75		3.03	
<i>SD</i>	0.96		0.88		0.92		1.02		0.98	

<sup>13</sup> Distributions of Rater 2 scores were not included since only 25% of the responses received second reads.

Table J19. Distributions of Scores for SAWS Grade 5 by Trait – 4-point Prompt

Form	Score	Response-to-Text		Holistic	
		<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
4	0	61	6.9	10	1.1
	1	272	30.9	410	46.5
	2	548	62.2	461	52.3
	<i>N</i>	881		881	
	MEAN	1.55		1.51	
	<i>SD</i>	0.62		0.52	
5	0	35	4.0	6	0.7
	1	214	24.3	388	44.0
	2	632	71.7	487	55.3
	<i>N</i>	881		881	
	MEAN	1.68		1.55	
	<i>SD</i>	0.55		0.51	
6	0	42	4.8	5	0.6
	1	237	27.2	401	46.0
	2	592	68.0	465	53.4
	<i>N</i>	871		871	
	MEAN	1.63		1.53	
	<i>SD</i>	0.57		0.51	
7	0	114	13.0	6	0.7
	1	357	40.8	499	57.0
	2	404	46.2	370	42.3
	<i>N</i>	875		875	
	MEAN	1.33		1.42	
	<i>SD</i>	0.70		0.51	
8	0	71	8.1	7	0.8
	1	254	29.1	439	50.3
	2	548	62.8	427	48.9
	<i>N</i>	873		873	
	MEAN	1.55		1.48	
	<i>SD</i>	0.64		0.52	

Table J20. Distributions of Rater 1 Scores for SAWS Grade 5 – 8-point Prompt<sup>14</sup>

Score	Form									
	<u>4</u>		<u>5</u>		<u>6</u>		<u>7</u>		<u>8</u>	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	5	0.6	9	1.0	5	0.6	8	0.9	4	0.5
1	10	1.1	9	1.0	18	2.1	13	1.5	8	0.9
2	31	3.5	29	3.3	63	7.2	35	4.0	28	3.2
3	82	9.3	83	9.4	123	14.1	126	14.4	63	7.2
4	147	16.7	107	12.2	170	19.5	149	17.0	112	12.8
5	172	19.5	162	18.4	176	20.2	197	22.5	199	22.8
6	205	23.3	209	23.7	151	17.3	188	21.5	277	31.7
7	137	15.6	175	19.9	104	11.9	119	13.6	148	17.0
8	92	10.4	98	11.1	61	7.0	40	4.6	34	3.9
<i>N</i>	881		881		871		875		873	
MEAN	5.32		5.47		4.82		4.94		5.34	
<i>SD</i>	1.69		1.74		1.77		1.66		1.48	

<sup>14</sup> Distributions of Rater 2 scores were not included since only 25% of the responses received second reads.

Table J21. Distributions of Scores for SAWS Grade 5 by Trait – 8-point Prompt

Form	Score	Response-to-Text		Holistic	
		<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
4	0	54	6.1	5	0.6
	1	237	26.9	27	3.1
	2	590	67.0	132	15.0
	3			225	25.5
	4			248	28.1
	5			148	16.8
	6			96	10.9
	<i>N</i>		881		881
	MEAN		1.61		3.72
	<i>SD</i>		0.60		1.32
5	0	41	4.7	9	1.0
	1	199	22.6	30	3.4
	2	641	72.8	122	13.8
	3			200	22.7
	4			239	27.1
	5			180	20.4
	6			101	11.5
	<i>N</i>		881		881
	MEAN		1.68		3.79
	<i>SD</i>		0.56		1.37
6	0	226	25.9	5	0.6
	1	314	36.1	30	3.4
	2	331	38.0	117	13.4
	3			225	25.8
	4			259	29.7
	5			166	19.1
	6			69	7.9
	<i>N</i>		871		871
	MEAN		1.12		3.70
	<i>SD</i>		0.79		1.27
7	0	105	12.0	8	0.9
	1	292	33.4	22	2.5
	2	478	54.6	133	15.2
	3			273	31.2
	4			267	30.5
	5			132	15.1
	6			40	4.6
	<i>N</i>		875		875
	MEAN		1.43		3.51
	<i>SD</i>		0.70		1.18
8	0	52	6.0	4	0.5
	1	147	16.8	19	2.2
	2	674	77.2	106	12.1
	3			256	29.3
	4			302	34.6
	5			152	17.4

Form	<u>Response-to-Text</u>		<u>Holistic</u>	
	Score	<i>N</i>	<i>Pct N</i>	
	6			
	<i>N</i>	873		873
	MEAN	1.71		3.63
	<i>SD</i>	0.57		1.12

Table J22. Distributions of Rater 1 Scores for SAWS Grade 7 – 12-point Prompt

Score	Form					
	<u>1</u>	<u>2</u>	<u>3</u>			
	<i>N</i>	<i>Pct N</i>	<i>N</i>	<i>Pct N</i>	<i>N</i>	<i>Pct N</i>
0	7	0.8	5	0.6	8	1.0
1	0	0.0	1	0.1	2	0.2
2	1	0.1	3	0.4	10	1.2
3	3	0.4	5	0.6	8	1.0
4	127	15.1	62	7.4	122	14.5
5	57	6.8	58	6.9	61	7.3
6	76	9.0	78	9.3	77	9.2
7	73	8.7	109	13.0	99	11.8
8	234	27.8	250	29.9	200	23.8
9	47	5.6	54	6.5	44	5.2
10	51	6.1	64	7.6	39	4.6
11	51	6.1	54	6.5	67	8.0
12	114	13.6	94	11.2	102	12.2
<i>N</i>	841		837		839	
MEAN	7.73		7.93		7.56	
<i>SD</i>	2.66		2.40		2.75	



Table J23. Distributions of Rater 1 Scores for SAWS Grade 7 by Trait – 12-point Prompt

Form	Score	Idea Development		Organization		Voice		Conventions	
		<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>
1	0	9	1.1	10	1.2	7	0.8	9	1.1
	1	195	23.2	230	27.3	243	28.9	255	30.3
	2	421	50.1	386	45.9	418	49.7	411	48.9
	3	216	25.7	215	25.6	173	20.6	166	19.7
	<i>N</i>	841		841		841		841	
	MEAN	2.00		1.96		1.90		1.87	
2	<i>SD</i>	0.73		0.76		0.72		0.73	
	0	6	0.7	14	1.7	8	1.0	8	1.0
	1	130	15.5	192	22.9	177	21.1	212	25.3
	2	493	58.9	444	53.0	472	56.4	466	55.7
	3	208	24.9	187	22.3	180	21.5	151	18.0
	<i>N</i>	837		837		837		837	
3	MEAN	2.08		1.96		1.98		1.91	
	<i>SD</i>	0.65		0.72		0.68		0.68	
	0	10	1.2	19	2.3	17	2.0	23	2.7
	1	220	26.2	253	30.2	255	30.4	250	29.8
	2	401	47.8	364	43.4	387	46.1	412	49.1
	3	208	24.8	203	24.2	180	21.5	154	18.4
	<i>N</i>	839		839		839		839	
	MEAN	1.96		1.90		1.87		1.83	
	<i>SD</i>	0.75		0.79		0.76		0.75	

Table J24. Distributions of Scores for SAWS Grade 7 – 4-point Prompt<sup>15</sup>

Score	Form									
	4		5		6		7		8	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	13	1.5	5	0.6	4	0.5	12	1.4	5	0.6
1	57	6.7	29	3.4	33	3.9	42	5.0	56	6.6
2	209	24.5	159	18.6	189	22.4	170	20.2	201	23.5
3	233	27.3	254	29.8	229	27.1	208	24.7	223	26.1
4	341	40.0	406	47.6	389	46.1	409	48.6	370	43.3
<i>N</i>	853		853		844		841		855	
MEAN	2.98		3.20		3.14		3.14		3.05	
<i>SD</i>	1.02		0.90		0.93		1.00		0.99	

<sup>15</sup> Distributions of Rater 2 scores were not included since only 25% of the responses received second reads.

Table J25. Distributions of Scores for SAWS Grade 7 by Trait – 4-point Prompt

Form	Score	Response-to-Text		Holistic	
		<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
4	0	85	10.0	13	1.5
	1	358	42.0	320	37.5
	2	410	48.1	520	61.0
	<i>N</i>	853		853	
	MEAN	1.38		1.59	
	<i>SD</i>	0.66		0.52	
5	0	45	5.3	5	0.6
	1	270	31.7	309	36.2
	2	538	63.1	539	63.2
	<i>N</i>	853		853	
	MEAN	1.58		1.63	
	<i>SD</i>	0.59		0.50	
6	0	42	5.0	4	0.5
	1	288	34.1	342	40.5
	2	514	60.9	498	59.0
	<i>N</i>	844		844	
	MEAN	1.56		1.59	
	<i>SD</i>	0.59		0.50	
7	0	59	7.0	12	1.4
	1	274	32.6	306	36.4
	2	508	60.4	523	62.2
	<i>N</i>	841		841	
	MEAN	1.53		1.61	
	<i>SD</i>	0.62		0.52	
8	0	68	8.0	5	0.6
	1	291	34.0	376	44.0
	2	496	58.0	474	55.4
	<i>N</i>	855		855	
	MEAN	1.50		1.55	
	<i>SD</i>	0.64		0.51	

Table J26. Distributions of Scores for SAWS Grade 7 – 8-point Prompts<sup>16</sup>

Score	Form									
	<u>4</u>		<u>5</u>		<u>6</u>		<u>7</u>		<u>8</u>	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	13	1.5	5	0.6	7	0.8	8	1.0	8	0.9
1	7	0.8	11	1.3	20	2.4	10	1.2	8	0.9
2	22	2.6	28	3.3	51	6.0	18	2.1	26	3.0
3	85	10.0	100	11.7	94	11.1	64	7.6	70	8.2
4	142	16.7	168	19.7	124	14.7	117	13.9	121	14.2
5	183	21.5	166	19.5	158	18.7	189	22.5	187	21.9
6	193	22.6	175	20.5	183	21.7	201	23.9	219	25.6
7	120	14.1	128	15.0	138	16.4	139	16.5	154	18.0
8	88	10.3	72	8.4	69	8.2	95	11.3	62	7.3
<i>N</i>	853		853		844		841		855	
MEAN	5.26		5.15		5.10		5.46		5.35	
<i>SD</i>	1.72		1.69		1.82		1.67		1.62	

<sup>16</sup> Distributions of Rater 2 scores were not included since only 25% of the responses received second reads.

Table J27. Distributions of Scores for SAWS Grade 7 by Trait – 8-point Prompt

Form	Score	Response-to-Text		Holistic	
		<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
4	0	41	4.8	13	1.5
	1	282	33.1	24	2.8
	2	530	62.1	109	12.8
	3			241	28.3
	4			237	27.8
	5			138	16.2
	6			91	10.7
	<i>N</i>	853		853	
	MEAN	1.57		3.69	
	<i>SD</i>	0.58		1.34	
5	0	44	5.2	5	0.6
	1	340	39.9	27	3.2
	2	469	55.0	123	14.4
	3			253	29.7
	4			219	25.7
	5			151	17.7
	6			75	8.8
	<i>N</i>	853		853	
	MEAN	1.50		3.65	
	<i>SD</i>	0.60		1.29	
6	0	122	14.5	7	0.8
	1	169	20.0	30	3.6
	2	553	65.5	142	16.8
	3			232	27.5
	4			215	25.5
	5			147	17.4
	6			71	8.4
	<i>N</i>	844		844	
	MEAN	1.51		3.59	
	<i>SD</i>	0.74		1.32	
7	0	28	3.3	8	1.0
	1	233	27.7	20	2.4
	2	580	69.0	91	10.8
	3			244	29.0
	4			223	26.5
	5			159	18.9
	6			96	11.4
	<i>N</i>	841		841	
	MEAN	1.66		3.80	
	<i>SD</i>	0.54		1.30	
8	0	48	5.6	8	0.9
	1	189	22.1	18	2.1
	2	618	72.3	123	14.4
	3			236	27.6
	4			235	27.5
	5			170	19.9

Form	Score	<u>Response-to-Text</u>		<u>Holistic</u>	
		<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
	6			65	7.6
	<i>N</i>	855		855	
	MEAN	1.67		3.69	
	<i>SD</i>	0.58		1.26	

## Appendix K: Classical Item Statistics for 2014 Operational Items

### Reading

Table K1. Reading Grade 3 Classical Statistics for Operational Items

Accession Number	N	Average item Score	Point Biserial Corr.
VF394056	7365	0.65	0.36
VF394053	7365	0.78	0.48
VF394041	7365	0.71	0.30
VF394054	7365	0.91	0.39
VF394046	7365	0.91	0.47
VF394050	7365	0.84	0.34
VF394049	7365	0.80	0.35
VF394051	7365	0.78	0.26
VF389477	7365	0.66	0.51
VF389949	7365	0.74	0.41
VF389470	7365	0.68	0.49
VF389620	7365	0.62	0.43
VF389457	7365	0.49	0.35
VF389467	7365	0.66	0.41
VF389165	7365	0.69	0.52
VF497660	7365	0.85	0.47
VF497668	7365	0.77	0.48
VF497700	7365	0.63	0.39
VF497705	7365	0.42	0.29
VF497671	7365	0.64	0.44
VF497695	7365	0.77	0.44
VF497696	7365	0.75	0.43
VF497690	7365	0.64	0.48
VF497684	7365	0.52	0.39
VF497676	7365	0.78	0.51
VF497818	7365	0.72	0.53
VF497815	7365	0.73	0.35
VF497822	7365	0.39	0.30
VF497820	7365	0.67	0.44
VF497783	7365	0.86	0.53
VF497793	7365	0.72	0.49
VF497812	7365	0.75	0.37
VF494759	7365	0.79	0.51
VF494915	7365	0.73	0.43

Accession Number	<i>N</i>	Average item Score	Point Biserial Corr.
VF494661	7365	0.70	0.49
VF494732	7365	0.83	0.59
VF494764	7365	0.77	0.53
VF494956	7365	0.51	0.41
VF494909	7365	0.57	0.39
VF494745	7365	0.49	0.45
VF493383	7365	0.81	0.52
VF493480	7365	0.73	0.54
VF494098	7365	0.68	0.49
VF497716	7365	0.68	0.39
VF497751	7365	0.65	0.40
VF497761	7365	0.71	0.52
VF497758	7365	0.56	0.46
VF497767	7365	0.52	0.43
VF497766	7365	0.41	0.25
VF497731	7365	0.84	0.53

Table K2. Reading Grade 4 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average item	
		Score	Point Biserial Corr.
VF495028	7022	0.42	0.29
VF495644	7022	0.69	0.30
VF494993	7022	0.47	0.37
VF495021	7022	0.68	0.32
VF495015	7022	0.77	0.48
VF495003	7022	0.59	0.49
VF497359	7022	0.64	0.49
VF497361	7022	0.81	0.51
VF497384	7022	0.52	0.39
VF497390	7022	0.72	0.58
VF497356	7022	0.35	0.31
VF497354	7022	0.84	0.43
VF497365	7022	0.44	0.27
VF497381	7022	0.79	0.52
VF497387	7022	0.85	0.50
VF494842	7022	0.76	0.26
VF494914	7022	0.75	0.40
VF494852	7022	0.82	0.49
VF494964	7022	0.81	0.40
VF494863	7022	0.89	0.44
VF494937	7022	0.89	0.41
VF497147	7022	0.86	0.49
VF497155	7022	0.68	0.49
VF497162	7022	0.48	0.46
VF497220	7022	0.74	0.52
VF497215	7022	0.34	0.32
VF497188	7022	0.93	0.47
VF497212	7022	0.71	0.39
VF497159	7022	0.87	0.44
VF497270	7022	0.47	0.32
VF497265	7022	0.88	0.44
VF497247	7022	0.69	0.42
VF497261	7022	0.70	0.28
VF497243	7022	0.70	0.36
VF497233	7022	0.70	0.34
VF497311	7022	0.44	0.17
VF497318	7022	0.81	0.40
VF497297	7022	0.90	0.50
VF497322	7022	0.66	0.49
VF497334	7022	0.72	0.45
VF497338	7022	0.60	0.24
VF497326	7022	0.62	0.50
VF497327	7022	0.71	0.41
VF407243	7022	0.73	0.43
VF407287	7022	0.83	0.50



Accession Number	<i>N</i>	Average item Score	Point Biserial Corr.
VF407232	7022	0.81	0.56
VF407235	7022	0.54	0.28
VF407297	7022	0.67	0.42
VF407282	7022	0.87	0.54
VF407298	7022	0.64	0.36

Table K3. Reading Grade 5 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF497182	7075	0.95	0.29
VF497177	7075	0.76	0.44
VF497174	7075	0.93	0.40
VF497172	7075	0.84	0.38
VF497056	7075	0.77	0.41
VF497170	7075	0.85	0.35
VF497052	7075	0.59	0.21
VF496101	7075	0.74	0.38
VF496032	7075	0.63	0.44
VF496188	7075	0.65	0.38
VF496085	7075	0.65	0.32
VF496185	7075	0.71	0.31
VF496024	7075	0.75	0.50
VF496115	7075	0.92	0.49
VF407319	7075	0.66	0.34
VF407388	7075	0.72	0.42
VF407329	7075	0.79	0.46
VF407332	7075	0.76	0.37
VF407355	7075	0.82	0.43
VF407322	7075	0.63	0.36
VF407360	7075	0.91	0.40
VF496211	7075	0.95	0.37
VF496865	7075	0.68	0.46
VF496879	7075	0.64	0.38
VF496213	7075	0.75	0.33
VF496209	7075	0.80	0.44
VF496201	7075	0.94	0.46
VF496206	7075	0.65	0.38
VF495924	7075	0.79	0.45
VF495921	7075	0.68	0.50
VF495800	7075	0.90	0.47
VF495780	7075	0.53	0.39
VF495943	7075	0.86	0.41
VF496875	7075	0.83	0.51
VF496872	7075	0.70	0.43
VF496878	7075	0.45	0.33
VF496882	7075	0.67	0.37
VF496884	7075	0.71	0.45
VF496869	7075	0.41	0.43
VF496886	7075	0.72	0.48
VF497284	7075	0.80	0.46
VF497278	7075	0.85	0.50
VF497273	7075	0.88	0.27
VF497282	7075	0.48	0.20
VF497285	7075	0.35	0.30

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF497287	7075	0.75	0.44
VF497274	7075	0.48	0.40
VF497272	7075	0.43	0.23
VF497288	7075	0.82	0.37
VF497037	7075	0.87	0.46
VF497039	7075	0.49	0.45
VF497030	7075	0.67	0.47
VF497028	7075	0.53	0.51
VF497012	7075	0.68	0.42

Table K4. Reading Grade 6 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF497042	6756	0.87	0.49
VF497035	6756	0.79	0.37
VF497034	6756	0.83	0.41
VF497033	6756	0.53	0.37
VF497047	6756	0.80	0.38
VF497041	6756	0.83	0.48
VF496873	6756	0.67	0.39
VF496204	6756	0.63	0.48
VF496208	6756	0.63	0.44
VF496863	6756	0.63	0.47
VF496191	6756	0.57	0.50
VF496164	6756	0.72	0.34
VF496172	6756	0.65	0.45
VF496055	6756	0.49	0.46
VF496036	6756	0.55	0.40
VF496065	6756	0.68	0.43
VF496071	6756	0.80	0.46
VF496100	6756	0.82	0.48
VF496051	6756	0.64	0.43
VF496029	6756	0.84	0.42
VF496087	6756	0.79	0.47
VF495908	6756	0.81	0.37
VF495961	6756	0.63	0.41
VF495968	6756	0.51	0.42
VF495990	6756	0.80	0.43
VF495938	6756	0.48	0.40
VF495954	6756	0.74	0.35
VF388881	6756	0.88	0.41
VF388912	6756	0.76	0.42
VF388853	6756	0.70	0.54
VF388848	6756	0.75	0.42
VF388868	6756	0.89	0.44
VF388851	6756	0.52	0.43
VF497084	6756	0.52	0.40
VF497082	6756	0.46	0.42
VF497087	6756	0.56	0.47
VF497079	6756	0.58	0.42
VF497078	6756	0.57	0.30
VF497083	6756	0.42	0.40
VF497077	6756	0.40	0.33
VF497076	6756	0.41	0.38
VF497074	6756	0.44	0.44
VF523861	6756	0.66	0.50
VF523801	6756	0.61	0.43
VF523846	6756	0.71	0.51

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF523825	6756	0.73	0.53
VF523818	6756	0.66	0.42
VF523813	6756	0.50	0.26
VF523863	6756	0.80	0.56
VF523804	6756	0.59	0.41
VF523786	6756	0.64	0.48
VF497071	6756	0.49	0.39
VF497069	6756	0.57	0.35
VF497053	6756	0.37	0.34
VF497073	6756	0.69	0.53
VF497059	6756	0.65	0.39

Table K5. Reading Grade 7 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF496937	6463	0.81	0.28
VF496901	6463	0.76	0.25
VF496913	6463	0.88	0.44
VF496906	6463	0.89	0.39
VF496895	6463	0.50	0.48
VF496900	6463	0.86	0.46
VF497975	6463	0.51	0.34
VF497958	6463	0.74	0.50
VF497951	6463	0.69	0.41
VF497969	6463	0.48	0.42
VF497955	6463	0.79	0.50
VF497961	6463	0.59	0.51
VF497978	6463	0.49	0.39
VF497974	6463	0.56	0.41
VF497941	6463	0.67	0.38
VF497950	6463	0.80	0.44
VF497938	6463	0.64	0.41
VF497943	6463	0.58	0.41
VF497935	6463	0.64	0.23
VF497931	6463	0.73	0.39
VF497930	6463	0.71	0.50
VF497862	6463	0.49	0.37
VF497882	6463	0.63	0.36
VF497879	6463	0.68	0.35
VF497893	6463	0.71	0.30
VF497890	6463	0.73	0.41
VF497876	6463	0.75	0.39
VF497868	6463	0.45	0.40
VF497873	6463	0.71	0.43
VF497883	6463	0.51	0.27
VF498058	6463	0.70	0.38
VF497877	6463	0.55	0.28
VF498030	6463	0.77	0.37
VF498018	6463	0.78	0.52
VF497980	6463	0.47	0.41
VF498062	6463	0.58	0.38
VF498051	6463	0.62	0.37
VF498064	6463	0.49	0.42
VF498054	6463	0.59	0.46
VF498057	6463	0.64	0.43
VF498034	6463	0.69	0.49
VF498047	6463	0.83	0.36
VF498032	6463	0.80	0.42
VF498052	6463	0.43	0.33
VF497224	6463	0.62	0.37

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF497211	6463	0.72	0.38
VF497175	6463	0.78	0.43
VF497190	6463	0.77	0.42
VF497198	6463	0.45	0.37
VF497205	6463	0.67	0.43
VF497281	6463	0.81	0.32
VF497301	6463	0.51	0.36
VF497299	6463	0.67	0.42
VF497291	6463	0.40	0.32
VF497260	6463	0.81	0.52
VF497263	6463	0.73	0.43

Table K6. Reading Grade 8 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF497427	6467	0.94	0.37
VF497441	6467	0.85	0.36
VF497443	6467	0.91	0.35
VF497446	6467	0.75	0.36
VF497436	6467	0.88	0.39
VF497445	6467	0.79	0.45
VF497444	6467	0.96	0.42
VF497199	6467	0.56	0.27
VF497180	6467	0.93	0.35
VF497203	6467	0.66	0.41
VF497196	6467	0.82	0.39
VF497178	6467	0.62	0.30
VF497193	6467	0.63	0.35
VF497209	6467	0.72	0.38
VF497257	6467	0.79	0.26
VF497229	6467	0.67	0.32
VF497259	6467	0.56	0.31
VF497244	6467	0.82	0.35
VF497242	6467	0.51	0.37
VF497235	6467	0.93	0.47
VF497266	6467	0.55	0.31
VF497252	6467	0.84	0.40
VF497095	6467	0.88	0.43
VF497113	6467	0.43	0.38
VF497114	6467	0.80	0.42
VF497101	6467	0.67	0.46
VF497098	6467	0.53	0.32
VF497094	6467	0.77	0.45
VF497115	6467	0.79	0.41
VF497148	6467	0.62	0.42
VF497161	6467	0.60	0.34
VF497164	6467	0.50	0.36
VF497137	6467	0.66	0.52
VF497166	6467	0.82	0.62
VF497139	6467	0.81	0.47
VF497120	6467	0.52	0.42
VF497127	6467	0.83	0.51
VF497116	6467	0.57	0.35
VF497117	6467	0.66	0.49
VF497132	6467	0.51	0.31
VF497130	6467	0.73	0.45
VF497123	6467	0.75	0.39
VF497329	6467	0.80	0.49
VF497349	6467	0.89	0.44
VF497353	6467	0.63	0.48



Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF497328	6467	0.66	0.44
VF497325	6467	0.67	0.48
VF497363	6467	0.57	0.36
VF497355	6467	0.92	0.47
VF497298	6467	0.86	0.51
VF497316	6467	0.82	0.43
VF497302	6467	0.57	0.39
VF497309	6467	0.81	0.55
VF497313	6467	0.64	0.51
VF497305	6467	0.79	0.45
VF497306	6467	0.91	0.45

Mathematics

Table K7. Mathematics Grade 3 Classical Statistics for Operational Items

Accession Number	N	Average Item Score	Point Biserial Corr.
VF393959	7369	0.94	0.33
VF387496	7369	0.88	0.36
VF393772	7369	0.62	0.54
VF494670	7369	0.34	0.39
VF494103	7369	0.48	0.45
VF406339	7369	0.81	0.46
VF387500	7369	0.63	0.47
VF406297	7369	0.48	0.31
VF394355	7369	0.36	0.35
VF387498	7369	0.13	0.27
VF406327	7369	0.24	0.48
VF406204	7369	0.61	0.35
VF394252	7369	0.73	0.35
VF494820	7369	0.37	0.44
VF493146	7369	0.70	0.50
VF394250	7369	0.59	0.43
VF393782	7369	0.90	0.37
VF394361	7369	0.80	0.50
VF394339	7369	0.64	0.34
VF493415	7369	0.53	0.49
VF394382	7369	0.90	0.32
VF394375	7369	0.80	0.41
VF394362	7369	0.40	0.45
VF394369	7369	0.68	0.51
VF493287	7369	0.54	0.47
VF394368	7369	0.65	0.45
VF394376	7369	0.58	0.47
VF393748	7369	0.56	0.55
VF394221	7369	0.46	0.38
VF494693	7369	0.43	0.43
VF494895	7369	0.37	0.48
VF394378	7369	0.40	0.24
VF394381	7369	0.49	0.43
VF406343	7369	0.82	0.42
VF494880	7369	0.93	0.32
VF406295	7369	0.70	0.19
VF493127	7369	0.54	0.48
VF393824	7369	0.44	0.37
VF394239	7369	0.56	0.60
VF494690	7369	0.51	0.39
VF494750	7369	0.69	0.41
VF493461	7369	0.48	0.31
VF393786	7369	0.72	0.43

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF493124	7369	0.52	0.48
VF394356	7369	0.62	0.52
VF394229	7369	0.51	0.37
VF493153	7369	0.80	0.42
VF493387	7369	0.80	0.40
VF387502	7369	0.97	0.26
VF494756	7369	0.67	0.44

Table K8. Mathematics Grade 4 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF492346	7026	0.97	0.22
VF492315	7026	0.90	0.44
VF393675	7026	0.54	0.46
VF492332	7026	0.54	0.38
VF492333	7026	0.61	0.36
VF493356	7026	0.89	0.27
VF492358	7026	0.89	0.32
VF493249	7026	0.42	0.30
VF493349	7026	0.42	0.50
VF492311	7026	0.91	0.36
VF493284	7026	0.24	0.43
VF492390	7026	0.93	0.28
VF493334	7026	0.29	0.39
VF497391	7026	0.73	0.50
VF493344	7026	0.82	0.40
VF493373	7026	0.83	0.32
VF493140	7026	0.60	0.50
VF492392	7026	0.73	0.37
VF492353	7026	0.54	0.42
VF492320	7026	0.89	0.47
VF493238	7026	0.94	0.24
VF492330	7026	0.58	0.35
VF493228	7026	0.52	0.57
VF492312	7026	0.89	0.42
VF497395	7026	0.60	0.53
VF492334	7026	0.77	0.49
VF492343	7026	0.22	0.34
VF492370	7026	0.63	0.27
VF493154	7026	0.87	0.37
VF493303	7026	0.67	0.43
VF493219	7026	0.80	0.35
VF393726	7026	0.77	0.32
VF493257	7026	0.53	0.45
VF493312	7026	0.33	0.28
VF492373	7026	0.72	0.41
VF493223	7026	0.47	0.51
VF493366	7026	0.34	0.28
VF493143	7026	0.64	0.25
VF493377	7026	0.66	0.40
VF492338	7026	0.71	0.36
VF493295	7026	0.35	0.45
VF493301	7026	0.49	0.35
VF493126	7026	0.90	0.36
VF493135	7026	0.54	0.48
VF393648	7026	0.31	0.37

Accession Number	<i>N</i>	Average Item	Point Biserial Corr.
		Score	
VF493142	7026	0.65	0.49
VF493288	7026	0.28	0.43
VF492386	7026	0.65	0.41
VF493318	7026	0.32	0.28
VF493130	7026	0.52	0.26
VF492352	7026	0.97	0.23
VF493329	7026	0.41	0.27
VF492306	7026	0.92	0.36
VF493242	7026	0.59	0.40
VF493262	7026	0.54	0.26
VF493361	7026	0.52	0.15
VF493294	7026	0.23	0.16
VF492337	7026	0.61	0.42
VF493371	7026	0.56	0.34

Table K9. Mathematics Grade 5 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF491951	7076	0.86	0.41
VF491924	7076	0.66	0.44
VF491941	7076	0.52	0.40
VF492083	7076	0.37	0.48
VF492203	7076	0.85	0.48
VF492088	7076	0.54	0.48
VF492027	7076	0.48	0.19
VF491963	7076	0.36	0.27
VF491626	7076	0.71	0.46
VF492000	7076	0.62	0.35
VF491900	7076	0.45	0.50
VF492313	7076	0.63	0.57
VF492048	7076	0.54	0.47
VF492235	7076	0.52	0.35
VF492120	7076	0.42	0.40
VF492031	7076	0.53	0.46
VF492298	7076	0.55	0.53
VF491636	7076	0.53	0.57
VF492255	7076	0.62	0.49
VF491967	7076	0.51	0.43
VF492007	7076	0.67	0.37
VF492003	7076	0.51	0.35
VF492296	7076	0.40	0.29
VF492214	7076	0.36	0.51
VF492174	7076	0.60	0.53
VF492532	7076	0.56	0.29
VF491948	7076	0.48	0.50
VF492427	7076	0.46	0.51
VF492099	7076	0.59	0.54
VF491627	7076	0.84	0.32
VF491771	7076	0.69	0.51
VF492248	7076	0.66	0.48
VF492186	7076	0.44	0.31
VF491937	7076	0.54	0.37
VF491895	7076	0.73	0.45
VF492423	7076	0.67	0.38
VF491916	7076	0.58	0.60
VF491635	7076	0.42	0.51
VF492374	7076	0.72	0.36
VF492528	7076	0.63	0.36
VF491804	7076	0.61	0.63
VF491630	7076	0.47	0.39
VF492397	7076	0.43	0.46
VF492095	7076	0.57	0.53
VF491783	7076	0.36	0.50

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF492304	7076	0.71	0.53
VF492091	7076	0.50	0.22
VF491939	7076	0.80	0.47
VF491794	7076	0.71	0.47
VF491932	7076	0.56	0.51
VF491905	7076	0.67	0.48
VF491753	7076	0.63	0.29
VF492010	7076	0.51	0.44
VF491761	7076	0.68	0.49
VF492001	7076	0.67	0.34
VF491727	7076	0.46	0.44
VF491821	7076	0.47	0.49
VF491927	7076	0.73	0.40
VF492281	7076	0.67	0.59

Table K10. Mathematics Grade 6 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF491953	6759	0.69	0.39
VF492542	6759	0.63	0.53
VF492181	6759	0.63	0.29
VF491930	6759	0.65	0.46
VF492399	6759	0.68	0.39
VF492732	6759	0.39	0.26
VF492287	6759	0.52	0.32
VF492722	6759	0.43	0.38
VF492572	6759	0.46	0.34
VF492759	6759	0.38	0.31
VF492721	6759	0.46	0.36
VF492383	6759	0.61	0.51
VF492593	6759	0.68	0.57
VF492030	6759	0.58	0.43
VF492192	6759	0.51	0.36
VF492053	6759	0.54	0.51
VF492709	6759	0.43	0.44
VF492240	6759	0.57	0.22
VF492412	6759	0.40	0.38
VF492660	6759	0.26	0.45
VF492533	6759	0.72	0.50
VF492577	6759	0.71	0.59
VF491996	6759	0.52	0.28
VF491960	6759	0.70	0.47
VF492078	6759	0.56	0.43
VF492388	6759	0.59	0.38
VF491879	6759	0.88	0.40
VF492931	6759	0.38	0.53
VF493058	6759	0.63	0.48
VF491874	6759	0.61	0.58
VF493013	6759	0.33	0.30
VF493089	6759	0.54	0.50
VF492582	6759	0.67	0.49
VF423647	6759	0.49	0.11
VF492280	6759	0.45	0.45
VF492879	6759	0.47	0.54
VF492025	6759	0.43	0.43
VF491931	6759	0.87	0.44
VF492716	6759	0.43	0.43
VF492711	6759	0.67	0.35
VF492290	6759	0.84	0.39
VF423146	6759	0.68	0.48
VF492284	6759	0.83	0.39
VF493001	6759	0.64	0.52
VF491787	6759	0.67	0.35



Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF493003	6759	0.61	0.48
VF491966	6759	0.60	0.36
VF493068	6759	0.39	0.47
VF491894	6759	0.52	0.47
VF492941	6759	0.29	0.39
VF423225	6759	0.46	0.46
VF492890	6759	0.48	0.46
VF491940	6759	0.61	0.53
VF493092	6759	0.62	0.57
VF493062	6759	0.59	0.46
VF491976	6759	0.46	0.32
VF493002	6759	0.52	0.45
VF492415	6759	0.63	0.44
VF491837	6759	0.83	0.33

Table K11. Mathematics Grade 7 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF492966	6467	0.72	0.49
VF492597	6467	0.37	0.38
VF492307	6467	0.71	0.48
VF492394	6467	0.78	0.44
VF492967	6467	0.50	0.32
VF492878	6467	0.36	0.42
VF492672	6467	0.89	0.37
VF492888	6467	0.48	0.40
VF492871	6467	0.76	0.46
VF492640	6467	0.36	0.36
VF492836	6467	0.46	0.33
VF492853	6467	0.57	0.39
VF492835	6467	0.72	0.43
VF492419	6467	0.37	0.42
VF492666	6467	0.47	0.46
VF492760	6467	0.32	0.38
VF492653	6467	0.12	0.20
VF493021	6467	0.37	0.33
VF492578	6467	0.28	0.39
VF493038	6467	0.64	0.59
VF492357	6467	0.35	0.32
VF492663	6467	0.61	0.54
VF493057	6467	0.82	0.25
VF492665	6467	0.52	0.44
VF492413	6467	0.46	0.31
VF492973	6467	0.43	0.35
VF492696	6467	0.67	0.50
VF493061	6467	0.45	0.58
VF492864	6467	0.68	0.54
VF492302	6467	0.86	0.44
VF493046	6467	0.46	0.44
VF492425	6467	0.52	0.48
VF492951	6467	0.50	0.23
VF492720	6467	0.45	0.56
VF492765	6467	0.27	0.20
VF492538	6467	0.72	0.55
VF493019	6467	0.44	0.42
VF492673	6467	0.13	0.20
VF493067	6467	0.42	0.41
VF492830	6467	0.64	0.49
VF492929	6467	0.90	0.23
VF492531	6467	0.87	0.33
VF492955	6467	0.45	0.25
VF493071	6467	0.32	0.30
VF492780	6467	0.60	0.46

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF493015	6467	0.48	0.42
VF492567	6467	0.43	0.33
VF493077	6467	0.75	0.46
VF492546	6467	0.42	0.31
VF493052	6467	0.49	0.39
VF493036	6467	0.37	0.33
VF492861	6467	0.59	0.43
VF492589	6467	0.25	0.45
VF492778	6467	0.45	0.52
VF492259	6467	0.43	0.47
VF493043	6467	0.31	0.19
VF492901	6467	0.39	0.42
VF492748	6467	0.65	0.48
VF493064	6467	0.71	0.37

Table K12. Mathematics Grade 8 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF491923	6470	0.86	0.33
VF493115	6470	0.77	0.20
VF491907	6470	0.62	0.38
VF493150	6470	0.46	0.50
VF492845	6470	0.45	0.35
VF491824	6470	0.39	0.35
VF494699	6470	0.56	0.44
VF492863	6470	0.60	0.51
VF493113	6470	0.67	0.40
VF492712	6470	0.43	0.49
VF491857	6470	0.56	0.48
VF492726	6470	0.35	0.39
VF493157	6470	0.46	0.30
VF491873	6470	0.25	0.43
VF492874	6470	0.39	0.35
VF493121	6470	0.51	0.36
VF491915	6470	0.35	0.28
VF494120	6470	0.27	0.21
VF492856	6470	0.55	0.53
VF493159	6470	0.57	0.33
VF492438	6470	0.73	0.38
VF491991	6470	0.68	0.45
VF491965	6470	0.62	0.36
VF492289	6470	0.57	0.45
VF493034	6470	0.53	0.49
VF494727	6470	0.48	0.48
VF492410	6470	0.52	0.43
VF494928	6470	0.38	0.32
VF492278	6470	0.74	0.53
VF492345	6470	0.39	0.42
VF494751	6470	0.46	0.41
VF493040	6470	0.66	0.50
VF491975	6470	0.72	0.49
VF492907	6470	0.73	0.51
VF492430	6470	0.56	0.35
VF493107	6470	0.39	0.33
VF492563	6470	0.60	0.32
VF492579	6470	0.49	0.31
VF492028	6470	0.52	0.46
VF492272	6470	0.64	0.57
VF494801	6470	0.56	0.44
VF492420	6470	0.74	0.49
VF494769	6470	0.68	0.45
VF492920	6470	0.64	0.60
VF492592	6470	0.68	0.36

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF492024	6470	0.33	0.41
VF492268	6470	0.37	0.21
VF491949	6470	0.67	0.48
VF492258	6470	0.50	0.30
VF493045	6470	0.64	0.45
VF492212	6470	0.55	0.33
VF492414	6470	0.57	0.51
VF494776	6470	0.52	0.50
VF492008	6470	0.50	0.41
VF493011	6470	0.75	0.53
VF492400	6470	0.49	0.37
VF492917	6470	0.75	0.52
VF494819	6470	0.36	0.35
VF493088	6470	0.66	0.51
VF492436	6470	0.74	0.38
VF493097	6470	0.32	0.19
VF492231	6470	0.33	0.22
VF494760	6470	0.58	0.52
VF492393	6470	0.53	0.48
VF492440	6470	0.73	0.45

Science

Table K13. Science Grade 4 Classical Statistics for Operational Items

Accession Number	N	Average Item	Point Biserial Corr.
		Score	
VF431075	7022	0.58	0.47
VF431078	7022	0.50	0.45
VF483448	7022	0.90	0.35
VF484396	7022	0.56	0.21
VF430688	7022	0.62	0.39
VF430686	7022	0.49	0.38
VF294929	7022	0.69	0.43
VF296821	7022	0.68	0.38
VF484935	7022	0.42	0.23
VF290777	7022	0.76	0.44
VF431081	7022	0.80	0.38
VF283606	7022	0.87	0.39
VF283022	7022	0.69	0.46
VF311559	7022	0.64	0.40
VF311567	7022	0.60	0.42
VF311586	7022	0.67	0.43
VF311548	7022	0.75	0.36
VF431027	7022	0.69	0.42
VF431028	7022	0.84	0.40
VF287722	7022	0.47	0.40
VF287717	7022	0.37	0.38
VF284006	7022	0.50	0.46
VF284002	7022	0.52	0.35
VF431125	7022	0.56	0.30
VF431127	7022	0.74	0.44
VF431129	7022	0.63	0.36
VF293507	7022	0.82	0.42
VF292879	7022	0.73	0.36
VF294472	7022	0.64	0.40
VF407152	7022	0.41	0.28
VF407138	7022	0.44	0.32
VF406427	7022	0.51	0.45
VF483424	7022	0.68	0.46
VF483437	7022	0.93	0.25
VF287740	7022	0.47	0.36
VF287742	7022	0.47	0.35
VF287745	7022	0.50	0.31
VF393724	7022	0.78	0.53
VF393699	7022	0.57	0.47
VF393721	7022	0.59	0.38
VF431046	7022	0.52	0.32
VF269846	7022	0.53	0.34
VF269841	7022	0.41	0.21

Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF296839	7022	0.48	0.39
VF282661	7022	0.60	0.45
VF282670	7022	0.82	0.43
VF386811	7022	0.59	0.40
VF386826	7022	0.59	0.41
VF287864	7022	0.68	0.40
VF287870	7022	0.47	0.30

Table K14. Science Grade 8 Classical Statistics for Operational Items

Accession Number	<i>N</i>	Average Item	
		Score	Point Biserial Corr.
VF484958	6455	0.59	0.43
VF484974	6455	0.46	0.33
VF486678	6455	0.27	0.29
VF486675	6455	0.51	0.32
VF394477	6455	0.53	0.49
VF394502	6455	0.50	0.36
VF484993	6455	0.45	0.43
VF484999	6455	0.56	0.31
VF407356	6455	0.88	0.33
VF407330	6455	0.58	0.43
VF394777	6455	0.56	0.38
VF394780	6455	0.53	0.37
VF394809	6455	0.56	0.38
VF394814	6455	0.49	0.30
VF431421	6455	0.60	0.36
VF431423	6455	0.36	0.33
VF394561	6455	0.64	0.44
VF394565	6455	0.74	0.44
VF313289	6455	0.49	0.39
VF313291	6455	0.64	0.43
VF313300	6455	0.57	0.41
VF486847	6455	0.65	0.42
VF486858	6455	0.36	0.38
VF486815	6455	0.63	0.47
VF486821	6455	0.63	0.45
VF308868	6455	0.84	0.45
VF308871	6455	0.53	0.46
VF308869	6455	0.60	0.30
VF486166	6455	0.63	0.32
VF486163	6455	0.62	0.49
VF407480	6455	0.55	0.47
VF407483	6455	0.57	0.26
VF407155	6455	0.69	0.40
VF407242	6455	0.75	0.26
VF431683	6455	0.75	0.38
VF431688	6455	0.51	0.40
VF486771	6455	0.55	0.38
VF486782	6455	0.58	0.41
VF486765	6455	0.60	0.52
VF486914	6455	0.36	0.32
VF486941	6455	0.48	0.39
VF313274	6455	0.34	0.29
VF313280	6455	0.48	0.46
VF313281	6455	0.56	0.44
VF485018	6455	0.58	0.33



Accession Number	<i>N</i>	Average Item Score	Point Biserial Corr.
VF485023	6455	0.51	0.34

SAWS

Table K15. Distributions of Rating 1 Scores for SAWS

Score	<u>Grade 3-Prompt 1</u>		<u>Grade 3-Prompt 2</u>		<u>Grade 5</u>		<u>Grade 7</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	47	0.64	19	0.26	15	0.21	31	0.46
1	102	1.39	71	0.97	28	0.40	7	0.10
2	169	2.31	106	1.45	54	0.77	32	0.47
3	290	3.97	92	1.26	58	0.83	35	0.52
4	1038	14.19	1294	17.69	1289	18.47	833	12.33
5	823	11.25	617	8.44	626	8.97	475	7.03
6	937	12.81	871	11.91	830	11.89	650	9.62
7	1027	14.04	982	13.43	817	11.71	734	10.86
8	1279	17.49	1724	23.57	1759	25.20	1925	28.49
9	464	6.34	373	5.10	355	5.09	462	6.84
10	376	5.14	320	4.38	298	4.27	425	6.29
11	324	4.43	319	4.36	281	4.03	358	5.30
12	437	5.98	525	7.18	569	8.15	789	11.68
<i>N</i>	7313		7313		6979		6756	
MEAN	6.75		6.98		7.09		7.72	
<i>SD</i>	2.63		2.53		2.5		2.52	

Table K16. Distributions of Trait Scores for SAWS – Grade 3 – 12-point Prompt 1

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	98	1.34	663	9.07	249	3.40	362	4.95
1	2425	33.16	2236	30.58	2926	40.01	3012	41.19
2	3608	49.34	3207	43.85	3212	43.92	3051	41.72
3	1182	16.16	1207	16.50	926	12.66	888	12.14
<i>N</i>	7313		7313		7313		7313	
MEAN	1.8		1.68		1.66		1.61	
<i>SD</i>	0.71		0.85		0.74		0.76	

Table K17. Distributions of Trait Scores for SAWS - Grade 3 - 12-point Prompt 2

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	31	0.42	149	2.04	173	2.37	287	3.92
1	2145	29.33	2456	33.58	2765	37.81	2975	40.68
2	3896	53.27	3668	50.16	3313	45.30	3199	43.74
3	1241	16.97	1040	14.22	1062	14.52	852	11.65
<i>N</i>	7313		7313		7313		7313	
MEAN	1.87		1.77		1.72		1.63	
<i>SD</i>	0.68		0.71		0.73		0.74	

Table K18. Distributions of Trait Scores for SAWS - Grade 5 - 12-point Prompt

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	26	0.37	97	1.39	86	1.23	132	1.89
1	2104	30.15	2496	35.76	2565	36.75	2687	38.50
2	3667	52.54	3330	47.71	3308	47.40	3255	46.64
3	1182	16.94	1056	15.13	1020	14.62	905	12.97
<i>N</i>	6979		6979		6979		6979	
MEAN	1.86		1.77		1.75		1.71	
<i>SD</i>	0.68		0.71		0.71		0.71	

Table K19. Distributions of Trait Scores for SAWS - Grade 7 - 12-point Prompt

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	47	0.70	79	1.17	49	0.73	83	1.23
1	1560	23.09	1753	25.95	1840	27.24	1884	27.89
2	3661	54.19	3368	49.85	3556	52.63	3504	51.87
3	1488	22.02	1556	23.03	1311	19.40	1285	19.02
<i>N</i>	6756		6756		6756		6756	
MEAN	1.98		1.95		1.91		1.89	
<i>SD</i>	0.69		0.73		0.70		0.71	

Table K20. Distributions of Scores for SAWS Grade 5 - 4-point Prompt

Score	<u>4-point</u>	
	<i>N</i>	Pct <i>N</i>
0	49	0.70
1	480	6.88
2	1416	20.29
3	2016	28.89
4	3018	43.24
<i>N</i>	6979	
MEAN	3.07	
<i>SD</i>	0.98	

Table K21. Distributions of Trait Scores for SAWS - Grade 5 – 4-point Prompt

Score	<u>Response-to-Text</u>		<u>Holistic</u>	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	587	8.41	49	0.70
1	2093	29.99	3119	44.69
2	4299	61.60	3811	54.61
<i>N</i>	6979		6979	
MEAN	1.53		1.54	
<i>SD</i>	0.65		0.51	

Table K22. Distributions of Rater 1 Scores for SAWS Grade 5 - 8-point Prompt

Score	<u>8-point</u>	
	<i>N</i>	Pct <i>N</i>
0	42	0.60
1	78	1.12
2	382	5.47
3	821	11.76
4	1376	19.72
5	1705	24.43
6	1538	22.04
7	790	11.32
8	247	3.54
<i>N</i>	6979	
MEAN	4.88	
<i>SD</i>	1.57	

Table K23. Distributions of Scores for SAWS Grade 5 by Trait – 8-point Prompt

Score	<u>Response-to-Text</u>		<u>Holistic</u>	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	1213	17.38	42	0.60
1	2214	31.72	120	1.72
2	3552	50.90	964	13.81
3			2296	32.90
4			2246	32.18
5			1036	14.84
6			275	3.94
<i>N</i>	6979		6979	
MEAN	1.34		3.55	
<i>SD</i>	0.76		1.11	

Table K24. Distributions of Scores for SAWS Grade 7 - 4-point Prompt

Score	<u>4-point</u>	
	<i>N</i>	Pct <i>N</i>
0	50	0.74
1	765	11.32
2	1626	24.07
3	1742	25.78
4	2573	38.08
<i>N</i>	6756	
MEAN	2.89	
<i>SD</i>	1.06	

Table K25. Distributions of Trait Scores for SAWS - Grade 5 – 4-point Prompt

Score	<u>Response-to-Text</u>		<u>Holistic</u>	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	966	14.30	50	0.74
1	2360	34.93	3097	45.84
2	3430	50.77	3609	53.42
<i>N</i>	6756		6756	
MEAN	1.36		1.53	
<i>SD</i>	0.72		0.51	

Table K26. Distributions of Rater 1 Scores for SAWS Grade 7 - 8-point Prompt

Score	<u>8-point</u>	
	<i>N</i>	Pct <i>N</i>
0	51	0.75
1	63	0.93
2	245	3.63
3	692	10.24
4	1233	18.25
5	1645	24.35
6	1508	22.32
7	867	12.83
8	452	6.69
<i>N</i>	6756	
MEAN	5.11	
<i>SD</i>	1.61	

Table K27. Distributions of Scores for SAWS Grade 7 by Trait – 8-point Prompt

Score	<u>Response-to-Text</u>		<u>Holistic</u>	
	<i>N</i>	Pct <i>N</i>	<i>N</i>	Pct <i>N</i>
0	628	9.30	52	0.77
1	2262	33.48	161	2.38
2	3866	57.22	915	13.54
3			2008	29.72
4			2068	30.61
5			1074	15.90
6			478	7.08
<i>N</i>	6756		6756	
MEAN	1.48		3.63	
<i>SD</i>	0.66		1.22	

# Appendix L: Rasch Difficulty, Standard Error, Fit Statistics, and *N*-counts for 2014 Field Test Items

## Reading

Table L1. Reading Grade 3 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF815015	776	-0.589	0.096	0.90	0.79
VF815011	776	-0.674	0.098	1.20	1.65
VF815010	776	-1.396	0.117	0.85	0.69
VF815012	776	0.243	0.085	0.96	0.93
VF815022	776	-1.083	0.108	0.96	0.88
VF815020	776	0.112	0.086	1.24	1.28
VF798298	1508	-0.594	0.070	1.00	0.96
VF798266	776	0.808	0.081	1.05	1.08
VF798299	776	0.060	0.086	0.91	0.83
VF798292	776	-0.179	0.089	0.93	0.86
VF798282	776	2.252	0.089	1.01	1.34
VF798239	776	-0.762	0.100	0.98	1.09
VF885220	776	-0.428	0.093	0.79	0.67
VF885209	776	1.515	0.082	1.12	1.32
Form 2					
VF885358	728	0.184	0.088	0.92	0.83
VF885388	728	-0.769	0.104	0.87	0.68
Form 3					
VF815562	734	-1.584	0.134	0.81	0.56
VF815556	734	-1.859	0.147	0.93	0.80
VF815598	734	2.061	0.087	1.16	1.51
VF815575	734	1.369	0.082	1.26	1.50
VF815528	734	-0.963	0.112	0.89	0.73
VF815537	734	-1.755	0.142	0.88	0.68
VF885405	734	-1.430	0.128	0.88	0.84
VF885201	734	1.449	0.082	1.34	1.52

Form 4

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF815017	732	0.031	0.090	1.00	0.97
VF815018	732	-1.684	0.137	0.81	0.44
VF815014	732	-0.001	0.090	0.96	0.90
VF815009	732	0.652	0.084	1.35	1.47
VF815019	732	-1.274	0.121	0.84	0.55
VF815021	732	-1.043	0.113	0.85	0.73
VF798274	732	2.277	0.090	1.21	1.96
VF798297	732	2.334	0.091	1.28	1.81
VF798301	732	0.915	0.083	1.02	1.03
VF798290	732	0.680	0.084	0.95	0.93
VF798300	732	1.804	0.085	1.12	1.33
VF885322	732	-1.018	0.112	0.88	0.73
VF885192	732	-1.394	0.125	0.81	0.51
Form 5					
VF814997	723	-1.029	0.110	0.86	0.73
VF814980	723	-0.877	0.106	0.91	0.97
VF814974	723	-0.543	0.099	0.83	0.70
VF814982	723	-0.877	0.106	1.09	1.41
VF814989	723	0.024	0.089	1.15	1.27
VF814966	723	-1.830	0.140	0.79	0.45
VF884228	723	1.557	0.084	1.19	1.35
VF884215	723	1.071	0.083	1.19	1.33
VF884430	1460	2.384	0.066	1.13	1.64
VF884415	723	-0.553	0.099	1.05	0.97
VF884522	723	-0.164	0.092	0.97	0.91
VF884498	723	0.669	0.084	1.13	1.15
VF885399	723	-0.048	0.090	0.94	0.89
VF885379	723	0.739	0.084	1.08	1.12
Form 6					
VF814976	737	-0.297	0.095	0.96	0.99
VF814978	737	-1.072	0.112	0.79	0.58
VF814983	737	-0.124	0.092	0.99	1.03
VF814992	737	-1.175	0.115	0.90	1.05
VF814971	737	1.258	0.083	1.17	1.36
VF814994	737	-2.114	0.155	0.78	0.39
VF884237	737	0.373	0.086	0.93	0.88
VF884239	737	0.548	0.085	1.25	1.32
VF884250	737	-1.479	0.126	0.86	0.85
VF884503	737	1.964	0.087	1.16	1.61
VF884518	737	0.569	0.085	1.10	1.10
VF885412	737	-0.445	0.097	0.97	0.99
VF885214	737	0.097	0.089	1.01	1.00
Form 7					
VF814724	740	1.318	0.081	1.13	1.22
VF814758	740	1.074	0.081	0.98	0.96



Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF814839	740	0.216	0.087	1.10	1.09
VF814748	740	0.638	0.083	1.18	1.26
VF814762	740	-1.229	0.120	0.88	0.64
VF814688	740	0.290	0.086	1.18	1.23
VF883326	740	0.320	0.086	1.17	1.24
VF883330	740	0.962	0.081	1.01	1.01
VF883549	740	-0.614	0.102	0.87	0.81
VF883561	740	0.935	0.082	1.11	1.14
VF883619	740	-0.141	0.092	0.95	0.88
VF883622	740	-0.253	0.094	1.06	1.30
VF885434	740	0.093	0.088	0.83	0.78
VF885162	740	-1.038	0.114	0.92	0.82

Form 8					
VF814737	732	-0.407	0.096	0.86	0.73
VF814753	732	1.342	0.083	1.11	1.16
VF814821	732	2.122	0.089	1.22	1.64
VF814829	732	0.898	0.083	1.10	1.13
VF814673	732	-1.592	0.131	0.89	0.73
VF814681	732	-0.570	0.100	0.91	0.81
VF882884	732	1.069	0.083	0.97	0.98
VF882936	732	0.294	0.087	1.02	1.07
VF883543	732	0.047	0.089	1.11	1.06
VF883364	732	-1.978	0.149	0.83	0.49
VF883614	732	-0.398	0.096	1.04	1.20
VF883610	732	0.760	0.083	1.01	1.01
VF885187	732	0.499	0.085	1.02	1.04
VF885218	732	-0.142	0.092	0.78	0.70

Form 9					
VF821218	725	-0.497	0.098	1.02	1.05
VF821123	725	-0.212	0.093	1.08	1.08
VF821312	725	0.540	0.085	1.07	1.09
VF821362	725	0.533	0.085	0.96	0.97
VF821338	725	0.304	0.087	0.96	0.93
VF821088	725	-1.043	0.112	1.03	1.10
VF821030	725	0.468	0.086	1.10	1.10
VF821078	725	1.669	0.085	1.03	1.14
VF821006	725	-1.212	0.117	0.89	0.69
VF821011	725	0.497	0.085	1.08	1.09
VF821070	725	-1.268	0.119	0.80	0.66
VF885423	725	-0.556	0.100	0.90	0.81
VF885198	725	-0.689	0.103	0.91	0.76

Form 10					
VF821120	731	-0.190	0.093	0.92	0.85
VF821206	731	2.269	0.091	1.02	1.24
VF821272	731	-1.294	0.121	0.79	0.47
VF821320	731	1.015	0.083	1.05	1.09

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF821332	731	-0.006	0.090	0.88	0.80
VF821360	731	1.137	0.082	0.91	0.89
VF821072	731	1.137	0.082	1.18	1.26
VF821062	731	0.105	0.089	1.09	1.12
VF821037	731	-0.461	0.098	0.91	0.80
VF821055	731	-0.461	0.098	1.03	1.00
VF821065	731	-0.122	0.092	0.94	0.92
VF821024	731	1.124	0.082	1.20	1.30
VF885384	731	-2.123	0.160	0.84	0.46
VF885340	731	-1.167	0.116	0.85	0.65

Table L2. Reading Grade 4 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF822267	759	-0.94	0.12	0.91	0.77
VF822250	759	1.18	0.08	1.15	1.25
VF822284	759	0.31	0.09	1.12	1.21
VF822291	759	0.88	0.08	1.08	1.12
VF822301	759	0.08	0.09	1.10	1.21
VF822303	759	0.42	0.09	0.91	0.86
VF862927	759	-0.48	0.10	0.93	0.89
VF862890	759	-1.48	0.14	0.91	0.87
VF862920	759	-1.38	0.13	0.86	0.55
VF862909	759	0.67	0.09	1.04	1.02
VF862957	759	0.76	0.08	0.89	0.82
VF862882	759	0.02	0.09	0.99	0.89
VF885009	759	-0.74	0.11	0.84	0.70
VF885043	759	-0.65	0.11	0.99	0.80
Form 2					
VF885156	702	1.801	0.084	1.18	1.30
VF885173	702	-0.612	0.111	0.88	0.76
Form 3					
VF885200	705	-0.758	0.119	0.80	0.67
VF885215	705	1.164	0.084	1.22	1.25
Form 4					
VF822261	699	-1.522	0.153	0.81	0.42
VF822269	699	2.061	0.085	1.30	1.60
VF822292	699	-0.066	0.100	0.86	0.75
VF822294	699	0.720	0.087	1.15	1.21
VF822302	699	0.855	0.086	1.05	1.04
VF822298	699	1.812	0.084	1.32	1.48
VF862893	699	0.387	0.091	1.06	1.12
VF862946	699	-1.190	0.136	0.91	0.73
VF862897	699	-1.266	0.140	0.88	0.74
VF862965	699	0.115	0.096	1.01	0.98
VF862952	699	0.788	0.087	0.96	0.93
VF862870	699	-0.375	0.107	0.98	0.83
VF885233	699	-0.157	0.102	0.90	0.79
VF885037	699	0.705	0.087	0.98	0.93
Form 5					
VF884830	700	0.712	0.087	0.99	0.99

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF884836	700	1.189	0.084	0.93	0.94
VF884906	700	1.026	0.084	1.15	1.21
VF884910	700	-0.247	0.103	1.10	1.10
VF884918	1402	0.037	0.069	1.03	1.05
VF884913	700	2.107	0.086	1.05	1.17
VF880683	700	0.651	0.088	1.00	0.96
VF880649	700	1.653	0.083	1.14	1.28
VF880694	700	-0.036	0.098	1.20	1.37
VF880689	700	-1.896	0.172	0.90	0.63
VF880676	700	-0.788	0.118	0.92	0.79
VF880576	700	0.113	0.095	0.95	0.91
VF885059	700	1.849	0.084	0.96	1.02
VF885226	700	-0.859	0.121	0.86	0.77
Form 6					
VF884828	702	0.262	0.093	1.00	0.98
VF884843	702	-0.825	0.120	0.82	0.62
VF884900	702	-0.686	0.116	0.89	0.76
VF884896	702	-0.214	0.103	1.07	1.14
VF884925	702	0.101	0.096	1.04	1.11
VF880629	702	-0.607	0.113	0.97	0.96
VF880664	702	-0.121	0.101	0.95	0.82
VF880611	702	-0.471	0.109	1.01	0.92
VF880672	702	1.419	0.084	1.22	1.31
VF880678	702	0.390	0.091	1.10	1.11
VF880686	702	-1.349	0.142	0.96	1.00
VF885205	702	0.356	0.092	0.85	0.74
VF885195	702	-0.152	0.101	0.94	0.88
Form 7					
VF884587	695	-0.547	0.111	0.99	0.84
VF884603	695	0.417	0.091	1.15	1.15
VF884592	695	-0.572	0.112	1.04	1.18
VF884608	695	1.283	0.084	0.99	1.00
VF884605	695	-0.404	0.107	0.90	0.83
VF884561	695	0.844	0.086	1.05	1.00
VF864776	695	1.212	0.084	1.10	1.13
VF864822	695	0.655	0.088	0.92	0.85
VF864878	695	1.704	0.084	1.05	1.13
VF864868	695	1.163	0.084	0.91	0.88
VF864893	695	1.276	0.084	1.14	1.15
VF864887	695	1.423	0.084	1.08	1.11
VF885219	695	0.640	0.088	0.89	0.83
VF885232	695	-0.842	0.121	0.88	0.66
Form 8					
VF884582	682	-0.532	0.113	0.92	0.77
VF884602	682	1.166	0.085	1.04	1.07

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF884611	682	1.992	0.086	1.30	1.54
VF884593	682	1.064	0.086	1.07	1.11
VF884577	682	2.059	0.086	1.00	1.18
VF864828	682	2.633	0.093	1.15	1.52
VF864786	682	1.166	0.085	1.00	0.98
VF864876	682	1.773	0.085	1.12	1.21
VF864861	682	0.340	0.093	0.93	0.85
VF864895	682	2.202	0.088	1.06	1.25
VF864889	682	1.345	0.085	0.93	0.92
VF885028	682	-0.570	0.114	0.93	0.76
VF885064	682	1.302	0.085	1.10	1.20
Form 9					
VF884781	690	-0.453	0.109	0.95	1.02
VF884769	690	-0.218	0.103	1.06	1.08
VF884813	690	1.670	0.084	1.15	1.25
VF884777	690	0.344	0.092	1.06	1.11
VF884817	690	1.106	0.084	1.08	1.09
VF884734	690	1.656	0.083	1.17	1.29
VF880215	690	0.293	0.093	0.96	0.87
VF880200	690	0.285	0.093	0.91	0.86
VF880326	690	0.580	0.089	1.07	1.07
VF880314	690	2.452	0.090	1.09	1.52
VF880345	690	0.897	0.085	0.97	0.96
VF880343	690	0.875	0.086	1.04	1.05
VF885078	690	-0.564	0.112	0.95	0.85
VF885092	690	0.493	0.090	0.98	1.02
Form 10					
VF884743	684	0.807	0.087	0.98	1.00
VF884766	684	1.346	0.084	1.08	1.07
VF884773	684	-0.248	0.106	0.98	1.00
VF884802	684	1.833	0.084	1.32	1.61
VF884723	684	-2.044	0.191	0.84	0.46
VF884807	684	0.523	0.091	1.17	1.25
VF880204	684	1.642	0.084	0.94	0.94
VF880210	684	1.473	0.084	1.10	1.16
VF880311	684	1.261	0.084	1.00	1.02
VF880321	684	1.712	0.084	1.04	1.07
VF880350	684	0.994	0.086	1.07	1.10
VF885228	684	-0.140	0.103	0.89	0.77
VF885166	684	-0.601	0.116	0.87	0.76

Table L3. Reading Grade 5 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF884413	737	-1.532	0.156	0.84	0.60
VF884409	737	1.020	0.084	1.11	1.17
VF884354	737	-0.805	0.122	0.90	0.83
VF884360	737	2.524	0.086	1.20	1.39
VF888390	737	0.174	0.095	1.06	1.04
VF884312	737	1.153	0.083	0.93	0.89
VF884476	737	1.471	0.082	0.99	1.03
VF884481	737	0.183	0.095	0.93	0.84
VF884509	737	1.221	0.083	0.92	0.89
VF884517	737	0.165	0.095	0.84	0.69
VF884556	737	1.438	0.082	1.11	1.14
VF884535	737	0.913	0.085	1.05	1.04
VF885335	737	-0.609	0.115	0.82	0.59
VF885180	737	1.317	0.082	1.06	1.08
Form 2					
VF885197	705	-0.018	0.103	0.89	0.75
VF885191	705	-0.301	0.110	0.79	0.65
Form 3					
VF885224	700	0.422	0.094	0.98	1.02
VF885329	700	1.271	0.085	0.99	0.96
Form 4					
VF884420	701	3.259	0.099	1.28	1.90
VF884348	701	1.906	0.083	1.20	1.28
VF884341	701	0.392	0.095	0.90	0.80
VF884405	701	0.054	0.102	1.06	1.42
VF884336	701	1.181	0.085	1.00	0.99
VF884333	701	0.583	0.092	0.93	0.88
VF909884	701	1.544	0.083	1.00	1.01
VF884489	701	-1.218	0.151	0.87	0.54
VF884524	701	1.217	0.084	1.08	1.10
VF884520	701	0.624	0.091	1.08	1.06
VF884567	701	1.510	0.083	1.01	1.00
VF885345	701	1.468	0.083	1.08	1.12
VF885142	701	1.252	0.084	1.14	1.21

Form 5

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF880876	703	2.303	0.085	1.24	1.56
VF882773	703	1.372	0.084	1.29	1.41
VF909893	703	1.013	0.086	1.09	1.11
VF882794	703	1.511	0.083	1.19	1.25
VF822491	703	-0.073	0.105	0.87	0.71
VF822534	703	-0.197	0.108	0.92	1.04
VF822545	703	-0.04	0.104	1.00	0.99
VF822556	703	0.622	0.091	0.91	0.88
VF822548	703	1.766	0.083	1.03	1.08
VF822551	703	0.205	0.098	1.01	1.14
VF885161	703	0.426	0.094	1.02	0.97
VF885146	703	-0.084	0.105	0.87	0.74
Form 6					
VF880864	711	-1.191	0.145	0.92	0.76
VF881653	711	1.429	0.083	1.26	1.36
VF882769	711	0.334	0.094	1.09	1.15
VF882762	711	1.463	0.083	1.06	1.09
VF882790	711	1.381	0.083	1.08	1.12
VF882786	711	0.555	0.091	1.11	1.16
VF822538	711	1.484	0.083	1.01	1.04
VF822463	711	-0.666	0.122	0.94	0.78
VF822542	711	-0.203	0.106	1.00	1.02
VF822571	711	0.628	0.089	1.07	1.10
VF822496	711	1.214	0.084	1.08	1.12
VF822549	711	0.438	0.092	1.02	0.93
VF885134	711	-0.030	0.102	0.92	0.80
VF885204	711	2.002	0.083	1.00	1.07
Form 7					
VF822290	705	1.486	0.082	1.13	1.14
VF822271	705	3.797	0.118	1.10	1.59
VF822280	705	-0.542	0.116	1.01	1.05
VF822287	705	2.518	0.087	1.07	1.21
VF822278	705	1.850	0.082	1.17	1.27
VF822285	705	1.775	0.082	1.21	1.29
VF814960	705	-0.201	0.105	0.96	1.02
VF814977	705	0.738	0.087	0.94	0.87
VF814970	705	0.107	0.098	0.93	0.81
VF814962	705	1.526	0.082	1.03	1.06
VF814973	705	1.053	0.084	1.14	1.18
VF814958	705	1.081	0.084	1.09	1.10
VF885217	705	0.009	0.100	0.93	0.83
VF885212	705	-0.569	0.117	0.87	0.67
Form 8					
VF822288	697	-1.344	0.154	0.96	1.03
VF822275	697	2.479	0.088	1.09	1.13

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF822282	697	0.243	0.097	1.05	1.01
VF822276	697	0.660	0.090	1.09	1.25
VF822283	697	1.793	0.083	1.13	1.26
VF822259	697	-0.068	0.104	1.01	1.08
VF814959	697	2.494	0.088	1.15	1.35
VF814961	697	0.841	0.088	1.01	1.05
VF814963	697	3.018	0.095	1.11	1.39
VF814968	697	0.787	0.088	1.01	0.97
VF814975	697	-0.750	0.126	0.84	0.63
VF814956	697	1.557	0.083	1.06	1.08
VF885167	697	1.431	0.084	1.12	1.18
VF885154	697	0.243	0.097	0.94	0.84
Form 9					
VF822723	706	-1.196	0.146	0.87	0.56
VF822757	706	0.560	0.092	1.13	1.26
VF822776	706	-1.801	0.183	0.89	0.65
VF822832	706	0.509	0.093	1.03	1.02
VF822823	706	0.943	0.087	1.09	1.12
VF884191	706	0.209	0.098	0.92	0.88
VF884224	706	2.101	0.084	1.23	1.46
VF884208	706	1.899	0.083	1.16	1.25
VF884240	706	1.733	0.083	1.01	1.04
VF884152	706	1.540	0.083	1.16	1.21
VF884231	706	-0.519	0.118	0.88	0.69
VF885221	706	1.187	0.085	1.08	1.21
VF885202	706	1.602	0.083	1.12	1.21
Form 10					
VF822732	708	-1.177	0.145	0.98	0.95
VF822718	708	-0.090	0.104	1.05	1.14
VF822785	708	1.074	0.086	1.00	0.97
VF822797	708	-0.036	0.103	1.04	0.99
VF822829	708	-0.685	0.123	0.98	1.00
VF822821	708	1.059	0.086	1.17	1.28
VF884196	708	1.191	0.085	1.03	1.04
VF884218	708	0.343	0.095	0.93	0.82
VF884226	708	1.404	0.084	0.83	0.79
VF884213	708	1.565	0.083	1.11	1.18
VF884236	708	0.715	0.089	0.92	0.86
VF884158	708	2.953	0.093	1.01	1.25
VF885158	708	0.097	0.100	0.84	0.70
VF885314	708	0.361	0.095	0.92	0.84



Table L4. Reading Grade 6 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF883357	800	0.785	0.083	1.02	0.96
VF883356	800	1.618	0.079	1.04	1.05
VF883348	800	1.343	0.079	1.00	1.00
VF883351	800	2.498	0.082	1.32	1.58
VF883334	800	1.711	0.079	1.08	1.12
VF883365	800	1.835	0.079	1.10	1.13
VF884733	800	1.362	0.079	1.10	1.11
VF884751	800	1.139	0.080	1.19	1.23
VF884844	800	1.318	0.080	1.02	1.02
VF884814	800	2.643	0.084	1.40	1.74
VF884886	800	2.279	0.081	1.16	1.35
VF884880	800	0.819	0.083	1.09	1.04
VF885006	800	2.137	0.080	1.14	1.21
VF884659	800	0.874	0.082	0.93	0.86
Form 2					
VF884676	655	0.602	0.097	0.94	0.92
VF884630	655	1.242	0.090	1.13	1.21
Form 3					
VF884677	654	2.731	0.092	0.95	1.03
VF884693	654	0.454	0.101	0.94	0.81
Form 4					
VF883345	676	1.976	0.086	1.13	1.21
VF883367	676	0.687	0.096	0.96	0.96
VF883338	676	2.226	0.086	1.10	1.16
VF883354	676	0.575	0.098	0.96	0.88
VF883361	676	-0.668	0.135	0.86	0.60
VF883331	676	-0.401	0.124	0.98	1.40
VF884740	676	1.837	0.086	1.07	1.10
VF884772	676	2.278	0.086	1.02	1.05
VF884808	676	2.435	0.087	1.23	1.35
VF884876	676	2.642	0.088	1.18	1.43
VF884857	676	1.139	0.090	1.02	0.95
VF884624	676	0.239	0.105	1.02	0.93

Form 5

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF885203	679	1.909	0.085	1.06	1.06
VF885189	679	1.311	0.087	1.16	1.27
VF885141	679	1.765	0.085	1.13	1.18
VF885148	679	0.684	0.094	0.93	0.92
VF885178	679	1.073	0.089	1.02	1.01
VF885098	679	1.243	0.087	1.37	1.60
VF805047	679	0.475	0.097	1.06	1.07
VF805061	679	1.009	0.089	0.95	0.93
VF805054	679	1.415	0.086	1.15	1.19
VF805825	679	1.981	0.085	1.25	1.39
VF805822	679	1.460	0.086	1.06	1.07
VF804276	679	2.060	0.085	1.25	1.37
VF884658	679	2.017	0.085	1.05	1.15
VF884665	679	-0.152	0.112	0.91	0.80
Form 6					
VF885211	645	2.610	0.090	1.11	1.24
VF885199	645	2.700	0.091	1.06	1.10
VF885144	645	1.958	0.087	1.04	1.11
VF885152	645	-0.388	0.125	0.95	1.16
VF885193	645	1.686	0.087	1.28	1.37
VF885113	645	1.268	0.089	1.03	1.04
VF804289	645	1.064	0.091	1.02	1.07
VF805824	645	0.893	0.093	1.12	1.31
VF805055	645	2.232	0.088	1.05	1.09
VF805049	645	2.340	0.088	1.01	1.05
VF805052	645	0.390	0.102	1.03	1.03
VF804261	645	-0.373	0.125	0.96	0.93
VF884669	645	1.072	0.091	0.98	0.94
VF884654	645	2.011	0.087	1.08	1.10
Form 7					
VF820258	648	-0.457	0.129	1.05	1.28
VF820467	648	1.293	0.090	0.96	0.95
VF820394	648	-0.729	0.141	0.98	1.00
VF820332	648	0.635	0.098	1.00	1.07
VF820193	648	2.956	0.094	0.99	1.09
VF820442	648	1.437	0.089	1.16	1.21
VF821684	648	3.298	0.099	1.06	1.18
VF821664	648	0.683	0.098	0.90	0.82
VF821580	648	0.886	0.095	1.07	1.06
VF821704	648	-0.139	0.118	0.92	0.74
VF821619	648	1.325	0.090	1.08	1.09
VF821542	648	-0.709	0.140	0.92	0.82
VF885013	648	3.774	0.110	1.12	1.89
VF884657	648	2.113	0.087	0.95	0.96
Form 8					
VF820463	656	1.962	0.086	1.15	1.18
VF820281	656	0.311	0.104	1.10	1.26

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF820354	656	0.979	0.092	0.92	0.86
VF820218	656	1.171	0.090	1.00	1.05
VF820310	656	0.693	0.097	1.04	1.09
VF820457	656	0.479	0.101	1.01	0.98
VF821558	656	0.142	0.109	1.04	1.16
VF821572	656	0.211	0.107	1.03	1.15
VF821721	656	0.598	0.098	0.95	0.90
VF821673	656	1.798	0.087	1.08	1.08
VF821709	656	0.848	0.094	1.12	1.41
VF821429	656	0.712	0.096	0.94	0.88
VF884974	656	1.072	0.091	0.93	0.89
VF884689	656	2.756	0.091	1.13	1.21
Form 9					
VF814337	662	1.531	0.086	1.07	1.09
VF814311	662	-0.001	0.110	0.91	0.72
VF814388	662	0.962	0.091	0.97	0.92
VF814382	662	0.520	0.098	0.85	0.71
VF814394	662	2.116	0.086	1.25	1.37
VF814391	662	0.678	0.095	0.96	0.93
VF883112	662	2.190	0.087	1.02	1.07
VF883100	662	0.482	0.098	1.19	1.52
VF883095	662	0.227	0.104	0.98	1.00
VF883152	662	0.227	0.104	0.93	0.86
VF883106	662	1.249	0.088	1.05	1.08
VF883066	662	-0.950	0.148	0.92	0.77
VF884626	662	1.179	0.089	1.07	1.08
Form 10					
VF814327	680	0.630	0.096	0.88	0.77
VF814300	680	1.136	0.089	0.87	0.83
VF814384	680	2.553	0.088	1.25	1.39
VF814358	680	-0.605	0.130	0.93	0.79
VF814392	680	1.238	0.088	1.10	1.16
VF814393	680	2.553	0.088	1.09	1.22
VF883144	680	0.867	0.092	1.02	0.97
VF883088	680	1.268	0.088	1.20	1.24
VF883158	680	1.882	0.085	1.22	1.27
VF883061	680	0.798	0.093	0.86	0.78
VF883052	680	0.093	0.107	0.98	0.89
VF884988	680	1.268	0.088	0.98	0.93
VF884628	680	1.868	0.085	1.03	1.02

Table L5. Reading Grade 7 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF820419	767	1.559	0.081	0.96	0.95
VF820422	767	1.295	0.083	0.90	0.87
VF820444	767	-0.405	0.121	0.88	0.69
VF820435	767	2.800	0.083	1.09	1.20
VF820404	767	2.438	0.081	1.02	1.09
VF820464	767	1.638	0.081	1.17	1.20
VF864796	767	0.878	0.088	0.98	0.93
VF864756	767	2.517	0.081	1.02	1.03
VF864677	767	1.218	0.084	0.95	0.92
VF864684	767	1.037	0.086	1.05	1.07
VF864681	767	2.269	0.080	1.16	1.21
VF864667	767	0.600	0.093	1.03	1.07
VF885647	767	1.371	0.083	0.86	0.77
VF885607	767	1.840	0.080	1.15	1.23
Form 2					
VF885786	660	0.135	0.119	1.00	1.18
VF885813	660	2.349	0.086	1.07	1.10
Form 3					
VF885820	1319	2.252	0.061	0.86	0.85
VF885815	668	1.304	0.091	0.89	0.82
Form 4					
VF820466	671	2.364	0.084	1.09	1.12
VF820430	671	1.121	0.092	0.94	0.85
VF820412	671	0.758	0.099	1.00	1.08
VF820449	671	1.271	0.090	1.08	1.14
VF820438	671	-0.067	0.124	0.85	0.60
VF820391	671	1.230	0.091	1.03	1.01
VF864785	671	2.329	0.084	1.07	1.13
VF864676	671	0.288	0.111	0.92	0.84
VF864750	671	2.174	0.084	1.12	1.16
VF864685	671	0.844	0.097	0.91	0.84
VF864668	671	2.039	0.084	1.05	1.06
VF885757	671	1.163	0.092	1.08	1.12
VF885485	671	0.361	0.109	0.97	0.95

Form 5

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF814792	679	0.505	0.103	0.93	0.80
VF814759	679	1.762	0.085	1.11	1.16
VF814742	679	1.258	0.089	1.06	1.11
VF814788	679	2.809	0.086	1.05	1.05
VF814770	679	2.193	0.084	0.85	0.83
VF814720	679	0.794	0.097	0.97	0.92
VF865166	679	0.669	0.099	0.99	1.06
VF865194	679	2.922	0.087	1.25	1.37
VF865169	679	2.095	0.084	1.15	1.19
VF865185	679	2.116	0.084	0.88	0.87
VF865189	679	0.966	0.094	1.11	1.26
VF865164	679	4.050	0.109	1.15	1.79
VF885809	679	0.921	0.094	1.00	0.97
VF885612	679	2.334	0.084	1.04	1.07
Form 6					
VF814809	651	2.253	0.086	1.20	1.27
VF814766	651	1.671	0.088	0.91	0.91
VF814826	651	0.694	0.101	0.96	0.93
VF814781	651	0.578	0.104	1.07	1.09
VF814702	651	-0.560	0.147	0.94	0.68
VF865186	651	2.732	0.088	1.11	1.21
VF865195	651	1.702	0.087	1.19	1.30
VF865182	651	1.763	0.087	1.03	1.01
VF865187	651	2.364	0.086	1.07	1.09
VF865165	651	0.337	0.110	0.94	0.91
VF865141	651	0.814	0.099	0.95	0.94
VF885443	651	1.484	0.089	0.85	0.79
Form 7					
VF865426	675	0.628	0.100	0.94	0.93
VF865456	675	2.262	0.084	1.30	1.36
VF865482	675	0.727	0.098	1.09	1.34
VF865473	675	2.405	0.085	0.94	0.97
VF865624	675	1.478	0.087	1.05	1.05
VF865614	675	2.290	0.084	1.05	1.09
VF883991	675	1.883	0.085	1.09	1.13
VF883986	675	1.493	0.087	1.13	1.15
VF884006	675	0.812	0.096	0.98	0.91
VF884003	675	2.233	0.084	1.03	1.04
VF883998	675	1.455	0.088	1.07	1.06
VF883976	675	0.688	0.099	0.85	0.83
VF885659	675	2.878	0.088	1.06	1.09
VF910031	675	0.903	0.095	0.86	0.77
Form 8					
VF865413	666	1.203	0.092	1.02	1.02
VF865388	666	-0.328	0.134	0.86	0.58

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF865477	666	0.531	0.105	1.01	0.96
VF865494	666	1.938	0.085	1.10	1.16
VF865627	666	0.967	0.095	0.97	0.89
VF906623	666	2.466	0.085	1.10	1.19
VF883995	666	1.352	0.090	1.22	1.36
VF883997	666	-0.762	0.156	0.89	0.57
VF884008	666	1.408	0.089	0.91	0.84
VF884005	666	2.039	0.085	1.04	1.04
VF883972	666	0.430	0.107	0.97	0.90
VF883999	666	1.487	0.088	1.01	1.02
VF885797	666	0.079	0.118	0.86	0.64
VF885440	666	2.473	0.085	1.15	1.20
Form 9					
VF884891	679	1.867	0.085	1.14	1.19
VF885076	679	1.500	0.087	1.01	1.01
VF885063	679	3.673	0.099	1.17	1.68
VF884887	679	1.346	0.089	0.98	0.98
VF884846	679	-0.426	0.137	0.89	0.66
VF864898	679	-0.169	0.125	0.99	1.01
VF864910	679	3.309	0.092	1.31	1.70
VF865078	679	2.462	0.084	1.10	1.13
VF865063	679	2.024	0.084	1.10	1.15
VF865100	679	2.370	0.084	1.17	1.24
VF865094	679	2.649	0.085	1.04	1.03
VF885398	679	1.686	0.086	0.99	1.00
VF885385	679	0.603	0.101	0.84	0.69
Form 10					
VF884878	678	1.467	0.088	0.88	0.85
VF884859	678	2.056	0.084	1.04	1.05
VF885046	678	1.627	0.087	1.05	1.04
VF885072	678	1.913	0.085	1.01	1.02
VF885060	678	2.333	0.084	1.20	1.28
VF884855	678	-0.456	0.139	0.86	0.54
VF864902	678	0.951	0.095	0.88	0.79
VF865004	678	2.656	0.085	1.27	1.41
VF865057	678	1.244	0.091	0.83	0.73
VF865072	678	0.850	0.097	0.95	0.92
VF865104	678	1.899	0.085	0.91	0.87
VF865088	678	1.612	0.087	1.03	1.05
VF885375	678	0.783	0.098	0.99	1.12
VF885333	678	0.744	0.099	0.93	0.91

Table L6. Reading Grade 8 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF865107	793	1.639	0.082	1.12	1.17
VF864994	793	0.438	0.100	0.90	0.73
VF865075	793	2.435	0.079	1.24	1.36
VF865060	793	2.203	0.079	1.07	1.08
VF994816	793	2.712	0.080	1.19	1.30
VF997001	793	3.079	0.082	1.15	1.35
VF819971	793	2.712	0.080	1.19	1.26
VF820236	793	2.757	0.080	1.11	1.22
VF820165	793	2.241	0.079	1.05	1.09
VF820159	793	1.167	0.087	1.11	1.30
VF820261	793	1.842	0.080	1.02	1.00
VF883743	793	1.029	0.089	1.02	0.95
VF883621	793	1.504	0.083	0.96	0.97
Form 2					
VF820781	659	2.054	0.087	1.00	0.98
VF820771	659	2.562	0.086	1.03	1.09
VF820720	659	2.129	0.086	1.04	1.08
VF820786	659	1.538	0.091	1.20	1.31
VF820796	659	2.592	0.086	1.15	1.19
VF820792	659	1.684	0.089	1.11	1.12
VF883642	659	2.047	0.087	1.07	1.11
VF883685	659	2.173	0.086	1.19	1.26
Form 3					
VF820777	662	0.645	0.111	0.94	0.74
VF820740	662	1.112	0.101	1.00	0.99
VF820727	662	1.983	0.089	0.90	0.84
VF820734	662	-0.266	0.145	0.91	0.66
VF820750	662	1.152	0.100	1.07	1.15
VF820801	662	-0.204	0.142	0.91	0.72
VF883708	662	2.015	0.089	1.11	1.13
VF883653	662	0.020	0.132	0.89	0.73

Form 4

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF865101	668	0.853	0.106	1.09	1.16
VF865111	668	0.428	0.118	0.98	1.04
VF865050	668	2.584	0.085	0.99	1.00
VF865091	668	1.006	0.103	1.01	0.96
VF865171	668	1.637	0.092	1.10	1.20
VF865178	668	3.406	0.089	1.18	1.46
VF820011	668	1.793	0.090	1.19	1.28
VF819976	668	2.365	0.086	0.92	0.90
VF820170	668	2.758	0.085	0.99	1.03
VF820174	668	1.217	0.098	0.90	0.81
VF820025	668	0.898	0.105	0.94	0.84
VF820249	668	1.168	0.099	1.14	1.23
VF883823	668	0.995	0.103	0.97	0.87
VF883674	668	2.372	0.086	1.04	1.05
Form 5					
VF866201	679	-0.025	0.128	0.99	0.91
VF866195	679	1.127	0.097	0.94	0.91
VF866296	679	1.940	0.087	1.08	1.14
VF866316	679	-0.564	0.152	0.99	0.98
VF866341	679	0.668	0.106	0.94	0.90
VF867326	679	1.210	0.095	0.93	0.85
VF867246	679	1.299	0.094	0.94	0.88
VF867293	679	1.794	0.088	0.95	0.93
VF867267	679	2.736	0.086	1.26	1.45
VF867355	679	3.096	0.088	1.05	1.23
VF867368	679	1.909	0.087	1.14	1.15
VF883554	679	2.429	0.085	1.15	1.20
VF883624	679	1.360	0.093	0.95	0.90
Form 6					
VF866325	655	1.302	0.097	1.02	0.94
VF866173	655	0.677	0.111	0.95	0.77
VF866186	655	1.860	0.089	1.28	1.37
VF866307	655	-0.587	0.165	0.82	0.48
VF866228	655	1.311	0.097	1.04	1.00
VF866331	655	0.222	0.125	0.90	0.77
VF867197	655	1.118	0.100	0.93	0.78
VF867239	655	1.697	0.091	1.02	0.97
VF867333	655	1.492	0.094	0.91	0.87
VF867305	655	1.955	0.089	1.18	1.24
VF867338	655	1.630	0.092	1.00	0.96
VF867274	655	1.293	0.097	1.02	1.03
VF883655	655	2.133	0.087	0.99	0.99
VF883680	655	1.689	0.091	1.23	1.41
Form 7					
VF813900	657	1.260	0.097	0.99	0.95
VF813664	657	1.513	0.093	1.21	1.30



Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF813654	657	-0.637	0.166	0.88	0.54
VF813874	657	0.737	0.108	0.87	0.69
VF813924	657	1.902	0.089	0.99	0.95
VF813668	657	-0.038	0.135	0.92	0.95
VF813648	657	0.553	0.114	0.90	0.78
VF813641	657	0.260	0.123	1.04	1.12
VF813649	657	0.761	0.108	0.90	0.79
VF813671	657	2.058	0.088	1.07	1.04
VF813673	657	1.806	0.090	1.06	1.10
VF883695	657	3.281	0.090	1.43	1.85
VF883716	657	0.749	0.108	1.01	0.97
Form 8					
VF813601	666	2.036	0.087	1.03	1.05
VF813639	666	1.837	0.088	0.99	0.97
VF813842	666	0.903	0.101	0.98	0.96
VF813879	666	1.767	0.089	1.01	1.03
VF864943	666	0.522	0.111	0.87	0.69
VF813904	666	0.051	0.126	0.94	0.93
VF813667	666	2.096	0.086	1.02	1.00
VF813653	666	3.521	0.093	1.25	1.68
VF813655	666	1.003	0.099	1.01	0.95
VF813657	666	0.316	0.117	0.97	0.81
VF813645	666	3.141	0.089	1.03	1.19
VF813670	666	0.754	0.105	1.10	1.42
VF883726	666	4.029	0.103	1.05	1.36
Form 9					
VF812806	662	0.644	0.110	1.03	0.96
VF812982	662	2.404	0.085	1.13	1.16
VF812809	662	1.864	0.088	0.95	0.95
VF812818	662	2.120	0.086	0.97	0.97
VF812971	662	1.676	0.090	1.10	1.10
VF812965	662	-0.767	0.177	0.96	0.66
VF884543	662	3.105	0.087	1.12	1.29
VF884544	662	2.801	0.085	1.16	1.30
VF884581	662	3.586	0.093	1.20	1.49
VF884613	662	2.324	0.085	0.97	0.96
VF884609	662	2.793	0.085	1.06	1.16
VF883701	662	1.771	0.089	1.08	1.12
VF883817	662	3.386	0.090	1.04	1.16
Form 10					
VF812960	679	2.213	0.084	1.14	1.16
VF812799	679	1.294	0.093	0.97	0.94
VF812800	679	2.283	0.084	1.08	1.12
VF812796	679	2.663	0.084	0.98	1.05

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF812988	679	1.622	0.089	1.17	1.24
VF884547	679	3.552	0.092	1.27	1.65
VF884552	679	2.305	0.084	1.14	1.22
VF884583	679	0.808	0.103	0.92	0.89
VF884606	679	3.095	0.086	1.26	1.41
VF883629	679	2.613	0.084	1.01	1.04
VF883736	679	2.819	0.085	0.93	0.93

Mathematics

Table L7. Mathematics Grade 3 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	RaschSE	Infit	Outfit
Form 1					
VF803080	786	-0.691	0.089	0.94	0.89
VF867016	786	0.304	0.080	0.92	0.88
VF865397	786	-0.771	0.090	0.97	0.92
VF822811	786	0.841	0.080	0.98	0.96
VF867203	786	1.210	0.082	1.11	1.14
VF867073	786	0.777	0.080	1.12	1.21
VF866360	786	-1.470	0.106	0.82	0.56
VF867061	786	0.017	0.081	0.95	0.95
VF866931	786	0.336	0.080	1.08	1.12
VF865404	786	1.438	0.084	1.12	1.25
VF865420	786	1.230	0.082	0.81	0.77
VF866941	786	1.796	0.089	1.20	1.48
VF803183	786	0.253	0.080	0.97	0.94
Form 2					
VF803121	722	0.607	0.083	1.12	1.18
VF821403	722	-1.287	0.110	0.99	0.91
VF822819	722	2.076	0.094	1.18	1.62
VF867001	722	-0.181	0.088	0.94	0.94
VF819629	722	-3.009	0.204	1.04	1.32
VF866364	722	-0.775	0.097	0.96	1.09
VF867181	722	2.325	0.099	1.19	1.80
VF818296	722	-0.637	0.095	1.04	1.05
Form 3					
VF803161	724	0.030	0.085	1.01	0.98
VF821680	724	0.770	0.083	1.01	1.00
VF737752	724	0.887	0.083	0.83	0.80
VF867075	724	-0.130	0.086	1.04	1.04
VF866952	724	0.777	0.083	0.96	0.95
VF865570	724	0.449	0.083	0.88	0.83
VF821770	724	0.894	0.083	1.13	1.17
VF803199	724	0.214	0.084	0.92	0.88
Form 4					
VF803172	720	0.539	0.084	0.98	0.95
VF866981	720	-0.703	0.095	0.91	0.83
VF737767	720	0.203	0.085	0.89	0.84
VF740960	720	0.737	0.084	0.99	1.01
VF866256	720	1.246	0.086	0.97	1.00
VF819315	720	0.072	0.086	0.93	0.98
VF740954	720	-0.493	0.092	0.89	0.81
VF865462	720	-0.256	0.089	0.95	0.86
Form 5					

Accession Number	N	Rasch Difficulty	RaschSE	Infit	Outfit
VF866354	716	-1.187	0.104	0.88	0.74
VF866961	716	0.526	0.084	0.94	0.90
VF865381	716	-0.254	0.088	0.86	0.79
VF740917	716	0.804	0.085	0.91	0.90
VF740830	716	0.519	0.084	1.20	1.41
VF821698	716	0.941	0.085	1.15	1.22
VF822773	716	3.273	0.132	1.21	2.63
VF819577	716	-0.675	0.094	0.95	0.87
VF387508	716	-1.801	0.122	0.99	0.91
VF865285	716	-1.463	0.111	0.92	1.05
VF740915	716	2.044	0.097	0.95	1.09
VF819363	716	1.521	0.089	1.12	1.37
VF822685	716	0.189	0.085	0.92	0.84
Form 6					
VF866235	739	-0.248	0.086	0.96	0.91
VF867176	739	2.622	0.108	0.99	1.22
VF821665	739	-0.248	0.086	1.10	1.08
VF819348	739	-0.769	0.094	1.08	1.23
VF866906	739	0.285	0.083	0.97	0.96
VF819555	739	0.077	0.084	0.91	0.93
VF865389	739	-0.001	0.084	0.93	0.89
VF740959	739	1.550	0.088	1.31	1.64
VF865371	739	0.666	0.082	0.98	0.97
VF821723	739	0.920	0.083	1.07	1.08
VF867224	739	-0.743	0.093	0.93	0.80
VF821767	739	1.208	0.085	1.12	1.16
VF822716	739	0.961	0.083	1.10	1.17
Form 7					
VF866264	744	-0.483	0.087	0.94	1.01
VF821729	744	0.252	0.081	0.97	0.94
VF865449	744	-1.751	0.118	1.01	1.02
VF819660	744	1.341	0.085	1.08	1.20
VF803266	744	0.765	0.081	0.91	0.89
VF819299	744	0.476	0.081	1.00	1.01
VF865405	744	0.219	0.081	0.98	0.94
VF865478	744	0.858	0.082	1.08	1.13
VF737761	744	2.037	0.095	1.38	2.08
VF740949	744	-0.261	0.085	1.15	1.25
VF821407	744	-2.127	0.134	0.92	0.73
VF803242	744	0.607	0.081	0.93	0.92
VF822742	744	1.088	0.083	1.30	1.41
Form 8					
VF866946	739	0.196	0.083	0.84	0.77
VF866996	739	1.552	0.088	1.13	1.17

Accession Number	<i>N</i>	Rasch Difficulty	RaschSE	Infit	Outfit
VF819669	739	0.938	0.083	1.21	1.33
VF821481	739	-1.181	0.102	0.95	0.89
VF819375	739	1.042	0.084	0.88	0.84
VF822709	739	-0.224	0.086	1.01	0.96
VF803290	739	-0.086	0.085	1.04	1.14
VF818374	739	0.140	0.083	0.83	0.76
VF865302	739	1.750	0.090	1.18	1.44
VF821738	739	-2.077	0.133	0.92	0.67
VF819676	739	-1.366	0.107	1.11	1.32
VF818365	739	-0.791	0.094	1.06	1.15
VF822822	739	-0.629	0.091	0.92	0.82
Form 9					
VF822784	735	1.661	0.088	1.21	1.37
VF866988	735	-1.436	0.111	0.86	0.78
VF867009	735	0.979	0.082	1.01	1.04
VF822725	735	1.930	0.092	1.11	1.34
VF865323	735	1.081	0.083	1.07	1.13
VF819654	735	0.972	0.082	1.07	1.12
VF819622	735	-2.914	0.190	0.96	0.77
VF867066	735	-1.306	0.107	0.85	0.61
VF865468	735	1.319	0.085	1.02	1.12
VF819598	735	2.060	0.095	1.17	1.59
VF866898	735	-0.658	0.092	0.99	0.99
VF740957	735	2.227	0.098	1.12	1.47
VF821652	735	0.478	0.082	1.04	1.04
Form 10					
VF740890	730	0.590	0.083	1.02	0.98
VF821745	730	0.652	0.083	1.07	1.11
VF819675	730	-1.856	0.124	0.96	0.95
VF819337	730	0.797	0.083	0.90	0.88
VF803307	730	1.008	0.084	0.92	0.91
VF865488	730	0.597	0.083	1.20	1.25
VF866888	730	1.418	0.087	0.98	1.03
VF819543	730	0.342	0.083	0.95	0.95
VF867023	730	1.542	0.089	1.21	1.39
VF819639	730	2.755	0.113	0.92	1.07
VF865414	730	-0.269	0.087	1.00	0.97
VF865316	730	-1.341	0.107	0.87	0.68
VF821387	730	0.902	0.084	0.96	0.95

Table L8. Mathematics Grade 4 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF816041	765	0.815	0.081	1.00	1.07
VF865554	765	0.663	0.082	1.03	1.10
VF880261	765	2.289	0.086	1.01	1.08
VF867083	765	1.059	0.080	1.07	1.10
VF741942	765	2.341	0.086	1.19	1.35
VF800697	765	0.513	0.083	1.20	1.50
VF866686	765	1.772	0.081	1.13	1.22
VF880294	765	0.995	0.080	0.85	0.78
VF816026	765	1.553	0.080	0.90	0.87
VF816159	765	0.976	0.080	0.83	0.81
VF801227	765	2.082	0.083	1.07	1.09
VF823081	765	1.502	0.080	0.98	0.98
VF823371	765	0.308	0.086	1.04	1.02
VF866870	765	2.188	0.085	1.11	1.24
Form 2					
VF816048	704	0.366	0.090	0.91	0.88
VF823138	704	0.486	0.089	0.86	0.76
VF880252	704	0.215	0.093	0.87	0.74
VF880336	704	2.498	0.090	0.94	0.99
Form 3					
VF822848	699	0.046	0.096	0.90	0.85
VF741944	699	-0.483	0.110	1.01	0.96
VF823036	699	2.890	0.098	1.06	1.13
VF880325	699	1.071	0.083	1.05	1.08
Form 4					
VF866662	700	0.920	0.085	0.95	0.91
VF880413	700	1.370	0.083	0.88	0.84
VF815936	700	1.948	0.084	0.99	1.00
VF741929	700	2.417	0.089	1.06	1.12
Form 5					
VF816151	692	-0.572	0.113	0.93	0.80
VF880305	692	2.217	0.088	1.09	1.23
VF823330	692	2.519	0.091	0.96	0.98
VF880421	692	1.629	0.084	0.81	0.77
VF866416	692	2.133	0.087	1.11	1.17
VF867088	692	1.862	0.085	0.96	0.97
VF816057	692	0.700	0.087	0.98	0.97
VF801810	692	1.770	0.084	1.00	0.98
VF815880	692	0.484	0.089	1.03	1.09
VF816028	692	0.708	0.087	0.94	0.92
VF864158	692	3.742	0.121	1.05	1.31
VF815975	692	-2.649	0.255	0.99	0.74
VF866672	692	2.186	0.087	0.93	0.96
VF866699	692	1.218	0.084	0.87	0.89
Form 6					

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF822854	687	-1.197	0.138	0.91	0.63
VF815849	687	2.649	0.094	1.18	1.33
VF866677	687	0.295	0.091	0.93	0.86
VF801214	687	1.798	0.084	0.96	0.95
VF815888	687	1.088	0.084	1.11	1.16
VF815962	687	0.480	0.089	0.96	0.91
VF880334	687	2.477	0.092	1.19	1.43
VF741950	687	0.132	0.094	0.91	0.81
VF864145	687	1.636	0.084	1.11	1.21
VF863975	687	-0.360	0.105	0.99	0.89
VF815942	687	-0.119	0.099	0.99	1.02
VF866847	687	2.329	0.089	1.23	1.42
VF823410	687	1.986	0.086	0.90	0.89
VF880341	687	2.768	0.097	0.91	1.00
Form 7					
VF822864	690	2.297	0.089	1.07	1.10
VF815875	690	0.677	0.087	0.89	0.85
VF823304	690	0.553	0.088	1.16	1.37
VF880328	690	2.258	0.088	0.95	0.95
VF800875	690	0.285	0.092	0.84	0.74
VF741919	690	1.157	0.084	1.05	1.05
VF741945	690	2.120	0.087	0.93	0.89
VF801835	690	0.506	0.089	1.10	1.20
VF864141	690	2.817	0.097	1.19	1.57
VF867091	690	2.353	0.089	1.27	1.38
VF864104	690	2.052	0.086	0.81	0.79
VF864153	690	2.377	0.090	0.89	0.90
VF815948	690	0.233	0.093	1.00	0.93
VF815909	690	-0.781	0.120	0.85	0.67
Form 8					
VF822870	697	1.600	0.084	0.98	1.08
VF865651	697	1.396	0.084	0.95	0.94
VF880443	697	2.121	0.087	0.89	0.92
VF867078	697	2.450	0.090	0.89	0.94
VF741948	697	2.098	0.087	1.22	1.31
VF866702	697	4.824	0.174	1.08	2.31
VF864149	697	1.529	0.084	1.05	1.06
VF822874	697	2.758	0.095	1.14	1.34
VF866381	697	0.928	0.085	0.87	0.83
VF815957	697	0.737	0.087	0.98	0.93
VF866410	697	3.035	0.100	1.09	1.37
VF864100	697	0.393	0.090	0.97	0.95
VF800889	697	1.480	0.084	0.84	0.81
VF866368	697	3.149	0.103	1.13	1.40
Form 9					
VF864111	688	1.889	0.085	0.97	1.01
VF864078	688	-0.398	0.108	0.93	0.74

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF742706	688	0.783	0.086	0.99	1.02
VF866714	688	-0.410	0.108	0.88	0.72
VF866709	688	1.018	0.085	1.00	0.98
VF816162	688	1.379	0.084	0.97	0.96
VF866402	688	0.086	0.096	1.02	0.95
VF823145	688	-0.220	0.103	0.99	1.00
VF866392	688	3.039	0.101	1.19	1.59
VF741947	688	2.217	0.088	1.09	1.34
VF866857	688	1.217	0.084	0.94	0.89
VF864035	688	3.267	0.107	0.90	1.16
VF741936	688	3.703	0.121	1.12	1.35
VF815303	688	3.733	0.122	1.10	1.74
Form 10					
VF823141	700	2.085	0.087	1.08	1.14
VF864051	700	-0.021	0.097	0.88	0.75
VF867086	700	0.682	0.087	0.90	0.83
VF866691	700	-0.510	0.109	1.12	1.47
VF867084	700	2.116	0.088	1.00	1.01
VF815900	700	1.290	0.084	1.10	1.13
VF866696	700	1.861	0.086	1.10	1.20
VF823000	700	0.466	0.089	0.87	0.73
VF866830	700	0.904	0.085	0.92	0.87
VF880274	700	1.191	0.084	1.07	1.04
VF741924	700	0.151	0.094	0.93	0.87
VF741949	700	2.248	0.089	0.93	1.00
VF880269	700	-0.330	0.104	0.85	0.72
VF823196	700	1.346	0.084	1.02	1.02



Table L9. Mathematics Grade 5 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF802791	757	4.341	0.119	1.05	1.22
VF801975	757	3.391	0.095	1.03	1.30
VF798083	757	1.077	0.086	0.91	0.85
VF819978	757	1.348	0.084	0.94	0.90
VF866083	757	1.773	0.082	1.05	1.04
VF736258	757	3.455	0.096	1.10	1.41
VF740925	757	1.334	0.084	0.99	1.03
VF823764	757	0.805	0.088	0.88	0.82
VF880826	757	2.045	0.083	1.20	1.25
VF741570	757	0.790	0.089	1.05	1.11
VF816137	757	2.619	0.085	1.23	1.37
VF823490	757	1.760	0.082	0.98	0.96
VF864609	757	2.341	0.084	0.88	0.85
VF819989	757	2.348	0.084	0.98	0.96
Form 2					
VF802051	704	2.175	0.084	1.26	1.37
VF741381	704	3.348	0.095	1.10	1.22
VF864581	704	2.816	0.088	1.28	1.47
VF865968	704	3.312	0.095	1.11	1.38
Form 3					
VF823809	703	0.683	0.095	1.42	2.17
VF819955	703	3.407	0.096	1.10	1.20
VF823498	703	5.061	0.153	1.13	2.24
VF880803	703	1.696	0.085	0.92	0.89
Form 4					
VF823474	716	1.757	0.084	0.93	0.89
VF740894	716	0.195	0.106	0.92	0.86
VF802821	716	3.491	0.096	1.01	1.20
VF736492	716	1.438	0.086	0.91	0.88
Form 5					
VF864587	711	1.441	0.086	1.03	0.98
VF741081	711	1.615	0.085	1.09	1.13
VF741507	711	3.131	0.092	1.20	1.35
VF880721	711	2.442	0.086	1.01	1.00
VF802763	711	2.696	0.087	1.28	1.43
VF864638	711	3.748	0.104	1.16	1.43
VF823729	711	1.849	0.084	1.03	1.00
VF823759	711	2.964	0.090	0.98	1.04
VF801992	711	1.015	0.089	0.91	0.84
VF736482	711	-0.258	0.119	0.99	1.00
VF816183	711	2.673	0.087	1.02	1.04
VF736475	711	2.884	0.089	1.26	1.45
VF816152	711	0.934	0.090	1.07	1.08
VF741093	711	2.696	0.087	1.01	1.09
Form 6					

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF802778	707	2.837	0.088	1.11	1.24
VF741450	707	0.541	0.096	0.95	0.91
VF816021	707	1.912	0.084	0.96	0.97
VF866061	707	2.199	0.084	0.97	0.99
VF864521	707	1.801	0.084	0.81	0.76
VF741371	707	0.211	0.104	0.94	0.80
VF864548	707	3.331	0.095	1.02	1.04
VF816005	707	3.313	0.095	0.94	1.04
VF823504	707	2.892	0.089	1.35	1.59
VF823790	707	1.475	0.085	1.15	1.25
VF741416	707	2.341	0.084	1.20	1.35
VF864628	707	1.255	0.086	1.23	1.39
VF736633	707	0.970	0.090	0.90	0.83
VF864671	707	3.396	0.097	1.38	1.93
Form 7					
VF823638	695	0.942	0.090	0.95	0.88
VF797033	695	2.046	0.084	1.31	1.43
VF865989	695	2.733	0.089	0.80	0.73
VF864536	695	3.302	0.097	1.00	0.99
VF741106	695	0.630	0.094	1.01	1.19
VF802069	695	2.917	0.091	1.32	1.44
VF823779	695	2.311	0.085	1.31	1.37
VF880813	695	1.784	0.084	0.91	0.86
VF864590	695	2.348	0.085	1.16	1.18
VF815866	695	2.282	0.085	1.26	1.38
VF819994	695	1.678	0.084	1.03	1.01
VF864614	695	1.332	0.086	1.03	0.96
VF866022	695	3.331	0.097	1.06	1.20
VF797963	695	1.699	0.084	0.99	0.97
Form 8					
VF741941	693	1.509	0.085	1.00	0.99
VF865997	693	3.520	0.099	1.06	1.04
VF866065	693	1.831	0.084	1.08	1.09
VF741573	693	0.798	0.093	1.05	1.10
VF802894	693	1.938	0.084	1.13	1.17
VF880786	693	2.661	0.087	0.91	0.92
VF866009	693	0.062	0.109	0.94	0.82
VF802089	693	3.084	0.092	1.18	1.46
VF741405	693	3.109	0.092	1.23	1.47
VF802032	693	2.257	0.085	1.34	1.45
VF736503	693	0.026	0.110	0.90	0.88
VF741052	693	1.248	0.087	0.96	0.89
VF823652	693	1.767	0.084	0.87	0.84
VF802860	693	2.460	0.086	1.46	1.61
Form 9					
VF864604	690	2.066	0.086	1.06	1.08
VF815846	690	1.934	0.086	1.08	1.13

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF741193	690	0.568	0.097	1.01	1.10
VF815902	690	2.507	0.088	1.02	1.06
VF801897	690	1.839	0.086	0.78	0.71
VF864641	690	4.732	0.138	1.11	2.22
VF741382	690	1.192	0.089	1.07	1.18
VF823838	690	-0.082	0.113	1.00	1.43
VF797110	690	1.942	0.086	1.07	1.05
VF797938	690	1.355	0.088	0.98	0.94
VF802014	690	1.301	0.088	0.84	0.75
VF736495	690	0.742	0.094	1.05	1.11
VF866037	690	2.670	0.089	1.04	1.11
VF815982	690	2.798	0.090	1.00	1.00
Form 10					
VF819900	695	0.030	0.110	1.04	1.16
VF823819	695	2.253	0.085	1.30	1.37
VF802870	695	1.690	0.085	1.21	1.38
VF736438	695	3.289	0.096	1.16	1.32
VF741539	695	0.774	0.093	0.92	0.85
VF740936	695	0.817	0.093	0.93	0.87
VF736524	695	0.462	0.099	0.94	0.90
VF741389	695	2.079	0.085	1.10	1.15
VF864618	695	3.103	0.093	1.37	1.63
VF815953	695	1.245	0.088	1.04	1.06
VF802847	695	4.804	0.140	0.96	1.10
VF866034	695	2.311	0.086	1.43	1.57
VF864557	695	3.068	0.092	0.97	1.09
VF866103	695	1.647	0.085	1.11	1.18

Table L10. Mathematics Grade 6 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF862699	806	1.774	0.079	0.94	0.89
VF741557	806	2.682	0.079	0.94	0.92
VF741723	806	2.276	0.077	1.12	1.19
VF810665	806	3.363	0.085	1.06	1.15
VF883019	806	1.933	0.078	1.16	1.26
VF741771	806	2.324	0.077	1.11	1.14
VF797171	806	2.883	0.080	1.29	1.43
VF810667	806	3.542	0.088	0.96	1.02
VF797964	806	3.229	0.083	1.12	1.19
VF865682	806	3.481	0.087	1.09	1.18
VF822031	806	2.414	0.078	1.08	1.08
VF803280	806	2.150	0.077	0.96	0.91
VF741572	806	2.414	0.078	0.98	0.99
VF797954	806	4.043	0.097	1.53	2.26
Form 2					
VF741728	651	2.423	0.087	0.87	0.84
VF803302	651	2.333	0.087	0.93	0.93
VF865661	651	3.241	0.092	1.25	1.43
VF821954	651	3.459	0.095	1.14	1.32
Form 3					
VF862885	664	3.088	0.089	1.28	1.42
VF803311	664	2.622	0.086	1.24	1.32
VF865649	664	1.942	0.088	0.97	0.96
VF882800	664	2.347	0.086	1.12	1.12
Form 4					
VF882956	679	0.860	0.102	1.00	0.92
VF803328	679	2.807	0.086	1.09	1.11
VF865678	679	2.192	0.086	1.02	1.00
VF821946	679	2.126	0.086	1.09	1.10
Form 5					
VF741574	681	2.426	0.085	1.18	1.27
VF862858	681	1.893	0.086	1.10	1.14
VF741711	681	2.097	0.085	1.02	1.05
VF741566	681	3.465	0.093	1.27	1.56
VF809034	681	1.345	0.091	0.97	0.94
VF741928	681	0.957	0.097	0.93	0.84
VF812407	681	2.952	0.087	1.05	1.10
VF803399	681	2.832	0.086	1.05	1.08
VF797120	681	3.535	0.094	1.32	1.54
VF741533	681	2.700	0.085	1.06	1.08
VF821929	681	1.113	0.094	1.01	1.03
VF797970	681	0.938	0.098	0.85	0.73
VF741515	681	3.544	0.094	1.08	1.13
VF866206	681	2.952	0.087	1.31	1.42
Form 6					

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF883002	662	2.112	0.087	1.36	1.52
VF809062	662	1.677	0.090	0.95	0.90
VF862786	662	2.210	0.086	0.89	0.84
VF882803	662	3.878	0.098	1.41	1.95
VF741578	662	2.388	0.086	1.09	1.12
VF797163	662	2.773	0.086	1.25	1.32
VF741576	662	3.602	0.094	1.18	1.42
VF803386	662	3.481	0.092	1.04	1.02
VF741934	662	2.841	0.087	1.09	1.15
VF865621	662	3.859	0.098	1.09	1.35
VF821920	662	1.757	0.089	1.07	1.12
VF797977	662	1.302	0.095	0.89	0.86
VF822004	662	1.435	0.093	0.89	0.80
VF797944	662	4.684	0.119	1.29	2.24
Form 7					
VF741690	637	1.878	0.089	1.20	1.33
VF810689	637	1.957	0.089	0.86	0.83
VF862804	637	0.722	0.108	1.01	1.15
VF812185	637	1.965	0.089	1.03	1.09
VF882963	637	3.080	0.091	1.23	1.37
VF821992	637	2.926	0.089	1.23	1.33
VF865650	637	1.878	0.089	1.08	1.11
VF741935	637	2.799	0.089	1.04	1.07
VF866265	637	0.890	0.104	1.04	1.15
VF865654	637	3.586	0.097	1.11	1.34
VF882811	637	2.838	0.089	0.98	1.00
VF821988	637	1.634	0.091	1.09	1.18
VF821963	637	1.846	0.089	0.94	0.90
VF866221	637	1.782	0.090	0.97	0.95
Form 8					
VF811515	654	2.955	0.089	0.95	0.95
VF741538	654	1.177	0.097	0.87	0.70
VF882993	654	2.413	0.087	1.06	1.10
VF809839	654	1.887	0.088	0.80	0.80
VF741699	654	4.570	0.118	1.08	1.29
VF883062	654	3.537	0.095	1.28	1.58
VF882780	654	3.010	0.089	1.36	1.54
VF865635	654	2.466	0.087	0.93	0.93
VF882808	654	3.195	0.091	1.12	1.20
VF797981	654	2.179	0.087	1.14	1.15
VF883067	654	1.895	0.088	1.13	1.20
VF797996	654	1.196	0.097	0.93	0.83
VF866290	654	2.285	0.087	1.09	1.10
VF821976	654	2.662	0.087	0.98	0.99
Form 9					
VF741668	664	3.313	0.091	1.03	1.09
VF741549	664	1.108	0.097	0.89	0.80

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF741781	664	0.223	0.122	1.02	1.71
VF810696	664	2.352	0.086	0.91	0.91
VF809076	664	4.459	0.114	1.18	2.01
VF821906	664	0.952	0.100	1.01	1.00
VF882789	664	3.901	0.100	1.30	1.65
VF803317	664	2.190	0.086	1.00	1.00
VF883071	664	0.731	0.105	0.97	0.91
VF865668	664	2.528	0.086	1.09	1.12
VF803393	664	1.108	0.097	0.89	0.88
VF821939	664	3.983	0.102	1.16	1.60
VF822023	664	1.552	0.090	0.97	0.94
VF866230	664	2.131	0.086	0.96	0.97
Form 10					
VF862813	657	2.505	0.086	1.16	1.25
VF810701	657	3.191	0.090	1.13	1.20
VF741692	657	3.364	0.092	1.02	1.14
VF741562	657	3.356	0.092	1.38	1.76
VF811529	657	3.627	0.096	1.22	1.54
VF865671	657	2.351	0.086	0.99	1.02
VF882795	657	2.137	0.086	1.11	1.19
VF821998	657	1.425	0.092	0.95	0.90
VF866278	657	2.728	0.086	1.06	1.10
VF803324	657	2.002	0.087	0.97	0.95
VF741859	657	3.017	0.088	1.28	1.34
VF803293	657	2.698	0.086	0.99	1.01
VF822007	657	3.072	0.089	1.04	1.08
VF866301	657	4.646	0.122	1.23	1.90

Table L11. Mathematics Grade 7 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF880308	779	2.238	0.079	0.98	0.96
VF880331	779	3.854	0.089	1.03	1.16
VF866890	779	2.175	0.080	0.89	0.85
VF880323	779	2.660	0.079	1.17	1.21
VF823091	779	3.784	0.088	1.22	1.36
VF736963	779	3.427	0.083	1.12	1.17
VF883244	779	2.512	0.079	1.02	0.99
VF866547	779	4.331	0.098	1.17	1.65
VF736931	779	3.010	0.080	1.02	1.03
VF882920	779	3.685	0.086	0.98	0.99
VF867243	779	2.809	0.079	1.08	1.13
VF883150	779	2.326	0.079	1.08	1.33
VF880171	779	4.111	0.093	1.16	1.34
VF736947	779	2.112	0.080	0.92	0.86
Form 2					
VF818173	657	4.203	0.100	1.05	1.26
VF799825	657	1.235	0.102	0.90	0.95
VF882559	657	3.636	0.091	1.24	1.39
VF866826	657	3.612	0.091	1.28	1.45
Form 3					
VF866499	661	3.107	0.087	0.99	0.99
VF822880	661	3.069	0.087	1.02	1.05
VF800078	661	2.898	0.086	1.26	1.37
VF822986	661	2.344	0.086	1.08	1.06
Form 4					
VF813483	674	3.682	0.091	0.87	0.87
VF819351	674	3.922	0.094	1.04	1.14
VF867315	674	3.221	0.086	1.20	1.29
VF882910	674	2.621	0.084	1.03	1.05
Form 5					
VF736959	681	3.486	0.087	1.04	1.05
VF867219	681	3.854	0.092	1.13	1.27
VF818177	681	3.547	0.088	0.94	0.94
VF880250	681	3.373	0.086	1.10	1.20
VF736957	681	3.754	0.090	1.27	1.54
VF867395	681	2.418	0.084	1.04	1.05
VF870864	681	2.827	0.084	1.10	1.14
VF882715	681	3.905	0.093	1.09	1.19
VF867292	681	4.604	0.108	1.22	1.78
VF818184	681	2.946	0.084	1.07	1.09
VF736938	681	2.325	0.085	0.88	0.82
VF822889	681	4.675	0.110	0.95	1.10
VF866506	681	3.914	0.093	1.09	1.23
VF819294	681	2.961	0.084	0.95	1.03
Form 6					

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF822884	648	4.985	0.122	1.08	1.44
VF817427	648	3.374	0.089	1.18	1.27
VF800136	648	2.109	0.089	1.03	0.99
VF813502	648	2.349	0.087	0.88	0.81
VF813530	648	3.421	0.089	1.14	1.17
VF819535	648	4.415	0.105	1.13	1.33
VF813096	648	3.037	0.087	0.96	0.94
VF736961	648	3.768	0.093	1.06	1.15
VF867377	648	4.106	0.098	1.35	1.70
VF818347	648	1.014	0.109	1.12	1.62
VF866386	648	3.280	0.088	1.10	1.17
VF881807	648	1.972	0.090	0.85	0.75
VF868691	648	3.666	0.092	1.13	1.26
VF883264	648	3.343	0.088	0.95	0.94
Form 7					
VF800144	686	2.192	0.086	0.94	0.90
VF823026	686	3.977	0.094	1.11	1.28
VF818181	686	3.457	0.087	0.97	1.00
VF867610	686	3.699	0.090	1.05	1.17
VF866421	686	1.859	0.089	1.03	1.11
VF818335	686	2.934	0.084	1.20	1.37
VF880897	686	1.614	0.092	0.88	0.77
VF866491	686	4.132	0.097	0.86	0.87
VF800055	686	1.179	0.101	0.88	0.78
VF867323	686	3.133	0.085	1.33	1.45
VF736940	686	3.549	0.088	0.97	1.01
VF883129	686	2.408	0.084	0.94	0.91
VF866539	686	3.683	0.090	1.11	1.23
VF882732	686	1.747	0.090	0.96	1.00
Form 8					
VF818182	659	3.478	0.089	1.03	1.09
VF866376	659	4.569	0.109	1.27	1.98
VF800133	659	3.795	0.093	0.94	0.92
VF867183	659	4.197	0.100	1.14	1.28
VF866963	659	3.658	0.092	0.89	0.83
VF800103	659	3.048	0.086	0.89	0.86
VF867307	659	1.985	0.089	1.09	1.22
VF883138	659	3.375	0.088	1.18	1.23
VF869623	659	2.484	0.086	0.88	0.84
VF822997	659	1.423	0.097	0.92	0.89
VF819696	659	4.258	0.102	1.25	1.38
VF882691	659	2.820	0.086	0.95	0.94
VF882746	659	3.567	0.090	1.14	1.23
VF866531	659	3.244	0.087	1.02	1.02
Form 9					
VF736954	674	2.59	0.084	0.87	0.83
VF818180	674	3.275	0.086	1.14	1.24



Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF867060	674	4.000	0.096	0.97	1.03
VF818174	674	1.861	0.088	0.96	0.98
VF813104	674	3.442	0.088	1.03	1.11
VF819358	674	3.253	0.086	1.07	1.11
VF867260	674	1.976	0.087	0.82	0.73
VF866401	674	3.480	0.088	1.25	1.37
VF819694	674	3.320	0.087	1.40	1.54
VF882739	674	2.155	0.086	1.09	1.13
VF813100	674	2.935	0.084	1.24	1.29
VF883156	674	2.906	0.084	0.90	0.91
VF819306	674	1.884	0.088	0.87	0.79
VF867401	674	3.201	0.086	1.07	1.08
Form 10					
VF823079	677	5.085	0.127	1.04	1.43
VF880312	677	3.437	0.088	1.21	1.33
VF867038	677	2.980	0.084	1.13	1.16
VF818183	677	2.902	0.084	1.02	1.02
VF813490	677	2.678	0.084	0.86	0.82
VF866423	677	3.944	0.095	1.09	1.19
VF736941	677	3.899	0.094	1.08	1.19
VF882946	677	1.365	0.096	0.89	0.80
VF867365	677	3.890	0.094	1.40	1.54
VF867256	677	2.804	0.084	0.95	0.93
VF880886	677	2.320	0.084	1.12	1.30
VF799837	677	1.498	0.093	0.88	0.77
VF818361	677	2.284	0.084	1.13	1.19
VF883220	677	3.492	0.088	1.38	1.48

Table L12. Mathematics Grade 8 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF802927	828	3.457	0.079	1.03	1.02
VF885497	828	3.420	0.078	0.99	1.01
VF802937	828	3.031	0.077	1.09	1.11
VF812962	828	2.673	0.077	0.84	0.78
VF810708	828	1.782	0.084	0.98	0.92
VF810643	828	4.088	0.085	1.18	1.36
VF823784	828	2.296	0.078	1.02	1.01
VF880641	828	4.110	0.086	0.96	1.02
VF865981	828	3.501	0.079	1.25	1.32
VF883670	828	3.055	0.077	1.00	0.99
VF812445	828	4.710	0.097	1.21	1.58
VF809001	828	3.513	0.079	0.90	0.89
VF883722	828	4.095	0.085	1.13	1.27
Form 2					
VF803463	650	0.921	0.124	0.89	0.71
VF885510	650	3.017	0.086	0.92	0.89
VF865673	650	3.922	0.090	1.12	1.21
VF809017	650	4.446	0.098	1.11	1.19
VF865996	650	3.622	0.088	0.96	0.96
VF883593	650	2.437	0.089	0.99	0.98
VF823806	650	2.927	0.086	1.06	1.09
VF885529	650	4.217	0.094	1.16	1.28
Form 3					
VF823449	661	4.877	0.107	1.28	1.92
VF885500	661	3.087	0.086	1.06	1.05
VF880669	661	3.332	0.087	0.90	0.89
VF885549	661	3.883	0.090	1.01	1.05
VF863266	661	4.016	0.092	1.22	1.37
VF803474	661	2.045	0.093	0.95	0.88
VF866052	661	3.139	0.086	1.08	1.09
VF880638	661	4.224	0.095	1.16	1.25
Form 4					
VF812762	657	3.600	0.087	1.01	1.02
VF880849	657	3.221	0.086	1.13	1.15
VF866191	657	4.579	0.099	1.17	1.37
VF880501	657	2.506	0.089	0.95	0.93
VF863351	657	3.051	0.086	1.23	1.26
VF883663	657	3.607	0.087	1.13	1.17
VF883692	657	4.014	0.090	1.21	1.32
VF822454	657	2.715	0.087	0.92	0.88
Form 5					

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF812728	664	5.379	0.123	1.15	1.83
VF880420	664	3.465	0.086	1.00	1.00
VF802935	664	2.638	0.086	1.17	1.19
VF802924	664	3.681	0.087	1.22	1.29
VF883687	664	4.036	0.091	1.20	1.33
VF883715	664	2.325	0.089	0.93	0.90
VF823932	664	3.828	0.088	1.22	1.33
VF804251	664	2.806	0.085	1.00	0.99
VF866035	664	4.303	0.095	1.29	1.48
VF804267	664	3.340	0.085	1.12	1.15
VF822412	664	2.630	0.086	0.90	0.84
VF811990	664	4.986	0.110	1.03	1.38
VF812983	664	3.531	0.086	1.00	1.00
Form 6					
VF802939	637	4.585	0.103	1.14	1.38
VF880528	637	2.808	0.087	1.07	1.07
VF823336	637	4.963	0.112	1.27	2.11
VF802931	637	1.972	0.095	0.86	0.75
VF883657	637	3.026	0.087	1.19	1.23
VF883648	637	1.583	0.103	0.94	0.94
VF823206	637	3.778	0.090	0.93	0.96
VF880628	637	3.357	0.087	0.95	0.95
VF863290	637	4.214	0.095	1.10	1.19
VF809061	637	3.244	0.087	1.12	1.16
VF863242	637	2.785	0.087	1.03	1.02
VF812970	637	3.918	0.091	1.16	1.25
VF804260	637	2.785	0.087	0.92	0.88
Form 7					
VF885577	652	3.116	0.086	0.94	0.92
VF880798	652	2.461	0.089	0.96	0.97
VF823406	652	5.021	0.110	1.28	1.87
VF802936	652	1.995	0.095	0.89	0.87
VF866220	652	3.800	0.089	1.13	1.23
VF809049	652	2.824	0.087	0.97	0.95
VF823366	652	4.012	0.091	1.00	1.12
VF863280	652	3.824	0.089	0.94	0.94
VF823921	652	2.641	0.088	1.08	1.10
VF822425	652	3.064	0.086	1.03	1.02
VF804282	652	4.261	0.094	1.29	1.52
VF885555	652	2.649	0.088	1.08	1.10
VF880680	652	2.930	0.087	0.99	1.01
Form 8					
VF885519	688	2.082	0.090	1.01	1.02
VF880697	688	3.695	0.086	1.18	1.27

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF802934	688	2.225	0.088	0.95	0.90
VF823444	688	3.008	0.083	0.93	0.92
VF883698	688	3.001	0.083	1.03	1.02
VF883641	688	2.917	0.084	1.10	1.17
VF805819	688	3.564	0.085	1.02	1.04
VF823294	688	3.008	0.083	0.91	0.87
VF880525	688	3.492	0.085	1.16	1.19
VF822402	688	4.432	0.096	1.04	1.14
VF822441	688	3.336	0.084	0.93	0.92
VF810683	688	3.470	0.084	1.21	1.25
VF880559	688	3.907	0.088	1.29	1.48
Form 9					
VF802938	662	3.474	0.086	1.05	1.07
VF885483	662	2.931	0.086	1.05	1.08
VF880493	662	3.048	0.085	1.09	1.17
VF823307	662	2.887	0.086	0.94	0.92
VF804256	662	2.901	0.086	1.09	1.12
VF880675	662	2.551	0.087	0.86	0.81
VF823748	662	3.874	0.089	1.26	1.40
VF863346	662	2.761	0.086	0.92	0.89
VF809838	662	2.356	0.089	0.97	0.92
VF866181	662	4.045	0.091	1.14	1.26
VF812997	662	3.326	0.086	1.02	1.02
VF883707	662	3.422	0.086	0.97	0.95
VF866064	662	3.245	0.086	0.99	1.01
Form 10					
VF812743	683	4.087	0.091	1.12	1.26
VF880512	683	3.349	0.084	0.96	0.97
VF802932	683	3.180	0.084	1.30	1.34
VF823432	683	2.365	0.087	0.90	0.85
VF880646	683	2.801	0.084	0.94	0.92
VF885561	683	4.046	0.090	1.25	1.33
VF823736	683	3.623	0.086	1.05	1.10
VF880613	683	2.171	0.089	0.91	0.92
VF866019	683	4.170	0.092	1.06	1.20
VF863323	683	5.211	0.116	1.14	1.60
VF865675	683	3.307	0.084	1.04	1.05
VF823848	683	2.942	0.084	1.05	1.02
VF822465	683	3.173	0.084	1.03	1.02

Science

Table L13. Science Grade 4 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF800147	750	-0.179	0.083	1.10	1.13
VF800163	750	0.043	0.081	1.06	1.14
VF800182	750	-0.030	0.082	0.99	0.96
VF800193	750	-0.057	0.082	0.90	0.84
VF801233	750	1.532	0.085	1.02	1.03
VF656005	750	-1.153	0.100	0.94	0.90
VF656072	750	0.359	0.080	1.01	1.01
VF671249	750	-0.186	0.083	0.91	0.87
VF671205	750	1.583	0.086	1.16	1.32
VF671215	750	1.191	0.082	1.28	1.39
VF671241	750	2.101	0.095	1.18	1.48
Form 2					
VF800026	691	0.313	0.084	1.03	1.03
VF800044	691	1.175	0.085	1.16	1.22
VF800096	691	-0.131	0.086	1.07	1.08
VF800109	691	-1.035	0.101	0.91	0.80
VF800118	691	-0.473	0.090	1.17	1.46
VF656085	691	0.137	0.084	1.01	0.99
VF656089	691	3.688	0.161	1.08	2.38
VF671330	691	-0.994	0.100	0.85	0.66
VF671338	691	0.868	0.084	0.95	0.97
VF671340	691	1.461	0.088	1.09	1.13
VF671344	691	1.182	0.086	1.15	1.22
VF671357	691	0.008	0.085	0.99	1.02
Form 3					
VF656139	685	0.849	0.084	1.00	0.99
VF656143	685	0.610	0.084	1.03	1.02
VF656150	685	-0.876	0.100	1.13	1.30
VF656157	685	-0.515	0.093	0.86	0.77
VF656175	685	0.013	0.087	1.18	1.25
VF656106	685	0.504	0.084	1.04	1.03
VF656109	685	0.419	0.084	1.01	1.00
VF656178	685	-0.194	0.089	0.80	0.74
VF656181	685	0.005	0.087	0.93	0.95
VF656190	685	0.899	0.084	0.96	0.96
VF656195	685	1.027	0.085	1.07	1.10
VF656227	685	0.065	0.086	0.90	0.85

Form 4

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF656180	691	-0.097	0.086	0.94	0.91
VF656183	691	-0.453	0.090	1.09	1.28
VF656202	691	0.585	0.082	0.94	0.92
VF656217	691	0.272	0.083	1.12	1.18
VF656228	691	0.706	0.082	1.06	1.08
VF800157	691	-1.708	0.126	0.90	0.74
VF800137	691	-3.249	0.235	0.94	0.61
VF815566	691	-0.105	0.086	1.03	1.06
VF815606	691	2.009	0.096	1.24	1.59
VF815613	691	2.037	0.096	1.20	1.46
VF815620	691	1.369	0.086	1.09	1.13
VF815623	691	1.928	0.094	1.12	1.39
Form 5					
VF814118	683	-0.433	0.089	1.03	1.03
VF814057	683	-2.356	0.157	0.94	0.69
VF814125	683	1.193	0.085	1.15	1.21
VF814129	683	1.551	0.090	1.10	1.17
VF814143	683	0.152	0.084	0.90	0.86
VF800175	683	1.794	0.093	1.08	1.16
VF801217	683	0.599	0.083	1.12	1.15
VF800030	683	0.298	0.083	1.02	1.03
VF799850	683	0.592	0.083	1.04	1.05
VF800090	683	-0.269	0.087	0.91	0.86
VF800059	683	0.592	0.083	1.06	1.07
VF800018	683	2.012	0.098	1.03	1.20
Form 6					
VF656211	704	0.455	0.082	1.11	1.14
VF656239	704	-0.512	0.091	1.03	1.09
VF656220	704	0.050	0.084	0.99	0.96
VF656177	704	-1.299	0.110	0.90	0.72
VF656245	704	-1.040	0.103	0.95	0.92
VF671126	704	-0.145	0.086	1.03	1.06
VF671189	704	1.434	0.086	1.03	1.11
VF656237	704	1.677	0.089	1.24	1.50
VF656226	704	0.238	0.083	0.93	0.89
VF656221	704	0.622	0.082	0.94	0.92
VF656218	704	1.797	0.091	1.14	1.25
VF656179	704	-0.190	0.087	0.87	0.79
Form 7					
VF671286	705	-1.068	0.103	1.06	1.17
VF671354	705	2.508	0.106	1.16	1.63

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF671349	705	0.548	0.082	0.98	0.99
VF671365	705	-0.665	0.094	0.93	0.85
VF671318	705	-0.343	0.088	1.11	1.19
VF815653	705	0.682	0.082	1.29	1.36
VF815662	705	-0.273	0.087	0.92	0.86
VF814054	705	1.018	0.083	0.92	0.98
VF814067	705	-0.390	0.089	0.88	0.87
VF814152	705	-1.015	0.102	0.88	0.91
VF814064	705	0.182	0.083	0.96	0.94
VF814112	705	-0.745	0.095	0.92	0.90
Form 8					
VF815516	703	-0.722	0.095	1.02	0.92
VF814294	703	-0.607	0.093	0.99	0.92
VF814286	703	0.595	0.082	0.97	0.94
VF814283	703	-0.050	0.085	0.91	0.84
VF814431	703	-0.805	0.097	0.90	0.78
VF801255	703	0.859	0.083	1.02	1.04
VF801247	703	0.989	0.083	1.09	1.10
VF815658	703	0.344	0.083	1.05	1.07
VF815665	703	1.135	0.084	1.03	1.03
VF815667	703	-0.297	0.088	1.03	1.06
VF815652	703	1.220	0.084	0.98	1.03
VF815668	703	0.690	0.082	1.11	1.15
Form 9					
VF656748	704	0.466	0.082	1.05	1.05
VF656812	704	0.350	0.083	1.07	1.06
VF656804	704	0.851	0.082	1.21	1.28
VF656846	704	1.553	0.087	1.20	1.34
VF656794	704	-3.388	0.248	0.98	0.62
VF814076	704	-0.293	0.088	0.96	0.93
VF814089	704	-0.285	0.088	1.03	1.17
VF814292	704	-0.178	0.087	0.85	0.76
VF815472	704	0.085	0.084	0.86	0.80
VF814288	704	0.049	0.085	0.89	0.85
VF814285	704	1.789	0.090	1.20	1.36
VF814290	704	0.851	0.082	0.96	0.95
Form 10					
VF656800	706	1.492	0.086	1.03	1.06
VF656785	706	-0.293	0.088	1.05	1.07
VF656808	706	1.231	0.084	1.26	1.35
VF656820	706	-0.143	0.086	1.07	1.09

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF656815	706	1.058	0.083	1.06	1.08
VF815601	706	1.127	0.083	1.03	1.05
VF815618	706	0.976	0.082	1.18	1.21
VF815654	706	1.253	0.084	0.89	0.87
VF815661	706	0.306	0.082	1.04	1.04
VF815666	706	1.092	0.083	1.17	1.23
VF815670	706	1.024	0.083	1.06	1.09
VF815664	706	0.594	0.082	1.07	1.12



Table L14. Science Grade 8 IRT Statistics for Field Test Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
Form 1					
VF671265	785	-1.725	0.099	1.01	0.99
VF671278	785	-1.085	0.086	1.03	1.16
VF671327	785	0.947	0.083	1.24	1.30
VF671334	785	0.195	0.078	0.91	0.89
VF671285	785	-0.899	0.083	0.96	0.92
VF813872	785	0.377	0.078	1.03	1.01
VF813811	785	0.488	0.079	1.17	1.20
VF735904	785	-0.831	0.082	0.89	0.83
VF735980	785	-1.063	0.086	1.09	1.29
VF736039	785	-1.493	0.094	0.96	0.91
VF735995	785	-0.070	0.078	1.09	1.08
VF735983	785	0.014	0.078	1.02	1.06
Form 2					
VF737468	656	-1.542	0.108	1.01	1.02
VF737466	656	-0.867	0.093	1.01	1.10
VF737445	656	-1.178	0.099	0.94	0.92
VF737472	656	0.918	0.087	1.33	1.47
VF737473	656	-0.304	0.086	0.96	0.95
VF813803	656	-1.638	0.112	0.93	0.90
VF813860	656	-0.194	0.085	1.07	1.09
VF735043	656	-0.547	0.088	1.19	1.32
VF734993	656	0.242	0.084	1.13	1.17
VF735020	656	1.019	0.089	1.21	1.27
VF735039	656	-0.042	0.085	0.98	0.99
VF735045	656	0.745	0.086	1.08	1.09
Form 3					
VF812720	671	-0.844	0.092	0.93	0.94
VF812729	671	0.783	0.086	1.34	1.49
VF812711	671	0.027	0.084	0.95	0.93
VF812684	671	-1.374	0.103	0.85	0.73
VF812703	671	-1.527	0.107	0.94	1.02
VF813905	671	-1.353	0.102	0.92	0.80
VF813827	671	-0.869	0.092	0.98	0.95
VF671280	671	-0.633	0.089	0.98	0.96
VF671343	671	-0.078	0.084	1.04	1.08
VF671294	671	-0.761	0.090	0.91	0.88
VF671350	671	-1.406	0.103	0.92	0.88
VF671352	671	-0.221	0.085	1.02	1.04
Form 4					
VF813963	664	-0.879	0.094	0.92	0.85

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF814074	664	1.270	0.091	1.24	1.37
VF814058	664	0.193	0.084	1.02	1.05
VF814043	664	-0.037	0.085	1.02	1.00
VF814084	664	0.021	0.085	1.06	1.06
VF813865	664	-0.102	0.086	1.03	1.04
VF813887	664	-1.426	0.107	0.83	0.66
VF812690	664	0.450	0.084	1.13	1.17
VF812745	664	-0.022	0.085	1.14	1.17
VF812741	664	-0.648	0.091	0.99	1.00
VF812733	664	0.286	0.084	1.09	1.10
VF812726	664	-0.132	0.086	1.06	1.06
Form 5					
VF735041	685	-0.207	0.084	1.02	1.01
VF735007	685	0.222	0.083	1.04	1.04
VF735827	685	2.062	0.107	0.96	1.10
VF735791	685	-1.385	0.102	0.95	0.93
VF735035	685	0.666	0.084	1.25	1.37
VF813878	685	0.428	0.083	1.15	1.18
VF813848	685	0.498	0.083	1.18	1.24
VF823979	685	-0.115	0.084	1.00	1.01
VF824038	685	-0.908	0.092	0.98	0.95
VF824046	685	-0.808	0.090	0.95	0.95
VF824033	685	0.993	0.087	1.01	1.03
VF823985	685	-0.752	0.090	0.90	0.91
Form 6					
VF813970	649	0.619	0.086	1.22	1.30
VF814052	649	-0.511	0.089	1.00	1.03
VF814047	649	0.151	0.085	1.15	1.19
VF814096	649	-0.657	0.091	0.97	0.99
VF814068	649	1.870	0.104	1.12	1.43
VF671317	649	-2.464	0.149	0.88	0.56
VF671342	649	0.806	0.087	1.11	1.17
VF736114	649	-2.940	0.180	0.97	0.88
VF736075	649	-0.009	0.086	1.06	1.10
VF735912	649	-0.463	0.089	0.96	0.91
VF735917	649	-0.640	0.091	0.89	0.81
VF735960	649	1.220	0.092	1.12	1.21
Form 7					
VF824043	657	0.960	0.088	1.13	1.25
VF823953	657	-0.102	0.086	1.06	1.07
VF823970	657	0.108	0.085	0.99	0.99

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF824010	657	0.883	0.087	1.16	1.22
VF824019	657	-2.129	0.130	0.90	0.64
VF736122	657	-0.425	0.088	1.18	1.38
VF736130	657	0.725	0.086	1.05	1.07
VF815578	657	0.353	0.085	0.93	0.92
VF815494	657	-0.132	0.086	0.97	0.97
VF815561	657	0.454	0.085	1.03	1.03
VF815521	657	-2.198	0.134	0.89	0.75
Form 8					
VF684529	684	-0.709	0.089	1.09	1.13
VF685187	684	-0.434	0.086	0.93	0.90
VF685871	684	-0.269	0.084	1.04	1.08
VF671386	684	0.143	0.082	0.98	0.99
VF686532	684	2.387	0.117	1.08	1.43
VF735828	684	0.649	0.083	1.04	1.06
VF735123	684	0.959	0.085	1.12	1.21
VF671359	684	0.795	0.084	1.09	1.10
VF671184	684	-0.508	0.086	0.90	0.88
VF671279	684	1.415	0.091	1.16	1.32
VF671247	684	0.197	0.082	1.02	1.03
VF671361	684	-0.701	0.089	1.07	1.19
Form 9					
VF684522	654	-0.761	0.092	0.97	0.95
VF684505	654	0.880	0.087	1.19	1.29
VF686540	654	0.255	0.085	1.19	1.23
VF685863	654	-0.813	0.093	1.00	1.06
VF687024	654	-1.577	0.111	0.88	0.76
F824049	654	-0.216	0.086	1.08	1.10
VF824029	654	0.176	0.085	0.95	0.95
VF862725	654	-0.937	0.095	0.98	0.93
VF862740	654	0.820	0.087	1.03	1.04
VF862778	654	1.202	0.091	1.05	1.10
VF862757	654	0.219	0.085	0.94	0.93
VF862752	654	-0.813	0.093	0.92	0.84
Form 10					
VF862697	665	0.392	0.084	0.97	0.94
VF862703	665	-0.725	0.091	1.08	1.12
VF862684	665	-1.522	0.108	0.93	0.84
VF862659	665	-0.077	0.085	1.01	1.00
VF862718	665	-0.842	0.093	1.07	1.12

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF671315	665	-1.323	0.103	0.94	0.86
VF671269	665	-0.611	0.089	0.94	0.96
VF815568	665	1.379	0.093	1.04	1.06
VF815593	665	-0.509	0.088	0.97	0.99
VF815536	665	-0.455	0.088	1.02	1.04
VF815541	665	-0.401	0.087	0.92	0.92
VF815587	665	0.850	0.087	1.06	1.06

# Appendix M: Rasch Difficulty, Standard Error, and Fit Statistics for 2014 Operational Items

## *Reading*

Table M1. Reading Grade 3 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF394056	7350	0.282	0.027	1.10	1.16
VF394053	7350	-0.479	0.031	0.94	0.87
VF394041	7350	-0.047	0.029	1.16	1.20
VF394054	7350	-1.808	0.044	0.92	0.91
VF394046	7350	-1.673	0.042	0.86	0.63
VF394050	7350	-0.934	0.034	1.04	1.16
VF394049	7350	-0.614	0.032	1.07	1.10
VF394051	7350	-0.481	0.031	1.17	1.45
VF389477	7350	0.243	0.028	0.93	0.87
VF389949	7350	-0.219	0.029	1.03	1.02
VF389470	7350	0.106	0.028	0.95	0.90
VF389620	7350	0.451	0.027	1.02	1.00
VF389457	7350	1.144	0.026	1.09	1.18
VF389467	7350	0.227	0.028	1.05	1.04
VF389165	7350	0.059	0.028	0.92	0.86
VF497660	7350	-1.067	0.036	0.91	0.77
VF497668	7350	-0.430	0.031	0.94	0.89
VF497700	7350	0.426	0.027	1.07	1.11
VF497705	7350	1.500	0.026	1.15	1.29
VF497671	7350	0.347	0.027	1.01	1.01
VF497695	7350	-0.445	0.031	0.99	0.94
VF497696	7350	-0.270	0.030	1.01	1.02
VF497690	7350	0.354	0.027	0.96	0.96
VF497684	7350	0.980	0.026	1.04	1.08
VF497676	7350	-0.485	0.031	0.90	0.84
VF497818	7350	-0.133	0.029	0.90	0.82
VF497815	7350	-0.171	0.029	1.10	1.19
VF497822	7350	1.635	0.027	1.11	1.26
VF497820	7350	0.207	0.028	1.01	1.03
VF497783	7350	-1.118	0.036	0.84	0.68
VF497793	7350	-0.088	0.029	0.95	0.91
VF497812	7350	-0.321	0.030	1.06	1.20

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF494759	7350	-0.581	0.032	0.90	0.83
VF494915	7350	-0.158	0.029	1.01	1.01
VF494661	7350	-0.003	0.028	0.95	0.88
VF494732	7350	-0.856	0.034	0.80	0.59
VF494764	7350	-0.406	0.031	0.89	0.81
VF494956	7350	1.044	0.026	1.03	1.08
VF494909	7350	0.743	0.026	1.07	1.11
VF494745	7350	1.134	0.026	0.96	0.99
VF493383	7350	-0.760	0.033	0.88	0.78
VF493480	7350	-0.168	0.029	0.89	0.82
VF494098	7350	0.128	0.028	0.95	0.93
VF497716	7350	0.109	0.028	1.07	1.13
VF497751	7350	0.307	0.027	1.06	1.07
VF497761	7350	-0.033	0.029	0.92	0.85
VF497758	7350	0.753	0.026	1.00	1.00
VF497767	7350	0.965	0.026	1.01	1.03
VF497766	7350	1.550	0.026	1.16	1.38
VF497731	7350	-0.944	0.034	0.86	0.75

Table M2. Reading Grade 4 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF495028	7013	1.873	0.027	1.11	1.23
VF495644	7013	0.470	0.029	1.14	1.21
VF494993	7013	1.593	0.026	1.05	1.09
VF495021	7013	0.513	0.028	1.12	1.14
VF495015	7013	-0.011	0.031	0.93	0.83
VF495003	7013	0.989	0.027	0.93	0.90
VF497359	7013	1.011	0.027	0.90	0.87
VF497361	7013	-0.553	0.035	1.03	0.84
VF497384	7013	1.235	0.027	1.03	1.06
VF497390	7013	0.445	0.029	0.80	0.71
VF497356	7013	2.226	0.028	1.07	1.16
VF497354	7013	-0.566	0.035	0.94	0.95
VF497365	7013	1.744	0.027	1.14	1.27
VF497381	7013	-0.134	0.032	0.89	0.74
VF497387	7013	-0.385	0.034	0.77	0.63
VF494842	7013	0.034	0.031	1.15	1.33
VF494914	7013	0.097	0.030	1.01	1.05
VF494852	7013	-0.393	0.034	0.90	0.78
VF494964	7013	-0.342	0.033	1.00	0.99
VF494863	7013	-1.031	0.040	0.90	0.73
VF494937	7013	-1.093	0.041	0.90	0.99
VF497147	7013	-0.793	0.038	0.87	0.68
VF497155	7013	0.518	0.028	0.94	0.87
VF497162	7013	1.566	0.026	0.94	0.95
VF497220	7013	0.202	0.030	0.90	0.80
VF497215	7013	2.289	0.028	1.05	1.20
VF497188	7013	-1.642	0.049	0.83	0.49
VF497212	7013	0.383	0.029	1.04	1.00
VF497159	7013	-0.890	0.039	0.90	0.82
VF497270	7013	1.622	0.026	1.10	1.19
VF497265	7013	-0.904	0.039	0.90	0.79
VF497247	7013	0.450	0.029	1.01	0.99
VF497261	7013	0.411	0.029	1.16	1.23
VF497243	7013	0.395	0.029	1.07	1.11
VF497233	7013	0.423	0.029	1.09	1.11
VF497311	7013	1.736	0.027	1.26	1.45
VF497318	7013	-0.653	0.036	1.20	1.35
VF497297	7013	-1.224	0.043	0.83	0.55

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF497322	7013	0.474	0.029	0.99	0.94
VF497334	7013	0.352	0.029	0.96	0.91
VF497338	7013	0.773	0.028	1.26	1.42
VF497326	7013	1.440	0.026	0.93	0.95
VF497327	7013	0.417	0.029	1.00	0.97
VF407243	7013	0.079	0.031	1.05	1.04
VF407287	7013	-0.598	0.036	0.92	0.80
VF407232	7013	-0.760	0.037	1.12	0.89
VF407235	7013	0.944	0.027	1.22	1.30
VF407297	7013	0.731	0.028	0.99	0.98
VF407282	7013	-1.088	0.041	0.98	0.69
VF407298	7013	0.745	0.028	1.08	1.10



Table M3. Reading Grade 5 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF497182	7063	-1.668	0.054	0.95	0.90
VF497177	7063	0.279	0.030	0.97	0.91
VF497174	7063	-1.295	0.047	0.89	0.71
VF497172	7063	-0.287	0.035	0.99	0.96
VF497056	7063	0.195	0.031	0.98	1.04
VF497170	7063	-0.429	0.036	1.00	1.09
VF497052	7063	1.218	0.027	1.24	1.35
VF496101	7063	0.373	0.030	1.03	1.00
VF496032	7063	1.011	0.027	0.98	0.96
VF496188	7063	0.912	0.028	1.05	1.06
VF496085	7063	0.944	0.027	1.11	1.18
VF496185	7063	0.609	0.029	1.11	1.19
VF496024	7063	0.333	0.030	0.91	0.82
VF496115	7063	-1.202	0.046	0.83	0.49
VF407319	7063	0.839	0.028	1.09	1.20
VF407388	7063	0.500	0.029	1.00	0.94
VF407329	7063	0.058	0.032	0.94	0.81
VF407332	7063	0.257	0.031	1.03	1.03
VF407355	7063	-0.153	0.033	0.94	0.94
VF407322	7063	1.049	0.027	1.07	1.13
VF407360	7063	-1.035	0.043	0.92	0.83
VF496211	7063	-1.767	0.057	0.89	0.69
VF496865	7063	0.765	0.028	0.96	0.94
VF496879	7063	0.953	0.027	1.05	1.04
VF496213	7063	0.353	0.030	1.07	1.15
VF496209	7063	0.016	0.032	0.96	0.88
VF496201	7063	-1.593	0.053	0.83	0.45
VF496206	7063	0.943	0.027	1.05	1.03
VF495924	7063	0.088	0.032	0.95	0.83
VF495921	7063	0.749	0.028	0.92	0.88
VF495800	7063	-0.892	0.041	0.86	0.65
VF495780	7063	1.534	0.026	1.02	1.04
VF495943	7063	-0.548	0.037	0.95	0.84
VF496875	7063	-0.253	0.034	0.88	0.69
VF496872	7063	0.711	0.028	0.98	0.94
VF496878	7063	1.865	0.026	1.09	1.17
VF496882	7063	0.562	0.029	1.14	1.23
VF496884	7063	0.136	0.031	1.19	1.17
VF496869	7063	2.350	0.027	0.99	1.08
VF496886	7063	0.711	0.028	0.88	0.81

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF497284	7063	-0.038	0.033	0.93	0.85
VF497278	7063	-0.433	0.036	0.86	0.71
VF497273	7063	-0.661	0.038	1.03	1.40
VF497282	7063	1.771	0.026	1.25	1.37
VF497285	7063	2.422	0.027	1.08	1.25
VF497287	7063	0.346	0.030	0.97	0.95
VF497274	7063	1.762	0.026	1.01	1.05
VF497272	7063	2.010	0.027	1.18	1.33
VF497288	7063	-0.131	0.033	1.00	1.07
VF497037	7063	-0.598	0.038	0.90	0.72
VF497039	7063	1.746	0.026	0.93	0.97
VF497030	7063	0.809	0.028	0.95	0.90
VF497028	7063	1.528	0.026	0.89	0.88
VF497012	7063	0.727	0.028	1.00	1.03

Table M4. Reading Grade 6 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF497042	6748	-0.365	0.038	0.85	0.63
VF497035	6748	0.289	0.032	1.03	1.04
VF497034	6748	-0.081	0.035	0.96	0.91
VF497033	6748	1.749	0.027	1.08	1.12
VF497047	6748	0.174	0.033	1.00	1.05
VF497041	6748	-0.037	0.035	0.89	0.76
VF496873	6748	0.982	0.029	1.05	1.13
VF496204	6748	1.223	0.028	0.96	0.91
VF496208	6748	1.243	0.028	1.00	0.96
VF496863	6748	1.228	0.028	0.97	0.95
VF496191	6748	1.546	0.027	0.94	0.92
VF496164	6748	0.726	0.030	1.08	1.26
VF496172	6748	1.100	0.028	0.98	0.94
VF496055	6748	1.917	0.027	0.97	0.98
VF496036	6748	1.635	0.027	1.05	1.06
VF496065	6748	0.815	0.029	1.03	1.10
VF496071	6748	0.040	0.034	0.99	0.92
VF496100	6748	-0.199	0.036	1.03	0.94
VF496051	6748	0.976	0.029	1.06	1.15
VF496029	6748	0.029	0.034	0.87	0.80
VF496087	6748	0.488	0.031	0.84	0.79
VF495908	6748	0.124	0.034	1.01	1.00
VF495961	6748	1.215	0.028	1.03	1.16
VF495968	6748	1.839	0.027	1.03	1.05
VF495990	6748	0.196	0.033	0.95	0.92
VF495938	6748	1.992	0.027	1.05	1.07
VF495954	6748	0.612	0.030	1.06	1.13
VF388881	6748	-0.526	0.040	0.90	0.86
VF388912	6748	0.468	0.031	0.99	1.00
VF388853	6748	0.848	0.029	0.88	0.81
VF388848	6748	0.541	0.031	0.99	1.04
VF388868	6748	-0.574	0.041	0.87	0.71
VF388851	6748	1.795	0.027	1.00	1.02
VF497084	6748	1.789	0.027	1.04	1.08
VF497082	6748	2.062	0.027	1.00	1.03
VF497087	6748	1.570	0.027	0.97	0.95
VF497079	6748	1.456	0.028	1.03	1.04
VF497078	6748	1.536	0.027	1.16	1.21
VF497083	6748	2.313	0.028	1.00	1.08
VF497077	6748	2.409	0.028	1.10	1.17

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF497076	6748	2.323	0.028	1.05	1.12
VF497074	6748	2.167	0.027	0.99	1.01
VF523861	6748	1.085	0.028	0.93	0.87
VF523801	6748	1.458	0.028	1.01	1.05
VF523846	6748	1.252	0.028	0.85	0.79
VF523825	6748	0.818	0.029	0.84	0.75
VF523818	6748	0.949	0.029	1.05	1.05
VF523813	6748	1.676	0.027	1.22	1.30
VF523863	6748	0.171	0.033	0.83	0.65
VF523804	6748	1.448	0.028	1.04	1.07
VF523786	6748	1.143	0.028	0.96	0.95
VF497071	6748	1.915	0.027	1.06	1.08
VF497069	6748	1.542	0.027	1.11	1.13
VF497053	6748	2.569	0.028	1.07	1.19
VF497073	6748	0.916	0.029	0.90	0.83
VF497059	6748	1.121	0.028	1.06	1.11

Table M5. Reading Grade 7 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF496937	6462	0.498	0.034	1.07	1.23
VF496901	6462	0.845	0.031	1.12	1.28
VF496913	6462	-0.120	0.040	0.90	0.66
VF496906	6462	-0.248	0.042	0.92	0.79
VF496895	6462	2.197	0.027	0.93	0.93
VF496900	6462	0.061	0.038	0.89	0.69
VF497975	6462	2.134	0.027	1.08	1.10
VF497958	6462	1.216	0.030	0.83	0.78
VF497951	6462	1.224	0.029	0.99	0.98
VF497969	6462	2.454	0.028	1.00	1.02
VF497955	6462	0.595	0.033	0.88	0.74
VF497961	6462	1.837	0.028	0.89	0.85
VF497978	6462	2.265	0.027	1.02	1.03
VF497974	6462	2.088	0.027	0.99	0.99
VF497941	6462	1.342	0.029	1.03	1.02
VF497950	6462	0.556	0.033	0.93	0.83
VF497938	6462	1.516	0.028	1.00	1.00
VF497943	6462	1.811	0.028	1.01	1.01
VF497935	6462	1.499	0.029	1.18	1.32
VF497931	6462	0.981	0.031	1.00	0.99
VF497930	6462	1.144	0.030	0.91	0.85
VF497862	6462	2.247	0.027	1.04	1.07
VF497882	6462	1.539	0.028	1.05	1.07
VF497879	6462	1.272	0.029	1.06	1.09
VF497893	6462	1.115	0.030	1.09	1.17
VF497890	6462	1.011	0.030	0.99	0.93
VF497876	6462	0.900	0.031	0.99	1.04
VF497868	6462	2.474	0.028	0.99	1.02
VF497873	6462	1.154	0.030	0.97	0.99
VF497883	6462	2.161	0.027	1.15	1.21
VF498058	6462	1.168	0.030	1.02	1.07
VF497877	6462	1.983	0.028	1.15	1.21
VF498030	6462	0.762	0.032	1.00	1.07
VF498018	6462	0.659	0.033	0.86	0.74
VF497980	6462	2.339	0.027	1.00	1.01
VF498062	6462	1.824	0.028	1.04	1.07
VF498051	6462	1.630	0.028	1.05	1.06
VF498064	6462	2.275	0.027	0.99	1.00
VF498054	6462	1.766	0.028	0.95	0.92
VF498057	6462	1.508	0.028	0.98	0.99

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF498034	6462	1.255	0.029	0.92	0.84
VF498047	6462	0.348	0.035	0.98	1.05
VF498032	6462	0.544	0.034	0.94	0.91
VF498052	6462	2.530	0.028	1.07	1.11
VF497224	6462	1.769	0.028	1.03	1.04
VF497211	6462	0.855	0.031	1.09	1.15
VF497175	6462	0.784	0.032	0.92	0.86
VF497190	6462	0.240	0.036	1.27	1.34
VF497198	6462	2.375	0.027	1.03	1.05
VF497205	6462	1.312	0.029	0.99	0.96
VF497281	6462	0.511	0.034	1.04	1.12
VF497301	6462	2.141	0.027	1.06	1.08
VF497299	6462	1.341	0.029	0.99	0.94
VF497291	6462	2.679	0.028	1.07	1.14
VF497260	6462	0.491	0.034	0.86	0.72
VF497263	6462	1.009	0.031	0.97	0.96

Table M6. Reading Grade 8 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF497427	6462	-0.811	0.056	0.92	0.59
VF497441	6462	0.351	0.038	1.01	1.01
VF497443	6462	-0.238	0.046	0.96	0.92
VF497446	6462	1.156	0.032	1.07	1.07
VF497436	6462	0.074	0.041	0.96	0.90
VF497445	6462	0.883	0.033	0.96	0.89
VF497444	6462	-1.120	0.063	0.86	0.40
VF497199	6462	2.221	0.028	1.18	1.25
VF497180	6462	-0.564	0.051	0.93	0.89
VF497203	6462	1.661	0.029	1.02	1.11
VF497196	6462	0.636	0.035	1.00	1.01
VF497178	6462	1.907	0.028	1.14	1.18
VF497193	6462	1.841	0.029	1.09	1.16
VF497209	6462	1.347	0.031	1.05	1.08
VF497257	6462	0.770	0.034	1.18	1.30
VF497229	6462	1.692	0.029	1.10	1.13
VF497259	6462	2.184	0.028	1.15	1.22
VF497244	6462	0.621	0.035	1.05	1.12
VF497242	6462	2.626	0.028	1.05	1.11
VF497235	6462	0.169	0.040	0.54	0.36
VF497266	6462	2.253	0.028	1.13	1.17
VF497252	6462	0.554	0.036	0.94	0.92
VF497095	6462	0.062	0.041	0.92	0.83
VF497113	6462	2.874	0.028	1.01	1.11
VF497114	6462	0.770	0.034	0.98	0.99
VF497101	6462	1.601	0.029	0.97	0.95
VF497098	6462	2.355	0.028	1.12	1.18
VF497094	6462	1.043	0.032	0.97	0.90
VF497115	6462	0.881	0.033	1.00	0.96
VF497148	6462	1.889	0.029	1.01	1.00
VF497161	6462	2.005	0.028	1.10	1.15
VF497164	6462	2.514	0.028	1.04	1.09
VF497137	6462	1.683	0.029	0.90	0.84
VF497166	6462	0.656	0.035	0.78	0.55
VF497139	6462	0.730	0.035	0.93	0.82
VF497120	6462	2.390	0.028	0.99	1.02
VF497127	6462	0.597	0.036	0.89	0.71
VF497116	6462	2.148	0.028	1.08	1.11
VF497117	6462	1.707	0.029	0.94	0.91
VF497132	6462	2.436	0.028	1.13	1.19

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF497130	6462	1.278	0.031	0.98	0.91
VF497123	6462	1.177	0.031	1.04	0.99
VF497329	6462	0.770	0.034	0.94	0.87
VF497349	6462	-0.559	0.051	1.42	1.36
VF497353	6462	1.804	0.029	0.95	0.92
VF497328	6462	1.578	0.030	1.03	1.07
VF497325	6462	1.635	0.029	0.94	0.88
VF497363	6462	2.126	0.028	1.08	1.11
VF497355	6462	-0.428	0.049	0.84	0.58
VF497298	6462	0.256	0.039	0.85	0.69
VF497316	6462	0.641	0.035	0.96	0.89
VF497302	6462	2.151	0.028	1.05	1.09
VF497309	6462	0.743	0.034	0.85	0.68
VF497313	6462	1.768	0.029	0.92	0.88
VF497305	6462	0.856	0.034	0.94	0.98
VF497306	6462	-0.211	0.045	0.88	0.66



Mathematics

Table M7. Mathematics Grade 3 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF393959	7350	-2.559	0.051	0.91	0.70
VF387496	7350	-1.788	0.039	0.95	0.80
VF393772	7350	0.030	0.027	0.88	0.85
VF494670	7350	1.488	0.028	1.04	1.13
VF494103	7350	0.755	0.026	1.00	1.01
VF406339	7350	-1.108	0.032	0.90	0.77
VF387500	7350	-0.009	0.027	0.97	0.93
VF406297	7350	0.746	0.026	1.17	1.25
VF394355	7350	1.375	0.027	1.10	1.17
VF387498	7350	3.026	0.038	1.05	1.41
VF406327	7350	2.079	0.030	0.91	0.85
VF406204	7350	0.070	0.027	1.09	1.14
VF394252	7350	-0.603	0.029	1.06	1.16
VF494820	7350	1.318	0.027	0.98	1.05
VF493146	7350	-0.377	0.028	0.91	0.86
VF394250	7350	0.222	0.026	1.02	1.04
VF393782	7350	-1.986	0.041	0.92	0.74
VF394361	7350	-1.005	0.031	0.86	0.76
VF394339	7350	-0.052	0.027	1.11	1.21
VF493415	7350	0.516	0.026	0.95	0.94
VF394382	7350	-2.000	0.041	0.96	0.94
VF394375	7350	-1.021	0.031	0.96	0.88
VF394362	7350	1.154	0.027	0.99	1.01
VF394369	7350	-0.260	0.028	0.91	0.84
VF493287	7350	0.446	0.026	0.97	0.99
VF394368	7350	-0.094	0.027	0.97	1.00
VF394376	7350	0.273	0.026	0.97	0.96
VF393748	7350	0.373	0.026	0.88	0.85
VF394221	7350	0.862	0.026	1.08	1.10
VF494693	7350	0.989	0.026	1.02	1.05
VF494895	7350	1.315	0.027	0.95	0.94
VF394378	7350	1.153	0.027	1.25	1.37
VF394381	7350	0.719	0.026	1.03	1.04
VF406343	7350	-1.205	0.033	0.93	0.89
VF494880	7350	-2.391	0.048	0.92	0.84
VF406295	7350	-0.404	0.028	1.25	1.40
VF493127	7350	0.446	0.026	0.96	0.95
VF393824	7350	0.968	0.026	1.08	1.15

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF394239	7350	0.362	0.026	0.83	0.78
VF494690	7350	0.597	0.026	1.07	1.07
VF494750	7350	-0.313	0.028	1.02	1.01
VF493461	7350	0.770	0.026	1.16	1.25
VF393786	7350	-0.535	0.029	0.98	0.94
VF493124	7350	0.562	0.026	0.96	0.96
VF394356	7350	0.046	0.027	0.91	0.85
VF394229	7350	0.590	0.026	1.09	1.12
VF493153	7350	-1.047	0.032	0.95	0.89
VF493387	7350	-1.023	0.031	0.98	0.89
VF387502	7350	-3.230	0.067	0.93	0.64
VF494756	7350	-0.239	0.027	0.98	0.96

Table M8. Mathematics Grade 4 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF492346	7019	-2.621	0.078	0.96	0.57
VF492315	7019	-1.138	0.042	0.85	0.52
VF393675	7019	1.249	0.026	0.95	0.94
VF492332	7019	0.874	0.027	1.11	1.11
VF492333	7019	1.202	0.026	1.03	1.03
VF493356	7019	-0.739	0.037	0.86	0.76
VF492358	7019	-0.948	0.040	0.96	0.93
VF493249	7019	1.818	0.027	1.12	1.22
VF493349	7019	1.827	0.027	0.91	0.90
VF492311	7019	-1.259	0.044	0.92	0.65
VF493284	7019	2.845	0.031	0.91	1.03
VF492390	7019	-1.424	0.047	0.95	0.90
VF493334	7019	2.389	0.028	0.95	0.99
VF497391	7019	-0.060	0.031	1.05	0.98
VF493344	7019	-0.865	0.039	1.33	1.25
VF493373	7019	-0.372	0.033	1.00	1.00
VF493140	7019	0.978	0.027	0.91	0.86
VF492392	7019	-0.010	0.031	1.13	1.26
VF492353	7019	1.505	0.026	1.00	1.01
VF492320	7019	-0.953	0.040	0.84	0.53
VF493238	7019	-1.679	0.052	0.97	1.13
VF492330	7019	1.068	0.026	1.06	1.07
VF493228	7019	1.357	0.026	0.84	0.81
VF492312	7019	-0.111	0.031	0.60	0.47
VF497395	7019	1.026	0.027	0.86	0.81
VF492334	7019	0.254	0.029	0.80	0.68
VF492343	7019	2.975	0.031	1.00	1.16
VF492370	7019	0.825	0.027	1.13	1.25
VF493154	7019	-0.783	0.038	0.93	0.85
VF493303	7019	0.613	0.028	0.97	0.91
VF493219	7019	-0.197	0.032	0.98	1.02
VF393726	7019	0.031	0.030	1.03	1.08
VF493257	7019	1.280	0.026	0.96	0.94
VF493312	7019	2.312	0.028	1.12	1.22
VF492373	7019	0.303	0.029	0.96	0.95
VF493223	7019	1.601	0.026	0.90	0.89
VF493366	7019	2.245	0.028	1.12	1.22
VF493143	7019	0.736	0.027	1.14	1.20
VF493377	7019	0.630	0.027	0.99	0.99
VF492338	7019	0.363	0.029	1.02	1.04

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF493295	7019	2.206	0.028	0.95	0.96
VF493301	7019	1.493	0.026	1.07	1.10
VF493126	7019	-0.889	0.039	0.79	0.60
VF493135	7019	1.404	0.026	0.94	0.91
VF393648	7019	2.441	0.028	1.02	1.07
VF493142	7019	0.750	0.027	0.90	0.85
VF493288	7019	2.597	0.029	0.92	1.04
VF492386	7019	0.694	0.027	0.98	0.98
VF493318	7019	2.357	0.028	1.11	1.25
VF493130	7019	1.369	0.026	1.16	1.22
VF492352	7019	-2.375	0.070	0.95	0.84
VF493329	7019	1.912	0.027	1.14	1.23
VF492306	7019	-1.397	0.046	0.90	0.67
VF493242	7019	1.024	0.027	1.01	1.00
VF493262	7019	1.260	0.026	1.15	1.21
VF493361	7019	1.371	0.026	1.26	1.37
VF493294	7019	2.888	0.031	1.19	1.54
VF492337	7019	0.924	0.027	0.98	0.94
VF493371	7019	1.152	0.026	1.07	1.09

Table M9. Mathematics Grade 5 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF491951	7062	-0.158	0.036	0.91	0.72
VF491924	7062	1.180	0.028	0.98	0.98
VF491941	7062	1.924	0.027	1.07	1.09
VF492083	7062	2.715	0.028	0.95	0.98
VF492203	7062	-0.047	0.035	0.85	0.60
VF492088	7062	1.973	0.027	0.98	0.98
VF492027	7062	1.956	0.027	1.31	1.43
VF491963	7062	2.319	0.027	1.18	1.27
VF491626	7062	0.881	0.029	0.95	0.88
VF492000	7062	1.397	0.027	1.10	1.16
VF491900	7062	2.260	0.027	0.95	0.93
VF492313	7062	1.328	0.027	0.85	0.77
VF492048	7062	1.811	0.027	0.98	0.95
VF492235	7062	1.922	0.027	1.13	1.19
VF492120	7062	2.403	0.027	1.06	1.08
VF492031	7062	1.879	0.027	1.00	0.99
VF492298	7062	1.782	0.027	0.91	0.88
VF491636	7062	1.838	0.027	0.87	0.83
VF492255	7062	1.420	0.027	0.95	0.90
VF491967	7062	1.980	0.027	1.03	1.05
VF492007	7062	1.132	0.028	1.06	1.10
VF492003	7062	1.976	0.027	1.13	1.17
VF492296	7062	2.513	0.027	1.18	1.28
VF492214	7062	2.745	0.028	0.90	0.93
VF492174	7062	1.502	0.027	0.91	0.87
VF492532	7062	1.730	0.027	1.19	1.36
VF491948	7062	2.075	0.027	0.94	0.91
VF492427	7062	2.173	0.027	0.94	0.92
VF492099	7062	1.552	0.027	0.89	0.85
VF491627	7062	-0.008	0.035	0.99	1.14
VF491771	7062	0.998	0.028	0.90	0.86
VF492248	7062	1.191	0.028	0.94	0.87
VF492186	7062	2.304	0.027	1.17	1.24
VF491937	7062	2.173	0.027	1.13	1.21
VF491895	7062	0.837	0.029	0.92	0.91
VF492423	7062	0.916	0.029	1.10	1.22
VF491916	7062	1.619	0.027	0.83	0.76
VF491635	7062	2.425	0.027	0.93	0.94
VF492374	7062	0.814	0.029	1.03	1.11
VF492528	7062	1.347	0.027	1.09	1.17

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF491804	7062	1.458	0.027	0.79	0.70
VF491630	7062	2.155	0.027	1.08	1.11
VF492397	7062	2.395	0.027	1.00	1.00
VF492095	7062	1.684	0.027	0.91	0.87
VF491783	7062	2.768	0.028	0.93	0.94
VF492304	7062	0.889	0.029	0.87	0.78
VF492091	7062	2.022	0.027	1.28	1.44
VF491939	7062	0.724	0.030	0.77	0.67
VF491794	7062	0.874	0.029	0.95	0.90
VF491932	7062	2.485	0.027	1.09	1.12
VF491905	7062	1.114	0.028	0.93	0.89
VF491753	7062	1.373	0.027	1.17	1.27
VF492010	7062	1.967	0.027	1.02	1.02
VF491761	7062	1.068	0.028	0.93	0.89
VF492001	7062	1.304	0.028	1.06	1.07
VF491727	7062	1.923	0.027	1.02	1.03
VF491821	7062	1.896	0.027	0.97	0.95
VF491927	7062	0.778	0.029	1.00	1.07
VF492281	7062	1.124	0.028	0.81	0.71

Table M10. Mathematics Grade 6 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF491953	6745	1.494	0.029	1.01	1.04
VF492542	6745	1.790	0.028	0.88	0.84
VF492181	6745	1.797	0.028	1.14	1.26
VF491930	6745	1.666	0.028	0.95	0.96
VF492399	6745	1.504	0.029	1.01	1.06
VF492732	6745	3.022	0.028	1.21	1.31
VF492287	6745	2.355	0.027	1.13	1.19
VF492722	6745	2.797	0.028	1.07	1.11
VF492572	6745	2.623	0.027	1.11	1.14
VF492759	6745	3.065	0.028	1.14	1.22
VF492721	6745	2.647	0.027	1.10	1.12
VF492383	6745	1.891	0.028	0.91	0.87
VF492593	6745	1.524	0.029	0.82	0.74
VF492030	6745	2.060	0.027	1.00	1.00
VF492192	6745	2.372	0.027	1.10	1.14
VF492053	6745	2.239	0.027	0.92	0.88
VF492709	6745	2.787	0.027	1.00	1.01
VF492240	6745	2.096	0.027	1.24	1.32
VF492412	6745	2.939	0.028	1.07	1.12
VF492660	6745	3.777	0.031	0.92	1.02
VF492533	6745	1.305	0.029	0.88	0.81
VF492577	6745	1.341	0.029	0.79	0.67
VF491996	6745	2.345	0.027	1.18	1.28
VF491960	6745	1.437	0.029	0.92	0.91
VF492078	6745	2.169	0.027	1.01	1.01
VF492388	6745	2.010	0.027	1.06	1.06
VF491879	6745	0.081	0.039	0.89	0.66
VF492931	6745	3.059	0.028	0.89	0.87
VF493058	6745	1.784	0.028	0.94	0.91
VF491874	6745	1.907	0.028	0.83	0.77
VF493013	6745	3.332	0.029	1.13	1.26
VF493089	6745	2.228	0.027	0.93	0.91
VF492582	6745	1.579	0.028	0.92	0.85
VF423647	6745	2.517	0.027	1.38	1.56
VF492280	6745	2.705	0.027	0.98	0.99
VF492879	6745	2.595	0.027	0.89	0.88
VF492025	6745	2.780	0.027	1.02	1.02
VF491931	6745	0.157	0.039	0.86	0.61
VF492716	6745	2.804	0.028	1.00	1.02
VF492711	6745	1.568	0.028	1.05	1.13

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF492290	6745	0.467	0.035	0.94	0.82
VF423146	6745	1.536	0.029	0.92	0.86
VF492284	6745	0.516	0.035	0.93	0.82
VF493001	6745	1.756	0.028	0.89	0.82
VF491787	6745	1.562	0.028	1.05	1.11
VF493003	6745	1.909	0.028	0.94	0.91
VF491966	6745	1.953	0.027	1.07	1.11
VF493068	6745	2.999	0.028	0.97	0.96
VF491894	6745	2.356	0.027	0.97	0.95
VF492941	6745	3.565	0.030	1.01	1.10
VF423225	6745	2.656	0.027	0.98	0.99
VF492890	6745	2.555	0.027	0.98	0.99
VF491940	6745	1.890	0.028	0.89	0.84
VF493092	6745	1.867	0.028	0.85	0.78
VF493062	6745	2.012	0.027	0.97	0.96
VF491976	6745	2.635	0.027	1.13	1.17
VF493002	6745	2.333	0.027	0.99	0.97
VF492415	6745	1.772	0.028	0.98	0.94
VF491837	6745	0.530	0.035	0.99	1.04



Table M11. Mathematics Grade 7 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF492966	6458	1.676	0.030	0.88	0.77
VF492597	6458	3.456	0.029	1.04	1.09
VF492307	6458	1.717	0.030	0.89	0.80
VF492394	6458	1.264	0.032	0.89	0.79
VF492967	6458	2.760	0.027	1.11	1.13
VF492878	6458	3.085	0.028	0.97	0.99
VF492672	6458	0.511	0.040	0.79	0.58
VF492888	6458	2.864	0.027	1.03	1.05
VF492871	6458	1.507	0.031	0.87	0.79
VF492640	6458	3.487	0.029	1.06	1.11
VF492836	6458	2.979	0.028	1.10	1.12
VF492853	6458	2.420	0.028	1.03	1.03
VF492835	6458	2.227	0.028	0.87	0.83
VF492419	6458	3.735	0.030	1.08	1.14
VF492666	6458	2.797	0.027	0.96	0.96
VF492760	6458	3.689	0.029	1.03	1.07
VF492653	6458	5.232	0.042	1.06	1.74
VF493021	6458	3.443	0.029	1.10	1.16
VF492578	6458	3.948	0.031	1.00	1.08
VF493038	6458	2.110	0.028	0.80	0.72
VF492357	6458	3.531	0.029	1.10	1.19
VF492663	6458	2.241	0.028	0.86	0.80
VF493057	6458	1.014	0.034	1.03	1.32
VF492665	6458	2.674	0.027	0.98	0.95
VF492413	6458	3.004	0.028	1.12	1.18
VF492973	6458	3.130	0.028	1.08	1.11
VF492696	6458	1.951	0.029	0.88	0.82
VF493061	6458	3.008	0.028	0.84	0.80
VF492864	6458	1.984	0.029	0.82	0.74
VF492302	6458	0.393	0.041	1.03	0.69
VF493046	6458	2.980	0.028	0.98	0.98
VF492425	6458	2.664	0.027	0.94	0.92
VF492951	6458	2.766	0.027	1.20	1.28
VF492720	6458	3.043	0.028	0.86	0.84
VF492765	6458	3.661	0.029	1.12	1.23
VF492538	6458	1.763	0.029	0.79	0.67
VF493019	6458	3.110	0.028	1.01	1.01
VF492673	6458	5.128	0.041	1.11	1.52
VF493067	6458	3.178	0.028	1.02	1.04
VF492830	6458	2.083	0.028	0.90	0.83

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF492929	6458	0.287	0.043	1.01	1.14
VF492531	6458	0.533	0.039	0.94	0.91
VF492955	6458	3.017	0.028	1.19	1.26
VF493071	6458	3.689	0.029	1.13	1.20
VF492780	6458	2.280	0.028	0.94	0.91
VF493015	6458	2.892	0.028	1.00	0.99
VF492567	6458	3.115	0.028	1.09	1.13
VF493077	6458	1.453	0.031	0.89	0.81
VF492546	6458	3.196	0.028	1.14	1.17
VF493052	6458	2.628	0.027	1.05	1.09
VF493036	6458	3.679	0.029	1.16	1.29
VF492861	6458	2.280	0.028	0.98	1.00
VF492589	6458	4.105	0.032	0.92	1.02
VF492778	6458	3.036	0.028	0.90	0.87
VF492259	6458	3.121	0.028	0.96	0.96
VF493043	6458	3.747	0.030	1.23	1.41
VF492901	6458	3.334	0.028	1.00	1.03
VF492748	6458	2.030	0.028	0.91	0.87
VF493064	6458	1.732	0.030	0.99	1.03

Table M12. Mathematics Grade 8 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF491923	6467	1.010	0.038	0.97	0.91
VF493115	6467	1.680	0.032	1.13	1.35
VF491907	6467	2.573	0.028	1.03	1.03
VF493150	6467	3.351	0.028	0.92	0.90
VF492845	6467	3.393	0.028	1.08	1.10
VF491824	6467	3.680	0.028	1.07	1.13
VF494699	6467	2.840	0.028	0.98	0.96
VF492863	6467	2.636	0.028	0.90	0.88
VF493113	6467	2.277	0.029	0.99	1.01
VF492712	6467	3.501	0.028	0.94	0.91
VF491857	6467	2.842	0.028	0.94	0.92
VF492726	6467	3.902	0.029	1.04	1.05
VF493157	6467	3.374	0.028	1.13	1.18
VF491873	6467	4.489	0.031	0.94	1.00
VF492874	6467	3.697	0.028	1.08	1.11
VF493121	6467	3.102	0.028	1.07	1.09
VF491915	6467	3.924	0.029	1.13	1.24
VF494120	6467	4.376	0.031	1.20	1.38
VF492856	6467	2.924	0.028	0.89	0.86
VF493159	6467	2.799	0.028	1.10	1.16
VF492438	6467	1.984	0.030	0.99	1.00
VF491991	6467	2.234	0.029	0.95	0.86
VF491965	6467	2.582	0.028	1.05	1.11
VF492289	6467	2.805	0.028	0.97	0.97
VF493034	6467	3.021	0.028	0.93	0.90
VF494727	6467	3.230	0.028	0.94	0.91
VF492410	6467	3.070	0.028	1.00	0.98
VF494928	6467	3.732	0.028	1.12	1.16
VF492278	6467	1.871	0.031	0.84	0.72
VF492345	6467	3.701	0.028	0.99	1.03
VF494751	6467	3.367	0.028	1.02	1.01
VF493040	6467	2.331	0.029	0.90	0.85
VF491975	6467	2.045	0.030	0.89	0.85
VF492907	6467	1.963	0.030	0.87	0.78
VF492430	6467	2.845	0.028	1.07	1.06
VF493107	6467	3.678	0.028	1.10	1.16
VF492563	6467	2.664	0.028	1.09	1.14
VF492579	6467	3.219	0.028	1.13	1.15
VF492028	6467	3.035	0.028	0.97	0.95
VF492272	6467	2.461	0.028	0.84	0.77

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF494801	6467	2.840	0.028	0.98	0.95
VF492420	6467	1.917	0.030	0.88	0.79
VF494769	6467	2.228	0.029	0.95	0.89
VF492920	6467	2.446	0.028	0.80	0.76
VF492592	6467	2.219	0.029	1.02	1.21
VF492024	6467	4.030	0.029	0.98	1.05
VF492268	6467	3.790	0.028	1.22	1.31
VF491949	6467	2.296	0.029	0.91	0.89
VF492258	6467	3.175	0.028	1.14	1.19
VF493045	6467	2.456	0.028	0.95	0.91
VF492212	6467	2.934	0.028	1.10	1.12
VF492414	6467	2.822	0.028	0.91	0.91
VF494776	6467	3.041	0.028	0.92	0.88
VF492008	6467	3.147	0.028	1.02	1.02
VF493011	6467	1.842	0.031	0.84	0.72
VF492400	6467	3.211	0.028	1.06	1.09
VF492917	6467	1.847	0.031	0.85	0.78
VF494819	6467	3.858	0.029	1.06	1.11
VF493088	6467	2.375	0.029	0.89	0.85
VF492436	6467	1.878	0.031	0.99	0.98
VF493097	6467	4.093	0.030	1.21	1.41
VF492231	6467	4.003	0.029	1.20	1.31
VF494760	6467	2.779	0.028	0.89	0.84
VF492393	6467	2.995	0.028	0.95	0.92
VF492440	6467	1.939	0.030	0.93	0.87

Science

Table M13. Science Grade 4 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF431075	7020	0.280	0.026	0.93	0.91
VF431078	7020	0.633	0.026	0.94	0.93
VF483448	7020	-1.681	0.039	0.81	0.68
VF484396	7020	0.493	0.026	1.18	1.23
VF430688	7020	0.072	0.027	1.00	0.98
VF430686	7020	0.703	0.026	1.02	1.04
VF294929	7020	-0.285	0.028	0.95	0.95
VF296821	7020	-0.219	0.028	1.00	1.01
VF484935	7020	1.201	0.027	1.21	1.29
VF290777	7020	-0.996	0.032	1.09	1.04
VF431081	7020	-0.974	0.032	0.95	0.99
VF283606	7020	-1.498	0.037	0.92	0.75
VF283022	7020	-0.294	0.028	0.93	0.88
VF311559	7020	-0.051	0.027	0.99	0.99
VF311567	7020	0.164	0.027	0.98	0.96
VF311586	7020	-0.200	0.027	0.96	0.92
VF311548	7020	-0.648	0.030	1.00	1.00
VF431027	7020	-0.278	0.028	0.95	0.95
VF431028	7020	-1.271	0.034	0.93	0.80
VF287722	7020	0.785	0.026	1.00	1.01
VF287717	7020	1.289	0.027	0.99	1.04
VF284006	7020	0.355	0.026	0.97	0.94
VF284002	7020	0.237	0.026	1.09	1.12
VF431125	7020	0.371	0.026	1.09	1.11
VF431127	7020	-0.596	0.029	0.92	0.88
VF431129	7020	0.018	0.027	1.03	1.07
VF293507	7020	-0.916	0.031	0.82	0.74
VF292879	7020	-0.332	0.028	0.95	0.90
VF294472	7020	-0.010	0.027	0.98	0.95
VF407152	7020	1.097	0.027	1.11	1.15
VF407138	7020	0.957	0.026	1.07	1.11
VF406427	7020	0.586	0.026	0.95	0.93
VF483424	7020	-0.183	0.027	0.92	0.87
VF483437	7020	-2.379	0.050	1.02	1.02
VF287740	7020	0.843	0.026	1.04	1.06
VF287742	7020	0.762	0.026	1.05	1.08
VF287745	7020	0.753	0.026	1.10	1.12
VF393724	7020	-0.825	0.031	0.84	0.68

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF393699	7020	0.331	0.026	0.93	0.92
VF393721	7020	0.237	0.026	1.02	1.02
VF431046	7020	0.543	0.026	1.08	1.10
VF269846	7020	0.586	0.026	1.06	1.06
VF269841	7020	0.916	0.026	1.18	1.24
VF296839	7020	0.832	0.026	1.01	1.03
VF282661	7020	0.160	0.027	0.95	0.93
VF282670	7020	-1.089	0.033	0.91	0.82
VF386811	7020	0.241	0.026	1.00	1.01
VF386826	7020	0.235	0.026	0.99	0.96
VF287864	7020	-0.205	0.028	0.97	0.96
VF287870	7020	0.722	0.026	1.10	1.12

Table M14. Science Grade 8 IRT Statistics for Operational Items

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF484958	6452	-0.002	0.027	0.96	0.93
VF484974	6452	0.472	0.027	1.06	1.11
VF486678	6452	1.360	0.030	1.05	1.10
VF486675	6452	0.195	0.027	1.07	1.09
VF394477	6452	0.113	0.027	0.91	0.89
VF394502	6452	0.298	0.027	1.04	1.05
VF484993	6452	0.464	0.027	0.96	0.96
VF484999	6452	-0.030	0.027	1.08	1.11
VF407356	6452	-1.969	0.039	0.95	0.88
VF407330	6452	-0.094	0.027	0.97	0.96
VF394777	6452	-0.002	0.027	1.01	1.02
VF394780	6452	0.135	0.027	1.03	1.05
VF394809	6452	-0.011	0.027	1.01	1.02
VF394814	6452	0.325	0.027	1.09	1.13
VF431421	6452	-0.184	0.028	1.03	1.04
VF431423	6452	0.968	0.028	1.05	1.11
VF394561	6452	-0.405	0.028	0.94	0.92
VF394565	6452	-0.968	0.030	0.92	0.88
VF313289	6452	0.301	0.027	1.00	1.01
VF313291	6452	-0.425	0.028	0.95	0.92
VF313300	6452	-0.069	0.027	0.98	0.97
VF486847	6452	-0.523	0.028	0.98	0.98
VF486858	6452	0.948	0.028	1.00	1.02
VF486815	6452	-0.329	0.028	0.92	0.88
VF486821	6452	-0.302	0.028	0.93	0.89
VF308868	6452	-1.654	0.036	0.87	0.72
VF308871	6452	0.123	0.027	0.94	0.91
VF308869	6452	-0.201	0.028	1.08	1.14
VF486166	6452	-0.275	0.028	1.04	1.07
VF486163	6452	-0.227	0.028	0.90	0.85
VF407480	6452	0.044	0.027	0.93	0.91
VF407483	6452	-0.075	0.027	1.12	1.14
VF407155	6452	-0.649	0.029	0.97	0.95
VF407242	6452	-0.990	0.031	1.06	1.21
VF431683	6452	-0.994	0.031	0.97	0.93
VF431688	6452	0.242	0.027	1.00	1.00
VF486771	6452	-0.098	0.027	1.03	1.04
VF486782	6452	-0.126	0.027	0.99	0.98
VF486765	6452	-0.182	0.028	0.88	0.83
VF486914	6452	0.961	0.028	1.06	1.10

Accession Number	<i>N</i>	Rasch Difficulty	Rasch <i>SE</i>	Infit	Outfit
VF486941	6452	0.413	0.027	1.01	1.02
VF313274	6452	1.083	0.029	1.08	1.16
VF313280	6452	0.366	0.027	0.94	0.92
VF313281	6452	-0.031	0.027	0.96	0.93
VF485018	6452	-0.177	0.028	1.07	1.08
VF485023	6452	0.300	0.027	1.06	1.08



## Appendix N: Frequency of Individual Accommodations for 2014 PAWS and SAWS Tests

Students received the same accommodations for all subjects (Reading, Mathematics, Science and SAWS). The only exceptions are for those accommodations shaded in yellow that were not allowed for the Reading test.

### Reading

Table N1. Frequency of IEP Student's Standard Accommodations: Presentation Accommodations

Code	Accommodation	3	4	5	6	7	8
1	Student uses a Braille Special Test Form.	0	0	0	0	1	0
2	Student uses a Large Print Special Test Form.	5	4	4	4	2	6
3	Student uses an Audio Special Test Form.	2	1	0	25	53	47
4	Student uses magnification devices.	3	1	1	1	2	1
5	Student uses color overlays to reduce glare or enhance text.	12	10	13	6	3	5
6	Student uses templates to reduce the amount of visible print.	21	8	7	6	7	11
7	Student uses tactile graphics.	0	0	0	0	0	0
8	Sign language interpreter signs directions in all content areas and/or signs test questions as written in all content areas EXCEPT Reading. The interpreter may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one interpreter be provided for each individual student.	1	3	0	1	1	0
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	4	11	4	4	2	6
10	A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.	145	171	144	81	75	73
11	Student asks for clarification of directions (not test questions or answer choices).	308	375	375	393	290	307
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	5	6	1	2	7	3
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	1	0	0	2	1	1

Table N2. Frequency of IEP Student's Standard Accommodations: Response Accommodations

Code	Accommodation	3	4	5	6	7	8
14	A certified staff member or access assistant scribes what a student dictates through alternate augmentative communications (AAC), pointing, sign language, or speech. The scribe may not edit or alter the student's work in any way and must record, word for word, exactly what the student has dictated. A scribe must allow the student to review and edit what he or she has written. The student's final response must be transcribed by a certified staff member or access assistant into the Student Test and Answer Book on the pages that the student's response is to be written.	18	45	26	24	18	17
15	A student types responses using a word processor. Dictionary and synonym/thesaurus devices MUST be disabled. The margins for word-processed documents should match the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	7	13	9	6	7	3
16	Student uses speech-to-text conversion or voice recognition in all content areas. The margins for this document should match as closely as possible the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	1	1	0	0	0
17	Student uses a Braille. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	0	1	0	0	0
18	Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	3	1	0	3	0	1
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	87	97	98	110	65	46
20	Student uses visual organizers including graph paper, place markers, and templates. Student uses a pencil to underline text. Highlighters CANNOT be used in the Student Test and Answer Book.	47	81	82	117	75	50

Table N3. Frequency of IEP Student's Standard Accommodations: Setting Accommodations

Code	Accommodation	3	4	5	6	7	8
21	Student takes the test in a different building location in a small group or individually. Changes can also be made to a student's location within a room to reduce distractions to the student or to other students, to increase physical access, or enable the use of special equipment. Students must be monitored by a certified staff member.	630	705	671	669	607	599

Table N4. Frequency of IEP Student's Standard Accommodations: Timing and Scheduling Accommodations

Code	Accommodation	3	4	5	6	7	8
22	Student is provided with extended time to complete the assessment.	480	531	514	528	401	398
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	453	483	427	435	309	294
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	144	118	107	121	52	57

Table N5. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations

Code	Accommodation	3	4	5	6	7	8
25	A certified staff member or access assistant translates written directions to the student.	1	5	7	6	4	2
26	A certified staff member or access assistant re-reads, simplifies, or clarifies directions in English or in the student's primary language (NOT test questions or answer choices) without clueing correct responses.	40	26	20	30	29	26
27	A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one reader be provided for each individual student.	9	11	12	2	4	0
28	Student uses a bilingual dictionary provided by the school.	10	5	19	11	5	8

Table N6. Frequency of English Language Learners Standard Accommodations: Setting Accommodations

Code	Accommodation	3	4	5	6	7	8
29	Student takes test in a different or individual location, or in a small group .	80	43	44	25	25	15

Table N7. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations

Code	Accommodation	3	4	5	6	7	8
30	Student is provided with multiple, individual breaks as needed.	48	32	23	24	22	11
31	Student is allowed to complete the test over multiple days.	27	13	17	12	10	8

Mathematics

Table N8. Frequency of IEP Student's Standard Accommodations: Presentation Accommodations

Code	Accommodation	3	4	5	6	7	8
1	Student uses a Braille Special Test Form.	0	0	0	0	0	0
2	Student uses a Large Print Special Test Form.	5	4	3	3	2	6
3	Student uses an Audio Special Test Form.	29	17	19	29	64	68
4	Student uses magnification devices.	1	1	1	1	2	0
5	Student uses color overlays to reduce glare or enhance text.	7	8	14	3	3	4
6	Student uses templates to reduce the amount of visible print.	10	5	9	3	6	10
7	Student uses tactile graphics.	0	0	0	0	0	0
8	Sign language interpreter signs directions in all content areas and/or signs test questions as written in all content areas EXCEPT Reading. The interpreter may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one interpreter be provided for each individual student.	0	4	0	1	0	0
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	6	9	4	4	4	6
10	A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.	458	497	473	419	336	303
11	Student asks for clarification of directions (not test questions or answer choices).	318	371	360	405	322	300
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	5	4	1	2	7	6
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	10	19	13	4	1	2

Table N9. Frequency of IEP Student's Standard Accommodations: Response Accommodations

Code	Accommodation	3	4	5	6	7	8
14	A certified staff member or access assistant scribes what a student dictates through alternate augmentative communications (AAC), pointing, sign language, or speech. The scribe may not edit or alter the student's work in any way and must record, word for word, exactly what the student has dictated. A scribe must allow the student to review and edit what he or she has written. The student's final response must be transcribed by a certified staff member or access assistant into the Student Test and Answer Book on the pages that the student's response is to be written.	16	46	27	24	15	15
15	A student types responses using a word processor. Dictionary and synonym/thesaurus devices MUST be disabled. The margins for word-processed documents should match the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	4	12	9	6	7	3
16	Student uses speech-to-text conversion or voice recognition in all content areas. The margins for this document should match as closely as possible the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	1	1	0	0	0
17	Student uses a Braille. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	0	0	0	0	0
18	Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	6	2	2	0	2
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	82	84	96	106	77	44
20	Student uses visual organizers including graph paper, place markers, and templates. Student uses a pencil to underline text. Highlighters CANNOT be used in the Student Test and Answer Book.	121	136	130	167	116	96

Table N10. Frequency of IEP Student's Standard Accommodations: Setting Accommodations

Code	Accommodation	3	4	5	6	7	8
21	Student takes the test in a different building location in a small group or individually. Changes can also be made to a student's location within a room to reduce distractions to the student or to other students, to increase physical access, or enable the use of special equipment. Students must be monitored by a certified staff member.	614	703	665	653	608	600

Table N11. Frequency of IEP Student's Standard Accommodations: Timing and Scheduling Accommodations

Code	Accommodation	3	4	5	6	7	8
22	Student is provided with extended time to complete the assessment.	451	515	517	501	384	391
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	415	459	424	422	307	295
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	139	112	106	112	55	56

Table N12. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations

Code	Accommodation	3	4	5	6	7	8
25	A certified staff member or access assistant translates written directions to the student.	4	4	6	10	6	1
26	A certified staff member or access assistant re-reads, simplifies, or clarifies directions in English or in the student's primary language (NOT test questions or answer choices) without clueing correct responses.	48	35	25	31	34	20
27	A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one reader be provided for each individual student.	132	70	49	43	39	15
28	Student uses a bilingual dictionary provided by the school.	58	28	16	16	11	9

Table N13. Frequency of English Language Learners Standard Accommodations: Setting Accommodations

Code	Accommodation	3	4	5	6	7	8
29	Student takes test in a different or individual location, or in a small group.	143	75	56	37	37	20

Table N14. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations

Code	Accommodation	3	4	5	6	7	8
30	Student is provided with multiple, individual breaks as needed.	58	41	23	28	28	14
31	Student is allowed to complete the test over multiple days.	28	15	24	17	15	10

Science

Table N15. Frequency of IEP Student's Standard Accommodations: Presentation Accommodations

Code	Accommodation	4	8
1	Student uses a Braille Special Test Form.	0	0
2	Student uses a Large Print Special Test Form.	4	6
3	Student uses an Audio Special Test Form.	20	65
4	Student uses magnification devices.	1	1
5	Student uses color overlays to reduce glare or enhance text.	10	4
6	Student uses templates to reduce the amount of visible print.	6	10
7	Student uses tactile graphics.	0	0
	Sign language interpreter signs directions in all content areas and/or signs test questions as written in all content areas EXCEPT Reading. The interpreter may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content.	4	0
8	It is recommended that one interpreter be provided for each individual student.		
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	8	4
	A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.	490	284
10			
11	Student asks for clarification of directions (not test questions or answer choices).	343	266
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	4	6
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	19	1

Table N16. Frequency of IEP Student's Standard Accommodations: Response Accommodations

Code	Accommodation	4	8
14	A certified staff member or access assistant scribes what a student dictates through alternate augmentative communications (AAC), pointing, sign language, or speech. The scribe may not edit or alter the student's work in any way and must record, word for word, exactly what the student has dictated. A scribe must allow the student to review and edit what he or she has written. The student's final response must be transcribed by a certified staff member or access assistant into the Student Test and Answer Book on the pages that the student's response is to be written.	47	15
15	A student types responses using a word processor. Dictionary and synonym/thesaurus devices MUST be disabled. The margins for word-processed documents should match the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	13	4
16	Student uses speech-to-text conversion or voice recognition in all content areas. The margins for this document should match as closely as possible the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	1
17	Student uses a Braille. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	0
18	Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	6	3
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	84	40
20	Student uses visual organizers including graph paper, place markers, and templates. Student uses a pencil to underline text. Highlighters CANNOT be used in the Student Test and Answer Book.	101	77

Table N17. Frequency of IEP Student's Standard Accommodations: Setting Accommodations

Code	Accommodation	4	8
21	Student takes the test in a different building location in a small group or individually. Changes can also be made to a student's location within a room to reduce distractions to the student or to other students, to increase physical access, or enable the use of special equipment. Students must be monitored by a certified staff member.	682	573



Table N18. Frequency of IEP Student's Standard Accommodations: Timing and Scheduling Accommodations

Code	Accommodation	4	8
22	Student is provided with extended time to complete the assessment.	487	371
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	461	281
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	111	62

Table N19. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations

Code	Accommodation	4	8
25	A certified staff member or access assistant translates written directions to the student.	4	2
26	A certified staff member or access assistant re-reads, simplifies, or clarifies directions in English or in the student's primary language (NOT test questions or answer choices) without clueing correct responses.	29	17
27	A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one reader be provided for each individual student.	64	14
28	Student uses a bilingual dictionary provided by the school.	33	4

Table N20. Frequency of English Language Learners Standard Accommodations: Setting Accommodations

Code	Accommodation	4	8
29	Student takes test in a different or individual location, or in a small group.	76	20

Table N21. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations

Code	Accommodation	4	8
30	Student is provided with multiple, individual breaks as needed.	41	13
31	Student is allowed to complete the test over multiple days.	15	6

SAWS

Table N22. Frequency of IEP Student's Standard Accommodations: Presentation Accommodations

Code	Accommodation	3	5	7
1	Student uses a Braille Special Test Form.	0	0	1
2	Student uses a Large Print Special Test Form.	4	1	3
3	Student uses an Audio Special Test Form.	3	1	1
4	Student uses magnification devices.	2	2	1
5	Student uses color overlays to reduce glare or enhance text.	6	10	3
6	Student uses templates to reduce the amount of visible print.	19	7	7
7	Student uses tactile graphics.	0	0	0
8	Sign language interpreter signs directions in all content areas and/or signs test questions as written in all content areas EXCEPT Reading. The interpreter may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one interpreter be provided for each individual student.	0	0	0
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	4	3	2
10	A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.	351	417	301
11	Student asks for clarification of directions (not test questions or answer choices).	290	377	321
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	7	1	3
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	0	0	3

Table N23. Frequency of IEP Student's Standard Accommodations: Response Accommodations

Code	Accommodation	3	5	7
14	A certified staff member or access assistant scribes what a student dictates through alternate augmentative communications (AAC), pointing, sign language, or speech. The scribe may not edit or alter the student's work in any way and must record, word for word, exactly what the student has dictated. A scribe must allow the student to review and edit what he or she has written. The student's final response must be transcribed by a certified staff member or access assistant into the Student Test and Answer Book on the pages that the student's response is to be written.	81	74	51
15	A student types responses using a word processor. Dictionary and synonym/thesaurus devices MUST be disabled. The margins for word-processed documents should match the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	17	25	55
16	Student uses speech-to-text conversion or voice recognition in all content areas. The margins for this document should match as closely as possible the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	0	0
17	Student uses a Braille. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	0	0
18	Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	0	1
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	55	82	34
20	Student uses visual organizers including graph paper, place markers, and templates. Student uses a pencil to underline text. Highlighters CANNOT be used in the Student Test and Answer Book.	78	129	54

Table N24. Frequency of IEP Student's Standard Accommodations: Setting Accommodations

Code	Accommodation	3	5	7
21	Student takes the test in a different building location in a small group or individually. Changes can also be made to a student's location within a room to reduce distractions to the student or to other students, to increase physical access, or enable the use of special equipment. Students must be monitored by a certified staff member.	496	588	507

Table N25. Frequency of IEP Student's Standard Accommodations: Timing and Scheduling Accommodations

Code	Accommodation	3	5	7
22	Student is provided with extended time to complete the assessment.	379	470	347
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	324	380	237
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	104	90	33

Table N26. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations

Code	Accommodation	3	5	7
25	A certified staff member or access assistant translates written directions to the student.	2	2	0
26	A certified staff member or access assistant re-reads, simplifies, or clarifies directions in English or in the student's primary language (NOT test questions or answer choices) without clueing correct responses.	33	14	17
27	A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one reader be provided for each individual student.	36	24	8
28	Student uses a bilingual dictionary provided by the school.	3	5	7

Table N27. Frequency of English Language Learners Standard Accommodations: Setting Accommodations

Code	Accommodation	3	5	7
29	Student takes test in a different or individual location, or in a small group .	55	26	21

Table N28. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations

Code	Accommodation	3	5	7
30	Student is provided with multiple, individual breaks as needed.	29	13	15
31	Student is allowed to complete the test over multiple days.	10	8	6

## Appendix O: Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

### Reading

Table O1. Reading Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	<i>N</i>	Mean Scale Score	<i>SD</i> Scale Score
3	Yes	1083	559.4	51.1
	No	6282	607.0	46.3
4	Yes	1084	572.9	48.3
	No	5938	623.9	42.0
5	Yes	1037	584.0	46.4
	No	6038	633.8	43.6
6	Yes	920	587.2	43.1
	No	5838	644.1	45.0
7	Yes	843	605.7	38.8
	No	5953	655.5	41.8
8	Yes	868	611.9	44.9
	No	5913	668.3	43.7

### Mathematics

Table O2. Mathematics Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	<i>N</i>	Mean Scale Score	<i>SD</i> Scale Score
3	Yes	1081	571.5	46.7
	No	6288	604.9	48.9
4	Yes	1085	604.7	42.0
	No	5941	640.2	43.5
5	Yes	1037	623.7	43.7
	No	6040	665.9	49.1
6	Yes	919	637.1	38.5
	No	5841	684.6	46.4
7	Yes	840	652.8	35.1
	No	5959	697.3	44.0
8	Yes	866	667.4	34.5
	No	5918	713.4	43.3

Science

Table O3. Science Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	<i>N</i>	Mean Scale Score	<i>SD</i> Scale Score
4	Yes	1081	644.9	41.8
	No	5941	674.3	45.8
8	Yes	863	616.6	37.0
	No	5907	655.9	44.5

SAWS

Table O4. SAWS Mean Raw Scores, Counts, and Raw Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	<i>N</i>	Mean Scale Score	<i>SD</i> Scale Score
3	Yes	1066	10.6	4.1
	No	6249	14.3	4.4
5	Yes	1029	12.1	3.8
	No	6004	15.5	3.9
7	Yes	834	11.6	3.6
	No	5929	16.3	4.1

## Appendix P: Scaled Score Descriptive Statistics by Demographic Subgroup

Table P1. Summary Statistics of Reading, Mathematics, and Science Scale Score by Grade

Grade	<i>N</i>	Mean	<i>SD</i>
Reading			
3	7365	600.0	50.0
4	7022	616.0	46.8
5	7075	626.5	47.4
6	6758	636.3	48.8
7	6796	649.3	44.6
8	6781	661.1	47.7
Mathematics			
3	7369	600.0	50.0
4	7026	634.7	45.1
5	7077	659.7	50.6
6	6760	678.1	48.2
7	6799	691.8	45.5
8	6784	707.5	45.0
Science			
4	7022	669.7	46.4
8	6770	650.9	45.5

Reading

Table P2. Summary Statistics of Reading Grade 3 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7365	600.0	50.0
Male	3709	595.8	50.7
Female	3636	604.3	48.8
Unknown	20	587.0	53.8
American Indian/Alaska Native	309	564.0	47.5
Asian	52	606.9	38.6
African American	92	589.8	52.4
Native Hawaiian or other/Pacific Islander	8	589.3	45.0
Hispanic/Latino	1068	582.8	47.6
White	5662	605.5	48.9
Multiracial	146	593.7	48.4
Unknown	28	596.3	70.0
Free/Reduced Lunch	2363	589.1	48.7
Not Free/Reduced Lunch	5002	605.1	49.8
Special Education	1083	559.4	51.1
Not Special Education	6282	607.0	46.3
English Language Learner	371	562.7	43.9
Not English Language Learner	6994	601.9	49.5

Table P3. Summary Statistics of Reading Grade 4 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7022	616.0	46.8
Male	3639	612.0	47.6
Female	3377	620.4	45.6
Unknown	6	582.2	55.7
American Indian/Alaska Native	267	584.9	45.3
Asian	72	623.9	40.9
African American	85	598.8	43.3
Native Hawaiian or other/Pacific Islander	6	609.2	32.6
Hispanic/Latino	968	597.9	45.4
White	5480	620.7	45.9
Multiracial	133	623.8	45.8
Unknown	11	608.1	43.7
Free/Reduced Lunch	2195	603.6	46.5
Not Free/Reduced Lunch	4827	621.7	45.9
Special Education	1084	572.9	48.3
Not Special Education	5938	623.9	42.0
English Language Learner	216	566.1	39.8
Not English Language Learner	6806	617.6	46.2



Table P4. Summary Statistics of Reading Grade 5 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7075	626.5	47.4
Male	3635	624.1	47.9
Female	3430	629.2	46.8
Unknown	10	605.2	41.9
American Indian/Alaska Native	268	589.5	44.8
Asian	74	630.3	50.3
African American	70	614.8	47.0
Native Hawaiian or other/Pacific Islander	7	579.9	27.7
Hispanic/Latino	933	608.8	44.6
White	5578	631.5	46.4
Multiracial	133	625.3	47.7
Unknown	12	602.2	43.5
Free/Reduced Lunch	2151	614.9	46.2
Not Free/Reduced Lunch	4924	631.6	47.1
Special Education	584.0	46.4	584.0
Not Special Education	633.8	43.6	633.8
English Language Learner	571.5	38.4	571.5
Not English Language Learner	627.7	46.9	627.7

Table P5. Summary Statistics of Reading Grade 6 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	6758	636.3	48.8
Male	3509	634.0	49.0
Female	3244	638.9	48.5
Unknown	5	609.8	43.0
American Indian/Alaska Native	269	602.7	43.7
Asian	53	650.6	46.4
African American	68	613.1	48.3
Native Hawaiian or other/Pacific Islander	10	642.1	59.1
Hispanic/Latino	912	622.1	44.3
White	5305	640.7	48.5
Multiracial	135	634.5	51.0
Unknown	6	635.2	59.3
Free/Reduced Lunch	2009	622.2	46.9
Not Free/Reduced Lunch	4749	642.3	48.4
Special Education	920	587.2	43.1
Not Special Education	5838	644.1	45.0
English Language Learner	155	581.0	39.0
Not English Language Learner	6603	637.6	48.3

Table P6. Summary Statistics of Reading Grade 7 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	6796	649.3	44.6
Male	3498	644.4	44.4
Female	3294	654.6	44.1
Unknown	4	589.8	21.8
American Indian/Alaska Native	245	610.1	37.3
Asian	58	659.2	45.3
African American	83	632.5	44.6
Native Hawaiian or other/Pacific Islander	15	640.6	51.0
Hispanic/Latino	916	633.7	40.7
White	5345	654.0	43.9
Multiracial	128	648.2	45.7
Unknown	6	611.3	30.9
Free/Reduced Lunch	2009	638.2	42.0
Not Free/Reduced Lunch	4787	653.9	44.8
Special Education	843	605.7	38.8
Not Special Education	5953	655.5	41.8
English Language Learner	163	596.5	29.8
Not English Language Learner	6633	650.6	44.1

Table P7. Summary Statistics of Reading Grade 8 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	6781	661.1	47.7
Male	3558	656.4	49.0
Female	3220	666.3	45.7
Unknown	3	603.3	42.6
American Indian/Alaska Native	222	631.5	48.6
Asian	57	670.5	45.2
African American	98	643.5	48.4
Native Hawaiian or other/Pacific Islander	11	647.8	70.3
Hispanic/Latino	891	642.1	43.9
White	5363	666.0	46.8
Multiracial	134	650.8	49.9
Unknown	5	624.4	45.5
Free/Reduced Lunch	1832	648.1	45.5
Not Free/Reduced Lunch	4949	665.9	47.6
Special Education	868	611.9	44.9
Not Special Education	5913	668.3	43.7
English Language Learner	145	601.4	34.9
Not English Language Learner	6636	662.4	47.1

Mathematics

Table P8. Summary Statistics of Mathematics Grade 3 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7369	600.0	50.0
Male	3718	602.3	51.0
Female	3639	597.7	48.9
Unknown	12	578.6	46.2
American Indian/Alaska Native	307	566.6	43.9
Asian	52	613.4	52.3
African American	94	577.4	48.5
Native Hawaiian or other/Pacific Islander	9	582.6	47.4
Hispanic/Latino	1076	580.1	44.1
White	5666	606.1	49.6
Multiracial	143	590.9	47.9
Unknown	22	588.9	48.0
Free/Reduced Lunch	2364	589.9	47.0
Not Free/Reduced Lunch	5005	604.8	50.7
Special Education	1081	571.5	46.7
Not Special Education	6288	604.9	48.9
English Language Learner	380	565.5	40.3
Not English Language Learner	6989	601.9	49.8

Table P9. Summary Statistics of Mathematics Grade 4 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7026	634.7	45.1
Male	3641	635.6	46.9
Female	3370	633.9	43.1
Unknown	15	607.8	41.8
American Indian/Alaska Native	268	609.2	41.5
Asian	74	645.8	53.7
African American	86	619.3	41.7
Native Hawaiian or other/Pacific Islander	6	616.8	43.5
Hispanic/Latino	969	618.2	40.5
White	5474	638.9	44.8
Multiracial	132	640.7	46.9
Unknown	17	619.2	41.9
Free/Reduced Lunch	2194	625.2	42.5
Not Free/Reduced Lunch	4832	639.0	45.6
Special Education	1085	604.7	42.0
Not Special Education	5941	640.2	43.5
English Language Learner	223	595.5	31.1
Not English Language Learner	6803	636.0	44.9

Table P10. Summary Statistics of Mathematics Grade 5 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7077	659.7	50.6
Male	3634	661.7	51.6
Female	3431	657.7	49.4
Unknown	12	660.4	45.8
American Indian/Alaska Native	267	627.7	40.5
Asian	75	683.9	61.0
African American	71	643.8	50.5
Native Hawaiian or other/Pacific Islander	7	622.1	24.6
Hispanic/Latino	933	643.6	44.5
White	5573	664.1	50.7
Multiracial	134	655.8	47.6
Unknown	17	622.5	41.4
Free/Reduced Lunch	2146	648.9	47.3
Not Free/Reduced Lunch	4931	664.4	51.2
Special Education	1037	623.7	43.7
Not Special Education	6040	665.9	49.1
English Language Learner	155	617.1	35.7
Not English Language Learner	6922	660.7	50.5

Table P11. Summary Statistics of Mathematics Grade 6 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	6760	678.1	48.2
Male	3511	677.2	48.7
Female	3244	679.1	47.6
Unknown	5	652.2	32.6
American Indian/Alaska Native	268	647.9	40.6
Asian	55	699.0	48.5
African American	69	662.2	41.3
Native Hawaiian or other/Pacific Islander	10	676.4	54.9
Hispanic/Latino	920	663.9	42.4
White	5298	682.3	48.5
Multiracial	134	672.0	49.0
Unknown	6	653.8	28.1
Free/Reduced Lunch	2011	664.9	44.3
Not Free/Reduced Lunch	4749	683.7	48.7
Special Education	919	637.1	38.5
Not Special Education	5841	684.6	46.4
English Language Learner	167	633.7	33.1
Not English Language Learner	6593	679.2	48.0

Table P12. Summary Statistics of Mathematics Grade 7 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	6799	691.8	45.5
Male	3504	691.5	45.8
Female	3291	692.3	45.1
Unknown	4	653.0	17.7
American Indian/Alaska Native	244	654.2	30.4
Asian	58	719.5	69.4
African American	84	668.5	40.9
Native Hawaiian or other/Pacific Islander	15	677.5	28.6
Hispanic/Latino	920	676.6	39.7
White	5345	696.4	45.2
Multiracial	127	688.4	46.3
Unknown	6	667.3	61.0
Free/Reduced Lunch	2011	680.4	40.1
Not Free/Reduced Lunch	4788	696.7	46.7
Special Education	840	652.8	35.1
Not Special Education	5959	697.3	44.0
English Language Learner	174	650.4	26.9
Not English Language Learner	6625	692.9	45.3

Table P13. Summary Statistics of Mathematics Grade 8 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	6784	707.5	45.0
Male	3561	706.8	46.0
Female	3220	708.3	43.8
Unknown	3	667.0	46.6
American Indian/Alaska Native	222	682.0	35.8
Asian	57	726.9	52.9
African American	98	685.3	44.5
Native Hawaiian or other/Pacific Islander	11	693.3	52.8
Hispanic/Latino	895	692.3	39.2
White	5362	711.5	45.1
Multiracial	134	702.3	44.3
Unknown	5	681.4	37.4
Free/Reduced Lunch	1836	695.9	40.6
Not Free/Reduced Lunch	4948	711.8	45.8
Special Education	866	667.4	34.5
Not Special Education	5918	713.4	43.3
English Language Learner	149	665.7	28.6
Not English Language Learner	6635	708.4	44.9

Science

Table P14. Summary Statistics of Science Grade 4 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7022	669.7	46.4
Male	3639	669.4	46.9
Female	3371	670.2	46.0
Unknown	12	645.6	26.3
American Indian/Alaska Native	267	633.3	38.3
Asian	74	672.3	47.9
African American	86	646.5	37.4
Native Hawaiian or other/Pacific Islander	6	654.5	24.4
Hispanic/Latino	973	650.6	40.1
White	5467	675.1	46.2
Multiracial	133	675.6	48.6
Unknown	16	665.0	32.9
Free/Reduced Lunch	2192	658.4	42.5
Not Free/Reduced Lunch	4830	674.9	47.2
Special Education	1081	644.9	41.8
Not Special Education	5941	674.3	45.8
English Language Learner	224	625.0	32.5
Not English Language Learner	6798	671.2	46.1

Table P15. Summary Statistics of Science Grade 8 Scale Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	6770	650.9	45.5
Male	3555	652.4	47.4
Female	3212	649.2	43.3
Unknown	3	610.7	19.1
American Indian/Alaska Native	221	621.3	38.3
Asian	56	659.4	51.6
African American	98	628.3	38.1
Native Hawaiian or other/Pacific Islander	11	631.4	57.8
Hispanic/Latino	892	632.6	40.0
White	5353	655.6	45.4
Multiracial	134	646.0	42.8
Unknown	5	627.6	34.7
Free/Reduced Lunch	1830	638.8	40.4
Not Free/Reduced Lunch	4940	655.3	46.5
Special Education	863	616.6	37.0
Not Special Education	5907	655.9	44.5
English Language Learner	149	605.5	27.9
Not English Language Learner	6621	651.9	45.3

SAWS

Table P16. Summary Statistics of SAWS Grade 3 Total Raw Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7315	13.7	4.6
Male	3690	12.7	4.4
Female	3621	14.8	4.5
Unknown	4	8.3	3.6
American Indian/Alaska Native	291	11.4	4.4
Asian	52	15.8	4.8
African American	89	13.0	4.9
Native Hawaiian or other/Pacific Islander	7	12.6	5.1
Hispanic/Latino	1065	12.8	4.3
White	5651	14.0	4.6
Multiracial	144	13.4	4.2
Unknown	16	12.1	3.8
Free/Reduced Lunch	2315	12.9	4.4
Not Free/Reduced Lunch	5000	14.1	4.6
Special Education	1066	10.6	4.1
Not Special Education	6249	14.3	4.4
English Language Learner	367	11.5	4.1
Not English Language Learner	6948	13.9	4.6

Table P17. Summary Statistics of SAWS Grade 5 Total Raw Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	7033	15.0	4.1
Male	3622	14.2	4.0
Female	3400	15.9	4.0
Unknown	11	13.8	3.9
American Indian/Alaska Native	254	12.6	3.8
Asian	75	16.2	4.3
African American	69	14.2	4.0
Native Hawaiian or other/Pacific Islander	6	13.0	3.8
Hispanic/Latino	925	14.1	3.9
White	5556	15.3	4.1
Multiracial	131	14.8	4.0
Unknown	17	13.5	4.8
Free/Reduced Lunch	2116	14.2	4.0
Not Free/Reduced Lunch	4917	15.4	4.1
Special Education	1029	12.1	3.8
Not Special Education	6004	15.5	3.9
English Language Learner	148	11.7	3.5
Not English Language Learner	6885	15.1	4.1

Table P18. Summary Statistics of SAWS Grade 7 Total Raw Score

Group	<i>N</i>	Mean	<i>SD</i>
Total	6763	15.7	4.3
Male	3474	14.4	4.1
Female	3276	17.1	4.0
Unknown	13	13.5	6.1
American Indian/Alaska Native	231	13.1	4.0
Asian	57	17.8	4.3
African American	81	13.9	4.6
Native Hawaiian or other/Pacific Islander	14	16.4	4.4
Hispanic/Latino	913	14.9	4.0
White	5318	16.0	4.3
Multiracial	126	15.6	4.1
Unknown	23	13.9	5.6
Free/Reduced Lunch	1965	14.9	4.1
Not Free/Reduced Lunch	4798	16.0	4.3
Special Education	834	11.6	3.6
Not Special Education	5929	16.3	4.1
English Language Learner	162	12.7	3.4
Not English Language Learner	6601	15.8	4.3



Table P19. Gender and Ethnicity Performance by SAWS Prompt and Trait – Grade 3 Prompt 1

Group	N	Percent	<u>Prompt Total</u>		<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	7315	100	6.75	2.63	1.80	0.71	1.68	0.85	1.66	0.74	1.61	0.76
Male	3690	50	6.22	2.55	1.69	0.69	1.54	0.84	1.52	0.72	1.47	0.74
Female	3621	50	7.29	2.61	1.92	0.71	1.82	0.85	1.80	0.73	1.75	0.76
Unknown	4	0	4.25	2.36	1.50	0.58	1.00	0.82	1.00	0.82	0.75	0.50
American Indian/Alaska Native	291	4	5.48	2.64	1.57	0.70	1.27	0.86	1.33	0.76	1.31	0.76
Asian	52	1	7.50	2.97	1.94	0.80	1.85	0.92	1.83	0.86	1.88	0.81
African American	89	1	6.43	2.80	1.80	0.73	1.47	0.97	1.60	0.73	1.56	0.84
Native Hawaiian or other/Pacific Islander	7	0	6.00	3.11	1.71	0.49	1.43	1.27	1.43	0.79	1.43	0.79
Hispanic/Latino	1065	15	6.29	2.49	1.71	0.69	1.54	0.84	1.53	0.71	1.50	0.73
White	5651	77	6.90	2.63	1.83	0.71	1.73	0.85	1.70	0.74	1.64	0.76
Multiracial	144	2	6.65	2.48	1.74	0.72	1.65	0.82	1.65	0.68	1.62	0.69
Unknown	16	0	6.25	2.41	1.69	0.70	1.69	0.70	1.44	0.63	1.44	0.73
Free/Reduced Lunch	2315	32	6.36	2.55	1.72	0.70	1.58	0.84	1.56	0.72	1.50	0.74
Not Free or Reduced Lunch	5000	68	6.93	2.65	1.84	0.72	1.73	0.86	1.70	0.75	1.66	0.77
Special Education	1066	15	5.22	2.47	1.44	0.66	1.32	0.82	1.28	0.70	1.18	0.73
Not Special Education	6249	85	7.01	2.57	1.86	0.70	1.74	0.85	1.72	0.73	1.68	0.74
English Language Learner	367	5	5.59	2.39	1.59	0.66	1.37	0.83	1.33	0.67	1.30	0.73
Not English Language Learner	6948	95	6.81	2.63	1.81	0.71	1.69	0.85	1.68	0.74	1.63	0.76

Table P20. Gender and Ethnicity Performance by SAWS Prompt and Trait – Grade 3 Prompt 2

Group	N	Percent	<u>Prompt Total</u>		<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	7315	100	6.98	2.53	1.87	0.68	1.77	0.71	1.72	0.73	1.63	0.74
Male	3690	50	6.44	2.45	1.74	0.66	1.63	0.69	1.57	0.72	1.50	0.72
Female	3621	50	7.55	2.49	2.00	0.68	1.91	0.70	1.87	0.72	1.77	0.73
Unknown	4	0	4.00	2.94	1.00	0.82	1.00	0.82	1.00	0.82	1.00	0.82
American Indian/Alaska Native	291	4	5.91	2.38	1.66	0.63	1.44	0.68	1.41	0.69	1.40	0.72
Asian	52	1	8.29	2.52	2.15	0.64	2.06	0.70	2.04	0.71	2.04	0.71
African American	89	1	6.56	2.60	1.80	0.68	1.61	0.76	1.61	0.75	1.55	0.77
Native Hawaiian or other/Pacific Islander	7	0	6.57	2.99	1.71	0.76	1.57	0.79	1.71	0.76	1.57	0.79
Hispanic/Latino	1065	15	6.54	2.42	1.76	0.67	1.67	0.70	1.59	0.71	1.52	0.69
White	5651	77	7.13	2.54	1.90	0.68	1.80	0.71	1.76	0.74	1.66	0.74
Multiracial	144	2	6.78	2.29	1.83	0.64	1.69	0.63	1.68	0.71	1.57	0.71
Unknown	16	0	5.88	2.58	1.56	0.73	1.56	0.73	1.44	0.73	1.31	0.70
Free/Reduced Lunch	2315	32	6.58	2.45	1.78	0.66	1.67	0.70	1.61	0.72	1.53	0.71
Not Free or Reduced Lunch	5000	68	7.17	2.55	1.91	0.69	1.81	0.71	1.77	0.74	1.68	0.75
Special Education	1066	15	5.38	2.25	1.50	0.61	1.38	0.67	1.28	0.65	1.21	0.67
Not Special Education	6249	85	7.26	2.48	1.93	0.67	1.83	0.70	1.79	0.72	1.70	0.72
English Language Learner	367	5	5.92	2.31	1.62	0.64	1.48	0.69	1.43	0.68	1.38	0.66
Not English Language Learner	6948	95	7.04	2.53	1.88	0.68	1.78	0.71	1.74	0.73	1.64	0.74

Table P21. Gender and Ethnicity Performance by SAWS 12-point Prompt and Trait – Grade 5

Group	N	Percent	<u>Prompt Total</u>		<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	7033	100	7.08	2.49	1.86	0.68	1.76	0.71	1.75	0.71	1.70	0.71
Male	3622	52	6.57	2.39	1.74	0.67	1.63	0.69	1.62	0.69	1.57	0.69
Female	3400	48	7.63	2.49	1.98	0.68	1.91	0.71	1.89	0.71	1.85	0.70
Unknown	11	0	6.55	2.62	1.82	0.87	1.55	0.69	1.73	0.79	1.45	0.52
American Indian/Alaska Native	254	4	5.68	2.25	1.50	0.63	1.37	0.64	1.40	0.64	1.41	0.66
Asian	75	1	7.71	2.73	1.95	0.70	1.87	0.78	1.93	0.70	1.96	0.80
African American	69	1	6.49	2.40	1.72	0.64	1.55	0.65	1.62	0.71	1.59	0.73
Native Hawaiian or other/Pacific Islander	6	0	6.50	1.64	1.83	0.41	1.50	0.55	1.50	0.55	1.67	0.52
Hispanic/Latino	925	13	6.67	2.37	1.77	0.66	1.63	0.69	1.66	0.69	1.60	0.67
White	5556	79	7.21	2.50	1.89	0.68	1.81	0.71	1.78	0.71	1.73	0.71
Multiracial	131	2	7.02	2.33	1.86	0.68	1.76	0.69	1.69	0.68	1.72	0.68
Unknown	17	0	6.71	2.76	1.82	0.88	1.53	0.72	1.82	0.73	1.53	0.62
Free/Reduced Lunch	2116	30	6.64	2.37	1.76	0.66	1.66	0.69	1.63	0.68	1.59	0.67
Not Free or Reduced Lunch	4917	70	7.27	2.52	1.90	0.69	1.81	0.72	1.80	0.71	1.75	0.72
Special Education	1029	15	5.46	2.17	1.50	0.62	1.36	0.66	1.35	0.63	1.24	0.64
Not Special Education	6004	85	7.36	2.44	1.92	0.67	1.83	0.70	1.82	0.70	1.78	0.69
English Language Learner	148	2	5.32	2.01	1.49	0.59	1.31	0.60	1.28	0.58	1.24	0.59
Not English Language Learner	6885	98	7.12	2.49	1.87	0.68	1.77	0.71	1.76	0.71	1.71	0.71

Table P22. Gender and Ethnicity Performance by SAWS 4-point Prompt and Trait – Grade 5

Group	N	Percent	<u>Prompt Total</u>		<u>Response-to-Text</u>		<u>Holistic</u>	
			Mean	SD	Mean	SD	Mean	SD
All	7033	100	3.06	0.99	1.52	0.65	1.54	0.52
Male	3622	52	2.96	1.00	1.50	0.66	1.46	0.52
Female	3400	48	3.17	0.98	1.55	0.64	1.62	0.50
Unknown	11	0	2.82	1.17	1.27	0.79	1.55	0.52
American Indian/Alaska Native	254	4	2.81	0.99	1.41	0.66	1.41	0.52
Asian	75	1	3.20	0.94	1.60	0.62	1.60	0.49
African American	69	1	2.96	0.98	1.52	0.72	1.43	0.50
Native Hawaiian or other/Pacific Islander	6	0	2.17	1.60	1.00	0.89	1.17	0.75
Hispanic/Latino	925	13	2.91	1.03	1.44	0.68	1.48	0.53
White	5556	79	3.10	0.98	1.54	0.64	1.55	0.51
Multiracial	131	2	3.15	1.00	1.56	0.66	1.58	0.50
Unknown	17	0	2.71	1.36	1.24	0.83	1.47	0.62
Free/Reduced Lunch	2116	30	2.92	1.01	1.46	0.68	1.46	0.52
Not Free or Reduced Lunch	4917	70	3.12	0.98	1.55	0.64	1.57	0.51
Special Education	1029	15	2.60	1.03	1.31	0.74	1.29	0.48
Not Special Education	6004	85	3.14	0.97	1.56	0.63	1.58	0.51
English Language Learner	148	2	2.55	1.03	1.28	0.75	1.28	0.48
Not English Language Learner	6885	98	3.07	0.99	1.53	0.65	1.54	0.51

Table P23. Gender and Ethnicity Performance by SAWS 8-point Prompt and Trait – Grade 5

Group	N	Percent	<u>Prompt Total</u>		<u>Response-to-Text</u>		<u>Holistic</u>	
			Mean	SD	Mean	SD	Mean	SD
All	7033	100	4.87	1.57	1.33	0.76	3.55	1.11
Male	3622	52	4.64	1.56	1.32	0.76	3.32	1.09
Female	3400	48	5.13	1.55	1.34	0.76	3.78	1.09
Unknown	11	0	4.45	1.29	1.18	0.75	3.27	0.79
American Indian/Alaska Native	254	4	4.11	1.55	1.03	0.85	3.08	0.99
Asian	75	1	5.25	1.58	1.33	0.83	3.92	1.05
African American	69	1	4.75	1.46	1.32	0.78	3.43	1.04
Native Hawaiian or other/Pacific Islander	6	0	4.33	1.51	1.33	0.82	3.00	0.89
Hispanic/Latino	925	13	4.57	1.55	1.21	0.79	3.35	1.07
White	5556	79	4.96	1.56	1.37	0.74	3.60	1.11
Multiracial	131	2	4.65	1.70	1.18	0.77	3.47	1.20
Unknown	17	0	4.12	1.45	1.18	0.73	2.94	0.97
Free/Reduced Lunch	2116	30	4.60	1.57	1.24	0.78	3.36	1.10
Not Free or Reduced Lunch	4917	70	4.99	1.56	1.37	0.75	3.63	1.11
Special Education	1029	15	4.00	1.56	1.12	0.81	2.88	1.06
Not Special Education	6004	85	5.02	1.53	1.36	0.75	3.66	1.08
English Language Learner	148	2	3.86	1.48	0.99	0.83	2.86	0.99
Not English Language Learner	6885	98	4.90	1.57	1.34	0.76	3.56	1.11

Table P24. Gender and Ethnicity Performance by SAWS Prompt and Trait – Grade 7

Group	N	Percent	<u>Prompt Total</u>		<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	6763	100	7.72	2.52	1.97	0.69	1.95	0.73	1.91	0.70	1.89	0.71
Male	3474	51	7.02	2.41	1.80	0.67	1.76	0.71	1.74	0.68	1.71	0.69
Female	3276	48	8.45	2.41	2.16	0.67	2.14	0.70	2.09	0.67	2.07	0.69
Unknown	13	0	6.77	3.09	1.85	0.80	1.77	0.83	1.69	0.85	1.46	0.97
American Indian/Alaska Native	231	3	6.51	2.35	1.64	0.65	1.60	0.71	1.61	0.67	1.65	0.69
Asian	57	1	8.91	2.50	2.25	0.66	2.32	0.69	2.25	0.71	2.11	0.70
African American	81	1	6.81	2.70	1.80	0.75	1.64	0.75	1.72	0.75	1.65	0.76
Native Hawaiian or other/Pacific Islander	14	0	8.14	2.88	1.93	0.92	2.07	0.92	2.00	0.68	2.14	0.77
Hispanic/Latino	913	13	7.32	2.29	1.89	0.67	1.83	0.70	1.80	0.64	1.80	0.66
White	5318	79	7.84	2.53	2.00	0.69	1.98	0.73	1.94	0.70	1.91	0.72
Multiracial	126	2	7.77	2.34	1.98	0.66	1.95	0.70	1.95	0.67	1.88	0.64
Unknown	23	0	7.30	3.15	1.91	0.79	1.83	0.94	1.83	0.89	1.74	0.86
Free/Reduced Lunch	1965	29	7.31	2.40	1.88	0.67	1.85	0.71	1.80	0.67	1.78	0.69
Not Free or Reduced Lunch	4798	71	7.88	2.54	2.01	0.70	1.99	0.74	1.95	0.70	1.93	0.71
Special Education	834	12	5.56	2.08	1.47	0.63	1.40	0.64	1.38	0.59	1.32	0.62
Not Special Education	5929	88	8.02	2.42	2.05	0.67	2.02	0.71	1.98	0.68	1.97	0.69
English Language Learner	162	2	6.27	2.09	1.62	0.63	1.56	0.63	1.52	0.61	1.56	0.66
Not English Language Learner	6601	98	7.75	2.52	1.98	0.69	1.96	0.73	1.92	0.70	1.89	0.71

Table P25. Gender and Ethnicity Performance by SAWS 4-point Prompt and Trait – Grade 7

Group	N	Percent	<u>Prompt Total</u>		<u>Response-to-Text</u>		<u>Holistic</u>	
			Mean	SD	Mean	SD	Mean	SD
All	6763	100	2.89	1.06	1.36	0.72	1.53	0.51
Male	3474	51	2.68	1.06	1.27	0.73	1.41	0.51
Female	3276	48	3.12	1.02	1.47	0.69	1.65	0.49
Unknown	13	0	2.46	1.05	1.08	0.49	1.38	0.65
American Indian/Alaska Native	231	3	2.29	1.03	0.95	0.74	1.34	0.49
Asian	57	1	3.23	1.04	1.54	0.66	1.68	0.47
African American	81	1	2.51	1.07	1.15	0.74	1.36	0.53
Native Hawaiian or other/Pacific Islander	14	0	3.14	1.03	1.50	0.65	1.64	0.50
Hispanic/Latino	913	13	2.70	1.07	1.22	0.75	1.47	0.51
White	5318	79	2.96	1.05	1.41	0.70	1.55	0.51
Multiracial	126	2	2.76	1.08	1.25	0.75	1.51	0.50
Unknown	23	0	2.43	1.24	1.09	0.73	1.35	0.65
Free/Reduced Lunch	1965	29	2.74	1.07	1.27	0.73	1.47	0.52
Not Free or Reduced Lunch	4798	71	2.95	1.05	1.40	0.71	1.55	0.51
Special Education	834	12	2.17	1.01	0.95	0.76	1.23	0.45
Not Special Education	5929	88	2.99	1.03	1.42	0.69	1.57	0.51
English Language Learner	162	2	2.21	1.01	0.88	0.72	1.33	0.47
Not English Language Learner	6601	98	2.91	1.06	1.38	0.72	1.53	0.51

Table P26. Gender and Ethnicity Performance by SAWS 8-point Prompt and Trait – Grade 7

Group	N	Percent	<u>Prompt Total</u>		<u>Response-to-Text</u>		<u>Holistic</u>	
			Mean	SD	Mean	SD	Mean	SD
All	6763	100	5.11	1.61	1.48	0.66	3.63	1.22
Male	3474	51	4.69	1.56	1.39	0.68	3.30	1.16
Female	3276	48	5.55	1.54	1.57	0.62	3.99	1.17
Unknown	13	0	4.23	2.45	1.23	0.83	3.00	1.73
American Indian/Alaska Native	231	3	4.26	1.57	1.23	0.71	3.03	1.14
Asian	57	1	5.63	1.67	1.56	0.60	4.07	1.32
African American	81	1	4.60	1.72	1.33	0.74	3.27	1.26
Native Hawaiian or other/Pacific Islander	14	0	5.14	1.61	1.50	0.65	3.64	1.15
Hispanic/Latino	913	13	4.89	1.57	1.44	0.68	3.44	1.16
White	5318	79	5.19	1.60	1.50	0.65	3.69	1.22
Multiracial	126	2	5.06	1.46	1.50	0.59	3.56	1.15
Unknown	23	0	4.17	2.29	1.35	0.78	2.83	1.75
Free/Reduced Lunch	1965	29	4.89	1.58	1.42	0.67	3.47	1.17
Not Free or Reduced Lunch	4798	71	5.20	1.61	1.50	0.65	3.70	1.23
Special Education	834	12	3.89	1.47	1.16	0.72	2.73	1.04
Not Special Education	5929	88	5.28	1.55	1.52	0.64	3.76	1.19
English Language Learner	162	2	4.25	1.44	1.22	0.75	3.04	1.00
Not English Language Learner	6601	98	5.13	1.61	1.49	0.66	3.64	1.22



## Appendix Q: Raw Score to Scaled Score Tables

### Reading

Table Q1. Reading Grade 3 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.872	375	63	1
1	-4.161	375	45	1
2	-3.434	402	32	1
3	-2.994	422	27	1
4	-2.672	436	23	1
5	-2.414	447	21	1
6	-2.197	457	20	1
7	-2.008	465	19	1
8	-1.839	472	18	1
9	-1.686	479	17	1
10	-1.545	485	16	1
11	-1.413	491	16	1
12	-1.290	497	15	1
13	-1.172	502	15	1
14	-1.060	507	15	1
15	-0.952	511	14	1
16	-0.849	516	14	1
17	-0.748	520	14	1
18	-0.650	525	14	1
19	-0.554	529	14	1
20	-0.459	533	13	1
21	-0.366	537	13	1
22	-0.275	541	13	1
23	-0.184	545	13	1
24	-0.093	549	13	1
25	-0.003	553	13	2
26	0.087	557	13	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
27	0.178	561	13	2
28	0.269	565	13	2
29	0.361	569	13	2
30	0.454	573	13	2
31	0.549	577	14	2
32	0.645	581	14	2
33	0.744	586	14	2
34	0.845	590	14	3
35	0.950	595	14	3
36	1.058	600	15	3
37	1.171	605	15	3
38	1.289	610	15	3
39	1.414	615	16	3
40	1.547	621	16	3
41	1.689	627	17	3
42	1.843	634	18	3
43	2.012	641	19	4
44	2.202	650	20	4
45	2.420	659	21	4
46	2.678	671	23	4
47	3.001	685	27	4
48	3.441	704	32	4
49	4.169	736	45	4
50	4.879	767	63	4

Table Q2. Reading Grade 4 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.702	400	63	1
1	-3.990	400	45	1
2	-3.259	410	32	1
3	-2.816	430	27	1
4	-2.490	444	24	1
5	-2.228	455	21	1
6	-2.007	465	20	1
7	-1.813	474	19	1
8	-1.640	481	18	1
9	-1.481	488	17	1
10	-1.335	495	17	1
11	-1.198	501	16	1
12	-1.069	506	16	1
13	-0.946	512	15	1
14	-0.828	517	15	1
15	-0.715	522	15	1
16	-0.605	527	14	1
17	-0.499	531	14	1
18	-0.395	536	14	1
19	-0.293	540	14	1
20	-0.192	545	14	1
21	-0.093	549	14	1
22	0.004	553	14	1
23	0.101	558	14	1
24	0.198	562	14	1
25	0.295	566	14	2
26	0.391	570	14	2
27	0.488	575	14	2
28	0.585	579	14	2
29	0.684	583	14	2
30	0.783	588	14	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.884	592	14	2
32	0.987	596	14	2
33	1.092	601	14	2
34	1.200	606	15	3
35	1.311	611	15	3
36	1.426	616	15	3
37	1.545	621	15	3
38	1.670	626	16	3
39	1.801	632	16	3
40	1.941	638	17	3
41	2.089	645	17	3
42	2.250	652	18	3
43	2.427	660	19	4
44	2.623	668	20	4
45	2.848	678	22	4
46	3.113	690	24	4
47	3.443	704	27	4
48	3.890	724	32	4
49	4.624	756	45	4
50	5.338	787	63	4

Table Q3. Reading Grade 5 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.778	425	63	1
1	-4.064	425	45	1
2	-3.328	425	32	1
3	-2.880	427	27	1
4	-2.550	441	24	1
5	-2.285	453	22	1
6	-2.060	463	20	1
7	-1.864	471	19	1
8	-1.689	479	18	1
9	-1.529	486	17	1
10	-1.382	493	17	1
11	-1.244	499	16	1
12	-1.115	504	16	1
13	-0.992	510	15	1
14	-0.875	515	15	1
15	-0.763	520	15	1
16	-0.655	524	14	1
17	-0.551	529	14	1
18	-0.450	533	14	1
19	-0.351	538	14	1
20	-0.254	542	14	1
21	-0.159	546	13	1
22	-0.066	550	13	1
23	0.026	554	13	1
24	0.118	558	13	1
25	0.208	562	13	1
26	0.298	566	13	1
27	0.387	570	13	1
28	0.477	574	13	1
29	0.567	578	13	2
30	0.657	582	13	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.748	586	13	2
32	0.839	590	13	2
33	0.932	594	13	2
34	1.026	598	14	2
35	1.121	602	14	2
36	1.219	607	14	2
37	1.319	611	14	2
38	1.422	616	14	2
39	1.529	620	14	3
40	1.639	625	15	3
41	1.754	630	15	3
42	1.874	635	15	3
43	2.001	641	16	3
44	2.136	647	16	3
45	2.281	653	17	3
46	2.438	660	18	3
47	2.610	668	19	4
48	2.803	676	20	4
49	3.024	686	21	4
50	3.286	697	24	4
51	3.612	712	27	4
52	4.056	731	32	4
53	4.787	763	45	4
54	5.500	795	63	4

Table Q4. Reading Grade 6 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-3.952	450	63	1
1	-3.244	450	45	1
2	-2.519	450	32	1
3	-2.083	462	27	1
4	-1.763	476	23	1
5	-1.509	487	21	1
6	-1.294	496	20	1
7	-1.108	505	18	1
8	-0.941	512	17	1
9	-0.791	518	17	1
10	-0.652	525	16	1
11	-0.523	530	15	1
12	-0.401	536	15	1
13	-0.287	541	15	1
14	-0.178	545	14	1
15	-0.073	550	14	1
16	0.028	554	14	1
17	0.125	559	14	1
18	0.219	563	13	1
19	0.311	567	13	1
20	0.401	571	13	1
21	0.488	575	13	1
22	0.575	578	13	1
23	0.660	582	13	1
24	0.744	586	13	1
25	0.827	589	13	2
26	0.910	593	13	2
27	0.992	597	13	2
28	1.074	600	13	2
29	1.156	604	13	2
30	1.238	608	13	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	1.320	611	13	2
32	1.403	615	13	2
33	1.486	618	13	2
34	1.570	622	13	2
35	1.656	626	13	2
36	1.743	630	13	3
37	1.832	634	13	3
38	1.922	638	13	3
39	2.015	642	13	3
40	2.111	646	14	3
41	2.210	650	14	3
42	2.313	655	14	3
43	2.421	659	15	3
44	2.533	664	15	3
45	2.653	670	15	3
46	2.780	675	16	3
47	2.916	681	17	4
48	3.065	688	17	4
49	3.229	695	18	4
50	3.413	703	19	4
51	3.625	712	21	4
52	3.877	723	23	4
53	4.194	737	26	4
54	4.628	756	32	4
55	5.349	788	45	4
56	6.056	819	63	4



Table Q5. Reading Grade 7 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-3.615	475	63	1
1	-2.907	475	45	1
2	-2.185	475	32	1
3	-1.751	476	26	1
4	-1.434	490	23	1
5	-1.181	501	21	1
6	-0.969	511	19	1
7	-0.785	519	18	1
8	-0.621	526	17	1
9	-0.473	532	17	1
10	-0.336	538	16	1
11	-0.209	544	15	1
12	-0.090	549	15	1
13	0.023	554	15	1
14	0.130	559	14	1
15	0.232	563	14	1
16	0.331	568	14	1
17	0.426	572	13	1
18	0.518	576	13	1
19	0.608	580	13	1
20	0.696	584	13	1
21	0.783	588	13	1
22	0.867	591	13	1
23	0.951	595	13	1
24	1.033	599	13	1
25	1.115	602	13	1
26	1.196	606	12	2
27	1.277	609	12	2
28	1.357	613	12	2
29	1.438	616	12	2
30	1.519	620	12	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	1.600	623	13	2
32	1.681	627	13	2
33	1.764	631	13	2
34	1.847	634	13	2
35	1.931	638	13	2
36	2.017	642	13	3
37	2.105	646	13	3
38	2.195	650	13	3
39	2.287	654	13	3
40	2.382	658	14	3
41	2.480	662	14	3
42	2.582	666	14	3
43	2.688	671	14	3
44	2.800	676	15	3
45	2.919	681	15	3
46	3.045	687	16	3
47	3.181	693	17	4
48	3.328	699	17	4
49	3.491	706	18	4
50	3.674	714	19	4
51	3.885	724	21	4
52	4.137	735	23	4
53	4.452	749	26	4
54	4.885	768	32	4
55	5.605	799	44	4
56	6.312	830	63	4

Table Q6. Reading Grade 8 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.034	500	63	1
1	-3.321	500	45	1
2	-2.587	500	32	1
3	-2.141	500	27	1
4	-1.814	500	24	1
5	-1.550	500	22	1
6	-1.328	500	20	1
7	-1.134	503	19	1
8	-0.960	511	18	1
9	-0.802	518	17	1
10	-0.657	524	16	1
11	-0.521	530	16	1
12	-0.394	536	15	1
13	-0.273	541	15	1
14	-0.158	546	15	1
15	-0.047	551	14	1
16	0.059	556	14	1
17	0.162	560	14	1
18	0.261	565	14	1
19	0.358	569	14	1
20	0.453	573	13	1
21	0.546	577	13	1
22	0.638	581	13	1
23	0.728	585	13	1
24	0.816	589	13	1
25	0.904	593	13	1
26	0.991	597	13	1
27	1.078	600	13	1
28	1.165	604	13	1
29	1.251	608	13	1
30	1.337	612	13	1

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	1.424	616	13	2
32	1.511	619	13	2
33	1.598	623	13	2
34	1.687	627	13	2
35	1.776	631	13	2
36	1.867	635	13	2
37	1.960	639	13	2
38	2.055	643	14	2
39	2.151	648	14	2
40	2.251	652	14	2
41	2.354	656	14	3
42	2.460	661	14	3
43	2.571	666	15	3
44	2.687	671	15	3
45	2.810	677	16	3
46	2.940	682	16	3
47	3.080	688	17	3
48	3.232	695	17	3
49	3.399	702	18	3
50	3.586	711	20	4
51	3.800	720	21	4
52	4.056	731	23	4
53	4.375	745	27	4
54	4.811	764	32	4
55	5.535	796	45	4
56	6.244	827	63	4

Mathematics

Table Q7. Mathematics Grade 3 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-5.396	375	63	1
1	-4.662	375	45	1
2	-3.888	402	33	1
3	-3.407	423	28	1
4	-3.046	438	25	1
5	-2.753	451	22	1
6	-2.503	462	21	1
7	-2.284	471	20	1
8	-2.087	480	19	1
9	-1.907	488	18	1
10	-1.741	495	17	1
11	-1.586	502	17	1
12	-1.441	508	16	1
13	-1.303	514	16	1
14	-1.171	520	16	1
15	-1.045	525	15	1
16	-0.924	530	15	1
17	-0.807	535	15	1
18	-0.693	540	15	1
19	-0.582	545	14	1
20	-0.474	550	14	2
21	-0.367	554	14	2
22	-0.263	559	14	2
23	-0.159	564	14	2
24	-0.057	568	14	2
25	0.045	572	14	2
26	0.146	577	14	2
27	0.247	581	14	2
28	0.349	586	14	2
29	0.451	590	14	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
30	0.553	594	14	2
31	0.657	599	14	3
32	0.763	604	14	3
33	0.871	608	14	3
34	0.981	613	14	3
35	1.095	618	15	3
36	1.212	623	15	3
37	1.334	628	15	3
38	1.461	634	16	3
39	1.595	640	16	3
40	1.736	646	17	3
41	1.888	652	17	3
42	2.053	660	18	4
43	2.234	667	19	4
44	2.436	676	20	4
45	2.668	686	22	4
46	2.942	698	24	4
47	3.283	713	27	4
48	3.744	733	32	4
49	4.498	766	45	4
50	5.224	797	62	4

Table Q8. Mathematics Grade 4 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-5.009	400	62	1
1	-4.281	400	45	1
2	-3.519	418	33	1
3	-3.048	438	27	1
4	-2.696	453	24	1
5	-2.411	466	22	1
6	-2.167	476	21	1
7	-1.954	486	19	1
8	-1.762	494	19	1
9	-1.586	502	18	1
10	-1.424	509	17	1
11	-1.272	515	17	1
12	-1.129	521	16	1
13	-0.994	527	16	1
14	-0.864	533	15	1
15	-0.740	538	15	1
16	-0.621	543	15	1
17	-0.506	548	15	1
18	-0.394	553	14	1
19	-0.286	558	14	1
20	-0.180	563	14	1
21	-0.077	567	14	1
22	0.024	571	14	1
23	0.123	576	14	1
24	0.221	580	13	1
25	0.317	584	13	2
26	0.412	588	13	2
27	0.506	592	13	2
28	0.599	596	13	2
29	0.692	600	13	2
30	0.784	604	13	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.876	608	13	2
32	0.967	612	13	2
33	1.059	616	13	2
34	1.151	620	13	2
35	1.244	624	13	2
36	1.337	628	13	2
37	1.432	633	13	2
38	1.527	637	13	3
39	1.624	641	14	3
40	1.723	645	14	3
41	1.823	650	14	3
42	1.926	654	14	3
43	2.032	659	14	3
44	2.141	663	14	3
45	2.254	668	15	3
46	2.372	673	15	3
47	2.495	679	15	3
48	2.625	684	16	3
49	2.762	690	16	3
50	2.910	697	17	4
51	3.069	704	18	4
52	3.244	711	19	4
53	3.439	720	20	4
54	3.662	729	21	4
55	3.925	741	23	4
56	4.253	755	27	4
57	4.698	774	32	4
58	5.431	806	44	4
59	6.143	837	62	4





Table Q9. Mathematics Grade 5 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-3.401	425	62	1
1	-2.693	454	44	1
2	-1.971	485	32	1
3	-1.536	504	26	1
4	-1.219	517	23	1
5	-0.967	528	21	1
6	-0.755	538	19	1
7	-0.572	546	18	1
8	-0.409	553	17	1
9	-0.262	559	16	1
10	-0.127	565	16	1
11	-0.001	570	15	1
12	0.116	575	15	1
13	0.226	580	14	1
14	0.331	585	14	1
15	0.431	589	14	1
16	0.527	593	13	1
17	0.619	597	13	1
18	0.708	601	13	1
19	0.795	605	13	1
20	0.879	609	13	2
21	0.962	612	12	2
22	1.043	616	12	2
23	1.122	619	12	2
24	1.200	622	12	2
25	1.277	626	12	2
26	1.354	629	12	2
27	1.429	632	12	2
28	1.505	636	12	2
29	1.580	639	12	2
30	1.654	642	12	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	1.729	645	12	2
32	1.804	649	12	2
33	1.880	652	12	3
34	1.955	655	12	3
35	2.032	659	12	3
36	2.109	662	12	3
37	2.188	665	12	3
38	2.268	669	12	3
39	2.349	672	12	3
40	2.432	676	13	3
41	2.517	680	13	3
42	2.605	683	13	3
43	2.696	687	13	3
44	2.790	692	13	3
45	2.887	696	14	3
46	2.990	700	14	3
47	3.098	705	14	3
48	3.212	710	15	3
49	3.334	715	15	3
50	3.466	721	16	3
51	3.610	727	17	4
52	3.769	734	18	4
53	3.949	742	19	4
54	4.156	751	21	4
55	4.403	762	23	4
56	4.715	775	26	4
57	5.145	794	31	4
58	5.861	825	44	4
59	6.566	855	62	4

Table Q10. Mathematics Grade 6 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-3.005	450	62	1
1	-2.295	471	44	1
2	-1.567	502	32	1
3	-1.127	521	26	1
4	-0.805	535	23	1
5	-0.548	547	21	1
6	-0.333	556	19	1
7	-0.145	564	18	1
8	0.022	571	17	1
9	0.173	578	16	1
10	0.311	584	16	1
11	0.439	589	15	1
12	0.560	595	15	1
13	0.673	600	14	1
14	0.780	604	14	1
15	0.883	609	14	1
16	0.981	613	13	1
17	1.076	617	13	1
18	1.168	621	13	1
19	1.257	625	13	1
20	1.343	629	13	2
21	1.428	632	13	2
22	1.510	636	12	2
23	1.592	640	12	2
24	1.672	643	12	2
25	1.751	646	12	2
26	1.829	650	12	2
27	1.907	653	12	2
28	1.984	657	12	2
29	2.060	660	12	2
30	2.137	663	12	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	2.213	666	12	2
32	2.290	670	12	2
33	2.367	673	12	2
34	2.445	677	12	3
35	2.523	680	12	3
36	2.602	683	12	3
37	2.682	687	12	3
38	2.764	690	12	3
39	2.847	694	13	3
40	2.932	698	13	3
41	3.019	701	13	3
42	3.109	705	13	3
43	3.201	709	13	3
44	3.297	714	14	3
45	3.397	718	14	3
46	3.502	722	14	3
47	3.612	727	15	3
48	3.729	732	15	3
49	3.853	738	16	3
50	3.987	743	16	4
51	4.134	750	17	4
52	4.296	757	18	4
53	4.478	765	19	4
54	4.688	774	21	4
55	4.938	785	23	4
56	5.253	798	26	4
57	5.686	817	32	4
58	6.406	848	44	4
59	7.112	879	62	4

Table Q11. Mathematics Grade 7 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-2.631	475	62	1
1	-1.914	487	44	1
2	-1.175	519	32	1
3	-0.724	539	27	1
4	-0.391	553	24	1
5	-0.124	565	21	1
6	0.103	575	20	1
7	0.300	583	19	1
8	0.477	591	18	1
9	0.637	598	17	1
10	0.785	604	16	1
11	0.922	610	16	1
12	1.050	616	15	1
13	1.172	621	15	1
14	1.288	626	15	1
15	1.398	631	14	1
16	1.504	636	14	1
17	1.606	640	14	1
18	1.705	644	14	1
19	1.800	649	13	1
20	1.894	653	13	2
21	1.985	657	13	2
22	2.074	660	13	2
23	2.161	664	13	2
24	2.248	668	13	2
25	2.332	672	13	2
26	2.416	675	13	2
27	2.500	679	13	2
28	2.582	682	12	2
29	2.664	686	12	2
30	2.746	690	12	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	2.828	693	12	2
32	2.910	697	12	3
33	2.992	700	12	3
34	3.075	704	13	3
35	3.158	707	13	3
36	3.243	711	13	3
37	3.328	715	13	3
38	3.415	719	13	3
39	3.503	722	13	3
40	3.593	726	13	3
41	3.685	730	13	3
42	3.780	734	13	3
43	3.878	739	14	3
44	3.980	743	14	3
45	4.086	748	14	3
46	4.196	753	15	4
47	4.313	758	15	4
48	4.436	763	15	4
49	4.567	769	16	4
50	4.709	775	17	4
51	4.864	782	17	4
52	5.034	789	18	4
53	5.226	797	20	4
54	5.446	807	21	4
55	5.709	818	23	4
56	6.037	832	27	4
57	6.484	852	32	4
58	7.220	884	44	4
59	7.935	915	62	4

Table Q12. Mathematics Grade 8 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-2.221	500	62	1
1	-1.515	505	44	1
2	-0.797	536	31	1
3	-0.366	555	26	1
4	-0.053	568	23	1
5	0.195	579	21	1
6	0.403	588	19	1
7	0.583	596	18	1
8	0.742	603	17	1
9	0.886	609	16	1
10	1.018	615	15	1
11	1.140	620	15	1
12	1.254	625	14	1
13	1.361	629	14	1
14	1.462	634	14	1
15	1.559	638	13	1
16	1.652	642	13	1
17	1.741	646	13	1
18	1.827	650	13	1
19	1.910	653	12	1
20	1.991	657	12	1
21	2.070	660	12	1
22	2.147	664	12	2
23	2.223	667	12	2
24	2.297	670	12	2
25	2.371	673	12	2
26	2.443	676	12	2
27	2.514	680	12	2
28	2.585	683	12	2
29	2.655	686	11	2
30	2.725	689	11	2



Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	2.794	692	11	2
32	2.863	695	11	2
33	2.932	698	11	2
34	3.002	701	11	2
35	3.071	704	11	2
36	3.141	707	11	3
37	3.211	710	12	3
38	3.282	713	12	3
39	3.353	716	12	3
40	3.425	719	12	3
41	3.499	722	12	3
42	3.573	726	12	3
43	3.649	729	12	3
44	3.726	732	12	3
45	3.805	736	12	3
46	3.886	739	12	3
47	3.970	743	13	3
48	4.056	746	13	3
49	4.145	750	13	3
50	4.238	754	13	3
51	4.334	759	14	3
52	4.436	763	14	4
53	4.543	768	14	4
54	4.657	773	15	4
55	4.779	778	15	4
56	4.910	784	16	4
57	5.054	790	17	4
58	5.213	797	18	4
59	5.393	805	19	4
60	5.601	814	21	4
61	5.849	824	23	4
62	6.161	838	26	4

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
63	6.591	857	31	4
64	7.309	888	44	4
65	8.014	918	62	4

Science

Table Q13. Science Grade 4 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.928	400	69	1
1	-4.211	434	49	1
2	-3.472	470	36	1
3	-3.022	492	30	1
4	-2.690	508	26	1
5	-2.424	521	24	1
6	-2.200	531	22	1
7	-2.004	541	21	1
8	-1.829	549	20	1
9	-1.670	557	19	1
10	-1.524	564	18	1
11	-1.387	571	18	1
12	-1.259	577	17	1
13	-1.138	583	17	1
14	-1.022	588	16	1
15	-0.911	594	16	1
16	-0.805	599	16	1
17	-0.701	604	15	1
18	-0.601	609	15	1
19	-0.503	613	15	2
20	-0.407	618	15	2
21	-0.312	622	15	2
22	-0.219	627	15	2
23	-0.127	631	15	2
24	-0.036	636	15	2
25	0.055	640	15	2
26	0.145	644	15	2
27	0.236	649	15	2
28	0.327	653	15	2
29	0.419	658	15	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
30	0.511	662	15	2
31	0.605	667	15	3
32	0.700	671	15	3
33	0.798	676	15	3
34	0.897	681	15	3
35	1.000	686	16	3
36	1.106	691	16	3
37	1.217	696	16	3
38	1.332	702	17	3
39	1.454	708	17	3
40	1.583	714	18	3
41	1.722	721	18	3
42	1.872	728	19	4
43	2.038	736	20	4
44	2.223	745	21	4
45	2.436	755	23	4
46	2.689	767	26	4
47	3.007	782	29	4
48	3.441	803	35	4
49	4.163	838	49	4
50	4.870	872	69	4

Table Q14. Science Grade 8 Raw Score to Scaled Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.772	407	69	1
1	-4.059	442	49	1
2	-3.328	477	35	1
3	-2.884	498	29	1
4	-2.559	514	26	1
5	-2.299	527	23	1
6	-2.079	537	22	1
7	-1.888	546	20	1
8	-1.718	555	19	1
9	-1.563	562	19	1
10	-1.420	569	18	1
11	-1.286	576	17	1
12	-1.161	582	17	1
13	-1.042	587	16	1
14	-0.929	593	16	1
15	-0.819	598	16	1
16	-0.714	603	16	1
17	-0.611	608	15	2
18	-0.511	613	15	2
19	-0.413	618	15	2
20	-0.317	622	15	2
21	-0.221	627	15	2
22	-0.126	631	15	2
23	-0.032	636	15	2
24	0.062	640	15	2
25	0.157	645	15	2
26	0.252	650	15	2
27	0.348	654	15	3
28	0.445	659	15	3
29	0.545	664	15	3
30	0.646	669	15	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.751	674	16	3
32	0.858	679	16	3
33	0.970	684	16	3
34	1.087	690	17	3
35	1.211	696	17	3
36	1.341	702	18	3
37	1.482	709	18	3
38	1.633	716	19	4
39	1.801	724	20	4
40	1.988	733	22	4
41	2.203	744	23	4
42	2.458	756	26	4
43	2.777	771	29	4
44	3.214	792	35	4
45	3.938	827	49	4
46	4.646	861	69	4

SAWS

Table Q15. SAWS Grade 3 Raw Score to Performance Level

Raw Score	Performance Level
0	1
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	2
10	2
11	2
12	2
13	2
14	3
15	3
16	3
17	3
18	3
19	3
20	3
21	4
22	4
23	4
24	4

Table Q16. SAWS Grade 5 Raw Score to Performance Level

Raw Score	Performance Level
0	1
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	2
10	2
11	2
12	2
13	2
14	3
15	3
16	3
17	3
18	3
19	3
20	3
21	4
22	4
23	4
24	4



Table Q17. SAWS Grade 7 Raw Score to Performance Level

Raw Score	Performance Level
0	1
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	2
10	2
11	2
12	2
13	2
14	3
15	3
16	3
17	3
18	3
19	3
20	3
21	4
22	4
23	4
24	4

## Appendix R: Performance Level Percentages by Demographic Subgroup

Table R1. Performance Levels of Reading, Mathematics, and Science by Grade Revised

Grade	Below Basic	Basic	Proficient	Advanced
Reading				
3	16.8	21.4	39.8	22.0
4	14.0	22.3	44.3	19.4
5	13.8	28.2	37.2	20.9
6	16.8	26.3	37.4	19.5
7	16.3	25.0	40.7	18.1
8	16.2	26.1	41.1	16.6
Mathematics				
3	14.2	35.4	37.6	12.8
4	10.4	42.9	36.6	10.0
5	14.9	31.0	43.2	10.8
6	13.8	37.7	38.5	10.1
7	18.4	38.8	32.2	10.6
8	16.5	34.3	37.0	12.2
Science				
4	9.7	37.9	39.9	12.6
8	16.4	36.8	37.6	9.2
SAWS				
3	13.3	36.0	43.3	7.4
5	4.9	31.5	53.6	10.0
7	4.6	26.3	54.4	14.6

Reading

Table R2. Performance Levels of Reading by Grade 3 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	16.8	21.4	39.8	22.0
Male	19.1	22.6	38.7	19.5
Female	14.5	20.1	40.9	24.5
Unknown	25.0	20.0	30.0	25.0
American Indian/Alaska Native	43.7	24.9	24.3	7.1
Asian	9.6	17.3	59.6	13.5
African American	25.0	21.7	34.8	18.5
Native Hawaiian or other/Pacific Islander	12.5	25.0	50.0	12.5
Hispanic/Latino	24.4	28.7	35.0	11.9
White	13.7	19.8	41.5	25.0
Multiracial	19.2	24.0	41.1	15.8
Unknown	28.6	14.3	25.0	32.1
Free/Reduced Lunch	22.6	25.2	37.2	15.0
Not Free/Reduced Lunch	14.1	19.6	41.0	25.3
Special Education	47.7	23.1	21.4	7.8
Not Special Education	11.5	21.1	43.0	24.5
English Language Learner	41.2	29.4	24.8	4.6
Not English Language Learner	15.5	20.9	40.6	22.9

Table R3. Performance Levels of Reading by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	14.0	22.3	44.3	19.4
Male	16.3	23.2	43.0	17.5
Female	11.5	21.4	45.8	21.4
Unknown	33.3	33.3	16.7	16.7
American Indian/Alaska Native	31.8	31.8	31.5	4.9
Asian	8.3	22.2	44.4	25.0
African American	20.0	34.1	35.3	10.6
Native Hawaiian or other/Pacific Islander	16.7	16.7	66.7	0.0
Hispanic/Latino	23.1	30.4	36.6	9.9
White	11.6	20.1	46.6	21.7
Multiracial	9.8	27.1	36.8	26.3
Unknown	18.2	27.3	54.5	0.0
Free/Reduced Lunch	20.3	26.9	40.1	12.8
Not Free/Reduced Lunch	11.2	20.2	46.2	22.4
Special Education	46.2	27.7	20.7	5.4
Not Special Education	8.1	21.3	48.6	21.9
English Language Learner	51.4	30.6	15.7	2.3
Not English Language Learner	12.8	22.0	45.2	19.9

Table R4. Performance Levels of Reading by Grade 5 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	13.8	28.2	37.2	20.9
Male	15.4	27.8	36.6	20.3
Female	12.0	28.5	37.9	21.7
Unknown	20.0	60.0	10.0	10.0
American Indian/Alaska Native	40.3	33.6	20.5	5.6
Asian	13.5	23.0	40.5	23.0
African American	15.7	37.1	27.1	20.0
Native Hawaiian or other/Pacific Islander	42.9	42.9	14.3	0.0
Hispanic/Latino	21.1	37.8	31.0	10.1
White	11.1	26.3	39.2	23.5
Multiracial	18.8	22.6	36.1	22.6
Unknown	16.7	58.3	16.7	8.3
Free/Reduced Lunch	19.3	33.9	33.0	13.8
Not Free/Reduced Lunch	11.3	25.6	39.0	24.0
Special Education	44.7	32.2	18.0	5.0
Not Special Education	8.4	27.5	40.4	23.7
English Language Learner	53.7	36.2	8.1	2.0
Not English Language Learner	12.9	28.0	37.8	21.3

Table R5. Performance Levels of Reading by Grade 6 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	16.8	26.3	37.4	19.5
Male	18.8	25.2	37.4	18.6
Female	14.7	27.4	37.4	20.5
Unknown	20.0	40.0	40.0	0.0
American Indian/Alaska Native	41.6	28.3	25.7	4.5
Asian	7.5	18.9	50.9	22.6
African American	32.4	32.4	26.5	8.8
Native Hawaiian or other/Pacific Islander	10.0	40.0	30.0	20.0
Hispanic/Latino	23.4	33.1	33.3	10.2
White	14.3	24.9	38.8	22.0
Multiracial	17.8	28.9	34.8	18.5
Unknown	16.7	33.3	16.7	33.3
Free/Reduced Lunch	24.5	31.2	33.0	11.3
Not Free/Reduced Lunch	13.6	24.2	39.3	22.9
Special Education	57.5	25.3	14.1	3.0
Not Special Education	10.4	26.4	41.1	22.1
English Language Learner	62.6	25.2	9.7	2.6
Not English Language Learner	15.8	26.3	38.1	19.9

Table R6. Performance Levels of Reading by Grade 7 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	16.3	25.0	40.7	18.1
Male	19.4	26.3	38.9	15.4
Female	12.9	23.5	42.7	20.9
Unknown	75.0	25.0	0.0	0.0
American Indian/Alaska Native	49.8	30.6	16.3	3.3
Asian	13.8	24.1	29.3	32.8
African American	32.5	19.3	37.3	10.8
Native Hawaiian or other/Pacific Islander	26.7	20.0	33.3	20.0
Hispanic/Latino	24.8	31.3	34.7	9.2
White	12.9	23.7	43.0	20.3
Multiracial	17.2	24.2	43.8	14.8
Unknown	66.7	16.7	16.7	0.0
Free/Reduced Lunch	21.6	29.3	38.7	10.5
Not Free/Reduced Lunch	14.0	23.1	41.6	21.3
Special Education	53.1	29.5	14.2	3.1
Not Special Education	11.0	24.3	44.5	20.2
English Language Learner	62.0	31.3	5.5	1.2
Not English Language Learner	15.1	24.8	41.6	18.5

Table R7. Performance Levels of Reading by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	16.2	26.1	41.1	16.6
Male	19.5	26.7	39.5	14.3
Female	12.5	25.4	42.8	19.2
Unknown	66.7	33.3	0.0	0.0
American Indian/Alaska Native	38.3	30.6	24.3	6.8
Asian	5.3	31.6	43.9	19.3
African American	23.5	37.8	33.7	5.1
Native Hawaiian or other/Pacific Islander	27.3	18.2	27.3	27.3
Hispanic/Latino	25.8	33.9	34.2	6.1
White	13.5	24.2	43.2	19.1
Multiracial	22.4	31.3	32.8	13.4
Unknown	40.0	40.0	20.0	0.0
Free/Reduced Lunch	21.8	31.9	36.8	9.4
Not Free/Reduced Lunch	14.1	23.9	42.6	19.3
Special Education	52.9	31.1	13.1	2.9
Not Special Education	10.8	25.4	45.2	18.7
English Language Learner	65.5	24.8	9.7	0.0
Not English Language Learner	15.1	26.1	41.7	17.0

Mathematics

Table R8. Performance Levels of Mathematics by Grade 3 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	14.2	35.4	37.6	12.8
Male	13.1	35.0	37.9	14.0
Female	15.3	35.8	37.3	11.7
Unknown	25.0	50.0	25.0	0.0
American Indian/Alaska Native	34.5	43.0	19.9	2.6
Asian	7.7	36.5	30.8	25.0
African American	33.0	28.7	34.0	4.3
Native Hawaiian or other/Pacific Islander	11.1	55.6	22.2	11.1
Hispanic/Latino	22.8	44.3	28.0	4.9
White	11.1	33.2	40.6	15.1
Multiracial	15.4	41.3	34.3	9.1
Unknown	22.7	36.4	36.4	4.5
Free/Reduced Lunch	18.0	39.9	34.0	8.2
Not Free/Reduced Lunch	12.4	33.3	39.3	15.0
Special Education	33.8	39.4	22.0	4.8
Not Special Education	10.8	34.7	40.3	14.2
English Language Learner	32.6	47.1	17.6	2.6
Not English Language Learner	13.2	34.8	38.7	13.4

Table R9. Performance Levels of Mathematics by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	10.4	42.9	36.6	10.0
Male	11.2	41.7	35.8	11.3
Female	9.6	44.2	37.5	8.7
Unknown	20.0	53.3	26.7	0.0
American Indian/Alaska Native	26.9	50.0	19.8	3.4
Asian	12.2	33.8	36.5	17.6
African American	18.6	48.8	29.1	3.5
Native Hawaiian or other/Pacific Islander	16.7	50.0	33.3	0.0
Hispanic/Latino	17.8	50.9	27.0	4.3
White	8.3	41.1	39.2	11.4
Multiracial	5.3	44.7	39.4	10.6
Unknown	5.9	52.9	41.2	0.0
Free/Reduced Lunch	14.0	48.5	30.9	6.5
Not Free/Reduced Lunch	8.8	40.4	39.2	11.6
Special Education	32.1	46.3	18.3	3.3
Not Special Education	6.5	42.3	39.9	11.2
English Language Learner	34.5	54.7	10.3	0.4
Not English Language Learner	9.7	42.5	37.5	10.3

Table R10. Performance Levels of Mathematics by Grade 5 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	14.9	31.0	43.2	10.8
Male	14.6	29.8	43.8	11.8
Female	15.4	32.4	42.6	9.7
Unknown	8.3	33.3	50.0	8.3
American Indian/Alaska Native	33.7	41.2	22.1	3.0
Asian	9.3	25.3	37.3	28.0
African American	29.6	31.0	29.6	9.9
Native Hawaiian or other/Pacific Islander	42.9	42.9	14.3	0.0
Hispanic/Latino	21.7	37.7	35.7	4.9
White	12.6	29.5	45.9	12.0
Multiracial	18.7	31.3	41.0	9.0
Unknown	41.2	29.4	29.4	0.0
Free/Reduced Lunch	20.2	34.9	38.4	6.4
Not Free/Reduced Lunch	12.7	29.3	45.3	12.7
Special Education	41.3	35.1	20.9	2.7
Not Special Education	10.4	30.3	47.1	12.2
English Language Learner	43.9	43.2	10.3	2.6
Not English Language Learner	14.3	30.8	44.0	11.0

Table R11. Performance Levels of Mathematics by Grade 6 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	13.8	37.7	38.5	10.1
Male	14.7	37.7	37.1	10.5
Female	12.8	37.6	39.9	9.6
Unknown	20.0	60.0	20.0	0.0
American Indian/Alaska Native	32.8	45.9	19.4	1.9
Asian	5.5	29.1	52.7	12.7
African American	18.8	42.0	37.7	1.4
Native Hawaiian or other/Pacific Islander	20.0	30.0	30.0	20.0
Hispanic/Latino	19.0	46.3	29.3	5.3
White	11.8	35.7	41.1	11.4
Multiracial	17.9	43.3	29.9	9.0
Unknown	16.7	66.7	16.7	0.0
Free/Reduced Lunch	19.2	44.3	30.7	5.8
Not Free/Reduced Lunch	11.5	34.9	41.7	11.9
Special Education	45.5	40.6	12.4	1.5
Not Special Education	8.8	37.2	42.6	11.4
English Language Learner	46.7	43.1	9.0	1.2
Not English Language Learner	13.0	37.6	39.2	10.3

Table R12. Performance Levels of Mathematics by Grade 7 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	18.4	38.8	32.2	10.6
Male	18.9	38.5	31.8	10.8
Female	17.8	39.2	32.5	10.5
Unknown	75.0	25.0	0.0	0.0
American Indian/Alaska Native	51.2	37.7	10.2	0.8
Asian	17.2	20.7	31.0	31.0
African American	35.7	35.7	26.2	2.4
Native Hawaiian or other/Pacific Islander	26.7	40.0	33.3	0.0
Hispanic/Latino	28.4	42.0	25.1	4.6
White	14.8	38.5	34.7	12.0
Multiracial	20.5	42.5	25.2	11.8
Unknown	50.0	33.3	0.0	16.7
Free/Reduced Lunch	24.4	42.5	27.4	5.7
Not Free/Reduced Lunch	15.9	37.2	34.2	12.7
Special Education	55.7	33.7	9.0	1.5
Not Special Education	13.2	39.5	35.4	11.9
English Language Learner	58.0	33.3	8.6	0.0
Not English Language Learner	17.4	38.9	32.8	10.9

Table R13. Performance Levels of Mathematics by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	16.5	34.3	37.0	12.2
Male	17.9	33.8	35.3	13.0
Female	14.8	35.0	38.9	11.3
Unknown	66.7	0.0	33.3	0.0
American Indian/Alaska Native	34.7	38.3	24.3	2.7
Asian	8.8	24.6	40.4	26.3
African American	33.7	39.8	22.4	4.1
Native Hawaiian or other/Pacific Islander	36.4	27.3	27.3	9.1
Hispanic/Latino	23.9	41.3	28.8	5.9
White	14.1	33.0	39.1	13.7
Multiracial	17.9	34.3	38.1	9.7
Unknown	40.0	20.0	40.0	0.0
Free/Reduced Lunch	21.8	40.0	31.5	6.7
Not Free/Reduced Lunch	14.5	32.2	39.0	14.2
Special Education	51.8	35.2	11.0	2.0
Not Special Education	11.3	34.2	40.8	13.7
English Language Learner	51.7	39.6	7.4	1.3
Not English Language Learner	15.7	34.2	37.7	12.4



Science

Table R14. Performance Levels of Science by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	9.7	37.9	39.9	12.6
Male	10.7	36.1	40.0	13.2
Female	8.5	39.7	39.8	12.0
Unknown	16.7	66.7	16.7	0.0
American Indian/Alaska Native	27.0	53.2	18.0	1.9
Asian	9.5	35.1	41.9	13.5
African American	18.6	52.3	27.9	1.2
Native Hawaiian or other/Pacific Islander	0.0	66.7	33.3	0.0
Hispanic/Latino	15.9	49.5	30.6	3.9
White	7.6	34.7	42.8	14.9
Multiracial	7.5	39.1	39.1	14.3
Unknown	0.0	56.3	37.5	6.3
Free/Reduced Lunch	13.0	44.7	34.7	7.6
Not Free/Reduced Lunch	8.1	34.7	42.2	14.9
Special Education	21.7	49.3	25.0	4.0
Not Special Education	7.5	35.8	42.6	14.2
English Language Learner	35.3	55.4	8.9	0.4
Not English Language Learner	8.8	37.3	40.9	13.0

Table R15. Performance Levels of Science by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	16.4	36.8	37.6	9.2
Male	17.0	34.6	37.7	10.7
Female	15.6	39.3	37.5	7.5
Unknown	33.3	66.7	0.0	0.0
American Indian/Alaska Native	36.7	41.6	20.4	1.4
Asian	16.1	35.7	28.6	19.6
African American	31.6	42.9	22.4	3.1
Native Hawaiian or other/Pacific Islander	27.3	36.4	27.3	9.1
Hispanic/Latino	25.6	46.0	25.6	2.9
White	13.7	34.9	40.8	10.6
Multiracial	17.2	41.0	35.1	6.7
Unknown	20.0	60.0	20.0	0.0
Free/Reduced Lunch	21.1	43.8	30.9	4.2
Not Free/Reduced Lunch	14.6	34.3	40.1	11.0
Special Education	43.0	40.6	14.6	1.9
Not Special Education	12.5	36.3	41.0	10.3
English Language Learner	47.0	49.7	3.4	0.0
Not English Language Learner	15.7	36.6	38.4	9.4

SAWS

Table R16. Performance Levels of SAWS by Grade 3 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	13.3	36.0	43.3	7.4
Male	18.7	40.0	37.0	4.3
Female	7.7	31.8	49.8	10.6
Unknown	50.0	50.0	0.0	0.0
American Indian/Alaska Native	26.1	42.3	29.6	2.1
Asian	7.7	21.2	57.7	13.5
African American	20.2	32.6	40.4	6.7
Native Hawaiian or other/Pacific Islander	14.3	42.9	42.9	0.0
Hispanic/Latino	16.1	41.5	37.8	4.6
White	12.1	34.7	44.9	8.3
Multiracial	12.5	38.2	46.5	2.8
Unknown	18.8	43.8	37.5	0.0
Free/Reduced Lunch	16.2	40.0	39.6	4.2
Not Free/Reduced Lunch	11.9	34.1	45.1	8.9
Special Education	34.4	43.1	20.8	1.7
Not Special Education	9.7	34.8	47.2	8.4
English Language Learner	24.3	45.0	29.7	1.1
Not English Language Learner	12.7	35.5	44.0	7.8

Table R17. Performance Levels of SAWS by Grade 5 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	4.9	31.5	53.6	10.0
Male	6.5	38.2	49.5	5.9
Female	3.2	24.3	58.1	14.4
Unknown	18.2	27.3	54.5	0.0
American Indian/Alaska Native	12.2	51.2	34.3	2.4
Asian	6.7	16.0	65.3	12.0
African American	4.3	44.9	43.5	7.2
Native Hawaiian or other/Pacific Islander	0.0	50.0	50.0	0.0
Hispanic/Latino	6.1	37.8	50.4	5.7
White	4.3	29.3	55.2	11.1
Multiracial	3.8	38.9	47.3	9.9
Unknown	23.5	29.4	41.2	5.9
Free/Reduced Lunch	6.7	37.5	49.7	6.1
Not Free/Reduced Lunch	4.1	28.8	55.4	11.7
Special Education	15.1	51.9	31.1	1.9
Not Special Education	3.2	27.9	57.5	11.4
English Language Learner	16.2	54.1	29.1	0.7
Not English Language Learner	4.7	31.0	54.2	10.2

Table R18. Performance Levels of SAWS by Grade 7 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	4.6	26.3	54.4	14.6
Male	6.8	35.8	49.2	8.2
Female	2.2	16.3	60.0	21.5
Unknown	23.1	23.1	38.5	15.4
American Indian/Alaska Native	11.7	43.7	41.1	3.5
Asian	3.5	12.3	52.6	31.6
African American	12.3	38.3	39.5	9.9
Native Hawaiian or other/Pacific Islander	7.1	21.4	50.0	21.4
Hispanic/Latino	5.9	30.2	55.4	8.4
White	3.9	25.0	54.9	16.2
Multiracial	4.8	23.0	62.7	9.5
Unknown	13.0	30.4	43.5	13.0
Free/Reduced Lunch	5.6	31.6	53.5	9.3
Not Free/Reduced Lunch	4.2	24.2	54.8	16.8
Special Education	18.0	53.4	27.9	0.7
Not Special Education	2.7	22.5	58.1	16.6
English Language Learner	13.0	48.8	36.4	1.9
Not English Language Learner	4.4	25.8	54.9	15.0

## Appendix S: SAWS Field Test Demographic Performance

Table S1. SAWS Field Test Gender and Ethnicity Performance by 8-point Prompt – Grade 3

Group	Form 1			Form 2			Form 3			Form 4			Form 5			Form 6		
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>
Total Group	1227	5.75	1.72	1218	5.61	1.80	1213	5.36	1.88	1212	5.35	1.77	1229	5.48	1.90	1216	5.34	1.84
Male	613	5.40	1.69	622	5.35	1.77	602	5.11	1.89	597	5.00	1.73	623	5.14	1.93	633	5.02	1.76
Female	613	6.11	1.67	596	5.88	1.80	611	5.60	1.85	614	5.70	1.73	605	5.83	1.81	582	5.69	1.86
Unknown	1	6.00	.							1	4.00	.	1	2.00	.	1	0.00	.
American Indian/Alaska Native	44	5.32	1.88	52	5.04	1.84	39	4.36	2.10	54	5.04	1.77	46	5.00	2.01	56	4.36	1.91
Asian	8	6.75	1.04	10	6.80	1.69	13	5.38	1.71	6	6.33	2.34	5	4.40	1.67	10	6.60	1.35
African American	12	5.50	2.28	21	5.24	2.05	14	4.86	2.32	16	5.63	1.50	14	5.57	2.38	12	5.67	2.06
Haw. Pac. Islander	2	4.00	0.00	1	6.00	.				3	4.67	2.31	1	4.00	.			
Hispanic/Latino	181	5.41	1.80	167	5.46	1.75	179	4.88	1.94	184	5.12	1.66	182	5.18	1.86	172	5.10	1.90
White	954	5.84	1.70	942	5.66	1.80	940	5.49	1.83	918	5.42	1.78	956	5.57	1.90	941	5.43	1.81
Multiracial	23	5.48	1.24	22	5.55	1.63	28	5.64	1.97	27	5.19	1.78	22	5.45	1.77	22	5.36	1.43
Unknown	3	6.00	0.00	3	6.67	1.15				4	4.50	1.00	3	4.00	2.00	3	3.33	3.06
Free Lunch	407	5.51	1.75	367	5.44	1.80	380	5.12	1.84	374	4.95	1.83	410	5.41	1.87	377	5.16	1.73
Not Free Lunch	820	5.88	1.69	851	5.68	1.80	833	5.47	1.90	838	5.53	1.71	819	5.51	1.92	839	5.42	1.88
Special Education	164	5.13	1.70	192	4.69	1.67	165	4.61	1.84	176	4.27	1.70	182	4.41	1.88	187	4.64	1.62
Not Special Education	1063	5.85	1.70	1026	5.78	1.77	1048	5.48	1.86	1036	5.54	1.71	1047	5.66	1.85	1029	5.46	1.85
English Lang. Learner	62	5.13	1.83	59	5.32	1.77	58	4.66	1.92	69	4.90	1.70	51	4.82	1.75	68	4.94	1.78
Not English Lang. Learner	1165	5.79	1.71	1159	5.63	1.80	1155	5.39	1.88	1143	5.38	1.77	1178	5.50	1.91	1148	5.36	1.84

Table S2. SAWS Field Test Gender and Ethnicity Performance by 12-point Prompt – Grade 5

Group	N	Form 1		N	Form 2		N	Form 3	
		Mean	SD		Mean	SD		Mean	SD
Total Group	892	7.64	2.62	880	7.29	2.64	880	7.33	2.55
Male	460	7.04	2.54	450	6.72	2.52	470	6.78	2.37
Female	430	8.29	2.55	430	7.88	2.63	409	7.97	2.60
Unknown	2	6.00	1.41				1	8.00	.
American Indian/Alaska Native	37	6.43	2.38	33	5.94	2.26	37	6.05	2.03
Asian	13	7.31	2.93	5	8.00	2.35	13	8.00	2.77
African American	7	7.00	2.65	10	6.30	2.11	8	5.88	2.64
Haw. Pac. Islander	2	6.00	2.83	1	4.00	.			
Hispanic/Latino	127	7.20	2.45	113	7.18	2.59	111	6.75	2.55
White	685	7.81	2.63	702	7.37	2.66	691	7.52	2.55
Multiracial	17	7.53	2.87	15	8.20	2.40	18	6.50	2.09
Unknown	4	6.00	1.83	1	4.00	.	2	8.00	0.00
Free Lunch	265	7.36	2.69	251	7.08	2.69	279	6.90	2.39
Not Free Lunch	627	7.76	2.58	629	7.37	2.62	601	7.53	2.59
Special Education	136	5.80	2.25	121	5.86	2.20	135	5.41	1.77
Not Special Education	756	7.97	2.54	759	7.51	2.63	745	7.68	2.51
English Lang. Learner	29	6.03	2.04	13	5.85	1.46	21	5.57	2.31
Not English Lang. Learner	863	7.69	2.62	867	7.31	2.65	859	7.37	2.54

Table S3. SAWS Field Test Gender and Ethnicity Performance by 4-point Prompt – Grade 5

Group	N	Form 4		N	Form 5		N	Form 6		N	Form 7		N	Form 8	
		Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD
Total Group	881	3.06	0.96	881	3.22	0.88	871	3.16	0.92	875	2.75	1.02	873	3.03	0.98
Male	459	2.89	0.96	442	3.07	0.88	471	3.03	0.90	447	2.59	0.98	423	2.86	0.98
Female	422	3.25	0.92	436	3.38	0.85	400	3.32	0.92	426	2.92	1.02	447	3.18	0.96
Unknown				3	2.67	0.58				2	2.00	2.83	3	3.33	0.58
American Indian/Alaska Native	35	2.83	1.01	32	2.81	0.97	25	3.08	0.86	29	2.59	1.02	26	2.62	1.02
Asian	9	3.78	0.67	11	3.55	0.69	4	3.50	1.00	9	3.22	0.83	11	3.55	0.52
African American	11	3.27	0.90	11	3.09	0.83	8	2.63	1.19	9	2.67	0.87	5	3.00	1.00
Haw. Pac. Islander							2	2.00	1.41				1	0.00	.
Hispanic/Latino	113	2.90	1.01	100	3.15	0.89	119	3.06	1.04	111	2.51	0.95	131	2.79	1.05
White	693	3.09	0.94	707	3.25	0.88	699	3.19	0.89	702	2.78	1.02	677	3.08	0.95
Multiracial	19	3.21	0.92	17	3.18	0.64	14	2.93	1.00	12	3.17	1.11	19	3.21	1.03
Unknown	1	0.00	.	3	2.67	0.58				3	2.33	2.08	3	3.67	0.58
Free Lunch	250	2.92	1.01	252	3.09	0.92	280	3.06	0.96	283	2.58	0.97	256	2.86	1.04
Not Free Lunch	631	3.12	0.93	629	3.28	0.86	591	3.20	0.90	592	2.83	1.03	617	3.10	0.95
Special Education	121	2.56	0.89	143	2.68	0.97	121	2.64	0.95	141	2.35	0.94	111	2.58	1.03
Not Special Education	760	3.14	0.94	738	3.33	0.82	750	3.24	0.89	734	2.82	1.01	762	3.09	0.96
English Lang. Learner	18	2.72	1.13	12	2.67	1.07	21	2.43	1.12	19	2.37	0.96	15	2.13	1.06
Not English Lang. Learner	863	3.07	0.95	869	3.23	0.87	850	3.18	0.91	856	2.76	1.02	858	3.04	0.97

Table S4. SAWS Field Test Gender and Ethnicity Performance by 8-point Prompt – Grade 5

Group	N	Form 4		N	Form 5		N	Form 6		N	Form 7		N	Form 8	
		Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD
Total Group	881	5.32	1.69	881	5.47	1.74	871	4.82	1.77	875	4.94	1.66	873	5.34	1.48
Male	459	4.94	1.66	442	5.02	1.74	471	4.46	1.74	447	4.63	1.62	423	4.97	1.51
Female	422	5.74	1.64	436	5.93	1.62	400	5.24	1.71	426	5.28	1.61	447	5.70	1.37
Unknown				3	4.33	2.31				2	2.50	3.54	3	5.67	1.15
American Indian/Alaska Native	35	4.71	1.60	32	4.16	1.90	25	4.20	1.87	29	3.76	1.24	26	4.65	1.90
Asian	9	6	2.50	11	6.27	1.49	4	5.75	0.96	9	5.44	1.42	11	6.27	1.01
African American	11	5.45	1.69	11	5.36	1.75	8	5.38	2.07	9	4.78	1.09	5	5.00	1.58
Haw. Pac. Islander							2	5.50	2.12				1	6.00	.
Hispanic/Latino	113	5.29	1.69	100	5.51	1.65	119	4.50	1.81	111	4.48	1.53	131	5.18	1.28
White	693	5.36	1.69	707	5.51	1.73	699	4.89	1.75	702	5.07	1.67	677	5.38	1.51
Multiracial	19	5.26	1.24	17	5.47	1.46	14	4.50	1.91	12	4.92	1.31	19	5.58	1.17
Unknown	1	0	.	3	4.33	2.31				3	3.33	2.89	3	6.33	1.15
Free Lunch	250	4.99	1.78	252	5.18	1.75	280	4.60	1.79	283	4.66	1.56	256	5.04	1.50
Not Free Lunch	631	5.46	1.64	629	5.58	1.73	591	4.92	1.75	592	5.07	1.69	617	5.47	1.46
Special Education	121	4.24	1.66	143	4.40	1.75	121	3.79	1.64	141	4.08	1.53	111	4.41	1.57
Not Special Education	760	5.50	1.64	738	5.67	1.66	750	4.98	1.74	734	5.11	1.63	762	5.48	1.42
English Lang. Learner	18	4.61	1.85	12	4.67	2.19	21	3.67	1.53	19	3.74	1.45	15	4.27	1.53
Not English Lang. Learner	863	5.34	1.69	869	5.48	1.73	850	4.84	1.77	856	4.97	1.65	858	5.36	1.48

Table S5. SAWS Field Test Gender and Ethnicity Performance by 12-point Prompt – Grade 7

Group	N	Form 1		N	Form 2		N	Form 3	
		Mean	SD		Mean	SD		Mean	SD
Total Group	841	7.73	2.66	837	7.93	2.40	839	7.56	2.75
Male	430	7.13	2.52	427	7.27	2.34	421	6.65	2.53
Female	410	8.38	2.66	406	8.64	2.25	414	8.50	2.66
Unknown	1	5.00	.	4	6.50	4.43	4	6.25	3.10
American Indian/Alaska Native	29	5.69	1.97	21	7.81	1.72	25	6.56	2.18
Asian	6	8.67	2.66	3	12.00	0.00	12	7.92	3.60
African American	9	6.89	2.32	11	7.73	3.20	9	6.67	2.55
Haw. Pac. Islander	4	8.50	1.00	3	10.00	3.46			
Hispanic/Latino	94	7.30	2.75	120	7.52	2.43	112	7.23	2.76
White	677	7.89	2.66	663	7.99	2.38	662	7.67	2.76
Multiracial	21	7.57	2.34	13	7.85	2.44	14	6.86	2.57
Unknown	1	4.00	.	3	7.33	5.03	5	7.40	2.19
Free Lunch	239	7.18	2.49	254	7.48	2.32	233	7.03	2.69
Not Free Lunch	602	7.96	2.70	583	8.13	2.41	606	7.76	2.75
Special Education	120	5.68	2.28	97	6.09	2.27	105	5.40	1.93
Not Special Education	721	8.08	2.57	740	8.17	2.32	734	7.87	2.71
English Lang. Learner	20	5.35	1.50	12	6.92	2.54	30	6.27	1.91
Not English Lang. Learner	821	7.79	2.66	825	7.95	2.40	809	7.61	2.77



Table S6. SAWS Field Test Gender and Ethnicity Performance by 4-point Prompt – Grade 7

Group	Form 4			Form 5			Form 6			Form 7			Form 8		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Total Group	853	2.98	1.02	853	3.20	0.90	844	3.14	0.93	841	3.14	1.00	855	3.05	0.99
Male	459	2.79	1.03	439	3.05	0.90	430	2.95	0.98	424	2.98	1.02	444	2.87	1.00
Female	394	3.20	0.97	414	3.37	0.87	412	3.34	0.83	416	3.31	0.94	410	3.25	0.93
Unknown							2	4.00	0.00	1	1.00	.	1	0.00	.
American Indian/Alaska Native	27	2.22	0.85	37	2.86	1.08	28	2.89	0.92	33	2.58	0.94	31	2.65	1.14
Asian	6	3.67	0.82	9	3.44	0.73	7	3.57	0.79	9	3.78	0.67	5	4.00	0.00
African American	10	2.90	1.10	7	3.00	0.82	10	2.70	1.42	9	3.56	0.53	16	3.06	1.00
Haw. Pac. Islander	1	1.00	.	1	4.00	.	3	3.33	1.15	2	1.50	0.71			
Hispanic/Latino	111	2.96	1.04	117	3.02	0.93	111	3.01	0.94	131	3.02	1.02	117	2.85	1.06
White	681	3.00	1.02	668	3.25	0.88	670	3.17	0.92	635	3.20	0.99	662	3.10	0.96
Multiracial	14	3.07	0.83	12	3.50	0.80	11	3.27	0.90	19	3.00	0.94	22	3.14	1.04
Unknown	3	3.00	1.00	2	3.50	0.71	4	3.75	0.50	3	1.67	0.58	2	1.50	2.12
Free Lunch	257	2.91	1.04	234	3.15	0.91	250	3.01	0.94	251	2.97	1.00	247	2.91	1.00
Not Free Lunch	596	3.00	1.02	619	3.22	0.90	594	3.20	0.92	590	3.22	0.99	608	3.11	0.98
Special Education	111	2.23	0.99	105	2.75	1.02	112	2.53	0.98	87	2.26	1.05	97	2.18	0.97
Not Special Education	742	3.09	0.98	748	3.27	0.86	732	3.24	0.89	754	3.24	0.94	758	3.16	0.94
English Lang. Learner	14	2.50	1.22	21	2.71	0.90	23	2.48	0.73	27	2.70	1.07	15	2.07	1.16
Not English Lang. Learner	839	2.98	1.02	832	3.22	0.90	821	3.16	0.93	814	3.16	0.99	840	3.07	0.98

Table S7. SAWS Field Test Gender and Ethnicity Performance by 8-point Prompt – Grade 7

Group	Form 4			Form 5			Form 6			Form 7			Form 8		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Total Group	853	5.26	1.72	853	5.15	1.69	844	5.10	1.82	841	5.46	1.67	855	5.35	1.62
Male	459	5.02	1.72	439	4.78	1.68	430	4.59	1.81	424	5.12	1.67	444	4.97	1.60
Female	394	5.55	1.68	414	5.53	1.62	412	5.62	1.69	416	5.81	1.59	410	5.79	1.51
Unknown							2	7.00	0.00	1	3.00	.	1	0.00	.
American Indian/Alaska Native	27	4.56	1.25	37	4.05	1.67	28	4.14	1.46	33	4.45	1.60	31	4.48	1.29
Asian	6	7.00	1.26	9	5.67	1.41	7	6.57	1.40	9	6.22	1.72	5	6.80	1.30
African American	10	4.10	1.79	7	4.43	1.51	10	4.20	2.25	9	5.89	1.54	16	4.44	2.10
Haw. Pac. Islander	1	2.00	.	1	5.00	.	3	5.00	1.00	2	4.00	2.83			
Hispanic/Latino	111	5.30	1.62	117	4.85	1.63	111	4.86	1.94	131	5.40	1.68	117	5.13	1.74
White	681	5.28	1.74	668	5.26	1.68	670	5.19	1.79	635	5.53	1.64	662	5.47	1.55
Multiracial	14	5.93	1.69	12	5.25	1.86	11	4.00	2.05	19	5.47	1.47	22	5.05	2.13
Unknown	3	6.33	1.53	2	5.50	0.71	4	6.75	0.50	3	2.00	1.00	2	2.00	2.83
Free Lunch	257	5.04	1.68	234	4.90	1.63	250	4.74	1.83	251	5.30	1.58	247	5.28	1.64
Not Free Lunch	596	5.36	1.73	619	5.24	1.71	594	5.26	1.80	590	5.52	1.70	608	5.38	1.62
Special Education	111	4.06	1.59	105	3.92	1.60	112	3.72	1.56	87	4.13	1.80	97	3.86	1.56
Not Special Education	742	5.44	1.67	748	5.32	1.63	732	5.31	1.77	754	5.61	1.58	758	5.54	1.53
English Lang. Learner	14	4.21	1.31	21	4.43	1.43	23	4.17	1.56	27	4.44	1.80	15	4.53	1.81
Not English Lang. Learner	839	5.28	1.72	832	5.17	1.69	821	5.13	1.82	814	5.49	1.65	840	5.37	1.62