

Technical Report
Proficiency Assessment for Wyoming Students (PAWS) and Student
Assessment of Writings Skills (SAWS)

Reading and Mathematics: Grades 3–8,

Science: Grades 4 and 8, and

Writing Grades 3–8

2012 - 2013 Administration

Submitted to Wyoming Department of Education

Prepared by

Educational Testing Service

October 23, 2013

FINAL

Copyright © 2013 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

Richard Crandall

Director of Wyoming Department of Education

Wyoming Department of Education Assessment Team

Deb Lindsay

Assessment Division Director

Stephen Marsh

PAWS Assessment Coordinator

Tamsin Schroeder

Test Development Program Manager

Educational Testing Service

Lauren Hebert

Wyoming Program Manager

Rebecca Rabin-Reed

SAWS Program Manager

Valerie Link

Senior Psychometrician

Junhui Liu

Associate Psychometrician

Mark Batleman

Principal Statistical Assoc - Level 1

William Bennett

Statistical Assistant-Senior

Table of Contents

Table of Tables	xii
Table of Figures	xix
1. OVERVIEW OF THE 2013 PAWS	1
1.1 Introduction.....	1
1.2 Background of PAWS and SAWS.....	2
1.3 Overview of PAWS and SAWS Test Components	4
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. PAWS TEST DESIGN AND DEVELOPMENT	1
2.1 Overview.....	1
2.2 Test Design and Blueprints.....	1
2.2.1 Purpose.....	1
2.2.2 Plan	2
2.3 Types of Items Used in PAWS and SAWS	2
2.3.1 PAWS Reading Tests.....	3
2.3.2. PAWS Mathematics Tests	9
2.3.3. PAWS Science Tests.....	16
2.3.4. SAWS Tests	20
2.3.4.1. Personal Narrative.....	20
2.3.4.2. Fictional Narrative	20
2.3.4.3. Expository	21
2.4 PAWS Test Development Process.....	21
2.5 Item and Test Form Development	23
2.5.1. Item Specifications.....	23
2.5.2. Item Difficulty Requirements	24
2.5.3. Item Graphics Requirements.....	24
2.6 Item Review	24

2.7 Field Testing	26
2.7.1 Classical Item Statistics	26
2.7.1.1. Item Difficulty	27
2.7.1.2. Item Discrimination	27
2.7.2. Differential Item Functioning	27
2.7.3. Item Response Theory (IRT) Analysis	29
2.7.4. Data Review Procedures	30
2.8 Test Form Construction	31
2.8.1. Construction of the Reading, Mathematics, and Science Forms	31
2.8.2. Construction of the SAWS Forms	32
2.8.3. Final Review of Assembled Operational Tests.....	33
3. TEST ADMINISTRATION	34
3.1 Test Materials.....	34
3.2 Materials Return.....	34
3.3 Directions for Administering and Training.....	34
3.3.1 Allowed Student Manipulatives.....	35
3.3.2. Test Security	35
3.4 Student Participation.....	36
3.5 PAWS and SAWS Standard Accommodations	36
3.5.1. Students Eligible for Test Accommodations	38
3.5.2. Requirements for Use of Test Accommodations	38
3.5.3. Description of Standard Accommodations for Students with Disabilities	39
3.5.3.1. Presentation Accommodations.....	40
3.5.3.2. Response Accommodations	41
3.5.3.3. Setting Accommodations	41
3.5.3.4. Timing and Scheduling Accommodations.....	42
3.5.4. Description of Standard Accommodations for English Language Learners (ELL) ...	42
3.5.4.1. Presentation Accommodations.....	42
3.5.4.2. Setting Accommodations	43
3.5.4.3. Timing and Scheduling Accommodations.....	43

3.5.5. PAWS 2013 Monitoring of Appropriate Accommodations	43
3.5.5.1. Empirical Analysis of Accommodations	43
3.5.6. Selection and Administration of Accommodations	44
4. PROCESSING AND SCORING OF PAWS and SAWS ITEMS	46
4.1 Overview	46
4.1.1 Multiple Choice Items.....	46
4.1.2 SAWS Prompts	46
4.2 Receipt Control, Processing, Scanning, Editing	46
4.2.1 Receipt Control	47
4.2.2 Processing	47
4.2.3 Scanning.....	47
4.2.4 Editing.....	48
4.3 Scorer Recruitment and Qualification	49
4.3.1 Eligibility	49
4.3.2 Rater Qualifications	49
4.3.2.1. Training.....	49
4.3.2.2 Certification Procedures.....	50
4.4 Rangefinding.....	50
4.4.1. Pre-Rangefinding Procedures	50
4.4.2 Rangefinding Procedures	50
4.5: Hand-Scoring Process.....	51
4.5.1. Scoring Responses to SAWS Writing Prompts	51
4.5.2 About SAWS Scoring.....	52
4.5.3 Reference Materials for Scoring	52
4.5.4. Guidance to Scorers: Points to Remember about SAWS Scoring.....	52
4.6: Procedures for Maintaining and Retrieving Individual Scores.....	53
4.7 Inter-Rater Reliability	53
4.8 Accuracy Monitoring	54
4.9 Blanks and Invalid Responses for SAWS.....	54
4.9.1 Blank (BL)	55

4.9.2 Copy of Prompt (CP)	55
4.9.3 Foreign Language (FL)	55
4.9.4 Illegible (IL).....	55
4.9.5 Incomprehensible (IN).....	56
4.9.6 Off-topic (OT).....	56
4.9.7 Refusal (RF).....	56
4.10 Reporting of PSC Alerts	56
4.10.1 Policy on the Reporting of Alerts	56
4.10.2 Reporting Procedure	57
5. EQUATING, AND SCALING PROCEDURES	58
5.1 Overview.....	58
5.2 Item and Forms Development.....	58
5.3 IRT Models and Calibrations.....	59
5.4 Fit Statistics for the Rasch Model.....	60
5.5 Equating Analyses	61
5.5.1 Calibration and Equating Process for the 2013 administration	62
5.6 Translating Raw Scores to Scaled Scores and Performance Levels	65
6. PAWS and SAWS REPORTING	66
6.1 Overview.....	66
6.2 Scale Scores	66
6.3 Performance Levels	67
6.4 Content Standard-Level Raw and Scale Scores.....	68
6.5 SAWS Raw Scores	69
6.6 Production of Printed Score Reports for PAWS and SAWS.....	70
6.7 Assessment Score Reports: Supplement Guide for Districts and Schools for PAWS and SAWS	70
7. RELIABILITY	71
7.1 Overview.....	71
7.2 Internal Consistency Reliability.....	71
7.3 Inter-Rater Reliability	72

7.4 Weighted Kappa.....	73
7.5 Classical and Conditional Standard Errors of Measurement	75
7.6 Accuracy and Consistency of Classifications	76
8. VALIDITY	81
8.1 Overview.....	81
8.1.1. Purpose of the PAWs and SAWS	81
8.1.2. The Constructs to Be Measured.....	82
8.1.3. The Interpretations and Uses of the Scores Generated	82
8.1.4. Intended Test Population(s)	83
8.2 Evidence Based on Content-related Validity.....	83
8.2.1. Description of the state standards	83
8.2.2. Specifications and Blueprints	84
8.2.3. Item development process.....	84
8.2.4. Item review process	84
8.2.5. Form construction process	84
8.2.6. Alignment study.....	84
8.3 Evidence Based on Relations to Other Variables	84
8.3.1. Correlations between Content Areas	85
8.3.2. Differential Item Functioning Analyses.....	85
8.4 Evidence Based on Response Processes	85
8.4.1. Evidence of Inter-rater Agreement	86
8.4.1.1. Inter-rater Reliability	86
8.4.1.2. Inter-rater Agreement.....	86
8.5 Evidence Based on Internal Structure.....	86
8.5.1. Classical Statistics.....	87
8.5.2. Reliability.....	87
8.5.3. Dimensionality	87
8.6 Evidence Based on Consequences of Testing.....	88
9. QUALITY CONTROL PROCEDURES	89
9.1 Quality Control of Item Development	89

9.1.1. Item and Prompt Specifications	89
9.1.2. Item Writers	89
9.1.3. Internal Contractor Reviews	89
9.1.4. Content Review.....	89
9.1.5. Editorial Review	90
9.1.6. Fairness Review	90
9.1.7. Assessment Director Review	90
9.1.8. Data Review of Field Tested items.....	90
9.1.9. Quality Control of the Item Bank	90
9.2 Quality Control of Test Materials	91
9.2.1. Collecting Test Materials.....	91
9.2.2. Processing Test Materials	91
9.3 Quality Control of Scanning.....	92
9.4 Quality Control of Psychometric Analyses.....	92
9.5 Quality Control of Reporting.....	92
9.5.1. Excluding Student Scores from Summary Reports	93
10. HISTORICAL COMPARISONS	94
11. REFERENCES	105
12. GLOSSARY OF TERMS	109
Appendix A: Classical Item Statistics for 2013 Field Test Items.....	115
Reading	115
Mathematics.....	139
Science	163
SAWS	171
Appendix B: DIF Results for Field Test 2013 Items	188
Reading	188
Mathematics.....	191
Science	194
Appendix C: Rasch Difficulty, Standard Error, Fit Statistics, and N-Counts for 2013 Field Test Items.....	195

Reading	195
Mathematics	219
Science	243
Appendix D: Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status	251
Reading	251
Mathematics	251
Science	252
SAWS	252
Appendix E: Frequency of Individual Accommodations for 2013 PAWS Tests	253
Reading	253
Mathematics	256
Science	259
Appendix F: Rasch Difficulty, Standard Error, and Fit Statistics for 2013 Operational Items ..	262
Reading	262
Mathematics	274
Science	286
Appendix G: Classical Item Statistics for 2013 Operational Items	290
Reading	290
Mathematics	302
Science	314
SAWS	318
Appendix H: Scale Score Descriptive Statistics by Demographic Subgroup	322
Appendix I: Raw Score to Scale Score Tables	336
Appendix J: Sample PAWS Reports	370
Appendix K: Sample SAWS Reports	375
Appendix L: Performance Level Percentages by Demographic Subgroup	378
Appendix M: SAWS Field Test Demographic Performance	386
Appendix N: SAWS Field Test Rater Reliability	396
Appendix O: SAWS Operational Rater Reliability	403

Appendix P: SAWS Field Test Subscale Correlations	404
Appendix Q: SAWS Operational Subscale Correlations.....	410

Table of Tables

Table 1. Statewide Participation in Reading PAWS.....	1
Table 2. Statewide Participation in Mathematics PAWS	2
Table 3. Statewide Participation in Science PAWS.....	3
Table 4. Statewide Participation in SAWS	4
Table 5. Percentages of Test Items Assessing each Skill-Reporting Category for 2013 PAWS Reading	5
Table 6. Paws 2013 Reading Test Blueprints Grade 3	5
Table 7. Paws 2013 Reading Test Blueprints Grade 4	6
Table 8. Paws 2013 Reading Test Blueprints Grade 5	6
Table 9. Paws 2013 Reading Test Blueprints Grade 6	7
Table 10. Paws 2013 Reading Test Blueprints Grade 7	7
Table 11. Paws 2013 Reading Test Blueprints Grade 8	8
Table 12. Percentages of Test Items Assessing each Skill-Reporting Category for 2013 PAWS Mathematics.....	10
Table 13. Paws 2013 Mathematics Test Blueprints Grade 3	10
Table 14. Paws 2013 Mathematics Test Blueprints Grade 4	11
Table 15. Paws 2013 Mathematics Test Blueprints Grade 5	12
Table 16. Paws 2013 Mathematics Test Blueprints Grade 6.....	13
Table 17. Paws 2013 Mathematics Test Blueprints Grade 7	14
Table 18. Paws 2013 Mathematics Test Blueprints Grade 8.....	15
Table 19. Percentages of Test Items Assessing each Skill-Reporting Category for 2013 PAWS Science	18
Table 20. Paws 2013 Science Test Blueprints Grade 4	18
Table 21. Paws 2013 Science Test Blueprints Grade 8	19
Table 22. SAWS Grades and Prompt Type	21
Table 23. DIF Categories for Multiple-Choice Items.....	29
Table 24. Criteria to Evaluate Mean-Square Fit Statistics.....	61
Table 25. PAWS Mathematics, Science, and Reading Scaling Constants, Lowest Obtainable Scale Scores, and Highest Obtainable Scale Scores	66
Table 26. Proficiency Level Ranges for Grades 3–8 Reading.....	67
Table 27. Proficiency Level Ranges for Grades 3–8 Mathematics	67
Table 28. Proficiency Level Ranges for Grades 4 and 8 Science.....	68
Table 29. 2013 State Level Normative Percentage Distributions for SAWS	69
Table 30. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade.....	72
Table 31. PAWS 2013 Overall Inter-Rater Reliability for SAWS Prompts and Traits.....	74

Table 32. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade.....	76
Table 33. PAWS 2013 Decision Accuracy and Consistency Indices–Grade 3	79
Table 34. PAWS 2013 Decision Accuracy and Consistency Indices – Grade 4	79
Table 35. PAWS 2013 Decision Accuracy and Consistency Indices – Grade 5	79
Table 36. PAWS 2013 Decision Accuracy and Consistency Indices – Grade 6	80
Table 37. PAWS 2013 Decision Accuracy and Consistency Indices – Grade 7	80
Table 38. PAWS 2013 Decision Accuracy and Consistency Indices – Grade 8	80
Table 39. Scale Scores Descriptive Statistics for the PAWS Reading Tests.....	95
Table 40. Percentage Proficient and Advanced for the PAWS Reading Tests.....	95
Table 41. Scale Scores Descriptive Statistics for the Mathematics Tests.....	96
Table 42. Percentage of Proficient + Advanced Students for the Mathematics Tests.....	96
Table 43. Scale Scores Descriptive Statistics for the Science Tests.....	97
Table 44. Percentage of Proficient + Advanced Students for the Science Tests	97
Table A1. Reading Grade 3 Classical Statistics for Field Test Items.....	115
Table A2. Reading Grade 4 Classical Statistics for Field Test Items.....	119
Table A3. Reading Grade 5 Classical Statistics for Field Test Items.....	123
Table A4. Reading Grade 6 Classical Statistics for Field Test Items.....	127
Table A5. Reading Grade 7 Classical Statistics for Field Test Items.....	131
Table A6. Reading Grade 8 Classical Statistics for Field Test Items.....	135
Table A7. Mathematics Grade 3 Classical Statistics for Field Test Items.....	139
Table A8. Mathematics Grade 4 Classical Statistics for Field Test Items.....	143
Table A9. Mathematics Grade 5 Classical Statistics for Field Test Items.....	147
Table A10. Mathematics Grade 6 Classical Statistics for Field Test Items.....	151
Table A11. Mathematics Grade 7 Classical Statistics for Field Test Items.....	155
Table A12. Mathematics Grade 8 Classical Statistics for Field Test Items.....	159
Table A13. Science Grade 4 Classical Statistics for Field Test Items.....	163
Table A14. Science Grade 8 Classical Statistics for Field Test Items.....	167
Table A15. Distributions of SAWS Grade 3 Field Test - 12-point Prompt.....	171
Table A16. Distributions of Rater 1 Scores for SAWS Grade 3 by Trait – 12-point Prompt	172
Table A17. Distributions of Scores for SAWS Grade 5 – 12-point Prompt.....	174
Table A18. Distributions of Scores for SAWS Grade 5 by Trait – 12-point Prompt.....	175
Table A19. Distributions of Scores for SAWS Grade 5 - 4-point Prompt	176
Table A20. Distributions of Scores for SAWS Grade 5 by Trait – 4-point Prompt.....	177
Table A21. Distributions of Rater 1 Scores for SAWS Grade 5 - 8-point Prompt.....	178
Table A22. Distributions of Rater 1 Scores for SAWS Grade 7 - 12-point Prompt.....	181
Table A23. Distributions of Rater 1 Scores for SAWS Grade 7 by Trait – 12-point Prompt	182
Table A24. Distributions of Scores for SAWS Grade 7 - 4-point Prompt	183
Table A25. Distributions of Scores for SAWS Grade 7 - 8-point Prompts.....	185

Table B1. Grade 3 Reading DIF Summary Statistics for Embedded Field Test Items	188
Table B2. Grade 4 Reading DIF Summary Statistics for Embedded Field Test Items	188
Table B3. Grade 5 Reading DIF Summary Statistics for Embedded Field Test Items	189
Table B4. Grade 6 Reading DIF Summary Statistics for Embedded Field Test Items	189
Table B5. Grade 7 Reading DIF Summary Statistics for Embedded Field Test Items	190
Table B6. Grade 8 Reading DIF Summary Statistics for Embedded Field Test Items	190
Table B7. Grade 3 Mathematics DIF Summary Statistics for Embedded Field Test Items	191
Table B8. Grade 4 Mathematics DIF Summary Statistics for Embedded Field Test Items	191
Table B9. Grade 5 Mathematics DIF Summary Statistics for Embedded Field Test Items	192
Table B10. Grade 6 Mathematics DIF Summary Statistics for Embedded Field Test Items	192
Table B11. Grade 7 Mathematics DIF Summary Statistics for Embedded Field Test Items	193
Table B12. Grade 8 Mathematics DIF Summary Statistics for Embedded Field Test Items	193
Table B13. Grade 4 Science DIF Summary Statistics for Embedded Field Test Items	194
Table B14. Grade 8 Science DIF Summary Statistics for Embedded Field Test Items	194
Table C1. Reading Grade 3 IRT Statistics for Field Test Items	195
Table C2. Reading Grade 4 IRT Statistics for Field Test Items	199
Table C3. Reading Grade 5 IRT Statistics for Field Test Items	203
Table C4. Reading Grade 6 IRT Statistics for Field Test Items	207
Table C5. Reading Grade 7 IRT Statistics for Field Test Items	211
Table C6. Reading Grade 8 IRT Statistics for Field Test Items	215
Table C7. Mathematics Grade 3 IRT Statistics for Field Test Items	219
Table C8. Mathematics Grade 4 IRT Statistics for Field Test Items	223
Table C9. Mathematics Grade 5 IRT Statistics for Field Test Items	227
Table C10. Mathematics Grade 6 IRT Statistics for Field Test Items	231
Table C11. Mathematics Grade 7 IRT Statistics for Field Test Items	235
Table C12. Mathematics Grade 8 IRT Statistics for Field Test Items	239
Table C13. Science Grade 4 IRT Statistics for Field Test Items	243
Table C14. Science Grade 8 IRT Statistics for Field Test Items	247
Table D1. Reading Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status	251
Table D2. Mathematics Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status	251
Table D3. Science Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status	252
Table D4. SAWS Mean Raw Scores, Counts, and Raw Score Standard Deviations for IEP Students by Accommodation Status	252
Table E1. Frequency of IEP Student’s Standard Accommodations: Presentation Accommodations	253

Table E2. Frequency of IEP Student’s Standard Accommodations: Response Accommodations	254
Table E3. Frequency of IEP Student’s Standard Accommodations: Setting Accommodations.	254
Table E4. Frequency of IEP Student’s Standard Accommodations: Timing and Scheduling Accommodations	255
Table E5. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations	255
Table E6. Frequency of English Language Learners Standard Accommodations: Setting Accommodations	255
Table E7. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations	255
Table E8. Frequency of IEP Student’s Standard Accommodations: Presentation Accommodations	256
Table E 9.Frequency of IEP Student’s Standard Accommodations: Response Accommodations	257
Table E10. Frequency of IEP Student’s Standard Accommodations: Setting Accommodations.....	258
Table E11. Frequency of IEP Student’s Standard Accommodations: Timing and Scheduling Accommodations	258
Table E12. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations	258
Table E13. Frequency of English Language Learners Standard Accommodations: Setting Accommodations	258
Table E14. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations	259
Table E15. Frequency of IEP Student’s Standard Accommodations: Presentation Accommodations	259
Table E 16.Frequency of IEP Student’s Standard Accommodations: Response Accommodations	260
Table E17. Frequency of IEP Student’s Standard Accommodations: Setting Accommodations.....	260
Table E18. Frequency of IEP Student’s Standard Accommodations: Timing and Scheduling Accommodations	261
Table E19. Frequency of English Language Learners Standard Accommodations: Presentation Accommodations	261
Table E20. Frequency of English Language Learners Standard Accommodations: Setting Accommodations	261
Table E21. Frequency of English Language Learners Standard Accommodations: Timing and Scheduling Accommodations	261
Table F1. Reading Grade 3 IRT Statistics for Operational Items.....	262
Table F2. Reading Grade 4 IRT Statistics for Operational Items.....	264

Table F3. Reading Grade 5 IRT Statistics for Operational Items.....	266
Table F4. Reading Grade 6 IRT Statistics for Operational Item	268
Table F5. Reading Grade 7 IRT Statistics for Operational Items.....	270
Table F6. Reading Grade 8 IRT Statistics for Operational Items.....	272
Table F7. Mathematics Grade 3 IRT Statistics for Operational Items	274
Table F8. Mathematics Grade 4 IRT Statistics for Operational Items	276
Table F9. Mathematics Grade 5 IRT Statistics for Operational Items	278
Table F10. Mathematics Grade 6 IRT Statistics for Operational Items	280
Table F11. Mathematics Grade 7 IRT Statistics for Operational Items	282
Table F12. Mathematics Grade 8 IRT Statistics for Operational Items	284
Table F13. Science Grade 4 IRT Statistics for Operational Items.....	286
Table F14. Science Grade 8 IRT Statistics for Operational Items.....	288
Table G1. Reading Grade 3 Classical Statistics for Operational Items	290
Table G2. Reading Grade 4 Classical Statistics for Operational Items	292
Table G3. Reading Grade 5 Classical Statistics for Operational Items	294
Table G4. Reading Grade 6 Classical Statistics for Operational Items	296
Table G5. Reading Grade 7 Classical Statistics for Operational Items	298
Table G6. Reading Grade 8 Classical Statistics for Operational Items	300
Table G7. Mathematics Grade 3 Classical Statistics for Operational Items.....	302
Table G8. Mathematics Grade 4 Classical Statistics for Operational Items.....	304
Table G9. Mathematics Grade 5 Classical Statistics for Operational Items.....	306
Table G10. Mathematics Grade 6 Classical Statistics for Operational Items.....	308
Table G11. Mathematics Grade 7 Classical Statistics for Operational Items.....	310
Table G12. Mathematics Grade 8 Classical Statistics for Operational Items.....	312
Table G13. Science Grade 4 Classical Statistics for Operational Items	314
Table G14. Science Grade 8 Classical Statistics for Operational Items	316
Table G15. Distributions of Rating 1 Scores for SAWS	318
Table G16. Distributions of Trait Scores for SAWS - Grade 3	319
Table G17. Distributions of Trait Scores for SAWS - Grade 4.....	319
Table G18. Distributions of Trait Scores for SAWS - Grade 5	320
Table G19. Distributions of Trait Scores for SAWS - Grade 6.....	320
Table G20. Distributions of Trait Scores for SAWS - Grade 7	321
Table G21. Distributions of Trait Scores for SAWS - Grade 8.....	321
Table H1. Summary Statistics of Reading, Mathematics, and Science Scale Score by Grade ..	322
Table H2. Summary Statistics of Reading Grade 3 Scale Score	323
Table H3. Summary Statistics of Reading Grade 4 Scale Score	323
Table H4. Summary Statistics of Reading Grade 5 Scale Score	324
Table H5. Summary Statistics of Reading Grade 6 Scale Score	324
Table H6. Summary Statistics of Reading Grade 7 Scale Score	325

Table H7. Summary Statistics of Reading Grade 8 Scale Score	325
Table H8. Summary Statistics of Mathematics Grade 3 Scale Score	326
Table H9. Summary Statistics of Mathematics Grade 4 Scale Score	326
Table H10. Summary Statistics of Mathematics Grade 5 Scale Score	327
Table H11. Summary Statistics of Mathematics Grade 6 Scale Score	327
Table H12. Summary Statistics of Mathematics Grade 7 Scale Score	328
Table H13. Summary Statistics of Mathematics Grade 8 Scale Score	328
Table H14. Summary Statistics of Science Grade 4 Scale Score	329
Table H15. Summary Statistics of Science Grade 8 Scale Score	329
Table H16. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 3.....	330
Table H17. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 4.....	331
Table H18. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 5.....	332
Table H19. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 6.....	333
Table H20. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 7.....	334
Table H21. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 8.....	335
Table I1. Reading Grade 3 Raw Score to Scale Score.....	336
Table I2. Reading Grade 4 Raw Score to Scale Score.....	338
Table I3. Reading Grade 5 Raw Score to Scale Score.....	340
Table I4. Reading Grade 6 Raw Score to Scale Score.....	342
Table I5. Reading Grade 7 Raw Score to Scale Score.....	344
Table I6. Reading Grade 8 Raw Score to Scale Score.....	346
Table I7. Mathematics Grade 3 Raw Score to Scale Score	348
Table I8. Mathematics Grade 4 Raw Score to Scale Score	350
Table I9. Mathematics Grade 5 Raw Score to Scale Score	353
Table I10. Mathematics Grade 6 Raw Score to Scale Score	356
Table I11. Mathematics Grade 7 Raw Score to Scale Score	359
Table I12. Mathematics Grade 8 Raw Score to Scale Score	362
Table I13. Science Grade 4 Raw Score to Scale Score.....	365
Table I14. Science Grade 8 Raw Score to Scale Score.....	367
Table I15. SAWS State Level Normative Percentage Distributions	369
Table L1. Performance Levels of Reading, Mathematics, and Science by Grade.....	378
Table L2. Performance Levels of Reading by Grade 3 Demographic Subgroup	379
Table L3. Performance Levels of Reading by Grade 4 Demographic Subgroup	379
Table L4. Performance Levels of Reading by Grade 5 Demographic Subgroup	380
Table L5. Performance Levels of Reading by Grade 6 Demographic Subgroup	380
Table L6. Performance Levels of Reading by Grade 7 Demographic Subgroup	381
Table L7. Performance Levels of Reading by Grade 8 Demographic Subgroup	381
Table L8. Performance Levels of Mathematics by Grade 3 Demographic Subgroup.....	382
Table L9. Performance Levels of Mathematics by Grade 4 Demographic Subgroup.....	382

Table L10. Performance Levels of Mathematics by Grade 5 Demographic Subgroup.....	383
Table L11. Performance Levels of Mathematics by Grade 6 Demographic Subgroup.....	383
Table L12. Performance Levels of Mathematics by Grade 7 Demographic Subgroup.....	384
Table L13. Performance Levels of Mathematics by Grade 8 Demographic Subgroup.....	384
Table L14. Performance Levels of Science by Grade 4 Demographic Subgroup.....	385
Table L15. Performance Levels of Science by Grade 8 Demographic Subgroup.....	385
Table M1. SAWS Field Test Gender and Ethnicity Frequency Representation – Grade 3.....	386
Table M2. SAWS Field Test Gender and Ethnicity Frequency Representation – Grade 5.....	387
Table M3. SAWS Field Test Gender and Ethnicity Frequency Representation – Grade 7.....	388
Table M4. SAWS Field Test Gender and Ethnicity Performance by 12-point Prompt - Grade 3	389
Table M5. Gender and Ethnicity Performance by 12-point Prompt - Grade 5.....	390
Table M6. Gender and Ethnicity Performance by 4-point Prompt – Grade 5.....	391
Table M7. Gender and Ethnicity Performance by 8-point Prompt – Grade 5.....	392
Table M8. Gender and Ethnicity Performance by 12-point Prompt – Grade 7.....	393
Table M9. Gender and Ethnicity Performance by 4-point Prompt – Grade 7.....	394
Table M10. Gender and Ethnicity Performance by 8-point Prompt – Grade 7.....	395
Table N1. Grade 3 Field Test Trait Rater Agreement and Weighted Kappa – 12-point Prompt	396
Table N2. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa– 12-point prompts	397
Table N3. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 4-point prompts	398
Table N4. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 8-point prompts	399
Table N5. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 12-point prompts	400
Table N6. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 4-point prompts	401
Table N7. Grade 7 Field Test Trait Rater Agreement and Weighted Kappa – 8-point prompts	402
Table O1. SAWS 2013 Overall Inter-Rater Reliability for Prompt and Trait Rater Agreement	403
Table P1. Grade 3 SAWS 2013 Field Test Subscale Correlations – 12-point prompt.....	404
Table P2. Grade 5 SAWS 2013 Field Test Subscale Correlations – 12-point prompt.....	405
Table P3. Grade 5 SAWS 2013 Field Test Subscale Correlations – 4-point prompt.....	406
Table P4. Grade 5 SAWS 2013 Field Test Subscale Correlations – 8-point prompt.....	406
Table P5. Grade 5 SAWS 2013 Field Test Subscale Correlations – 7-point prompt.....	407
Table P6. Grade 7 SAWS 2013 Field Test Subscale Correlations – 4-point prompt.....	408
Table P7. Grade 7 SAWS 2013 Field Test Subscale Correlations – 8-point prompt.....	409
Table Q1. SAWS 2013 Subscale Correlations	410

Table of Figures

Figure 1. Sample item characteristic curve.....	60
Figure 2. Grade 3 SAWS 2013 Normative Information Example for Each Individual Student Report.....	69
Figure 3. Accuracy or Consistency around Critical Cut Point.....	78
Figure 4. Percentage of Proficient and Advanced Students for Grade 3 Reading.....	98
Figure 5. Percentage of Proficient and Advanced Students for Grade 4 Reading.....	98
Figure 6. Percentage of Proficient and Advanced Students for Grade 5 Reading.....	99
Figure 7. Percentage of Proficient and Advanced Students for Grade 6 Reading.....	99
Figure 8. Percentage of Proficient and Advanced Students for Grade 7 Reading.....	100
Figure 9. Percentage of Proficient and Advanced Students for Grade 8 Reading.....	100
Figure 10. Percentage of Proficient and Advanced Students for Grade 3 Mathematics.....	101
Figure 11. Percentage of Proficient and Advanced Students for Grade 4 Mathematics.....	101
Figure 12. Percentage of Proficient and Advanced Students for Grade 5 Mathematics.....	102
Figure 13. Percentage of Proficient and Advanced Students for Grade 6 Mathematics.....	102
Figure 14. Percentage of Proficient and Advanced Students for Grade 7 Mathematics.....	103
Figure 15. Percentage of Proficient and Advanced Students for Grade 8 Mathematics.....	103
Figure 16. Percentage of Proficient and Advanced Students for Grade 4 Science.....	104
Figure 17. Percentage of Proficient and Advanced Students for Grade 8 Science.....	104

1. OVERVIEW OF THE 2013 PAWS

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 2012–2013 school year. Primary purposes of the PAWS and SAWS include supporting individual student proficiency and fostering program improvement at the school, district, and state levels in support of the teaching and learning in Wyoming public classrooms.

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 Reading, Writing, 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, 2012 Wyoming State Enrolled Act 65, removed grade 11 from the 2013 PAWS and SAWS future administrations.

In this Chapter, the policy decisions leading to the PAWS and SAWS are described followed by brief descriptions of the PAWS and SAWS as it was administered during the 2012–2013 school year. Following this overview chapter, technical information is provided in subsequent chapters on the following aspects of PAWS and SAWS:

- Test Design and Development
- Test Administration
- Processing and Scoring
- Linking, Equating, and Scaling Procedures
- Reporting
- Reliability
- Validity
- Statistical Summaries
- Quality Control Procedures
- Historical Comparisons

Finally, the Technical Manual concludes with a list of references, a glossary of terms, and Appendices.

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 statute requires 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. The 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 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 2012–2013 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, 2012 Wyoming State Enrolled Act 65, removed grade 11 from the 2013 SAWS and future administrations.

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 Wyoming Content and Performance¹ Standards in Reading and Mathematics in grades 3 through 8 & 11, Science at grades 4, 8, & 11 and SAWS in grades 3 through 8 & 11. 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 which 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 in Reading, Mathematics; Science and writing.
- 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 2012–2013 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 2013 administration of the PAWS and SAWS included content area assessments in reading, mathematics, science, and writing. For reading, mathematics, and science and writing, each test had two to three sessions. Multiple choice item and writing prompts were administered via pencil and paper in a consumable test booklet (i.e., students marked their responses in the booklet itself; a separate answer sheet was not used).

¹ 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.”

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, but also to inform instructional decision-making by Wyoming classroom teachers to address the specific academic needs of students.

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

The Wyoming Content and Performance Standards are organized by academic content area standards followed by benchmark statements.² Benchmarks are derived from the given content standards and specify skills within that content standard that students are expected to be able to demonstrate at the end of each grade level. Because NCLB makes the state assessment system central to holding schools and districts accountable for student achievement, content standards and benchmarks, while useful in guiding item development, are not immediately useful for interpretation of results because of the wide variety of types of information and varying levels of specificity they encompass.

A student's schema for organizing content knowledge is usually hierarchical with major concepts and principles subsuming more specific facts. To render the *results* of PAWS and SAWS assessments more instructionally useful to Wyoming teachers for addressing the academic needs of students, a delineation of more precise elements of knowledge and skills within each content standard at each grade level was needed. To this end, the Wyoming Assessment Descriptions were developed, drawing on the full range of the Wyoming Content and Performance Standards. 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.³ Thus, in the context of Wyoming Assessment Descriptions, skills:

- Organize the information in the standards into categories of knowledge that are highly related in terms of their use; and
- Lend themselves to a variety of instructional strategies by Wyoming teachers.

Subject matter is too often taught as a series of isolated facts, and students are unable to develop either an accurate schema or, a sense of the discipline (Rutherford & Ahlgren, 1989). While teachers need to break apart big ideas (standards) in order to teach some of the foundational

² For more information regarding Wyoming Content and Performance Standards and Benchmarks, refer to the Wyoming Content and Performance Standards (<http://edu.wyoming.gov/default.aspx>).

³ See Reading Specifications Interpretive Guide.

concepts, they and their students gain a depth of understanding by developing a sense of the organizing framework of the discipline. The concepts and skills can be more readily called upon for later use. Designing assessments that measure integrated concepts and skills is more demanding, but such assessments can better promote student learning of challenging academic content (WDE Assessment Handbook, 2001).

Thus, PAWS and SAWS serve two major purposes. First, it provides information about student attainment of Wyoming Content and Performance Standards in reading, mathematics, science, and writing over time. Second, and equally important, it 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, it 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 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 are 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, but are not exempt from the Mathematics and Science and SAWS tests. The Wyoming ELL assessment measures English language academic readiness.

Tables 1 through 4 provide data on the numbers of students tested in 2013. Additional information can be found on the WDE website: <http://edu.wyoming.gov/default.aspx>.

Table 1. Statewide Participation in Reading PAWS

	<u>Grade 3</u>		<u>Grade 4</u>		<u>Grade 5</u>		<u>Grade 6</u>		<u>Grade 7</u>		<u>Grade 8</u>	
	N	%	N	%	N	%	N	%	N	%	N	%
Total	7141	100.0	7158	100.0	6772	100.0	6800	100.0	6827	100.0	6774	100.0
Male	3703	51.9	3687	51.5	3498	51.7	3483	51.2	3577	52.4	3528	52.1
Female	3428	48.0	3461	48.4	3266	48.2	3313	48.7	3242	47.5	3240	47.8
Unknown	10	0.1	10	0.1	8	0.1	4	0.1	8	0.1	6	0.1
American Indian	274	3.8	278	3.9	264	3.9	241	3.5	221	3.2	219	3.2
Asian	67	0.9	67	0.9	46	0.7	53	0.8	51	0.7	59	0.9
Black	90	1.3	81	1.1	79	1.2	85	1.3	106	1.6	77	1.1
Hawaiian/Pacific Islander	15	0.2	11	0.2	17	0.3	17	0.3	15	0.2	10	0.1
Hispanic	989	13.8	936	13.1	893	13.2	889	13.1	873	12.8	831	12.3
White	5557	77.8	5651	78.9	5358	79.1	5391	79.3	5435	79.6	5469	80.7
Multiracial	128	1.8	110	1.5	90	1.3	103	1.5	117	1.7	94	1.4
Unknown	21	0.3	24	0.3	25	0.4	21	0.3	9	0.1	15	0.2
Free/Reduced Lunch	2669	37.4	2684	37.5	2471	36.5	2430	35.7	2310	33.8	2265	33.4
Not Free/Reduced Lunch	4472	62.6	4474	62.5	4301	63.5	4370	64.3	4517	66.2	4509	66.6
Special Education	1023	14.3	1015	14.2	934	13.8	856	12.6	869	12.7	824	12.2
Not Special Education	6118	85.7	6143	85.8	5838	86.2	5944	87.4	5958	87.3	5950	87.8
English Language Learner	349	4.9	195	2.7	164	2.4	149	2.2	153	2.2	127	1.9
Not English Language Learner	6792	95.1	6963	97.3	6608	97.6	6651	97.8	6674	97.8	6647	98.1

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	7133	100.0	7163	100.0	6766	100.0	6795	100.0	6825	100.0	6779	100.0
Male	3696	51.8	3686	51.5	3494	51.6	3475	51.1	3579	52.4	3532	52.1
Female	3431	48.1	3461	48.3	3265	48.3	3314	48.8	3242	47.5	3241	47.8
Unknown	6	0.1	16	0.2	7	0.1	6	0.1	4	0.1	6	0.1
American Indian	271	3.8	280	3.9	263	3.9	243	3.6	219	3.2	217	3.2
Asian	66	0.9	73	1.0	47	0.7	54	0.8	52	0.8	59	0.9
Black	89	1.2	81	1.1	79	1.2	84	1.2	105	1.5	78	1.2
Hawaiian/Pacific Islander	15	0.2	13	0.2	16	0.2	17	0.3	15	0.2	11	0.2
Hispanic	990	13.9	928	13.0	890	13.2	891	13.1	874	12.8	833	12.3
White	5562	78.0	5661	79.0	5360	79.2	5385	79.2	5429	79.5	5471	80.7
Multiracial	132	1.9	109	1.5	93	1.4	103	1.5	119	1.7	96	1.4
Unknown	8	0.1	18	0.3	18	0.3	18	0.3	12	0.2	14	0.2
Free/Reduced Lunch	2765	38.8	2610	36.4	2489	36.8	2422	35.6	2318	34.0	2285	33.7
Not Free/Reduced Lunch	4368	61.2	4553	63.6	4277	63.2	4373	64.4	4507	66.0	4494	66.3
Special Education	1053	14.8	981	13.7	923	13.6	838	12.3	856	12.5	827	12.2
Not Special Education	6080	85.2	6182	86.3	5843	86.4	5957	87.7	5969	87.5	5952	87.8
English Language Learner	356	5.0	194	2.7	168	2.5	155	2.3	153	2.2	132	1.9
Not English Language Learner	6777	95.0	6969	97.3	6598	97.5	6640	97.7	6672	97.8	6647	98.1

Table 3. Statewide Participation in Science PAWS

	Grade 4		Grade 8	
	N	%	N	%
Total	7157	100.0	6754	100.0
Male	3683	51.5	3523	52.2
Female	3459	48.3	3225	47.7
Unknown	15	0.2	6	0.1
American Indian	275	3.8	213	3.2
Asian	71	1.0	58	0.9
Black	82	1.1	77	1.1
Hawaiian/Pacific Islander	14	0.2	11	0.2
Hispanic	935	13.1	825	12.2
White	5652	79.0	5456	80.8
Multiracial	109	1.5	102	1.5
Unknown	19	0.3	12	0.2
Free/Reduced Lunch	2678	37.4	2315	34.3
Not Free/Reduced Lunch	4479	62.6	4439	65.7
Special Education	1000	14.0	836	12.4
Not Special Education	6157	86.0	5918	87.6
English Language Learner	199	2.8	133	2.0
Not English Language Learner	6958	97.2	6621	98.0

Table 4. Statewide Participation in SAWS

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	N	%	N	%	N	%	N	%	N	%	N	%
Total	7292	100.00	7282	100.00	6909	100.00	6926	100.00	6972	100.00	6914	100.00
Male	3786	51.92	3742	51.39	3575	51.74	3557	51.36	3664	52.55	3593	51.97
Female	3480	47.72	3515	48.27	3316	48.00	3353	48.41	3302	47.36	3300	47.73
Unknown	26	0.36	25	0.34	18	0.26	16	0.23	6	0.09	21	0.30
American Indian	283	3.88	285	3.91	274	3.97	247	3.57	234	3.36	232	3.36
Asian	65	0.89	73	1.00	50	0.72	57	0.82	52	0.75	58	0.84
Black	93	1.28	83	1.14	81	1.17	93	1.34	110	1.58	79	1.14
Hawaiian/Pacific Islander	16	0.22	12	0.17	17	0.25	16	0.23	14	0.20	13	0.19
Hispanic	1011	13.87	952	13.07	908	13.14	912	13.17	893	12.81	838	12.12
White	5646	77.43	5726	78.63	5451	78.90	5453	78.73	5532	79.35	5569	80.55
Multiracial	135	1.85	109	1.50	92	1.33	105	1.52	117	1.68	94	1.36
Unknown	43	0.59	42	0.58	36	0.52	43	0.62	20	0.29	31	0.45
Free/Reduced Lunch	2745	37.64	2649	36.38	2503	36.23	2476	35.75	2344	33.62	2302	33.30
Not Free/Reduced Lunch	4547	62.36	4633	63.62	4406	63.77	4450	64.25	4628	66.38	4612	66.71
Special Education	1041	14.28	991	13.61	945	13.68	872	12.59	895	12.84	839	12.14
Not Special Education	6251	85.72	6291	86.39	5964	86.32	6054	87.41	6077	87.16	6075	87.87
English Language Learner	353	4.84	185	2.54	161	2.33	153	2.21	161	2.31	134	1.94
Not English Language Learner	6939	95.16	7097	97.46	6748	97.67	6773	97.79	6811	97.69	6780	98.06

2. PAWS TEST DESIGN AND DEVELOPMENT

2.1 Overview

The Wyoming PAWS and SAWS statewide assessments adheres to the principles of sound and ethical test construction set forth in the *Standards for Educational and Psychological Testing* (1985, 1999). These assessments complies with the requirements of NCLB (P.L. 107–110).and was designed to provide teachers with information to improve instruction based on the Wyoming Content and Performance Standards.

2.2 Test Design and Blueprints

2.2.1 Purpose

Standards-based educational reform began in Wyoming in 1997–98, with adoption of rigorous academic content standards in language arts,⁴ mathematics, science, and social studies.⁵ Wyoming educators have continued the earlier efforts of others 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.⁶ 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

⁴ As previously noted, Wyoming tests only the Reading Language Arts Standards.

⁵ Social studies is not presently tested in the PAWS assessments.

⁶ Based on the recommendations of the Wyoming Statewide Task Force on Student Assessment and Education Accountability.

step was to empanel content experts from around the state to review and revise Dr. Marzano's work. The major purpose of this exercise was the support of an assessment design which 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. 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.

2.2.2 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

2.3 Types of Items Used in PAWS and SAWS

Consistent with Wyoming State law, (legislation passed in 2012 [Enrolled Act 90, The Wyoming Accountability in Education Act] modified this requirement; beginning with the 2013 administration, the PAWS assessments were composed of solely 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. The SAWS assessment was composed of a single writing prompt.

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.

2.3.1 PAWS Reading Tests

The Wyoming Language Arts Content and Performance Standards include an expectation that all students will become effective readers, 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 2013, a reconfigured writing test will be 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 2013 PAWS reading test was based on the 2013 PAWS reading Blueprint (Tables 5–11). The content of the test is aligned to the reading content standards of the Wyoming Language Arts Content and Performance Standards. Because functional, expository, and narrative types of texts are read for different purposes, 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 types of reading passages, functional, expository, and literary/narrative, were used in the reading assessment. The specific Wyoming reading Benchmarks are organized and assessed within eight skill categories. Below each type of reading passage, the benchmarks and skill categories are indicated as follows.

Functional and expository passages assessed the skills of *Relevance and Importance*, *Major Points and Detail*, and *Information Relationships* at all grade levels, with *Selection and Application* added at Grade 5 and *Organization* added at Grade 7. Narrative passages assessed *Story Elements* and *Plots* at all levels, with *Theme* added at Grade 6.

The approximate percentages of questions on the test at each grade level that assessed each skill reporting category are provided in Table 5. The 2012–2013 PAWS reading blueprints and reporting categories for each of the grade levels are provided in Tables 6 through 11. As noted in the tables below, the percentage of assessment coverage of text type reflects the emphasis of instruction in Wyoming classrooms across grades.

Table 5. Percentages of Test Items Assessing each Skill-Reporting Category for 2013 PAWS Reading

Grade	Functional Text		Expository Text			Literary/Narrative Text		
	Relevance and Importance	Selection and Application	Major Points and Details	Organization	Information Relationships	Story Elements	Plot	Theme
3	18		18-22		10-16	22-26	22-26	
4	18		14-20		14-20	24-26	24-26	
5	11-15	11-15	13-17		11-19	19-22	19-22	
6	11-14	14-18	13-18		13-18	11-14	11-14	14-18
7	11-14	14-18	11-14	11-14	11-14	11-14	11-14	11-14
8	11-14	14-18	11-14	11-14	11-14	11-14	11-14	11-14

Table 6. Paws 2013 Reading Test Blueprints Grade 3

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
One: Functional	Relevance and Importance	9	9	18%
Two: Expository	Major Points and Details	17	9-11	18-22%
	Information Relationships		5-8	10-16%
Three: Narrative	Story Elements	24	11-13	22-26%
	Plot		11-13	22-26%
Total		50		

Table 7. Paws 2013 Reading Test Blueprints Grade 4

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
One: Functional	Relevance and Importance	9	9	18%
Two: Expository	Major Points and Details	17	7-10	14-20%
	Information Relationships		7-10	14-20%
Three: Narrative	Story Elements	24	11-13	24-26%
	Plot		11-13	24-26%
Total		50		

Table 8. Paws 2013 Reading Test Blueprints Grade 5

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
One: Functional	Relevance and Importance	14	6-8	11-15%
	Selection and Application		6-8	11-15%
Two: Expository	Major Points and Details	18	7-9	13-17%
	Information Relationships		6-10	11-19%
Three: Narrative	Story Elements	22	10-12	19-22%
	Plot		10-12	19-22%
Total		54		

Table 9. Paws 2013 Reading Test Blueprints Grade 6

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
One: Functional	Relevance and Importance	16	6-8	11-14%
	Selection and Application		8-10	14-18%
Two: Expository	Major Points and Details	16	7-10	13-18%
	Information Relationships		7-10	13-18%
Three: Narrative	Story Elements	24	6-8	11-14%
	Plot		6-8	11-14%
	Theme		8-10	14-18%
Total		56		

Table 10. Paws 2013 Reading Test Blueprints Grade 7

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
One: Functional	Relevance and Importance	14	6-8	11-14%
	Selection and Application		8-10	14-18%
Two: Expository	Major Points and Details	20	6-8	11-14%
	Organization		6-8	11-14%
	Information Relationships		6-8	11-14%
Three: Narrative	Story Elements	22	6-8	11-14%
	Plot		6-8	11-14%
	Theme		6-8	11-14%
Total		56		

Table 11. Paws 2013 Reading Test Blueprints Grade 8

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
One: Functional	Relevance and Importance	14	6-8	11-14%
	Selection and Application		8-10	14-18%
Two: Expository	Major Points and Details	20	6-8	11-14%
	Organization		6-8	11-14%
	Information Relationships		6-8	11-14%
	Story Elements		6-8	11-14%
Three: Narrative	Plot	22	6-8	11-14%
	Theme		6-8	11-14%
Total		56		

2.3.2. 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 2013 PAWS Mathematics test is explicated in the 2013 PAWS Mathematics Blueprints (Tables 12–18). The content of the test is aligned to the five content standards within the Wyoming Mathematics Content and Performance Standards:

- Number Operations and Concepts
- Geometry
- Measurement
- Algebraic Concepts and Relationships
- Data Analysis and Probability

The benchmarks within each content standard are organized into skill-reporting categories.

- Within the Number Operations and Concepts Content standard are:
 - Number Representation, and Number Operations;
- Within the Geometry Content standard are:
 - Spatial Relationships, 2-D/3-D Shapes, and Transformations/Symmetry;
- Within the Measurement Content standard are:
 - Measurement Systems and Perimeter/Area/Volume;
- Within the Algebraic Concepts and Relationships Content standard are:
 - Patterns/Relations/Functions and Mathematical Representation; and
- Within the Data Analysis and Probability Content standard are:
 - Collect/Analyze Data and Inferences/Predictions.

The approximate percentages of test questions on the test at each grade level that assess each skill reporting category are provided in Table 12. The 2013 PAWS Mathematics blueprints and reporting categories for each of the grade levels are provided in Tables 13 through 18. 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 Algebraic Concepts and Relationships. Calculator use is not permitted for items that are in the Number Operations and Concepts standard for grades 4-8. Calculator use is not permitted at all for the grade 3 assessment.

Table 12. Percentages of Test Items Assessing each Skill-Reporting Category for 2013 PAWS Mathematics

Grade	<u>Number Operations and Concepts</u>		<u>Geometry</u>			<u>Measurement</u>		<u>Algebraic Concepts and Relationships</u>		<u>Data Analysis and Probability</u>	
	Number Representation	Number Operations	Spatial Relationships	2-D/3-D Shapes	Transformations & Symmetry	Measurement Systems	Perimeter, Area & Volume	Patterns, Relations and Functions	Mathematical Representation	Collect & Analyze Data	Inferences & Predictions
3	8	15		13	10	13	7	17		10	7
4	8	14	8	8	8	15	6	18		8	8
5	8	12	8	8	8	15	6	12	8	8	8
6	9	9	9	9	11	11	5	11	11	8	8
7	5	11	12	8	8	9	6	12	15	8	8
8	7	7	17		11	7	7	14	14	7	7

Table 13. Paws 2013 Mathematics Test Blueprints Grade 3

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
Number Operations and Concepts	Number Representation	14	5	8%
	Number Operations		9	15%
Geometry	2-D/3-D Shapes	14	8	13%
	Transformations & Symmetry		6	10%
Measurement	Measurement Systems	12	8	13%
	Perimeter, Area & Volume		4	7%
Algebraic Concepts and Relationships	Patterns, Relations and Functions	10	10	17%
Data Analysis and Probability	Collect & Analyze Data	10	6	10%
	Inferences & Predictions		4	7%
Total		60		

Table 14. Paws 2013 Mathematics Test Blueprints Grade 4

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
Number Operations and Concepts	Number Representation	14	5	8%
	Number Operations		9	14%
Geometry	Spatial Relationships	15	5	8%
	2-D/3-D Shapes		5	8%
	Transformations & Symmetry		5	8%
Measurement	Measurement Systems	14	10	15%
	Perimeter, Area & Volume		4	6%
Algebraic Concepts and Relationships	Patterns, Relations and Functions	12	12	18%
Data Analysis and Probability	Collect & Analyze Data	10	5	8%
	Inferences & Predictions		5	8%
Total		65		

Table 15. Paws 2013 Mathematics Test Blueprints Grade 5

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
Number Operations and Concepts	Number Representation	13	5	8%
	Number Operations		8	12%
Geometry	Spatial Relationships	15	5	8%
	2-D/3-D Shapes		5	8%
	Transformations & Symmetry		5	8%
Measurement	Measurement Systems	14	10	15%
	Perimeter, Area & Volume		4	6%
Algebraic Concepts and Relationships	Patterns, Relations and Functions	13	8	12%
	Mathematical Representation		5	8%
Data Analysis and Probability	Collect & Analyze Data	10	5	8%
	Inferences & Predictions		5	8%
Total		65		

Table 16. Paws 2013 Mathematics Test Blueprints Grade 6

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
Number Operations and Concepts	Number Representation	12	6	9%
	Number Operations		6	9%
Geometry	Spatial Relationships	19	6	9%
	2-D/3-D Shapes		6	9%
	Transformations & Symmetry		7	11%
Measurement	Measurement Systems	10	7	11%
	Perimeter, Area & Volume		3	5%
Algebraic Concepts and Relationships	Patterns, Relations and Functions	14	7	11%
	Mathematical Representation		7	11%
Data Analysis and Probability	Collect & Analyze Data	10	5	8%
	Inferences & Predictions		5	8%
Total		65		

Table 17. Paws 2013 Mathematics Test Blueprints Grade 7

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
Number Operations and Concepts	Number Representation	10	3	5%
	Number Operations		7	11%
Geometry	Spatial Relationships	18	8	12%
	2-D/3-D Shapes		5	8%
	Transformations & Symmetry		5	8%
Measurement	Measurement Systems	10	6	9%
	Perimeter, Area & Volume		4	6%
Algebraic Concepts and Relationships	Patterns, Relations and Functions	18	8	12%
	Mathematical Representation		10	15%
Data Analysis and Probability	Collect & Analyze Data	10	5	8%
	Inferences & Predictions		5	8%
Total		66		

Table 18. Paws 2013 Mathematics Test Blueprints Grade 8

Standard	Skill	Total Points per Standard	Total Points per Skill	Percentage of Test Items per Skill
Number Operations and Concepts	Number Representation	10	5	7%
	Number Operations		5	7%
Geometry	Spatial Relationships	20	12	17%
	Transformations & Symmetry		8	11%
Measurement	Measurement Systems	10	5	7%
	Perimeter, Area & Volume		5	7%
Algebraic Concepts and Relationships	Patterns, Relations and Functions	20	10	14%
	Mathematical Representation		10	14%
Data Analysis and Probability	Collect & Analyze Data	10	5	7%
	Inferences & Predictions		5	7%
Total		70		

2.3.3. 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 takes 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 student 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. 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 2013 PAWS science test was based on the 2012–2013 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 approximate percentages of questions on the test at each grade level that assessed each skill-reporting category are provided in Table 19. The number of items assessing each skill-reporting category and content standard is constant across all grade levels are provided in Tables 20 and 21.

Table 19. Percentages of Test Items Assessing each Skill-Reporting Category for 2013 PAWS Science

Grade	<u>Observe and Question</u>			<u>Design and Conduct a Scientific Investigation</u>			<u>Organize and Represent Data</u>			<u>Draw Conclusions and Make Connections</u>		
	Life Science	Earth & Space Science	Physical Science	Life Science	Earth & Space Science	Physical Science	Life Science	Earth & Space Science	Physical Science	Life Science	Earth & Space Science	Physical Science
4	4	3	4	4	8	6	3	2	4	5	3	4
8	4	2	3	7	4	6	2	4	5	3	6	4

Table 20. Paws 2013 Science Test Blueprints Grade 4

Standard	<u>Skill</u>			Total Points per Standard
	Life Science	Earth & Space Science	Physical Science	
Observe and Question	4	3	4	11
Design and Conduct a Scientific Investigation	4	8	6	18
Organize and Represent Data	3	2	4	9
Draw Conclusions and Make Connections	5	3	4	12
Total Points per Skill	16	16	18	
Percentage of Test Items per Skill	32%	32%	34%	
Total				50

Table 21. Paws 2013 Science Test Blueprints Grade 8

Standard	<u>Skill</u>			Total Points per Standard
	Life Science	Earth & Space Science	Physical Science	
Observe and Question	4	2	3	9
Design and Conduct a Scientific Investigation	7	4	6	17
Organize and Represent Data	2	4	5	11
Draw Conclusions and Make Connections	3	6	4	13
Total Points per Skill	16	16	18	
Percentage of Test Items per Skill	32%	32%	34%	
Total				50

2.3.4. SAWS Tests

The SAWS assessments are open-ended, requiring students to write an essay in response to a single prompt in Table 22. The writing assessment is designed to measure Wyoming Language Arts Content and Performance Standard II, which requires that “students use the writing process (generate ideas, draft, revise, and edit) to demonstrate expressive and expository writing.” To achieve this end, the writing prompt required a single session to administer. The students needed to initially generate their ideas and draft their essays. Students then revised and edited their draft, writing their final essay in the space provided in the test book for that purpose. Planning and drafting was done outside of the test booklet and was not scored.

Prompts were developed to be free of age, gender, geographic, ethnic, socioeconomic, religious, or physical disability stereotypes. They address experiences and interests common to the student’s age level. Although they offer the opportunity 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.

The prompts generally covered one of the following three writing modes:

2.3.4.1. *Personal Narrative*

Each writer recounts a personal experience. The paper has a clear, identifiable storyline that is easy to recognize and follow. All details work together in an integrated way to create a complete story with a beginning, a middle, and an end. There is a focus with a controlling idea, central impression, sense of change, or something learned or gained by the writer. Events move along, staying within that focus, with some sort of a narrative structure, often chronological. (For purposes of statewide assessment, narrative is distinguished from imaginative writing in that narrative prompts focus on real-life experiences, whereas imaginative prompts are meant to be fiction.)

2.3.4.2. *Fictional Narrative*

Each writer describes a situation or story that did not happen and is based on the writer’s imagination. The writer might create a scene, situation, and character(s); might predict what could happen under hypothetical circumstances; or might solve a hypothetical problem using a creative approach. Imaginative writing often, but not always, takes the form of a short story. In some of the most effective imaginative writing, the writer uses his or her knowledge of the world, people, or situations to make the situation or story seem realistic, but, as in all fictional writing, the writer is not bound by the constraints of reality. Strong imaginative writing may contain, as appropriate, elements of fantasy, drama, humor, the unusual, the unexpected, or suspense. Reader reactions often range from a sense of being challenged or intrigued to a sense of feeling delighted or amused.

2.3.4.3. Expository

Expository prompts (includes directions, report, and letters) require a writer to inform, explain, clarify or define. The writing informs or amplifies the reader’s understanding through a carefully crafted presentation of key points, explanations, and supportive detail. The writing contains clear ideas that are focused and fully explained. The writer uses personal knowledge to gather accurate, relevant information that provides a strong base of support in the form of facts, examples, illustrations, incidents, or explanations. Strong writers show a concern for audience and purpose by carefully selecting words, elaborative detail, and stylistic devices; they also recognize that greater stylistic distance may be required in a formal, academic paper than in an informal, personal paper, but that expository writing can be lively, engaging, and indicative of the writer’s commitment to the topic.

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

Table 22. SAWS Grades and Prompt Type

Grade	Assessment	12-point Operational Prompt
3	Writing	Expressive
4	Writing	Expository
5	Writing	Expressive
6	Writing	Expository
7	Writing	Expressive
8	Writing	Expository

2.4 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 is consistent with these documents, and adjustments were made as deemed appropriate by the state committees.⁷

⁷ These documents included the following publications:

- National Council of Teachers of English and International Reading Association,
- Standards for the English Language Arts;

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 (OSA), 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 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 WY educators (see Section 2.6 Item Review). Items approved at item

-
- 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.

review then became eligible for field testing, after which they were evaluated in the light of their statistics from field testing (see Section 2.7.4, 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 2013 forms considering both the content and psychometric properties of each item selected.

2.5 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, Science, or Mathematics content standard. These blueprints were initially developed in the fall of 2004 and have been refined during the course of the program, balancing the need to provide a high level of information about student ability in order 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.

2.5.1. 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.

2.5.2. 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.

2.5.3. 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.

2.6 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, bias, etc.) 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. Items field tested during the current administration were reviewed by this committee during meetings held in July 2012.

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.

2.7 Field Testing

During the 2013 PAWS and SAWS administration, reading, mathematics, science and writing field test items were embedded within the operational forms respectively. In 2013, 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 and will be used for equating future test forms after they are used as operational items.

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

2.7.1 Classical Item Statistics

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

- N-counts for each statistic;

- Item difficulty (or average item score);
- Item discrimination;
- 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.

2.7.1.1. 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 writing prompts expressed as item mean, for traits, this value ranged from 0 to 3, and for the prompt total, this value ranged from 0 to 12.

2.7.1.2. 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.

2.7.2. Differential Item Functioning

In addition to classical item analyses, Differential Item Functioning (DIF) analyses are also 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 2013 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, and White/Native American) where sample size was sufficient.

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

The formula for the estimate of constant odds ratio is:

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

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, 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. 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. No DIF analyses were performed for SAWS since a single writing prompt was administered for the 2013 SAWS assessment.

Based on these DIF statistics, items are classified into one of three categories and assigned values of A, B, or C. Category A items contain negligible DIF, Category B items exhibit slight or moderate DIF, and 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 are provided in Table 23.

Table 23. 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-”.

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 B summarizes the number and percentage of items by DIF category from the 2013 field test items for each grade and content area. The 2013 operational tests are comprised of items that were piloted in years prior to 2013, which were reviewed and approved by Content Review, Bias and Fairness Review, and Data Review Committees.

2.7.3. Item Response Theory (IRT) Analysis

Rasch IRT was used to scale the PAWS. 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 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. The probability of a correct response to an item increases as the student’s ability increases. See chapter 5 for further details on the Rasch model. The results of the Rasch IRT analyses of the field test items can be found in Appendix C.

2.7.4. Data Review Procedures

Following the spring 2013 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 with values 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 2013 administration were reviewed in July 2013 by a panel in Casper WY.

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. cluing 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 will also be given response or score distributions both for all examinees. In addition for multiple choice items they will receive distractor discrimination values. This information will be presented in both tabular format. Items with low or negative

discrimination and/or with distractors with positive discriminations will be called 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 will be 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 where 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 will be 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 will be 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 C, the classical analysis results appear in Appendix A, and DIF in Appendix B. Items accepted at data review from the 2013 administration are eligible for use as operational items beginning with the spring 2014 administration.

2.8 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 and step parameters. Therefore, when forms were constructed for the 2013 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 2012 base PAWS tests, as test characteristic curves (TCCs), information, and standard error curves could be evaluated to help ensure statistical comparability.

2.8.1. Construction of the Reading, Mathematics, and Science Forms

ETS utilized proprietary test construction software for the construction of the 2013 forms. The ETS psychometrician utilized the test content blueprint and the statistical targets into a

configuration file for each grade and subject test being constructed. The blueprints were unchanged from 2012 and can be found in Chapter 2. The TCCs, information and standard error curves from the 2012 administration constituted the statistical targets for the 2013 forms.

In addition, the targets for key balance (for multiple choice items, approximately 25% for each of options A-D, proportion of items from the 2012 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). In addition, limits were set based 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 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 to minimize the difference between the item's position on the 2013 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 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] 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.

2.8.2. Construction of the SAWS Forms

The test design for the spring 2013 SAWS assessment was composed solely of a census, stand-alone, single writing prompt either Expository or Expressive. ETS content specialists and psychometricians jointly selected a prompt according to test build specifications and test blueprints for the 2013 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 pulled during the test build. Test build was an iterative process to balance test content and its statistical properties. The 2013 Operational prompt was selected to ensure the writing prompts across administrations have difficulties that are as similar as possible. The selected prompt were provided to the WDE content specialists for approval.

2.8.3. 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.

3. TEST ADMINISTRATION

3.1 Test Materials

Test materials were sent to each Wyoming Building PAWS and SAWS Coordinator in shrink-wrapped packages within boxes that included school inventories. All students in grades 3–8 received scorable test booklets.

Building test coordinators were responsible for distributing the materials to test administrators. Materials were color-coded by grade and were printed with student identification and demographic codes. 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.

3.2 Materials Return

Once test administrations were completed, materials were collected and tabulated by Building PAWS Coordinators. In addition, the demographic information was hand gridded on the Test and Answer books if it was not pre-printed. 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. The tracking numbers were recorded on a Bill of Lading (included in the *2012-2013 Test Coordinator's Manual*) that was faxed to ETS after pickup by the carrier.

3.3 Directions for Administering and Training

The *PAWS and SAWS Directions for Administration Manual* and *PAWS and SAWS Test Coordinators Manual* provided the guidelines for planning and managing the PAWS SAWS administration for district and school administrators. The *PAWS and SAWS Directions for Administration Manual* by grade and test provided specific directions for test administrators from scheduling and timing for sessions and preparing students to testing students from special populations. Two half-day and two web-based comprehensive training sessions conducted jointly by the WDE and ETS were held in December 2012 prior to the spring 2013 testing window. In addition, several test administrator training videos were posted to the ETS Access webpage. 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 of Reading were administered in four untimed sessions. Grades 7 and 8 of Reading were administered in three untimed sessions. Grade 3 Mathematics was administered in two untimed sessions (this was the only grades which did not have separate calculator and non-calculator sessions). Grades 4–8 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.

SAWS assessment was administered in a single untimed session.

3.3.1 Allowed Student Manipulatives

Calculators were not allowed on the Number Operations and Concepts portions of the PAWS (session 1) in grades 4–8. Calculators were not permitted for the 3rd grade PAWS assessments. In addition, a *PAWS 2013 Allowable Resources* document was posted to the WDE webpage to assist test administrators in administering PAWS in a standardized manner.

3.3.2. 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 state law, violations were 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 Test 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 *2013 PAWS and SAWS Directions for Administration Manual*. Persons designated to administer the PAWS 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 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 *2013 PAWS and SAWS Directions for Administration Manual*. 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 PAWS and

SAWS Coordinator. The District PAWS and SAWS Coordinator subsequently reported all irregularities to the WDE Standards, Assessment and Accountability Unit.

3.4 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 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

3.5 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, reached a level sufficient to jeopardize the validity of interpretations. However, based on available evidence, the standard accommodations listed below 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 (1999).

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 *2013 PAWS Directions for Administration Manual*) 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, 2013 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 2013 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 2013 PAWS and SAWS Standard Accommodations* and the *2013 PAWS and SAWS Standard Accommodations Frequently Asked Questions (FAQ)*. The *Wyoming Statewide Assessment System 2013 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 & 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.

3.5.1. 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 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).

3.5.2. 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 *2013 Directions for Administration Manual* (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 2013 PAWS standard accommodations.
4. Standard accommodations were administered as described in the *Wyoming Statewide Assessment System 2013 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 cluing 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 any student without an IEP or 504 Plan or non-ELL students. 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 he/she:

1. Understands 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 2013 PAWS Training or has viewed the 2013 PAWS Training online and submitted record of the training to the building principal; and
4. Has completed the 2013 PAWS Accommodations Training online and submitted record of the training to the building principal.

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

3.5.3. 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 multi-sensory),

- 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:

3.5.3.1. 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. Readers 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. Readers 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.

3.5.3.2. 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 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.

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.

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.

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.

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.

19. A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.

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.

3.5.3.3. 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.

3.5.3.4. 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 he or she is most likely to demonstrate peak performance.

3.5.4. 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, 2013 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 and 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 2013 administration.

3.5.4.1. 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. Readers may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Readers 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.

3.5.4.2. 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.

3.5.4.3. 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.

3.5.5. PAWS 2013 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.

3.5.5.1. Empirical Analysis of Accommodations

IEP and 504 Plan students comprised approximately 11%–14% of students at each grade level, with between 60%–80% receiving testing accommodations (depending on grade and subject). 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 D.

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 2013 PAWS for Mathematics, Reading, and Science in grades three through eight and eleven are presented in Appendix E.

3.5.6. Selection and Administration of Accommodations

An important question regarding the use of accommodation in large-scale assessment is whether the resultant student scores mean the same thing as scores resulting from non-accommodated assessment (Kim, Wang, Zhao, & 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 calibration⁸, specifically in terms of item parameters and resulting test scores (Karkee, Lewis, & 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 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 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 2013 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 2013 PAWS and SAWS Standard Accommodations and 2013 PAWS and SAWS Standard Accommodations FAQ were distributed via a Superintendent's Memo.

⁸ 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 manual 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.

The training provided critical information regarding students eligible to receive standard accommodations, persons eligible to administered standard accommodations, standard and nonstandard accommodations, 2013 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 2013 PAWS Test Administrator Training Verification Form. All training materials and documents were available on the WDE website.

4. PROCESSING AND SCORING OF PAWS and SAWS ITEMS

4.1 Overview

This chapter describes the receipt control, scanning, and scoring procedures used at ETS for the 2013 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 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 scorers employed by ETS. The WDE had upfront oversight and control of training materials and audited scorer trainings at their discretion. For 2013 operational SAWS hand-scored items were scored in Concord, CA.

4.1.1 Multiple Choice Items

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

4.1.2 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.

4.2 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 districts. 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 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, school headers, and order headers for scoring.

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.

4.2.1 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. Information about schools for which receipts were incomplete or not received was communicated to the WDE by ETS's program manager.

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

4.2.2 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, district name, and a space to notate if any audio, large print or Braille returned scorable documents were present.

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

4.2.3 Scanning

In the scanning stage, ETS captured all the data from the student response forms, school headers, and order headers that were 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 scorers based upon rules established for the program. The data collected from the image scanners was 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 Scoring Center (PSC) 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 scorers who are experienced in scoring a wide range of open-ended assessments in writing. Scorers for the SAWS were drawn from this pool and received additional SAWS-specific training prior to their scoring the assessment.

4.2.4 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 was 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.

4.3 Scorer Recruitment and Qualification

4.3.1 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.

4.3.2 Rater Qualifications

4.3.2.1. Training

ETS raters were an integral part of the SAWS program, as they were responsible for evaluating the written portion of the Proficiency Assessments for Wyoming Students (PAWS) examination. 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 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.

4.3.2.2 Certification Procedures

After completing ETS’s Online Scoring Network (OSN)TM general training and SAWS specific training provided in the “LearnOSN” self tutorial website, prospective raters made arrangements with the Performing 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 he/she had successfully certified. No more than two attempts were permitted.

4.4 Rangefinding

Rangefinding was divided into two parts: pre-rangefinding and rangefinding. Rangefinding facilitators, representatives of the WDE, Lead Scoring Leaders (LSLs), 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 for each topic. 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.

4.4.1. Pre-Rangefinding Procedures

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

About fifty papers 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 papers selected for rangefinding. Any papers not used as anchor papers were used for calibration and monitor (validation) papers.

4.4.2 Rangefinding Procedures

The Performing Scoring Services (PASS) department prepared copies of all five sets for each participant. Facilitators then explained the purpose of and procedure for the rangefinding

process. Readers (representatives from WDE, AD, and LSLs) were then directed to read and score the first set papers. They were provided approximately twenty-five minutes per set to mimic the minimum reading rate required by scorers. Each reader was provided a score sheet to record his or her 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 essays, and encouraged the group to reach a consensus about the utility of papers 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 papers chosen. Because of this, the facilitators made careful notes of the discussion on each paper. Facilitators began with any papers 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 readers disagreed with the score, the facilitators initiated a discussion between a reader or readers in the majority and one of those in the minority by reading the essay aloud in order to fill as many slots as possible through the first set. After discussion, readers were permitted to modify their scores, thus producing a consensus. On the other hand, if a reader or readers 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 papers 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.

4.5: Hand-Scoring Process

All student responses to the writing prompts were scored in the OSN 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.

4.5.1. 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 OSN. As they scored the essays, raters referred to the appropriate SAWS

Scoring Guide and sample essays (“benchmarks” and “rangefinders”). A scoring leader guided and monitored the process to further ensure accuracy in scoring.

4.5.2 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 papers 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.

4.5.3 Reference Materials for Scoring

The following reference materials were easily accessible in OSN:

- Each SAWS writing prompt had 4 traits to score. Each grade had its own **Scoring Guide**, which explained the criteria for each score point. For example, Grade 3 had a range from 0 to 3, with 3 as the highest score.
- Each prompt had **Benchmarks**, which were papers 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.

4.5.4. Guidance to Scorers: 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.
- Scorers must recognize that papers 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.

4.6: Procedures for Maintaining and Retrieving Individual Scores

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

4.7 Inter-Rater 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 scorers and provided the results of their scoring on an ongoing basis. Information on these reports included the number of responses read by the scorers, 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 scorers, 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 with 25% of the responses randomly routed by ETS’s on-line system for a Reading by a second scorer to monitor inter-rater reliability. Non-adjacent scores received a third score or resolution score performed by a Supervisor, Scoring Director, or Content Specialist which was used as the operational score for the student’s response (i.e., the resolution score overrode both the initial and backread scores).

Scorers were expected to maintain a minimum cumulative perfect agreement rate of 70% agreement for extended-response items. Scorers who fell below this standard were targeted for additional training and backreading.

Section 7.3 presents the overall inter-rater reliability information for the 2013 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 which was used as the operational score for that item. The second scoring was used for inter-rater monitoring purposes only.

4.8 Accuracy Monitoring

The monitoring functions of the OSN provided a useful method for overseeing the accuracy of scoring and the performance of individual topics. The OSN produced a variety of reports with extensive data on both readers 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 readers while they are actively scoring a group of essays.

The OSN produced reports showing the degree to which readers 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 too low, too 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 OSN monitoring function also provided a statistical portrait of topic performance. Test development staff was able to see over time whether a given trait was performing well by considering:

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

Analytic evaluation was a procedure for scoring varied 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 inter-rater reliability calculations.

4.9 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 (BL), Copy of the Prompt (CP), Foreign Language (FL), Illegible (IL), Incomprehensible (IN), Off-Topic (OT), and Refusal (RF).

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). Scorers forwarded papers 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 he or she 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.

4.9.1 Blank (BL)

- A blank page was one that contained no writing or markings at all.
- A response with complete erasure was a Blank—the scorer was able to determine by the presence of smudges that the student had written and then erased something, but was unable to read any words or letters.
- A response with an incomplete erasure, where words and/or parts of words were still readable but it was obvious the student intended to erase the entire response was also scored as a Blank.
- A response was crossed out, but where the response was still readable was also scored as a Blank.

4.9.2 Copy of Prompt (CP)

- To receive a score of CP, the student's response consisted only of a word-for-word repetition of the test item or a substantial portion of it.
- A response that consisted of a word-for-word repetition of portions of the text was not a Copy of the Prompt but was a valid response and was scorable.

4.9.3 Foreign Language (FL)

- If ALL of the response was in a foreign language, the response was sent to the Review queue to be given an FL by a Scoring Supervisor, Scoring Director, or Content Specialist.
- If portions of a response were in a foreign language, scorers disregarded those portions and evaluated and scored what was written in English.

4.9.4 Illegible (IL)

- A response was Illegible only if all or a substantial portion of it was so illegible that the response could not be read. Scorers sent potentially illegible responses to a Review queue, where a Scoring Supervisor, Scoring Director, or Content Specialist determined whether the response was truly illegible. (Experienced scoring staff is often able to read responses that at first appear to be illegible.)
- A response was not considered to be illegible just because the student's handwriting was poor or sloppy.

4.9.5 *Incomprehensible (IN)*

- In an Incomprehensible response, a scorer was able to read words and/or letters but unable to make sense of them.
- Some students wrote responses in which all or a substantial portion of the words were misspelled. Before assigning an invalid score of Incomprehensible, every effort was made to decode the response. Students often write like they speak, so scorers tried to read it phonetically, and hear what they were attempting to say.

4.9.6 *Off-topic (OT)*

- The response bore no relationship or connection whatsoever to the prompt, nor was it a response to another prompt in the grade level. A response that is irrelevant is not necessarily Off-Topic. An Off-Topic response is usually considered “blue sky,” and well removed from being merely irrelevant.

4.9.7 *Refusal (RF)*

- Statements such as “I refuse to answer,” “No,” “I hate this test,” “I don’t care,” “I don’t know,” “I wasn’t taught this,” “X,” (large X on the page), and “?” (question mark) were all considered to be refusals.
- Any artwork (pictures, doodles, etc.) was considered to be a refusal
- A student may have written a refusal and then went on to provide additional writing that was not a refusal. In this case, scorers disregarded the invalid refusal portion and evaluated and scored the remainder of the response. If any portion of a response was scorable then the entire response was fully scorable.

4.10 *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 papers to the Program Manager who forwarded the copies to the WDE.

4.10.1 *Policy on the Reporting of Alerts*

ETS’s scorers 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 scorers were not instructed to flag and report any statements beyond the above three categories. The scorers 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.

4.10.2 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 fit 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.

5. EQUATING, AND SCALING PROCEDURES

5.1 Overview

This chapter covers:

- The equating of the 2013 PAWS assessments; and
- Translation of raw scores to scale scores along with descriptive statistics for all of the 2013 PAWS scales.

The equating analyses were carried out under the supervision of ETS's lead psychometrician for the WY PAWS assessment project, who conducted all of the primary analyses documented in this chapter. All analyses were independently replicated by a second ETS psychometrician, with supporting activities and analyses from an ETS Research Associate. After all analyses were concluded and documented, preliminary results (statewide mean scores, performance level percentages, and pass rates) were calculated over the student data sample used for equating (all grades and subjects had responses from > 99% 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.

5.2 Item and Forms Development

Kolen and Brennan (2004, p.3) state that, Equating adjusts for differences in difficulty, not for differences in content. Properly, then, a discussion of the equating of the PAWS assessment begins by noting that the development of the items and forms for the PAWS began in 2005 and has been an ongoing process. 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 2006 for Reading and Mathematics and in 2008 for Science. It should be noted that the writing assessment was discontinued as a component of PAWS beginning with the 2013 administration. Feedback from the school districts within the state in the early years of the program indicated that the test required too much time, and so the number of items and points on the Reading and Mathematics tests were reduced at several points from 2007 to 2009, but care was taken to ensure that the proportion of items and points addressing each of the standards and skills covered by the tests remained unchanged. The blueprints for the Science tests are the same as they were in 2008 (the first operational year of administration). Finally, a printing error on one form of the grade 5 Reading assessment resulted in the invalidation of one operational item for 2013.

As a consequence of the problems experienced during the 2010 administration of the PAWS (see the 2010 Technical Report for details), the WDE decided to move from the hybrid computer and

pencil and paper administration model used in 2010 and earlier (where the multiple choice items were administered via computer and the constructed response items were administered via paper and pencil) to an exclusively paper-based test. The WDE (after consultation with the TAC) decided to maintain the year-to-year scaling across this change in administration mode. Meta-analyses by Wang, Jiao, Young, Brooks & Olson (2007, 2008) have found that administration mode generally does not result in any statistically significant differences in Reading or Mathematics achievement scores of students, suggesting that the equating should not be affected by the change in administration mode from 2010 to 2012. There was one additional change in the test blueprints between the 2012 and 2013 administrations with the removal of constructed response items from the assessments.

5.3 IRT Models and Calibrations

The Item Response Theory (IRT) models used to calibrate the 2013 Wyoming PAWS assessments were the *Rasch model* (Rasch, 1980) 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, diagnosis, 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 2013 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 θ answering item i with difficulty D_i is shown in the equation below:

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

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 θ 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).

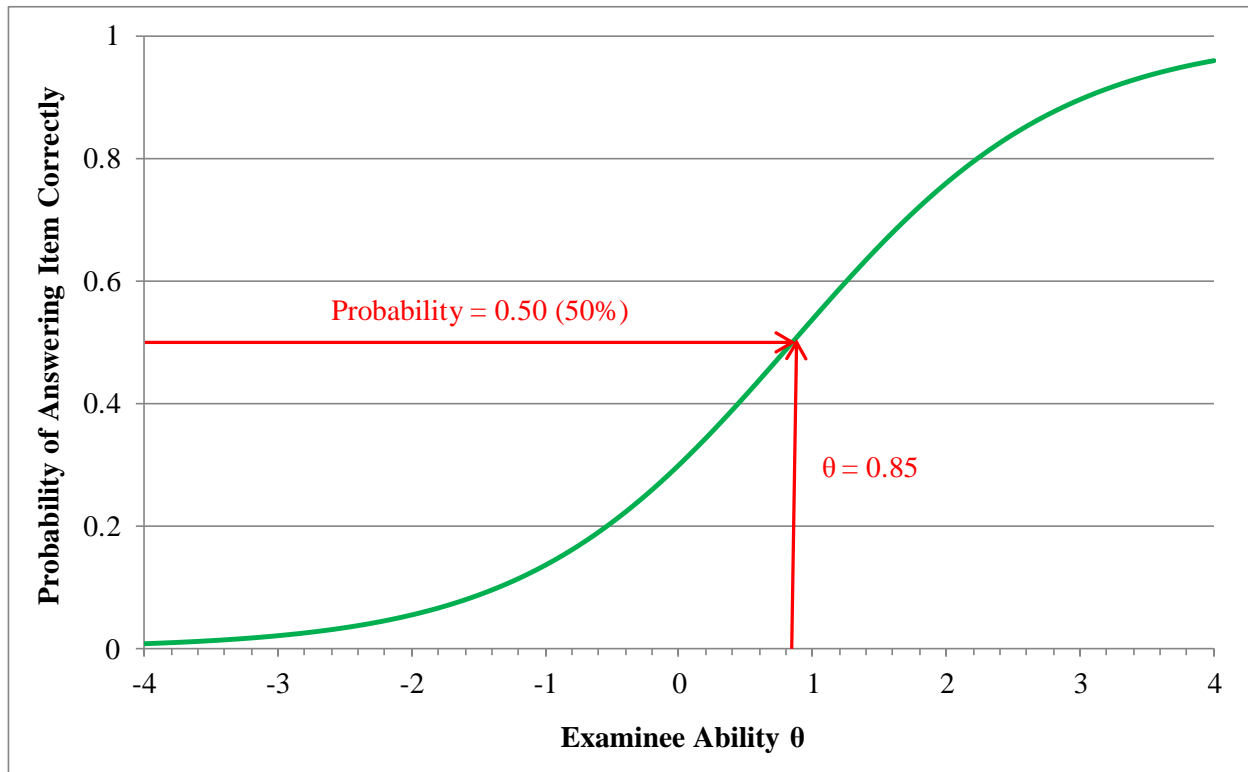


Figure 1. Sample item characteristic curve.

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 step 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.

5.4 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 *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.

Outfit mean-squares are influenced by outliers and are usually easy to diagnose and remedy. 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 are too predictable (redundancy, model overfit). Values greater than 1.0 indicate unpredictability (unmodeled noise, model underfit).

Appendix F provides Rasch difficulty estimates, standard errors, and infit and outfit statistics for 2013 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 2013 PAWS operational Mathematics, Reading, and Science tests. Operational classical item statistics are presented in Appendix G.

Appendix A provides IRT statistics and n-counts for items field-tested in 2013. 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.

5.5 Equating Analyses

As was previously mentioned, the PAWS assessments for 2013 were post-equated, meaning that the item and test statistics used to generate the scoring tables (i.e., tables explicating the relationship between specific raw scores and scale scores for a particular grade level and subject) came from the present (spring 2013) administration. All tests were equated to the pre-existing scale, and so scale scores on the 2013 administration use the same metric as scale scores for the same grade level and subject from previous administrations of the PAWS.

It should also be noted that for grades 3 through 8 of Reading and Mathematics, the Rasch statistical parameters of the item pools were vertically scaled in 2005. The Science tests are only given at grades 4 and 8 and therefore were not vertically scaled.

According to Young (2004), vertical scales have several important features. 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 do 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.

The vertical scales for Reading and Mathematics were developed in 2005 using the initial standalone field test for the PAWS assessments. Field test forms were designed so that they contained two kinds of linking items: Items that were common from one form to another at the same grade level (horizontal linking items) and items that were the same across forms from two adjacent grade levels (vertical linking items). Calibration of the item pool began by first calibrating items using the horizontal linking items to link the field test forms within each grade. This established a series of grade-by-grade item pools, each of which consisted of all of the field test items for a given content area. At this stage, all of the items within a grade and content area were on the same measurement scale; however, the measurement scales were not related across grade levels. This was addressed in the second stage of the item pool calibration when the vertical linking items were used to link the grade-level item pools. As a result, Reading and Mathematics test scores in grades 3–8 are directly comparable across adjacent grades.

5.5.1 Calibration and Equating Process for the 2013 administration

The procedures for equating the 2013 forms of the Reading, Mathematics, and Science forms to the preexisting scales were similar to those used in 2012. To establish a strong relationship between the 2012 and 2013 forms, each 2013 form had approximately 30% of its items drawn from the set of 2012 operational items. Other items were drawn from the item bank which was comprised of all items used operationally from 2006–2010 (with the exception of those items released publicly as sample PAWS items), and items field tested and accepted at data review from 2005 through 2012.

The tests were equated via common item equating to a calibrated item pool (Kolen and Brennan, 2004). All items (with a few exceptions) were drawn from previous years' administrations and can potentially function as anchor items with their parameters being drawn from their most

recent operational use⁹. Nevertheless, different from previous administration, only a core of items which had been operationally used in a previous form and deemed to have more reliable bank parameter values were selected for the anchor set. Other items which 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 standalone field tests for PAWS (2005 for Mathematics and Reading and 2007 for Science). 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 do students taking operational tests. Thus, items with statistics derived from these administrations were not used as linking items. In addition, a few items that were most recently used as operational items in the spring 2007 administration were identified as having possibly drifted during follow-up analyses after the administration (the 2007 administrations were pre-equated so only follow-up analyses were possible) and were also excluded as linking items. Finally, some items had been modified since their most recent use (mostly older items modified to bring them in line with current PAWS item style guidelines). Their previous statistics in 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 with respect to the parameters of the remaining anchor items.

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 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 2013 administration data differed significantly from their known values used for equating).

⁹ 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.

Anchor stability analysis was conducted to identify items that were not suitable for use as anchor items. Robust-Z statistic (Huynh and Meyer, 2010) was utilized to identify items that exhibited item parameter instability in their Rasch difficulties (multiple choice items) for the 2013 calibration as compared to their parameters from their most recent 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 parameters of anchor items in this run and their previous parameters 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 parameters of anchor items estimated from the free calibration and their bank parameter values, $MDN(d)$ 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 parameters 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 Wendy Yen's white paper. In the second round of anchor stability check, 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 to have 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 which 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 remained 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 explicated 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.

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

where

$P_{ij}(\theta)$ = the probability of a correct response for each of the items $i = 1, \dots, I$ given that the item categories are numbered $j=0, \dots, m_i$. And *True Score* is set to each achievable raw score point to find its theta equivalent.

These theta scores 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.

5.6 Translating Raw Scores to Scaled Scores and Performance Levels

Scale scores on the PAWS Reading, Mathematics, and Science tests ranged generally from 300 to 990 for grades 3–8; the specific minimum and maximum possible scale scores varied by grade and subject. As was discussed previously, the Reading and Mathematics scales for grades 3–8 were common and comparable across grades, while all grades of Science were separate. Appendix H provides scale score descriptive statistics for the 2013 PAWS operational Reading, Science, and Mathematics tests.

The following formulae were used to convert the underlying PAWS IRT Mathematics, Reading, and Science scales to the PAWS reporting scale:

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

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

where θ was the *IRT* ability estimate, and $SEM(\theta)$ was the conditional *SEM* of the ability estimate θ . Table 25 also contains the slope and intercept for the PAWS Mathematics, Science, and Reading scales. All subjects in grades 3–8 used a common slope and intercept. Science used the same constants as were used for the other subjects at the same grade level (4 and 8) to avoid confusion stemming from different scale metrics for different subjects at the same grade level.

The raw score to scale score conversion tables for the 2013 PAWS Reading, Mathematics, and Science tests can be found in Appendix I. Conditional standard errors and performance levels for the scale scores are also included in these tables.

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

Grade	Scaling constant
3 through 8	Scaled Score = $\Theta * 48.21 + 637.5$

6. PAWS and SAWS REPORTING

6.1 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 Appendix J for PAWS and Appendix K for SAWS.

The following reporting information is provided:

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

6.2 Scale 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 300 to 990 for both 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, 2004). Where each grade level test is based on a common blueprint design, the grade–level specifics from instruction as reflected in the test questions differs 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 meaningful.

6.3 Performance Levels

Performance classifications are determined by applying the appropriate scale score cuts established from the PAWS standard setting activities described in the 2006 (Reading and Mathematics) and 2008 (Science) PAWS Technical Manuals. Tables 26–28 provide the scale score ranges for the PAWS Reading, Mathematics, and Science tests.

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

Grade	Below Basic	Basic	Proficient	Advanced
3	300 - 519	520 - 583	584 - 660	661 - 975
4	300 - 569	570 - 633	634 - 699	700 - 975
5	300 - 586	587 - 638	639 - 706	707 - 975
6	300 - 593	594 - 649	650 - 717	718 - 975
7	300 - 609	610 - 667	668 - 745	746 - 975
8	300 - 623	624 - 675	676 - 748	749 - 975

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

Grade	Below Basic	Basic	Proficient	Advanced
3	300 - 557	558 - 599	600 - 679	680 - 975
4	300 - 583	584 - 619	620 - 697	698 - 975
5	300 - 606	607 - 644	645 - 720	721 - 975
6	300 - 631	632 - 662	663 - 740	741 - 975
7	300 - 652	653 - 686	687 - 757	758 - 975
8	300 - 675	676 - 705	706 - 776	777 - 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

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 Mathematics, Reading, 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 L.

6.4 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 scale and raw scores. The content standard-level scores were produced in the same way as the overall test scale scores—a raw score to scale 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 scale scores and associated error ranges (student scale score \pm one SEM) are graphically presented on the Student Score Report. The probability that student’s true score will be in the range indicated by the error bar is approximately 68%. For Mathematics, scale 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, scale 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 (more items = less error), the error ranges for the standard level 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 Chapter 2.

6.5 SAWS Raw Scores

The SAWS program is currently in transition with the reduction from two writing prompts to a single writing prompt for 2013. Therefore, a normative component will be reported on each student’s Individual Student Report (ISR). Table 30 provides the state level normative percentage distributions for SAWS for 2013. These percentages will be transformed into a graphic shown in Figure 2 unique to each grade and placed on the respective individual student reports.

Table 29. 2013 State Level Normative Percentage Distributions for SAWS

Grade	State N count	0-3 Points	4-7 Points	8-10 Points	11-12-points
3	7292	8	53	29	10
4	7282	6	58	28	8
5	6909	4	47	34	15
6	6926	5	52	32	10
7	6972	4	41	37	18
8	6914	4	36	39	21

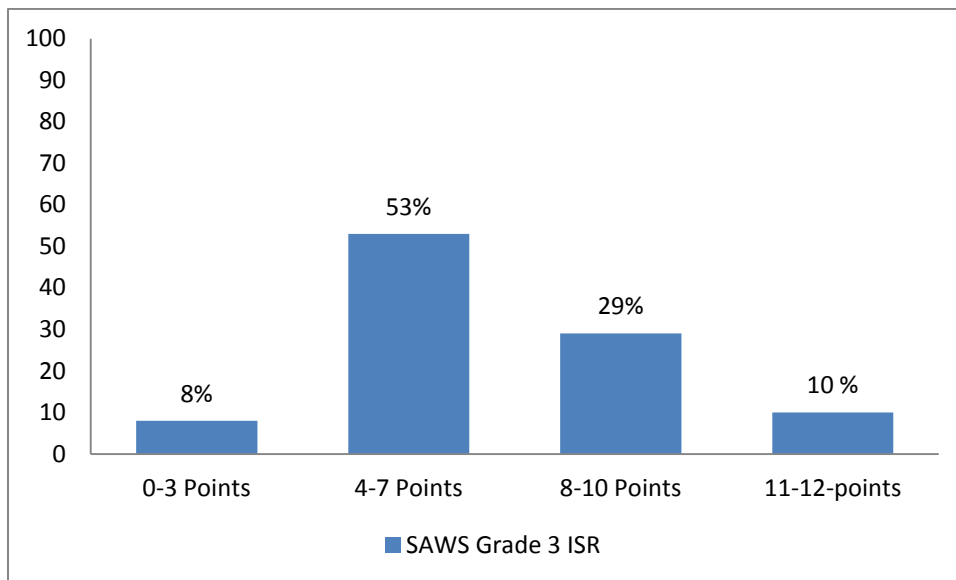


Figure 2. Grade 3 SAWS 2013 Normative Information Example for Each Individual Student Report

6.6 Production of Printed Score Reports for PAWS and SAWS

In final preparation for the production and printing of PAWS and SAWS 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 from all concerned parties, production reporting commenced.

Paper individual student reports were generated for distribution to WDE schools. 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 used data encryption methods (e.g., 256-bit) appropriate to the sensitivity of the data being transmitted. ETS provided secure user IDs and passwords to access the FTP site.

In the pre-mailing workstation, printed student reports were assembled and packed in color-keyed folders. 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. This was the final n-count verification checkpoint where the number of students reported was compared to the information recorded on the MFS and in the database. 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 individual student report is provided in Appendix J. A SAWS example can be found in Appendix K.

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

The 2013 Wyoming Assessment Score Reports: Supplement Guide for Districts and Schools for PAWS and 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.

7. RELIABILITY

7.1 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 feature of reliability.

This chapter describes the reliability analyses of the 2013 PAWS operational assessments. Internal consistency and inter-rater reliabilities, classical and conditional standard errors of measurement, and accuracy and consistency results are included.

7.2 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 investigates 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:

Cronbach's coefficient alpha is represented by:

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

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.

Overall alpha statistics suggest reasonable internal consistency reliability for PAWS assessments at all grades based on the total test. 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 30, including coefficient alpha and the standard error of measurement for each grade and content area.

Table 30. Summary Reliabilities, Standard Errors of Measurement, and Descriptive Statistics by Grade

Grade	N Counts	Possible Points	Cronbach's Alpha	SEM
Reading				
3	7141	50	0.91	3.00
4	7158	50	0.91	2.85
5	6772	54	0.90	3.23
6	6800	56	0.88	3.26
7	6827	56	0.89	3.28
8	6774	56	0.89	3.35
Mathematics				
3	7133	60	0.92	2.97
4	7163	65	0.92	3.37
5	6766	65	0.93	3.47
6	6795	65	0.92	3.54
7	6825	66	0.92	3.59
8	6779	70	0.93	3.62
Science				
4	7157	50	0.87	3.03
8	6754	50	0.88	3.19

7.3 Inter-Rater Reliability

Rater agreement or consistency is critical for valid test score interpretation for assessments comprised of constructed response items requiring human raters to score the performance of students. Inter-rater 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 SPE indicate well above 90% agreement. Table 31 provides the agreement rates for the SAWS operational prompts. Appendix M provides the SAWS field test demographic performance. Appendix N provides the agreement rates for all SAWS field test prompts.

7.4 Weighted Kappa

Also provided in the tables of Table 31 and Appendix M are the weighted kappas, an index of inter-rater 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 N categories, one can construct a N by N rating table with scores provided by two readers A and B. Define p_{ij} as the proportions of occurrence assigning scores i by reader A and score j by reader B, respectively for a given prompt, where $i = 1, 2, \dots, N; j = 1, 2, \dots, N$.

The weighted kappa coefficients k_{ij} is defined as

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

where

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

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

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

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

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

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

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.

Table 31. PAWS 2013 Overall Inter-Rater Reliability for SAWS Prompts and Traits

Grade	Trait	N	Rater 1		Rater 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
3	Prompt Total	1629	7.19	2.38	7.06	2.52	0.67	25.97	57.09	
	Idea Development	1629	1.89	0.69	1.87	0.70	0.56	61.57	98.77	0.56
	Organization	1629	1.83	0.69	1.81	0.71	0.55	60.04	98.59	0.55
	Voice	1629	1.76	0.69	1.72	0.75	0.54	57.58	98.04	0.53
	Conventions	1629	1.72	0.71	1.66	0.76	0.58	58.62	98.53	0.58
4	Prompt Total	1698	6.87	2.52	6.80	2.51	0.73	30.68	62.78	
	Idea Development	1698	1.77	0.73	1.77	0.72	0.59	62.01	98.29	0.59
	Organization	1698	1.61	0.79	1.57	0.78	0.69	70.32	97.11	0.69
	Voice	1698	1.74	0.73	1.73	0.73	0.59	60.78	98.53	0.59
	Conventions	1698	1.75	0.74	1.74	0.74	0.60	60.13	98.76	0.60
5	Prompt Total	1699	7.37	2.61	7.48	2.63	0.76	32.49	64.39	
	Idea Development	1699	1.95	0.72	2.01	0.72	0.66	67.57	98.82	0.65
	Organization	1699	1.84	0.75	1.88	0.77	0.67	66.39	98.29	0.66
	Voice	1699	1.83	0.76	1.86	0.76	0.66	63.74	99.18	0.66
	Conventions	1699	1.74	0.74	1.74	0.77	0.63	60.92	98.94	0.63
6	Prompt Total	1605	7.09	2.40	7.16	2.41	0.60	28.16	55.89	
	Idea Development	1605	1.80	0.69	1.84	0.67	0.50	58.19	98.69	0.50
	Organization	1605	1.79	0.68	1.79	0.69	0.52	60.00	98.44	0.52
	Voice	1605	1.79	0.69	1.82	0.69	0.50	58.57	98.01	0.50
	Conventions	1605	1.71	0.68	1.71	0.71	0.50	56.82	98.38	0.50
7	Prompt Total	1711	7.66	2.79	7.66	2.76	0.74	30.27	57.16	
	Idea Development	1711	1.96	0.77	1.97	0.75	0.60	59.44	98.07	0.60
	Organization	1711	1.91	0.78	1.90	0.79	0.64	62.36	98.01	0.64
	Voice	1711	1.95	0.78	1.97	0.77	0.65	61.13	98.89	0.65
	Conventions	1711	1.83	0.77	1.82	0.78	0.63	60.20	98.60	0.63
8	Prompt Total	1698	7.88	2.73	7.88	2.72	0.70	31.68	57.18	
	Idea Development	1698	2.05	0.75	2.02	0.74	0.61	60.54	98.59	0.61
	Organization	1698	1.98	0.76	2.00	0.75	0.63	62.60	98.35	0.63
	Voice	1698	1.96	0.74	1.97	0.75	0.60	60.78	98.23	0.60
	Conventions	1698	1.88	0.76	1.89	0.76	0.59	57.77	98.23	0.59

7.5 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 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 known as 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 the person'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 obtain the SEM from:

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

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 threshold level for many persons in the sample, the standard error is small (Embretson & Reise, 2000).

Overall Coefficient Alpha and SEM results for PAWS assessments are presented in Table 32. Conditional SEMs for all achievable scores on the assessment are included with the raw score to scale score tables in Appendix I for PAWS.

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

Grade	N Counts	Possible Points	Cronbach's Alpha	SEM
Reading				
3	7141	50	0.91	3.00
4	7158	50	0.91	2.85
5	6772	54	0.90	3.23
6	6800	56	0.88	3.26
7	6827	56	0.89	3.28
8	6774	56	0.89	3.35
Mathematics				
3	7133	60	0.92	2.97
4	7163	65	0.92	3.37
5	6766	65	0.93	3.47
6	6795	65	0.92	3.54
7	6825	66	0.92	3.59
8	6779	70	0.93	3.62
Science				
4	7157	50	0.87	3.03
8	6754	50	0.88	3.19

7.6 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 standard 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 results of decision consistency and accuracy analyses for the PAWS assessments administered in Spring 2013.

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 WCAP test, “below basic” and “basic” performance levels result in a “does not meet standard” classification denoted as A in Figure 3; “proficient” and “advanced” performance levels result in the “meets standard” classification indicated as B.

Figure 3. 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 33–38 present the results of the decision accuracy and consistency of the PAWS cut scores for Mathematics, Reading 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 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 0.95, with most results above 0.90. All false negative and false positive results are at or below 0.15. These results suggest acceptable levels of reliability at the cut points for all PAWS assessments.

Table 33. PAWS 2013 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		
Reading	7141	0.82	0.09	0.09	0.74	0.12	0.13	0.92	0.88
Mathematics	7133	0.86	0.07	0.08	0.80	0.10	0.10	0.95	0.94

Table 34. PAWS 2013 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		
Reading	7154	0.83	0.08	0.09	0.76	0.12	0.12	0.94	0.91
Mathematics	7163	0.83	0.08	0.09	0.77	0.11	0.12	0.94	0.92
Science	7157	0.78	0.11	0.11	0.70	0.15	0.15	0.89	0.85

Table 35. PAWS 2013 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		
Reading	6772	0.81	0.09	0.10	0.73	0.13	0.14	0.92	0.89
Mathematics	6766	0.84	0.08	0.08	0.78	0.11	0.11	0.94	0.92

Table 36. PAWS 2013 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		
Reading	6800	0.81	0.09	0.10	0.74	0.13	0.13	0.93	0.90
Mathematics	6795	0.83	0.08	0.09	0.77	0.11	0.12	0.94	0.91

Table 37. PAWS 2013 Decision Accuracy and Consistency Indices – Grade 7

Subject	N	Accuracy			Consistency			Cut Point Accuracy	Cut Point Consistency
		Overall	False Positive	False Negative	Overall	False Positive	False Negative		
Reading	6827	0.83	0.08	0.09	0.76	0.12	0.12	0.92	0.89
Mathematics	6825	0.82	0.08	0.10	0.75	0.12	0.13	0.92	0.89

Table 38. PAWS 2013 Decision Accuracy and Consistency Indices – Grade 8

Subject	N	Accuracy			Consistency			Cut Point Accuracy	Cut Point Consistency
		Overall	False Positive	False Negative	Overall	False Positive	False Negative		
Reading	6774	0.82	0.08	0.09	0.75	0.12	0.13	0.92	0.89
Mathematics	6779	0.81	0.10	0.10	0.74	0.13	0.13	0.92	0.89
Science	6754	0.78	0.11	0.11	0.69	0.15	0.15	0.90	0.86

8. VALIDITY

8.1 Overview

Validity refers to the degree to which each interpretation or use of a test score is supported by evidence that is gathered (AERA, APA, & NCME, 1999; 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, & NCME, 1999; 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 we present the evidence gathered 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, & NCME, 1999). 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.

8.1.1. Purpose of the PAWs and SAWS

The purposes of the PAWs and SAWS are multifold, as outlined previously in Chapters 1 and 2. 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.

8.1.2. 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 provide an operational definition of the construct and define the procedures used to measure the content standards provide an operational definition of the construct to which each set of standards refers. That is, they define, for each subject area 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, & 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 2 and on the WDE Web page at http://edu.wyoming.gov/programs/statewide_assessment_system/paws.aspx. ETS has developed all PAWS and SAWS test items to conform to the Wyoming content standards and test blueprints.

8.1.3. 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. On the basis of 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.

Reporting 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 6.

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, & NCME, 1999, Standard 1.4).

8.1.4. 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, mathematics, and writing. In addition, students in grades 4, and 8 take a grade-level science test.

8.2 Evidence Based on Content-related Validity

According to the AERA, APA, and NCME (1999), 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 appropriately be regarded as measures of the intended construct.

As noted in the AERA, APA, and NCME's *Test Standards* (1999), 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:

8.2.1. Description of the state standards

As was noted in Chapter 1, Wyoming adopted rigorous content standards in 2006 and 2008 in four major content areas: reading, mathematics, science, and writing. 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.

8.2.2. 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 2. 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 2 and on the WDE Web page at http://edu.wyoming.gov/programs/statewide_assessment_system/paws.aspx.

8.2.3. Item development process

A detailed description of the content and psychometric criteria applicable to the construction of the 2013 PAWS and SAWS is presented Chapter 2.

8.2.4. Item review process

Chapter 2 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.

8.2.5. 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 2.

8.2.6. 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. Alignment studies for the PAWS Reading, Science and Mathematics assessments were completed in years previous 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 manuals).

8.3 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, & NCME, 1999), 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.

8.3.1. 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. There are strong relationships between the PAWS Reading, Mathematics, and Science scaled 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. The 2012 PAWS Technical Report presents the intercorrelations for the subscales within the Reading, Mathematics, and Science tests respectively. 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.

8.3.2. 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 B. The vast majority of the items exhibited little or no significant DIF, suggesting that, in general, scores based on the PAWS items would have the same meaning for individuals who differed in their gender characteristics.

8.4 Evidence Based on Response Processes

As noted in the AERA, APA, and NCME's *Standards* (1999), 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.

8.4.1. Evidence of Inter-rater 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' essay responses. In order to evaluate the reliability of the student scores on the writing prompts administered in grades 3–8, two raters scored approximately 25% of the examinee responses. The data collected were used to evaluate inter-rater reliability and inter-rater agreement.

8.4.1.1. Inter-rater 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 may not assign the same score to a given response. The results showed moderate levels of agreement between raters that score examinees' written responses to the prompts administered in both grades 3–8. Chapter 7 provides the weighted kappas for all SAWS traits.

8.4.1.2. Inter-rater Agreement

As noted previously, 25% of the responses to the SAWS prompts were scored by two raters. The total score (12 points) showed approximately 26.0% exact agreement and 57.0% exact + adjacent agreement. The traits scores, having only 3 points, resulted in an exact score agreement from 57.8% to 70.0%. Exact + adjacent score agreement ranged from 98.2% to 99.2%. Chapter 7 provides the agreement rates for all SAWS grades.

8.5 Evidence Based on Internal Structure

As suggested by the *Standards* (AERA, APA, & NCME, 1999), evidence of validity can also be obtained from studies of the properties of the SPEs 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 2012 PAWS Technical Report.

8.5.1. 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 mean point biserials for the items in the PAWS are presented in Appendix G.

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 item *p*-values is given in Appendix G; the distributions of item *b*-values are given in Appendix F.

8.5.2. 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 level, as well as reliability results for the reporting clusters.

Overall reliability—The reliability analyses, on each of the PAWS assessments are presented in Chapter 7. 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 the Chapter 7; Classification and Accuracy. These levels of accuracy and consistency are high, and they are consistent with levels seen in previous years.

8.5.3. Dimensionality

Measurement using IRT implies order and magnitude on a single dimension (Andrich, 1989). Consequently, in the case of scholastic achievement, this requires a linear scale to reflect this idea of measurement. Such a test is considered to be unidimensional (Andrich, 1988, 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, & 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).

To assess the unidimensionality of an assessment form, Hattie (1985) suggested using the difference of eigenvalues between the first factor and the second factor divided by the difference of eigenvalues between the second factor and the third to evaluate unidimensionality. If the ratio is large (i.e., larger than 3), then the first factor is relatively strong. To check the unidimensionality of

the PAWS Reading and Mathematics tests, unrotated principle components analyses were performed at the item level to determine eigenvalues. The first three were then used to produce the ratio described above. The results of this study are provided in 2012 PAWS Technical Report.

8.6 Evidence Based on Consequences of Testing

As observed in the *Standards*, 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, evidence that the expected benefits accrue 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.

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 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 extended Wyoming content 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 items 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 were placed in the item bank with their statistics. ETS anticipates delivering 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 were logged; all change requests, including item bank updates for prompt availability status, were 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 the all testing related materials.

- **Publishing and Editing Review**—Three-way review of all project materials was performed. After this internal review, assessment materials were forwarded to WDE for review and approval.
- **Printing**—All external printing companies hired to print scannable and non-scannable 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. Post-testing, missing materials reports were provided to ETS for follow up.

9.2.1. Collecting Test Materials

After administration, school districts returned scorable materials within five working days after the last testing day of each test administration period. Districts were provided labels with bar-coded information identifying the school and district. School districts 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 School and Grade Identification (SGID) sheets with the number of 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 SGIDs to actual documents scanned. Schools were contacted by phone if there are any missing shipments or if the quantity of materials returned appears 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 answer documents, the Student Demographic Data Pages, 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 either scanned using a flatbed scanner or 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 form. 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 proceeds 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 percent of students meeting standards. Tables 39–44 provides a comparison percentage of the students classified as “Proficient + Advanced” from 2006 to 2013 for PAWS Reading, Mathematics, and Science respectively.

The PAWS Reading percentage of students in the equating population classified as “Proficient + Advanced” decreased for all grades from 2012. For Grade 3, the percentage of students “Proficient + Advanced” decreased from 69.6% in 2012 to 65.6% in 2013. Grade 4 had a decrease of 4.8% from 83.2% to 78.4%. Grade 5 had the largest drop in the percentage of students classified as “Proficient + Advanced,” from 79.1% to 72.5%, a decrease of 6.6%. For all grades, the percentage of “Proficient + Advanced” students is within previously observed values for the specific grade.

The percentage of Mathematic students in the equating population classified as “Proficient + Advanced” decreased for all grades from 2012. Grade 8 had the largest drop in the percentage of students classified as “Proficient + Advanced,” from 76.2% to 67.3%, a decrease of 5.3%. For all grades, the percentage of “Proficient + Advanced” students is within previously observed values for the specific grade.

The percentage of Science students in the equating population classified as “Proficient + Advanced” decreased for both grades from 2012. Grade 4 had a decrease of 5.8% from 63.3% to 57.5%. Grade 8 had the largest drop in the percentage of students classified as “Proficient + Advanced,” from 51.2% to 43.7%, a decrease of 7.5%. For both grades, the percentage of “Proficient + Advanced” students is within previously observed values for the specific grade.

Figures 4 through 7 display the PAWS percentage of students in the equating populations classified as “Proficient + Advanced” from the 2006 through 2013 for each grade level. The results for 2010 were not provided due to federal exemption for the reporting scores.

Table 39. Scale Scores Descriptive Statistics for the PAWS Reading Tests

Year	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
2006	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2007	6392	613.9	45.3	6256	662.8	48.9	6322	662.6	42.8	6160	674.8	45.5	6527	690.5	46.9	6662	702.0	43.4
2008	6467	595.7	54.8	6508	663.8	49.8	6387	653.9	46.0	6422	676.7	51.9	6290	682.1	50.2	6588	699.0	47.4
2009	6655	585.3	53.1	6631	659.7	53.0	6703	654.2	47.8	6528	681.2	56.5	6521	674.9	45.6	6339	693.4	48.6
2010	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2011	6761	600.9	48.3	6680	675.3	45.5	6721	674.9	46.3	6713	688.0	44.5	6700	690.3	42.0	6554	704.3	41.0
2012	7076	604.1	47.3	6771	680.3	50.2	6738	676.4	46.9	6787	692.4	43.2	6807	693.8	43.6	6752	707.2	43.2
2013	7141	602.8	52.7	7158	675.2	55.2	6772	670.5	48.5	6800	689.1	43.5	6827	697.2	44.4	6774	706.9	44.9

Table 40. Percentage Proficient and Advanced for the PAWS Reading Tests

Grade	2006	2007	2008	2009	2010	2011	2012	2013	Min	Max	Median	2013 Difference from Median	2013 Difference from 2012
3	61.6	76.5	61.7	53.7		65.5	69.6	65.6	53.7	76.5	65.5	0.1	-4.0
4	64.5	76.6	73.4	71.1		83.6	83.2	78.4	64.5	83.6	76.6	1.8	-4.8
5	66.2	75.2	65.7	62.8		77.1	79.1	72.5	62.8	79.1	72.5	0.0	-6.6
6	63.1	75.4	70.1	70.3		81.1	83.9	80.7	63.1	83.9	75.4	5.3	-3.1
7	63.4	72.7	64.9	57.1		70.4	75.0	74.1	57.1	75.0	70.4	3.7	-0.9
8	61.9	75.5	70.4	64.9		77.1	77.5	76.0	61.9	77.5	75.5	0.5	-1.5

Table 41. Scale Scores Descriptive Statistics for the Mathematics Tests

	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
2006	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2007	6392	664.0	49.8	6256	671.3	50.7	6322	677.7	46.9	6160	701.5	49.9	6527	710.2	45.4	6662	718.0	49.0
2008	6467	646.1	51.6	6508	660.8	57.7	6387	678.4	56.7	6422	705.3	53.2	6290	718.1	54.5	6588	731.3	55.2
2009	6655	648.1	56.0	6631	655.9	53.9	6703	680.3	55.6	6528	706.2	55.1	6521	716.8	52.5	6339	726.6	52.7
2010	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2011	6761	660.4	52.2	6680	668.0	53.1	6721	690.3	53.6	6713	710.5	51.9	6700	727.0	55.1	6554	736.5	49.3
2012	7076	668.3	54.3	6771	668.6	51.4	6738	694.2	53.9	6787	715.7	55.0	6807	726.5	52.4	6752	738.0	49.4
2013	7133	663.4	56.6	7163	668.1	55.7	6766	692.2	56.5	6795	712.2	54.2	6825	726.0	50.2	6779	734.4	50.4

Table 42. Percentage of Proficient + Advanced Students for the Mathematics Tests

Grade	2006	2007	2008	2009	2010	2011	2012	2013	Min	Max	Median	2013 Difference from Median	2013 Difference from 2012
3	74.8	91.6	83.4	80.9		88.4	90.4	87.3	74.8	91.6	87.3	0.0	-3.1
4	73.5	86.5	76.8	75.3		81.1	82.2	81.0	73.5	86.5	81.0	0.0	-1.2
5	66.2	77.3	71.0	72.0		78.4	81.9	79.5	66.2	81.9	77.3	2.2	-2.4
6	71.7	80.7	78.2	77.3		80.7	82.4	81.6	71.7	82.4	80.7	0.9	-0.8
7	60.6	71.4	72.2	72.5		74.1	76.0	75.8	60.6	76.0	72.5	3.3	-0.2
8	54.5	61.3	67.7	62.1		70.9	72.6	67.3	54.5	72.6	67.3	0.0	-5.3

Table 43. Scale Scores Descriptive Statistics for the Science Tests

Year	N	Grade 4		N	Grade 8	
		Mean	SD		Mean	SD
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

Table 44. Percentage of Proficient + Advanced Students for the Science Tests

Grade	2008	2009	2010	2011	2012	2013	Min	Max	Median	2013	2013
										Difference from Median	Difference from 2012
4	50.9	50.5		54.5	63.3	57.5	50.5	63.3	54.5	3.0	-5.8
8	46.4	42.9		50.7	51.2	43.7	42.9	51.2	46.4	-2.7	-7.5

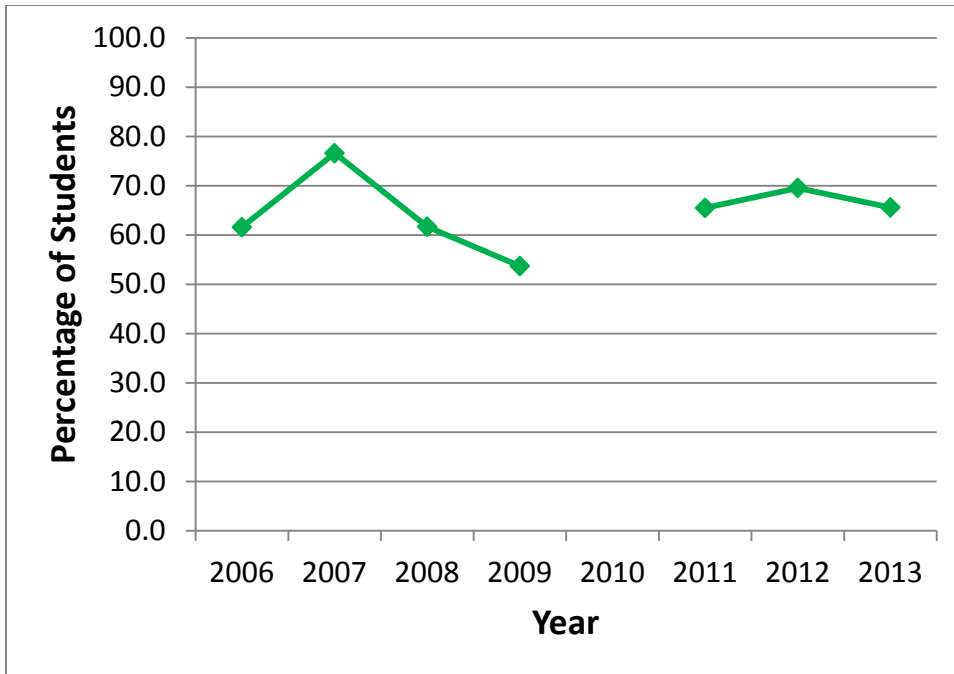


Figure 4. Percentage of Proficient and Advanced Students for Grade 3 Reading

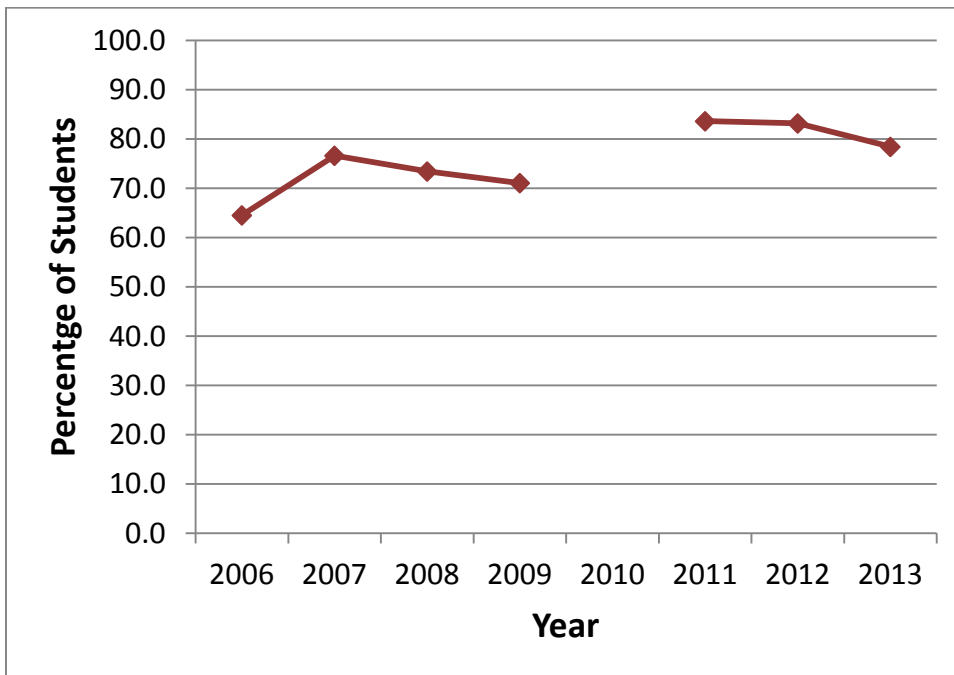


Figure 5. Percentage of Proficient and Advanced Students for Grade 4 Reading

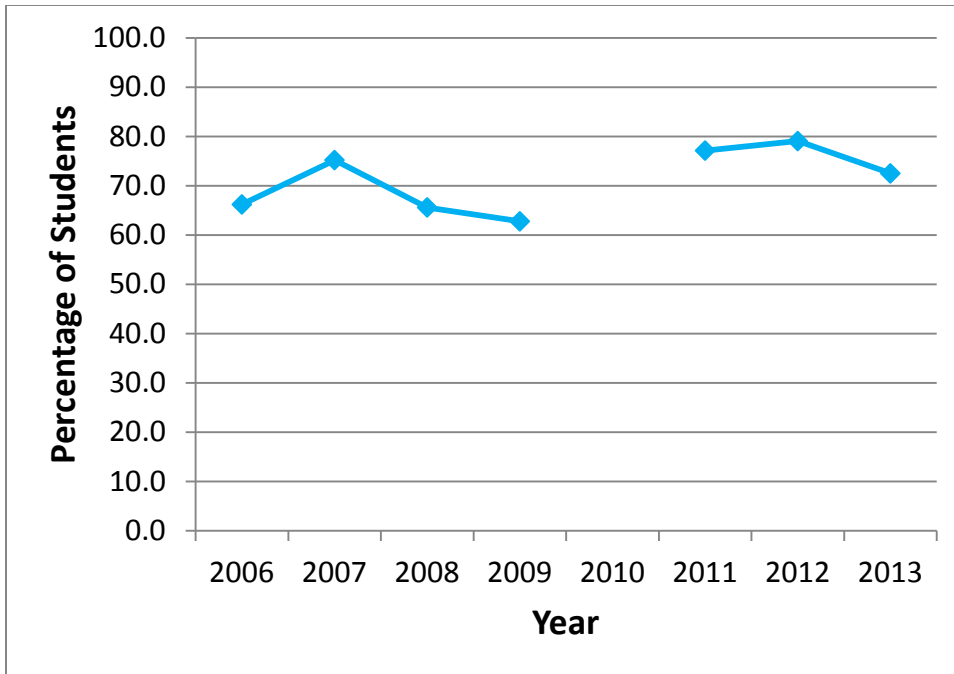


Figure 6. Percentage of Proficient and Advanced Students for Grade 5 Reading

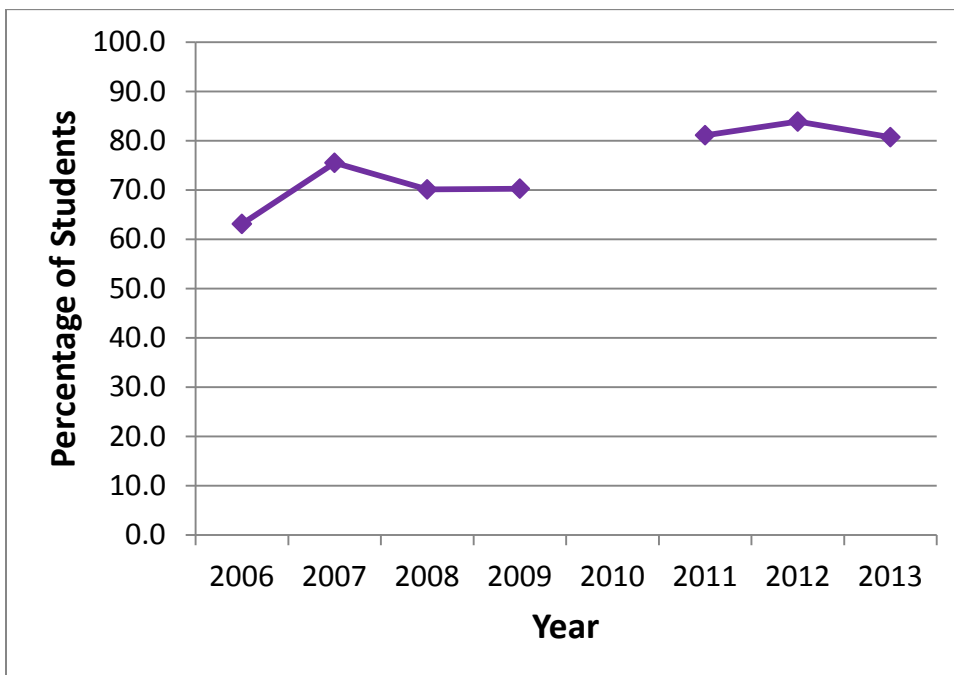


Figure 7. Percentage of Proficient and Advanced Students for Grade 6 Reading

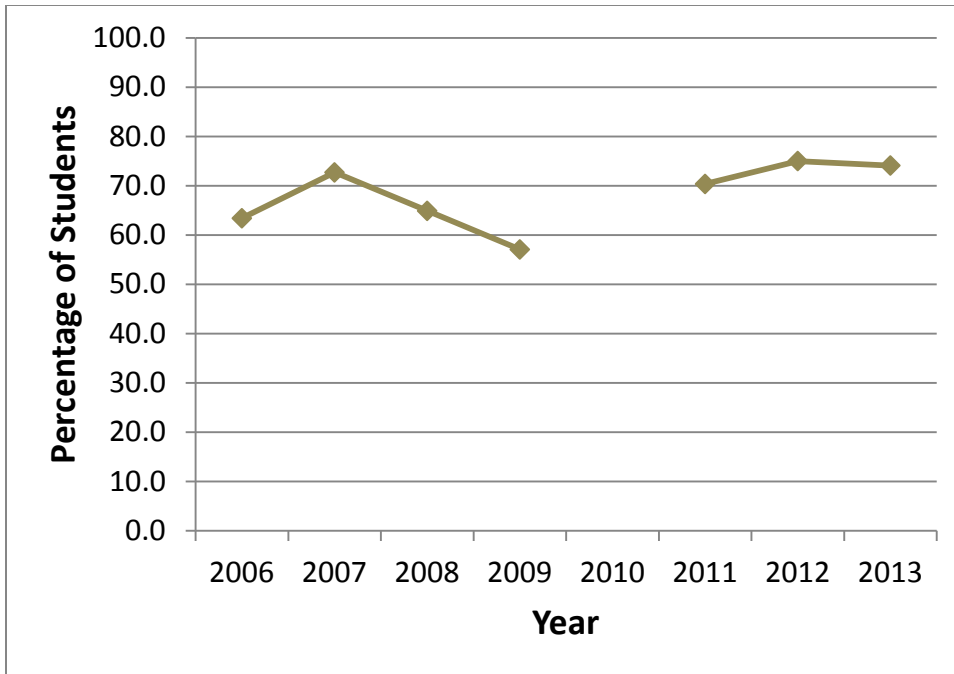


Figure 8. Percentage of Proficient and Advanced Students for Grade 7 Reading

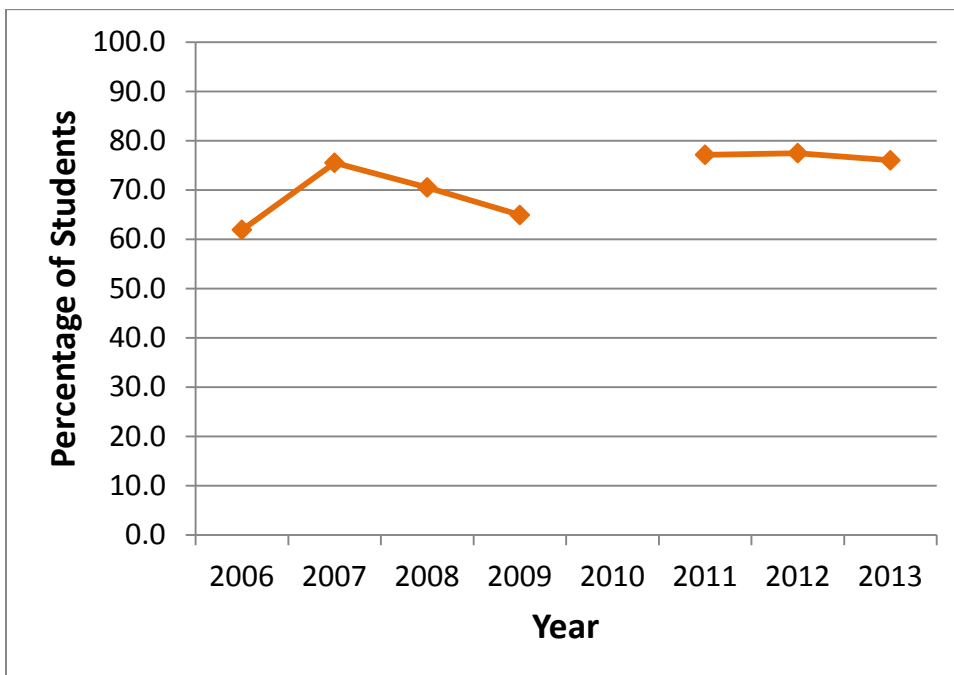


Figure 9. Percentage of Proficient and Advanced Students for Grade 8 Reading

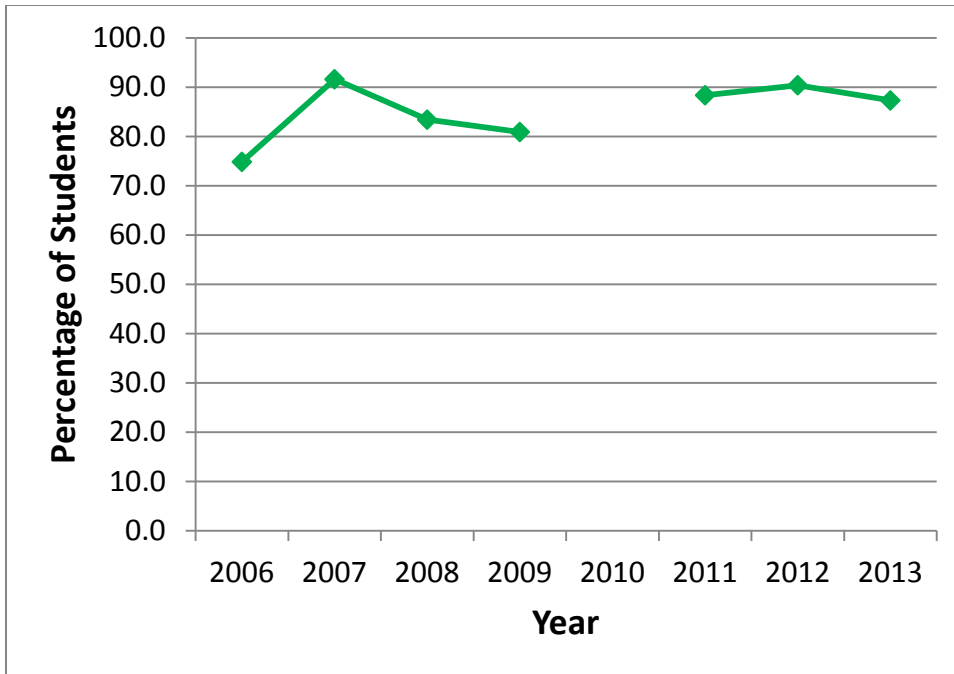


Figure 10. Percentage of Proficient and Advanced Students for Grade 3 Mathematics

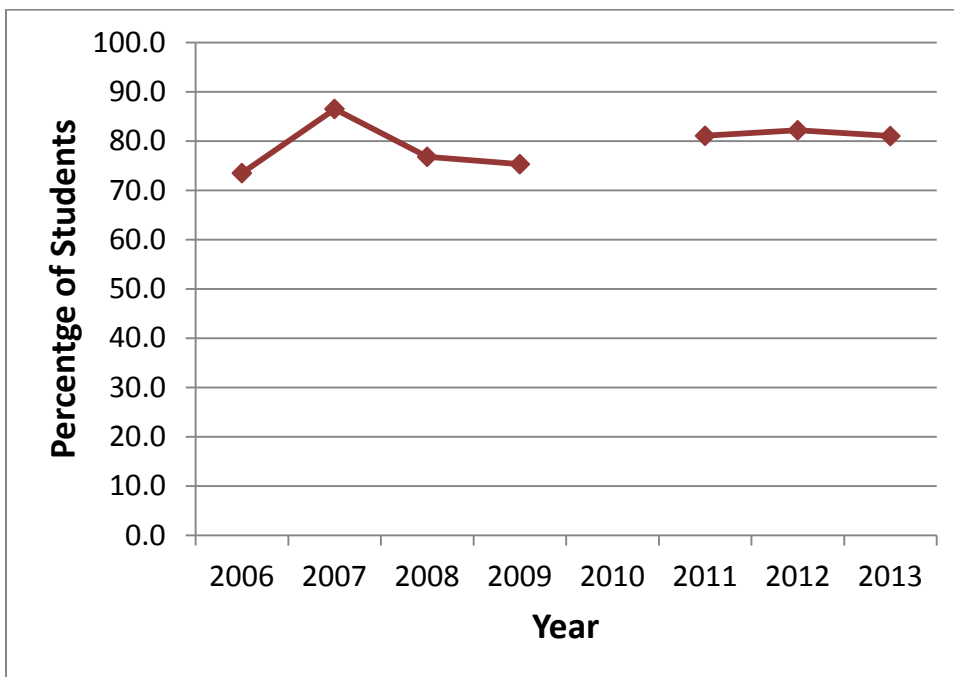


Figure 11. Percentage of Proficient and Advanced Students for Grade 4 Mathematics

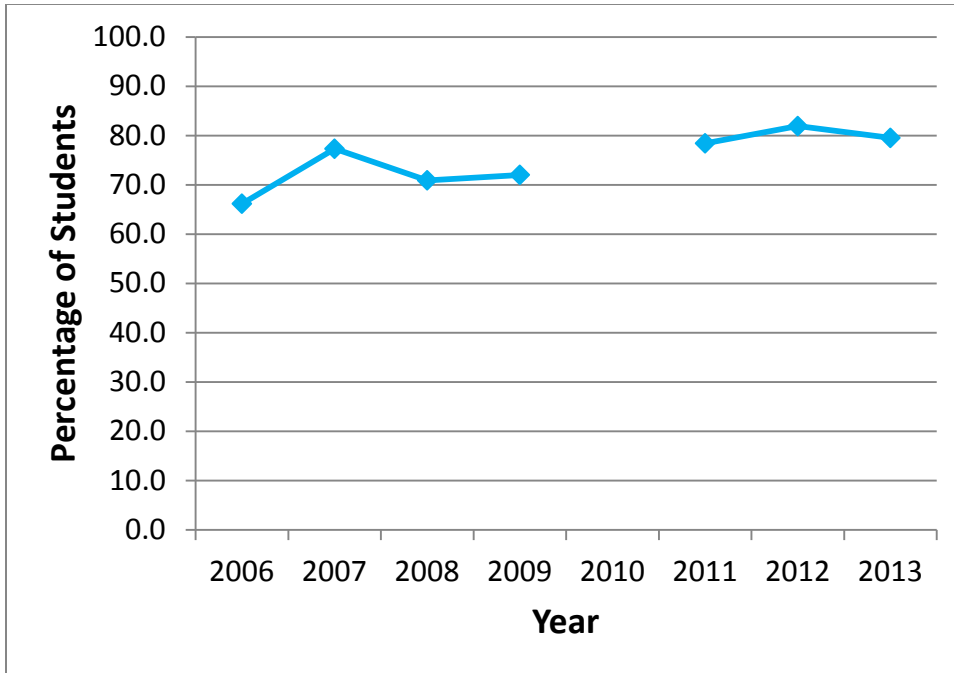


Figure 12. Percentage of Proficient and Advanced Students for Grade 5 Mathematics

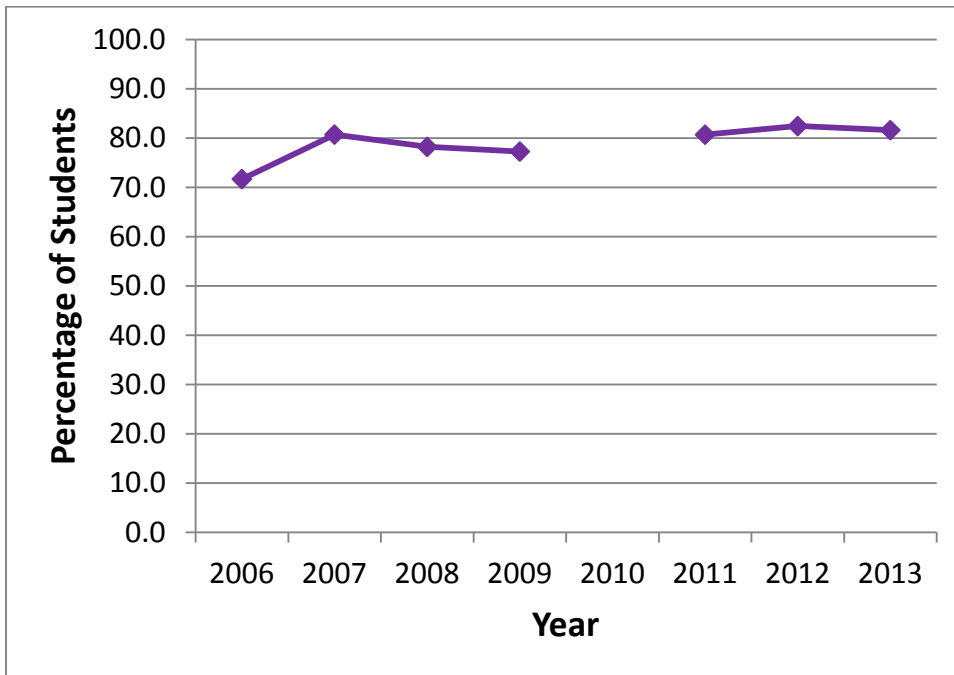


Figure 13. Percentage of Proficient and Advanced Students for Grade 6 Mathematics

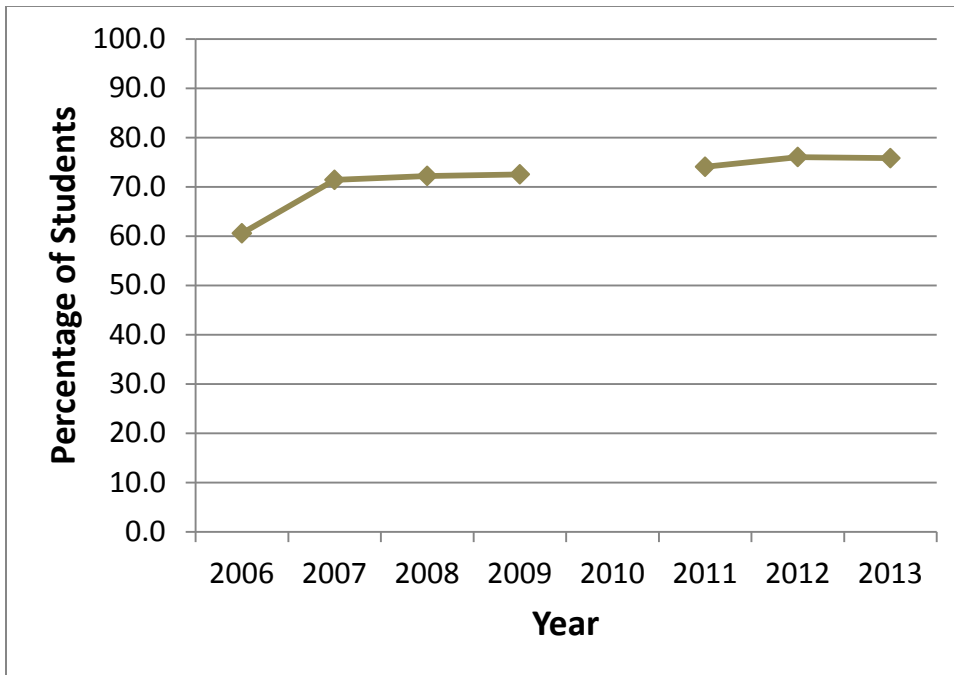


Figure 14. Percentage of Proficient and Advanced Students for Grade 7 Mathematics

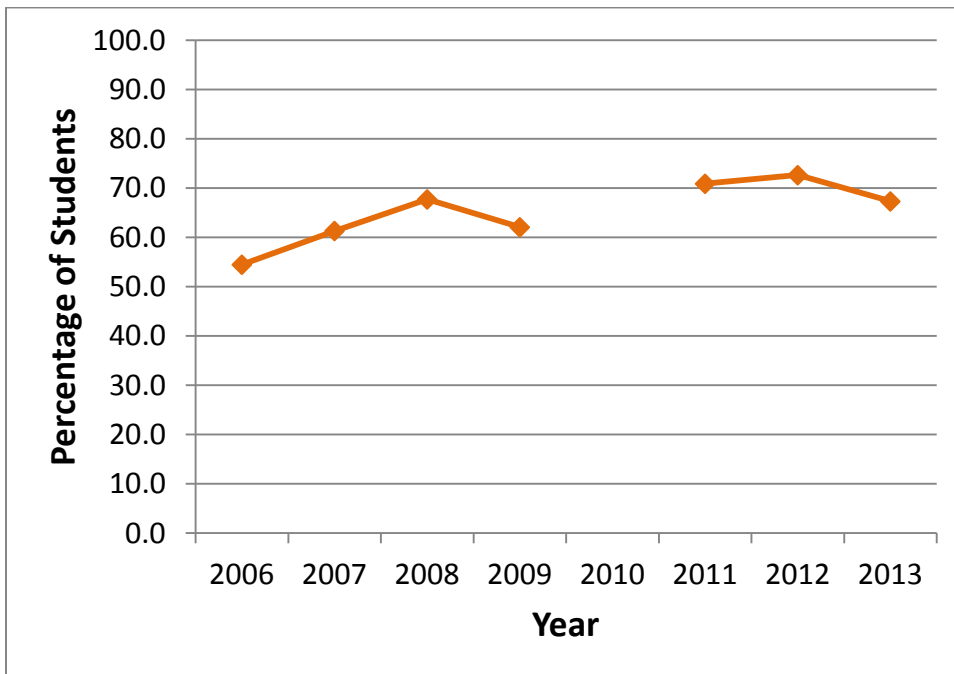


Figure 15. Percentage of Proficient and Advanced Students for Grade 8 Mathematics

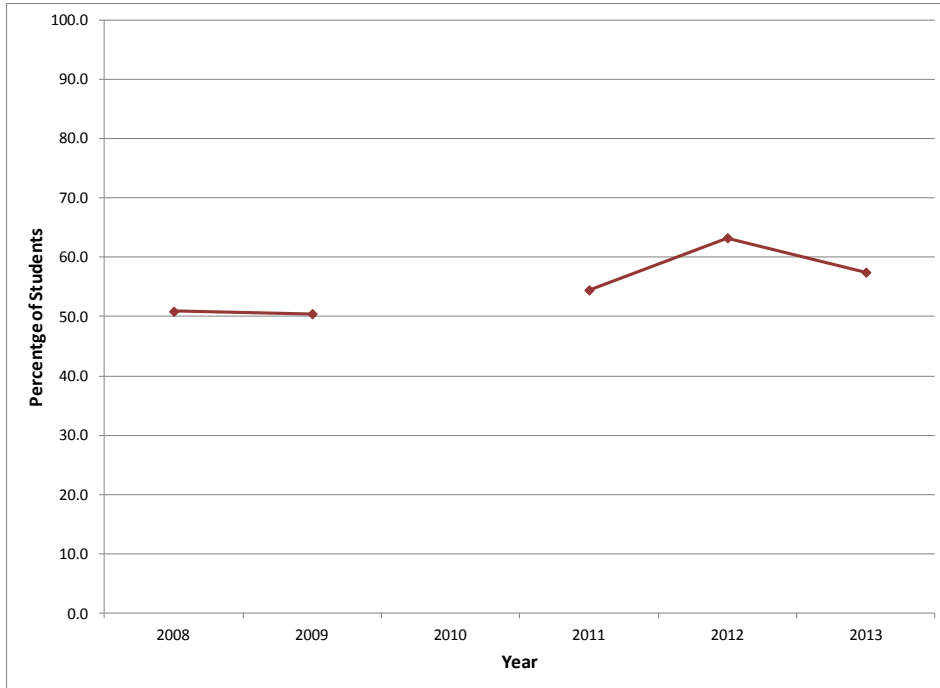


Figure 16. Percentage of Proficient and Advanced Students for Grade 4 Science

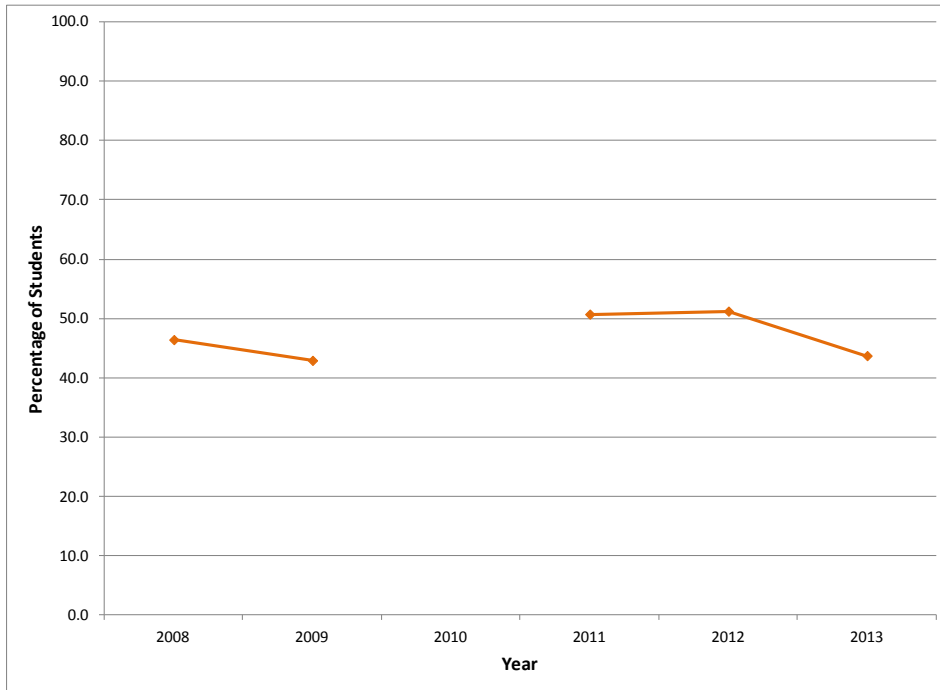


Figure 17. Percentage of Proficient and Advanced Students for Grade 8 Science

11. REFERENCES

- AERA, APA, & NCME (1999). *Standards for educational and psychological testing*. Washington, D.C.: Author.
- Andrich, A. (1988). *Rasch models for measurement*. Newbury Park, CA: SAGE Publications, Inc.
- Andrich, A. (1989). Distinctions between assumptions and requirements in measurement in the social Sciences, in J. A. Keats, R. Taft, R. A. Heath, & H. H. Lovibond (Eds.) *Mathematical and theoretical systems*. North-Holland: Elsevier Science Publisher B.V.
- Andrich, A., & Luo, G. (2004). *Modern measurement and analysis in social Science*. Murdoch University, Perth, Western Australia.
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: SAGE Publications.
- Chauncey, C. (2006). *In the news: proficiency for what?* Retrieved September 23, 2006, from the Harvard Education Letter website <http://www.edletter.org/inthenews/wyoming.shtml>
- Commission on Instructionally Supportive Assessment (2001). *Building Tests That Support Instruction and Accountability: A Guide for Policymakers*. Washington, D.C.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- Dorans, N. J., & Schmitt, A. P. (1991). *Constructed-response and differential item functioning: A pragmatic approach* (ETS Research Report No. 91-49). Princeton, NJ: Educational Testing Service.
- Embretson, S., & Reise, S. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Feldt, L.S., and Brennan, R.L. (1989). Reliability. In R. L. Linn (Ed.) *Educational Measurement* (3rd ed., pp.105–146). New York: Macmillan.
- Forte, E. and Popham, W. J. Red Light, Green Light. *Harvard Education Letter*, Vol. 22 (2), 2006, pp. 8, 6–7.
- Haertel, E. H. (1996). *Estimating the decision consistency from a single administration of a performance assessment battery. A report on the National Board of*

- Professional Teaching Standards McGEN Assessment*. Palo Alto, CA: Stanford University.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: SAGE Publications, Inc.
- Harvill, L. M. (1991). Standard error of measurement. *Educational Measurement: Issues and Practice*, 10, 181–189.
- Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9, 139–164.
- Harcourt Assessment. (2006). PAWS-ALT proficiency assessments for Wyoming students—Alternate 2006 directions for administering. Harcourt Assessment, Inc. San Antonio, TX.
- Harcourt Assessment (April, 2006). Proficiency assessment for Wyoming students (PAWS) technical report: Stand alone field test May 2005, Harcourt Assessment, Inc., San Antonio, TX.
- Huynh, H. & Meyer, P. (2010). Use of Robust z in Detecting Unstable Items in Item Response Theory Models. *Practical Assessment, Research & Evaluation*, 15(2). Retrieved July 13, 2013 from: <http://pareonline.net/getvn.asp?v=15&n=2>
- Karkee, T., Lewis, D. M., Barton, K. (2005). The effect of including or excluding students with testing accommodations on IRT calibrations. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.
- Kim, D., Wang, S., Zhao, Y., and Li, T. (2006). Validation and Invariance of Factor Structure of a Statewide Reading Comprehension Tests Under Accommodation and Non-Accommodation Conditions. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Kolen, M.J. & Brennan, R.L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). New York : Springer Science + Business Media.
- Linacre, J. M. (2007) *WINSTEPS Rasch measurement computer program and manual (PDF file)* v 3.64.2. Chicago: Winsteps.com
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179–197.
- Lord, F. M., & 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., & 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).
- Qualls, A. L. (1995). Estimating the reliability of a test containing multiple item formats, *Applied Measurement in Education*, 8, 111–120.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press.
- Rutherford, J.F. & Ahlgren, A. (1989). *Science for All Americans*. New York; 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*, San Francisco: Jossey-Bass.
- Suen, H.K. (1990). *Principles of Test Theories*, Lawrence Erlbaum, Hillsdale: NJ.
- Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, 51, 567–577.
- Tindal, G., & Fuchs, L. (1999). *A summary of research on test changes: An empirical basis for defining accommodations*. Lexington, KY: University of Kentucky, Mid-South Regional Resource Center.
- Wang, S., Jiao, H., Young, M. J., Brooks, T, & Olson, J. (2007). A Meta-Analysis of Testing Mode Effects in Grade K-12 Mathematics Tests. *Educational and Psychological Measurement*, 67, 219-238.
- Wang, S., Jiao, H., Young, M. J., Brooks, T, & Olson, J. (2008). Comparability of computer-based and paper-and-pencil testing in K–12 Reading assessments: A meta-analysis of testing mode effects. *Educational and Psychological Measurement*, 68, 5-24.

- Webb, N.L. (1999) *Alignment Between Standards and Assessment*, University of Wisconsin Center for Educational Research.
- Wyoming Department of Education (2001). *The Wyoming assessment handbook*. Cheyenne, WY: Author.
- Wyoming Department of Education (January, 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.
- Young, M.J. (2006). Vertical Scales, in S.M. Downing and T.M. Haladyna (eds.), *Handbook of test development*. Mahwah, N.J.: L. Erlbaum.
- Young, M. J., & Yoon, B. (1998, April). *Estimating the consistency and accuracy of classifications in a standards-referenced assessment*. (CSE Technical Report 475). Center for the Study of Evaluation, Standards, and Student Testing. Los Angeles, CA: University of California, Los Angeles.
- Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, 30, 233–251.
- Zwick, R., Thayer, D.T., Lewis, C. (1999) An Empirical Bayes Approach to Mantel-Haenszel DIF Analysis. *Journal of Educational Measurement*, 36, 1, 1–28.

12. GLOSSARY OF TERMS

The terms below are defined by their application in this document and their common uses in the Wyoming PAWS technical manual. 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, and Science 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 scorers assign scores consistently.

Answer Document - The form or document on which a student records answers to constructed response questions. Usually these are scannable and have grids for recording student name and demographic information.

Benchmarks - These statements specify what students are expected to know and 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 a blueprint, 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.

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

Content Standards - These statements define what students are expected to know and 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) - Is a procedure for helping detect if an item is unfair toward a particular group of test takers. In other words, 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). Equating produces a Raw Score-to-Scale Score conversion table.

Equivalent forms - Statistically insignificant differences between forms (i.e., the B form is not harder).

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 collection of items to approximate how a test form will work. Statistics produced will be used in interpreting item behavior/performance and allow for the calibration of item parameters used in equating tests.

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

Inter-Rater Reliability - A method of measuring the agreement among readers scoring the same responses. Computer programs compare the scores assigned by one reader to those of another for the same student. 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 percent of students responding to each answer choice as well as p-values, point-biserials, logit difficulties, theta, and DIF statistics for the items.

Item Bank - An item bank is a collection of test items in various stages of review, 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 parameters.

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 also 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., .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). This procedure is used to identify individual items for further 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 - Covers the same curricular material as other forms

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. The proficient level is required to meet the standards. These descriptors help teachers to 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 readers 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) as well as person ability (logit ability). It is used for scaling and equating test forms as well as producing item analysis.

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 .90's are considered good. In handscoring, reliability refers to agreement between readers 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.

Scale 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 dispersion of test scores around the mean. 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 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 specs ensures that equal 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.

Traffic Signal Report - Teachers receive per-skill category mastery reports for each of the Wyoming Content Standards in Reading, Science, and Mathematics. Therefore, teachers are able to discern the extent to which students have mastered all aspects of the Wyoming Content and Performance Standards.

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.

Appendix A: Classical Item Statistics for 2013 Field Test Items

Reading

Table A1. Reading Grade 3 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF389165	792	0.66	0.51	-0.40	0.51	-0.19	-0.16
VF389470	792	0.71	0.46	-0.25	-0.29	0.46	-0.15
VF389474	792	0.58	0.40	-0.25	-0.15	0.40	-0.15
VF389951	792	0.23	0.24	-0.11	0.02	-0.19	0.24
VF389473	792	0.56	0.33	-0.15	0.33	-0.19	-0.12
VF389457	792	0.44	0.26	-0.12	0.26	-0.10	-0.14
VF493383	792	0.77	0.46	-0.35	-0.13	-0.22	0.46
VF494661	792	0.68	0.48	-0.25	-0.16	-0.31	0.48
VF494909	792	0.50	0.35	-0.22	-0.21	-0.05	0.35
VF494764	792	0.72	0.48	-0.29	0.48	-0.22	-0.22
VF494944	792	0.38	0.23	-0.10	0.23	-0.23	0.03
VF494745	792	0.46	0.40	-0.15	0.40	-0.22	-0.18
Form 2							
VF389467	701	0.70	0.35	0.35	-0.23	-0.20	-0.19
VF389446	701	0.62	0.31	0.31	-0.32	-0.03	-0.20
VF389477	701	0.68	0.44	-0.15	-0.29	-0.23	0.44
VF389949	701	0.72	0.39	-0.24	-0.23	-0.13	0.39
VF389620	701	0.67	0.38	-0.22	-0.21	0.38	-0.19
VF389482	701	0.37	0.15	0.15	0.00	-0.08	-0.08
VF493480	701	0.71	0.44	-0.26	0.44	-0.24	-0.22
VF494098	701	0.68	0.46	-0.21	-0.28	0.46	-0.21
VF494956	701	0.55	0.42	-0.23	-0.22	0.42	-0.13
VF494732	701	0.83	0.52	-0.30	-0.30	-0.25	0.52
VF494759	701	0.78	0.54	0.54	-0.20	-0.34	-0.31
VF494915	701	0.73	0.45	-0.26	-0.31	0.45	-0.15
Form 3							
VF394046	694	0.85	0.49	-0.35	-0.27	0.49	-0.20
VF394049	694	0.73	0.39	-0.28	-0.26	0.39	-0.12
VF394057	694	0.78	0.34	0.34	-0.13	-0.27	-0.19
VF394054	694	0.90	0.38	-0.18	-0.26	-0.19	0.38
VF394052	694	0.70	0.27	0.27	-0.30	0.03	-0.25
VF394045	694	0.59	0.28	-0.15	0.28	-0.20	-0.08
VF389093	694	0.43	0.15	0.15	-0.03	-0.05	-0.12
VF389113	694	0.75	0.43	-0.22	-0.24	0.43	-0.21
VF389095	694	0.67	0.51	-0.28	0.51	-0.25	-0.24
VF389115	694	0.37	0.27	-0.08	-0.12	0.27	-0.21
VF389100	694	0.78	0.50	-0.37	-0.23	0.50	-0.16
VF389116	694	0.53	0.29	-0.15	0.29	-0.22	-0.05

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF394050	701	0.80	0.33	-0.24	0.33	-0.12	-0.18
VF394051	701	0.76	0.21	-0.17	-0.05	-0.14	0.21
VF394055	701	0.28	0.01	0.01	0.16	-0.15	-0.25
VF394056	701	0.57	0.26	-0.21	0.26	-0.24	0.04
VF394053	701	0.78	0.41	-0.23	-0.31	-0.11	0.41
VF394041	701	0.66	0.26	0.26	-0.20	-0.27	-0.07
VF389082	701	0.69	0.51	0.51	-0.32	-0.28	-0.19
VF389120	701	0.44	0.27	-0.07	0.27	-0.13	-0.18
VF389105	701	0.35	0.05	-0.11	-0.06	0.08	0.05
VF389074	701	0.43	0.31	0.31	0.03	-0.19	-0.32
VF389123	701	0.66	0.40	0.40	-0.29	-0.18	-0.17
VF389119	701	0.45	0.30	-0.10	-0.15	-0.16	0.30
Form 5							
VF497812	726	0.69	0.21	-0.05	0.21	-0.14	-0.20
VF497773	726	0.82	0.43	-0.29	-0.18	-0.22	0.43
VF497775	726	0.45	0.17	0.17	-0.12	-0.21	0.01
VF497822	726	0.42	0.29	0.29	-0.16	-0.21	-0.03
VF497827	726	0.60	0.38	-0.33	-0.13	0.38	-0.16
VF497820	726	0.74	0.42	-0.23	-0.28	-0.17	0.42
VF387075	726	0.54	0.26	-0.31	0.00	0.26	-0.21
VF387072	726	0.81	0.45	0.45	-0.17	-0.22	-0.35
VF387055	726	0.79	0.47	-0.33	-0.21	0.47	-0.21
VF387084	726	0.79	0.44	-0.29	0.44	-0.22	-0.22
VF387096	726	0.52	0.28	0.28	0.03	-0.22	-0.32
VF387089	726	0.73	0.48	-0.25	-0.32	-0.16	0.48
Form 6							
VF497783	704	0.85	0.50	-0.42	-0.20	0.50	-0.11
VF497793	704	0.71	0.45	-0.29	-0.22	-0.17	0.45
VF497818	704	0.71	0.47	-0.20	-0.34	-0.18	0.47
VF497781	704	0.62	0.47	-0.25	0.47	-0.28	-0.17
VF497807	704	0.68	0.41	-0.22	-0.27	0.41	-0.14
VF497815	704	0.70	0.36	0.36	-0.24	-0.17	-0.13
VF387092	704	0.55	0.35	-0.09	0.35	-0.27	-0.20
VF387071	704	0.62	0.22	0.22	-0.02	-0.09	-0.26
VF387050	704	0.50	0.26	0.26	-0.15	-0.23	0.00
VF387069	704	0.81	0.38	-0.29	0.38	-0.17	-0.17
VF387079	704	0.73	0.52	-0.30	-0.21	-0.29	0.52
VF387067	704	0.51	0.08	0.09	-0.08	-0.28	0.08

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF497676	696	0.75	0.41	0.41	-0.24	-0.28	-0.14
VF497660	696	0.82	0.42	-0.21	-0.20	0.42	-0.27
VF497690	696	0.61	0.33	-0.16	0.33	-0.16	-0.18
VF497705	696	0.40	0.26	0.26	-0.22	-0.18	-0.05
VF497668	696	0.74	0.40	-0.28	0.40	-0.21	-0.17
VF497671	696	0.60	0.43	-0.19	-0.24	0.43	-0.28
VF497573	696	0.91	0.49	-0.31	-0.26	-0.23	0.49
VF497419	696	0.59	0.34	-0.07	-0.25	0.34	-0.18
VF497594	696	0.64	0.48	-0.25	-0.26	0.48	-0.18
VF497592	696	0.66	0.54	-0.25	0.54	-0.31	-0.27
VF497611	696	0.46	0.40	0.40	-0.16	-0.15	-0.19
VF497614	696	0.46	0.36	0.36	-0.18	-0.15	-0.14
Form 8							
VF497684	701	0.51	0.35	0.35	-0.03	-0.23	-0.28
VF497700	701	0.61	0.39	-0.17	-0.11	-0.27	0.39
VF497665	701	0.35	0.06	0.17	-0.24	0.06	-0.26
VF497695	701	0.76	0.34	-0.25	0.34	-0.19	-0.14
VF497696	701	0.72	0.39	-0.15	-0.26	-0.22	0.39
VF497683	701	0.80	0.46	0.46	-0.21	-0.24	-0.28
VF497424	701	0.66	0.32	-0.27	0.32	-0.25	-0.11
VF497577	701	0.85	0.53	-0.30	-0.30	-0.28	0.53
VF497599	701	0.57	0.26	-0.12	-0.23	0.26	-0.04
VF497585	701	0.90	0.52	-0.31	-0.25	-0.30	0.52
VF497606	701	0.65	0.40	-0.23	-0.30	0.40	-0.10
VF497566	701	0.87	0.41	0.41	-0.27	-0.14	-0.25
Form 9							
VF394173	708	0.65	0.27	-0.28	0.27	-0.08	-0.22
VF394165	708	0.72	0.52	-0.33	-0.24	-0.24	0.52
VF394143	708	0.55	0.32	0.32	-0.17	-0.14	-0.15
VF394141	708	0.82	0.48	-0.23	0.48	-0.22	-0.32
VF394157	708	0.61	0.33	-0.25	-0.17	0.33	-0.04
VF394137	708	0.64	0.41	-0.23	0.41	-0.16	-0.21
VF497718	708	0.77	0.22	0.22	-0.08	-0.16	-0.11
VF497716	708	0.60	0.32	-0.27	-0.19	-0.04	0.32
VF497759	708	0.49	0.39	-0.23	0.39	-0.13	-0.18
VF497758	708	0.59	0.41	-0.28	-0.08	-0.22	0.41
VF497761	708	0.72	0.51	-0.41	-0.19	0.51	-0.15
VF497766	708	0.43	0.19	0.19	-0.20	-0.08	0.02

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF524523	718	0.82	0.42	-0.27	-0.24	0.42	-0.20
VF394145	718	0.68	0.43	-0.18	0.43	-0.25	-0.21
VF394134	718	0.50	0.28	-0.29	-0.27	0.28	0.04
VF394154	718	0.83	0.41	0.41	-0.16	-0.27	-0.25
VF394153	718	0.76	0.40	-0.20	-0.21	0.40	-0.23
VF394151	718	0.53	0.45	-0.06	-0.28	-0.34	0.45
VF497731	718	0.86	0.50	-0.32	-0.27	-0.25	0.50
VF497728	718	0.42	0.23	-0.13	0.23	-0.03	-0.13
VF522349	718	0.54	0.40	-0.19	0.40	-0.23	-0.19
VF497751	718	0.60	0.35	-0.11	-0.16	-0.31	0.35
VF497767	718	0.53	0.38	-0.31	0.38	-0.09	-0.12
VF497725	718	0.56	0.44	-0.22	-0.28	0.44	-0.13

Table A2. Reading Grade 4 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	<u>Option Discrimination (MC)</u>			
				A	B	C	D
Form 1							
VF495010	792	0.95	0.38	-0.26	-0.25	-0.11	0.38
VF495021	792	0.69	0.27	-0.27	0.27	-0.03	-0.12
VF495028	792	0.47	0.22	0.01	0.22	-0.23	-0.13
VF495001	792	0.83	0.45	-0.21	-0.31	0.45	-0.23
VF495647	792	0.49	0.09	0.09	0.09	-0.09	-0.21
VF494988	792	0.60	0.42	-0.15	-0.29	0.42	-0.17
VF497327	792	0.73	0.44	-0.25	-0.26	0.44	-0.19
VF497297	792	0.88	0.46	-0.25	-0.23	-0.31	0.46
VF497314	792	0.73	0.52	-0.26	-0.25	-0.31	0.52
VF497322	792	0.61	0.44	0.44	-0.21	-0.25	-0.19
VF497334	792	0.70	0.36	-0.20	0.36	-0.20	-0.20
VF497311	792	0.41	0.21	-0.05	0.21	-0.13	-0.13
Form 2							
VF495015	705	0.77	0.44	-0.17	-0.30	0.44	-0.22
VF494997	705	0.90	0.42	-0.26	0.42	-0.15	-0.28
VF495644	705	0.71	0.25	-0.27	-0.25	0.25	0.01
VF495003	705	0.63	0.45	-0.16	-0.22	-0.36	0.45
VF495030	705	0.73	0.12	0.12	-0.24	0.13	-0.24
VF494993	705	0.55	0.41	-0.21	-0.11	-0.25	0.41
VF497330	705	0.72	0.44	-0.25	-0.22	-0.27	0.44
VF497326	705	0.61	0.44	-0.25	0.44	-0.13	-0.27
VF497338	705	0.64	0.23	-0.26	0.23	-0.09	-0.05
VF497307	705	0.63	0.37	-0.19	-0.16	0.37	-0.17
VF497318	705	0.79	0.35	-0.32	-0.28	0.35	-0.07
VF497303	705	0.79	0.46	-0.14	0.46	-0.33	-0.22
Form 3							
VF497159	698	0.89	0.39	-0.32	-0.23	0.39	-0.06
VF497212	698	0.72	0.35	-0.14	-0.23	0.35	-0.21
VF497147	698	0.87	0.42	-0.19	-0.33	-0.16	0.42
VF497155	698	0.70	0.41	0.41	-0.25	-0.21	-0.22
VF497167	698	0.61	0.48	-0.33	-0.17	0.48	-0.21
VF497215	698	0.37	0.31	-0.07	-0.23	-0.20	0.31
VF497279	698	0.66	0.26	0.26	-0.15	-0.26	-0.09
VF497265	698	0.90	0.37	-0.13	0.37	-0.21	-0.26
VF497253	698	0.67	0.19	0.19	-0.15	-0.17	0.03
VF497233	698	0.68	0.30	0.30	-0.17	-0.14	-0.19
VF497250	698	0.50	0.28	-0.06	-0.13	-0.27	0.28
VF497256	698	0.85	0.43	-0.27	-0.20	-0.24	0.43

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF497188	707	0.95	0.41	-0.24	0.41	-0.29	-0.15
VF497206	707	0.93	0.34	-0.23	0.34	-0.15	-0.19
VF497195	707	0.54	0.27	0.27	-0.09	-0.11	-0.17
VF497162	707	0.44	0.33	-0.13	0.33	-0.31	-0.01
VF497151	707	0.46	0.23	-0.16	0.02	0.23	-0.21
VF497220	707	0.72	0.38	-0.16	-0.17	0.38	-0.31
VF497243	707	0.71	0.25	-0.10	-0.09	-0.25	0.25
VF497227	707	0.90	0.35	-0.20	0.35	-0.19	-0.21
VF497270	707	0.43	0.24	-0.13	-0.19	0.24	-0.02
VF497261	707	0.66	0.20	-0.05	-0.16	0.20	-0.15
VF497286	707	0.70	0.29	-0.19	0.29	-0.06	-0.24
VF497247	707	0.66	0.37	-0.17	0.37	-0.22	-0.22
Form 5							
VF496940	713	0.80	0.31	0.31	-0.12	-0.22	-0.18
VF496899	713	0.69	0.35	-0.16	-0.24	0.35	-0.18
VF496934	713	0.45	0.05	0.19	0.05	-0.12	-0.21
VF496894	713	0.57	0.40	-0.15	0.40	-0.26	-0.20
VF496883	713	0.53	0.29	-0.28	-0.04	0.29	-0.09
VF496921	713	0.49	0.14	-0.15	0.07	0.14	-0.22
VF495798	713	0.68	0.45	-0.28	-0.28	-0.13	0.45
VF495794	713	0.61	0.32	0.32	-0.12	-0.10	-0.27
VF495832	713	0.51	0.23	-0.18	-0.12	0.23	-0.04
VF495856	713	0.44	0.34	0.34	-0.16	-0.15	-0.13
VF495873	713	0.52	0.25	0.25	-0.25	-0.14	-0.02
VF495789	713	0.29	0.05	-0.04	-0.01	0.05	-0.02
Form 6							
VF496202	718	0.58	0.25	0.25	-0.13	-0.22	-0.08
VF496219	718	0.78	0.51	-0.29	0.51	-0.28	-0.26
VF496270	718	0.74	0.32	-0.28	-0.15	-0.07	0.32
VF496902	718	0.45	0.08	-0.27	-0.12	0.22	0.08
VF496864	718	0.75	0.43	-0.30	-0.14	0.43	-0.24
VF496917	718	0.42	0.31	-0.21	-0.11	-0.05	0.31
VF495786	718	0.90	0.43	-0.27	-0.29	0.43	-0.16
VF495812	718	0.49	0.36	0.36	-0.15	-0.20	-0.16
VF495868	718	0.51	0.36	0.36	-0.22	-0.12	-0.15
VF495823	718	0.41	0.25	-0.30	-0.03	0.25	-0.06
VF495843	718	0.60	0.46	-0.23	-0.18	0.46	-0.27
VF495775	718	0.54	0.31	-0.18	0.31	-0.25	-0.07

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF495975	714	0.34	0.25	-0.21	0.25	-0.08	-0.06
VF495970	714	0.76	0.45	0.45	-0.22	-0.23	-0.26
VF495987	714	0.53	0.24	-0.12	-0.09	-0.13	0.24
VF496010	714	0.41	0.14	0.03	0.14	-0.04	-0.16
VF495955	714	0.82	0.45	0.45	-0.28	-0.23	-0.22
VF496016	714	0.42	0.19	-0.20	0.03	0.19	-0.15
VF407282	714	0.88	0.44	0.44	-0.23	-0.32	-0.16
VF407237	714	0.80	0.31	0.31	-0.19	-0.21	-0.19
VF407298	714	0.64	0.28	-0.16	0.28	-0.17	-0.11
VF407232	714	0.86	0.49	-0.38	-0.24	0.49	-0.14
VF407247	714	0.89	0.45	-0.33	-0.19	0.45	-0.20
VF407239	714	0.57	0.31	-0.20	-0.06	-0.18	0.31
Form 8							
VF495963	699	0.76	0.27	0.27	-0.12	-0.23	-0.15
VF495947	699	0.38	0.13	0.13	0.08	-0.12	-0.18
VF496019	699	0.84	0.44	-0.30	0.44	-0.25	-0.16
VF495994	699	0.71	0.27	-0.18	0.27	-0.07	-0.19
VF495982	699	0.67	0.31	-0.14	-0.10	-0.21	0.31
VF495944	699	0.80	0.43	-0.26	-0.22	0.43	-0.22
VF407250	699	0.79	0.39	-0.14	-0.29	0.39	-0.20
VF407243	699	0.72	0.37	0.37	-0.26	-0.26	-0.11
VF407287	699	0.86	0.40	-0.35	0.40	-0.10	-0.15
VF407295	699	0.90	0.33	0.33	-0.18	-0.15	-0.23
VF407297	699	0.63	0.32	-0.14	-0.30	-0.11	0.32
VF407235	699	0.60	0.25	-0.31	-0.04	0.25	-0.02
Form 9							
VF494842	1412	0.78	0.24	0.24	-0.01	-0.26	-0.24
VF497387	696	0.86	0.42	-0.28	0.42	-0.26	-0.16
VF497361	696	0.85	0.43	-0.11	-0.17	0.43	-0.35
VF497390	696	0.77	0.51	-0.32	-0.23	-0.21	0.51
VF497365	696	0.47	0.20	-0.14	0.20	0.05	-0.25
VF497378	696	0.40	0.28	-0.08	0.28	-0.19	-0.10
VF497356	696	0.33	0.28	-0.11	0.28	-0.09	-0.14
VF494863	696	0.87	0.43	-0.25	0.43	-0.23	-0.25
VF494852	696	0.80	0.46	-0.25	-0.30	0.46	-0.23
VF494927	696	0.51	0.25	-0.17	0.25	-0.27	-0.03
VF494954	696	0.84	0.48	-0.38	-0.18	0.48	-0.16
VF494964	696	0.81	0.37	-0.20	-0.17	-0.28	0.37

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF497381	716	0.84	0.49	0.49	-0.36	-0.16	-0.26
VF497384	716	0.53	0.35	-0.14	0.35	-0.12	-0.21
VF497359	716	0.61	0.39	-0.22	-0.23	-0.16	0.39
VF497362	716	0.63	0.37	-0.24	-0.17	-0.18	0.37
VF497396	716	0.38	0.14	0.02	-0.07	0.14	-0.15
VF497354	716	0.86	0.37	0.37	-0.20	-0.27	-0.19
VF494842	1412	0.78	0.24	0.24	-0.01	-0.26	-0.24
VF494937	716	0.89	0.42	-0.26	0.42	-0.24	-0.21
VF494856	716	0.80	0.30	0.30	-0.21	-0.14	-0.15
VF494914	716	0.69	0.41	-0.26	-0.10	0.41	-0.34
VF494907	716	0.63	0.41	-0.13	-0.21	-0.30	0.41
VF494969	716	0.59	0.37	0.37	-0.20	-0.28	-0.08

Table A3. Reading Grade 5 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF495685	751	0.20	0.11	0.09	-0.05	-0.25	0.11
VF495826	751	0.35	0.14	0.14	-0.14	-0.10	-0.02
VF495879	751	0.35	0.14	-0.27	0.16	0.14	-0.11
VF495811	751	0.80	0.48	-0.20	-0.22	-0.34	0.48
VF495906	751	0.72	0.45	-0.08	-0.32	0.45	-0.29
VF495818	751	0.67	0.48	-0.36	0.48	-0.18	-0.14
VF497016	751	0.74	0.44	-0.22	-0.29	0.44	-0.18
VF497025	751	0.65	0.49	-0.23	-0.28	-0.24	0.49
VF497001	751	0.44	0.18	-0.03	0.18	-0.10	-0.10
VF497037	751	0.82	0.50	-0.21	0.50	-0.32	-0.27
VF497032	751	0.36	0.19	0.09	-0.14	-0.26	0.19
VF497039	751	0.56	0.42	-0.20	-0.20	-0.18	0.42
Form 2							
VF495793	659	0.74	0.42	-0.20	-0.23	0.42	-0.23
VF495802	659	0.87	0.47	-0.32	-0.22	0.47	-0.26
VF495840	659	0.77	0.52	0.52	-0.25	-0.27	-0.33
VF495886	659	0.29	0.13	-0.11	0.13	0.05	-0.26
VF495831	659	0.83	0.48	-0.28	-0.23	0.48	-0.28
VF495909	659	0.81	0.41	-0.21	0.41	-0.27	-0.19
VF497012	659	0.68	0.31	0.31	-0.16	-0.13	-0.21
VF497020	659	0.68	0.32	-0.12	-0.19	-0.18	0.32
VF497027	659	0.30	0.13	-0.02	-0.17	0.13	0.02
VF497009	659	0.56	0.39	-0.19	-0.20	0.39	-0.16
VF497028	659	0.50	0.46	0.46	-0.13	-0.27	-0.23
VF497030	659	0.63	0.40	0.40	-0.11	-0.21	-0.31
Form 3							
VF497274	668	0.46	0.38	-0.21	0.38	-0.17	-0.26
VF497277	668	0.44	0.35	0.35	-0.19	-0.16	-0.13
VF497287	668	0.73	0.43	-0.25	-0.24	0.43	-0.22
VF497280	668	0.89	0.26	-0.09	-0.17	-0.17	0.26
VF497282	668	0.46	0.15	0.15	-0.18	-0.20	0.11
VF497283	668	0.69	0.43	-0.27	-0.17	0.43	-0.22
VF496869	668	0.44	0.40	0.40	-0.05	-0.31	-0.29
VF496889	668	0.77	0.44	-0.36	-0.19	0.44	-0.16
VF496876	668	0.57	0.16	-0.22	0.16	-0.17	0.04
VF496884	668	0.74	0.48	-0.27	0.48	-0.24	-0.23
VF496878	668	0.46	0.24	-0.30	-0.08	0.24	-0.03
VF496875	668	0.78	0.50	-0.23	-0.26	0.50	-0.31

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF497272	670	0.43	0.14	0.14	-0.13	-0.02	-0.04
VF497288	670	0.74	0.37	-0.22	0.37	-0.15	-0.23
VF497284	670	0.76	0.43	-0.18	0.43	-0.30	-0.28
VF497278	670	0.80	0.52	0.52	-0.20	-0.31	-0.35
VF497273	670	0.87	0.33	-0.27	-0.15	0.33	-0.16
VF497285	670	0.38	0.35	0.35	-0.09	-0.09	-0.28
VF496871	670	0.85	0.27	-0.16	0.27	-0.17	-0.15
VF496886	670	0.79	0.34	-0.16	-0.15	0.34	-0.24
VF496887	670	0.84	0.47	0.47	-0.25	-0.29	-0.24
VF496882	670	0.72	0.35	0.35	-0.22	-0.18	-0.15
VF496872	670	0.68	0.34	-0.22	-0.08	0.34	-0.23
VF496868	670	0.66	0.29	-0.08	-0.23	0.29	-0.18
Form 5							
VF407329	677	0.79	0.44	-0.24	-0.26	0.44	-0.22
VF407357	677	0.94	0.22	0.22	-0.14	-0.10	-0.13
VF407332	677	0.78	0.28	-0.13	-0.16	0.28	-0.17
VF407338	677	0.60	0.37	0.37	-0.14	-0.24	-0.16
VF407386	677	0.76	0.41	-0.26	-0.24	-0.18	0.41
VF407336	677	0.72	0.29	0.29	-0.14	-0.15	-0.23
VF497226	677	0.89	0.37	-0.21	-0.23	0.37	-0.17
VF497237	677	0.66	0.40	-0.24	0.40	-0.16	-0.18
VF497231	677	0.68	0.38	0.38	-0.16	-0.24	-0.15
VF497241	677	0.75	0.44	-0.25	-0.23	-0.26	0.44
VF497251	677	0.28	0.17	-0.24	0.17	0.15	-0.08
VF497240	677	0.56	0.23	-0.10	-0.13	0.23	-0.09
Form 6							
VF407322	670	0.67	0.30	0.30	-0.23	-0.19	-0.06
VF407355	670	0.83	0.34	-0.25	-0.21	0.34	-0.16
VF407360	670	0.90	0.27	0.27	-0.14	-0.20	-0.12
VF407321	670	0.87	0.34	0.34	-0.16	-0.20	-0.21
VF407319	670	0.63	0.32	-0.08	-0.21	-0.19	0.32
VF407388	670	0.72	0.37	-0.28	0.37	-0.17	-0.17
VF497228	670	0.49	0.33	0.33	-0.12	-0.20	-0.12
VF497230	670	0.79	0.25	-0.10	0.25	-0.17	-0.16
VF497232	670	0.53	0.44	-0.21	-0.17	-0.23	0.44
VF497248	670	0.60	0.28	-0.18	0.28	-0.06	-0.16
VF497234	670	0.49	0.42	-0.20	-0.08	0.42	-0.28
VF497254	670	0.54	0.33	0.33	-0.15	-0.12	-0.19

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF497055	671	0.65	0.38	-0.18	-0.20	-0.21	0.38
VF497060	671	0.88	0.42	-0.22	-0.24	0.42	-0.23
VF497181	671	0.74	0.39	-0.28	-0.15	-0.24	0.39
VF497056	671	0.70	0.36	-0.25	0.36	-0.17	-0.11
VF497177	671	0.69	0.48	0.48	-0.18	-0.26	-0.31
VF497184	671	0.61	0.34	-0.17	-0.16	0.34	-0.20
VF496048	671	0.66	0.32	0.32	-0.09	-0.23	-0.20
VF496115	671	0.92	0.42	-0.19	-0.28	0.42	-0.24
VF496193	671	0.31	0.13	-0.20	-0.07	0.13	0.07
VF496185	671	0.67	0.30	0.30	-0.17	-0.14	-0.16
VF496032	671	0.61	0.37	0.37	-0.22	-0.15	-0.17
VF496195	671	0.49	0.29	-0.12	-0.27	-0.06	0.29
Form 8							
VF497052	662	0.54	0.21	0.21	-0.21	-0.13	0.08
VF497170	662	0.85	0.35	-0.11	0.35	-0.19	-0.25
VF497174	662	0.91	0.38	-0.23	-0.30	-0.10	0.38
VF497182	662	0.88	0.37	-0.21	-0.14	0.37	-0.27
VF497172	662	0.78	0.43	-0.29	-0.20	0.43	-0.18
VF497176	662	0.48	0.29	-0.06	-0.19	-0.22	0.29
VF496024	662	0.74	0.51	-0.34	-0.27	-0.21	0.51
VF496188	662	0.69	0.33	-0.18	-0.19	0.33	-0.13
VF496101	662	0.72	0.40	-0.19	-0.28	-0.19	0.40
VF496068	662	0.75	0.40	0.40	-0.09	-0.29	-0.27
VF496085	662	0.60	0.33	-0.21	-0.22	-0.06	0.33
VF496194	662	0.49	0.28	-0.02	0.28	-0.21	-0.15
Form 9							
VF495943	668	0.89	0.28	-0.13	0.28	-0.21	-0.13
VF495800	668	0.90	0.31	0.31	-0.11	-0.17	-0.23
VF495911	668	0.79	0.37	-0.29	-0.17	-0.22	0.37
VF495837	668	0.89	0.27	-0.19	0.27	-0.07	-0.19
VF495850	668	0.68	0.23	-0.18	-0.06	-0.23	0.23
VF495905	668	0.69	0.37	-0.19	-0.21	-0.20	0.37
VF496201	668	0.95	0.35	-0.29	-0.09	-0.18	0.35
VF496218	668	0.71	0.27	-0.19	0.27	-0.18	-0.09
VF496879	668	0.58	0.35	0.35	-0.21	-0.12	-0.16
VF496209	668	0.76	0.30	-0.18	-0.17	-0.13	0.30
VF496865	668	0.64	0.41	-0.26	-0.21	0.41	-0.12
VF496220	668	0.70	0.27	-0.23	-0.21	-0.10	0.27

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF495780	676	0.57	0.38	-0.16	-0.24	0.38	-0.13
VF495940	676	0.89	0.35	-0.07	-0.30	0.35	-0.19
VF495921	676	0.60	0.48	-0.24	0.48	-0.18	-0.29
VF495924	676	0.76	0.42	0.42	-0.24	-0.23	-0.25
VF495859	676	0.70	0.27	-0.27	0.27	-0.07	-0.29
VF495935	676	0.31	-0.06	0.19	-0.24	-0.06	-0.02
VF496200	676	0.85	0.28	0.28	-0.17	-0.28	-0.10
VF496206	676	0.64	0.32	-0.25	0.32	-0.06	-0.24
VF496213	676	0.72	0.24	0.24	-0.13	-0.23	-0.08
VF496212	676	0.82	0.39	-0.23	0.39	-0.23	-0.19
VF496211	676	0.93	0.39	-0.19	-0.25	0.39	-0.24
VF496221	676	0.92	0.35	0.35	-0.20	-0.21	-0.22

Table A4. Reading Grade 6 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF496164	769	0.69	0.33	0.33	-0.24	-0.23	-0.08
VF496187	769	0.36	0.23	0.23	0.01	-0.18	-0.13
VF496191	769	0.41	0.35	-0.10	-0.13	-0.18	0.35
VF496863	769	0.58	0.41	-0.19	-0.24	-0.14	0.41
VF496208	769	0.58	0.40	-0.16	0.40	-0.24	-0.16
VF496873	769	0.63	0.40	-0.21	-0.11	-0.31	0.40
VF497040	769	0.37	0.16	0.16	-0.24	0.09	-0.11
VF497034	769	0.78	0.43	-0.26	-0.25	-0.15	0.43
VF497042	769	0.78	0.52	-0.33	-0.22	0.52	-0.26
VF497046	769	0.55	0.17	0.17	-0.16	0.02	-0.25
VF497050	769	0.57	0.02	0.17	-0.25	0.02	-0.16
VF497031	769	0.47	0.20	0.20	-0.23	-0.11	-0.01
Form 2							
VF496172	656	0.66	0.42	-0.22	-0.30	0.42	-0.09
VF496180	656	0.33	0.31	-0.06	0.31	-0.16	-0.22
VF496880	656	0.35	0.11	0.06	0.11	-0.21	0.05
VF496204	656	0.56	0.36	-0.21	0.36	-0.21	-0.08
VF496415	656	0.61	0.36	-0.20	-0.14	-0.23	0.36
VF496867	656	0.67	0.46	0.46	-0.13	-0.26	-0.30
VF497035	656	0.74	0.41	-0.31	-0.24	0.41	-0.10
VF497041	656	0.82	0.50	-0.28	-0.33	0.50	-0.17
VF497038	656	0.92	0.41	-0.25	0.41	-0.22	-0.22
VF497049	656	0.77	0.28	-0.15	0.28	-0.12	-0.14
VF497033	656	0.55	0.32	-0.27	-0.14	0.32	-0.11
VF497047	656	0.78	0.37	-0.27	0.37	-0.24	-0.11
Form 3							
VF496916	666	0.45	0.30	-0.21	0.30	-0.08	-0.14
VF496925	666	0.40	0.20	-0.17	-0.26	0.04	0.20
VF496938	666	0.80	0.26	0.26	-0.08	-0.15	-0.19
VF496935	666	0.53	0.37	-0.21	0.37	-0.08	-0.21
VF496918	666	0.60	0.41	-0.25	-0.21	-0.16	0.41
VF496926	666	0.60	0.36	-0.15	-0.22	-0.18	0.36
VF523786	666	0.67	0.36	-0.33	-0.18	-0.07	0.36
VF523801	666	0.62	0.37	0.37	-0.08	-0.27	-0.28
VF523870	666	0.78	0.31	-0.27	-0.06	-0.22	0.31
VF523830	666	0.35	-0.12	-0.06	-0.12	-0.12	0.27
VF523863	666	0.82	0.52	-0.20	-0.30	0.52	-0.31
VF523813	666	0.49	0.25	0.25	-0.23	-0.22	0.05

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF496943	663	0.82	0.51	-0.17	-0.26	-0.36	0.51
VF496922	663	0.69	0.26	-0.25	-0.03	0.26	-0.16
VF496920	663	0.75	0.46	0.46	-0.28	-0.21	-0.24
VF496933	663	0.62	0.39	-0.27	-0.24	-0.09	0.39
VF496942	663	0.81	0.40	-0.19	0.40	-0.18	-0.28
VF496927	663	0.70	0.24	0.24	-0.01	-0.20	-0.17
VF523852	663	0.71	0.34	-0.25	-0.26	-0.08	0.34
VF523825	663	0.68	0.45	-0.25	-0.23	-0.20	0.45
VF523818	663	0.66	0.43	-0.28	0.43	-0.17	-0.17
VF523846	663	0.69	0.50	0.50	-0.34	-0.22	-0.20
VF523861	663	0.70	0.43	-0.19	0.43	-0.30	-0.19
VF523804	663	0.65	0.40	0.40	-0.23	-0.13	-0.23
Form 5							
VF496043	679	0.80	0.32	-0.19	0.32	-0.16	-0.17
VF496083	679	0.90	0.33	-0.24	-0.14	-0.18	0.33
VF496029	679	0.83	0.34	-0.16	0.34	-0.26	-0.16
VF496097	679	0.78	0.42	-0.29	0.42	-0.16	-0.22
VF496047	679	0.47	0.16	-0.14	-0.18	0.16	0.10
VF496071	679	0.81	0.47	-0.21	0.47	-0.33	-0.23
VF497073	679	0.69	0.36	0.36	-0.26	-0.27	-0.06
VF497053	679	0.38	0.25	0.01	-0.21	0.25	-0.10
VF497065	679	0.46	0.15	0.15	-0.05	-0.20	-0.16
VF497064	679	0.34	0.18	-0.01	0.18	-0.16	-0.03
VF497067	679	0.19	0.01	0.03	0.01	0.11	-0.18
VF497062	679	0.63	0.18	0.02	-0.21	0.18	-0.22
Form 6							
VF496055	678	0.50	0.43	-0.19	-0.20	0.43	-0.19
VF496087	678	0.81	0.40	0.40	-0.20	-0.26	-0.19
VF496036	678	0.57	0.34	0.34	-0.38	-0.13	-0.04
VF496065	678	0.67	0.41	-0.26	-0.22	0.41	-0.18
VF496100	678	0.78	0.40	0.40	-0.14	-0.26	-0.23
VF496051	678	0.58	0.36	-0.09	-0.16	-0.30	0.36
VF497058	678	0.48	0.35	-0.15	-0.20	-0.22	0.35
VF497059	678	0.68	0.33	0.33	-0.03	-0.33	-0.14
VF497070	678	0.41	0.08	-0.09	-0.15	0.08	0.04
VF497069	678	0.60	0.30	-0.11	0.30	-0.15	-0.18
VF497071	678	0.42	0.20	-0.08	-0.10	-0.09	0.20
VF497068	678	0.57	0.30	-0.24	-0.16	0.30	-0.12

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF497074	672	0.50	0.36	-0.15	-0.20	-0.15	0.36
VF497078	672	0.62	0.29	-0.13	0.29	-0.18	-0.15
VF497087	672	0.62	0.36	-0.13	-0.19	0.36	-0.20
VF497077	672	0.42	0.27	0.27	-0.12	-0.21	-0.03
VF497083	672	0.43	0.36	-0.09	-0.21	0.36	-0.15
VF497076	672	0.41	0.34	-0.20	-0.09	-0.13	0.34
VF496977	672	0.59	0.29	-0.18	-0.14	0.29	-0.07
VF496958	672	0.82	0.25	0.25	-0.22	-0.17	-0.06
VF496968	672	0.69	0.28	-0.14	-0.15	0.28	-0.14
VF496963	672	0.57	0.32	0.32	-0.12	-0.28	-0.12
VF496975	672	0.67	0.33	-0.33	0.33	-0.07	-0.22
VF496974	672	0.68	0.35	-0.31	0.35	-0.09	-0.14
Form 8							
VF497085	664	0.23	0.28	-0.08	0.28	-0.04	-0.18
VF497079	664	0.66	0.37	-0.26	-0.14	-0.18	0.37
VF497082	664	0.52	0.35	-0.08	0.35	-0.26	-0.17
VF497075	664	0.32	0.33	0.02	-0.22	-0.19	0.33
VF497084	664	0.51	0.32	-0.18	-0.19	0.32	-0.11
VF497086	664	0.34	-0.03	-0.03	0.02	0.01	0.01
VF496954	664	0.80	0.42	-0.16	0.42	-0.33	-0.16
VF496964	664	0.47	0.33	-0.04	-0.28	-0.17	0.33
VF496979	664	0.79	0.38	-0.24	-0.15	-0.21	0.38
VF496961	664	0.84	0.41	-0.22	0.41	-0.23	-0.22
VF496970	664	0.75	0.36	0.36	-0.22	-0.23	-0.12
VF496952	664	0.75	0.32	-0.24	-0.27	0.32	-0.09
Form 9							
VF388824	676	0.57	0.12	0.12	0.03	-0.17	-0.11
VF388868	676	0.91	0.43	-0.19	-0.31	0.43	-0.19
VF388909	676	0.41	0.18	-0.10	0.18	-0.08	-0.07
VF388911	676	0.67	0.46	-0.28	-0.30	-0.12	0.46
VF388907	676	0.35	0.22	-0.12	0.22	0.00	-0.21
VF388914	676	0.53	0.27	-0.16	0.27	-0.20	-0.06
VF495938	676	0.49	0.26	0.26	-0.19	-0.17	-0.05
VF495961	676	0.63	0.38	-0.30	0.38	-0.31	0.01
VF495954	676	0.76	0.33	-0.14	-0.16	-0.22	0.33
VF495945	676	0.56	0.31	0.31	-0.10	-0.24	-0.09
VF495990	676	0.76	0.50	0.50	-0.30	-0.28	-0.24
VF495925	676	0.72	0.35	-0.22	-0.18	0.35	-0.16

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF388848	676	0.78	0.24	-0.13	-0.19	-0.06	0.24
VF388851	676	0.55	0.37	-0.22	-0.17	0.37	-0.12
VF388881	676	0.88	0.30	0.30	-0.11	-0.13	-0.27
VF388853	676	0.70	0.40	-0.31	-0.17	0.40	-0.14
VF388905	676	0.24	0.12	-0.07	0.12	0.05	-0.10
VF388912	676	0.76	0.36	-0.22	-0.15	-0.22	0.36
VF495908	676	0.83	0.34	-0.25	0.34	-0.24	-0.07
VF495916	676	0.90	0.38	-0.18	-0.14	0.38	-0.26
VF495918	676	0.75	0.31	-0.18	-0.15	-0.14	0.31
VF495968	676	0.52	0.37	0.37	-0.21	-0.18	-0.10
VF495980	676	0.39	-0.01	-0.05	-0.01	-0.05	0.09
VF495978	676	0.34	0.06	0.06	-0.02	-0.07	0.02

Table A5. Reading Grade 7 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF498064	809	0.44	0.36	-0.14	-0.24	-0.07	0.36
VF498057	809	0.63	0.42	-0.21	-0.21	-0.21	0.42
VF498052	809	0.44	0.32	-0.14	-0.15	0.32	-0.12
VF498047	809	0.82	0.22	-0.22	-0.17	0.22	-0.07
VF498036	809	0.29	0.08	0.05	-0.06	-0.07	0.08
VF498034	809	0.72	0.40	0.40	-0.28	-0.16	-0.18
VF496895	809	0.43	0.41	-0.33	0.00	-0.29	0.41
VF496901	809	0.57	0.38	0.38	-0.25	-0.24	-0.11
VF496908	809	0.77	0.46	-0.22	-0.21	-0.29	0.46
VF496913	809	0.78	0.52	-0.25	-0.16	-0.38	0.52
VF496932	809	0.75	0.45	-0.28	-0.18	0.45	-0.23
VF496903	809	0.73	0.47	0.47	-0.24	-0.23	-0.26
Form 2							
VF498063	676	0.57	0.26	0.26	-0.16	-0.10	-0.10
VF498059	676	0.66	0.41	-0.19	0.41	-0.22	-0.22
VF498054	676	0.62	0.41	0.41	-0.15	-0.20	-0.26
VF498051	676	0.55	0.33	-0.07	0.33	-0.09	-0.32
VF498060	676	0.30	0.15	0.15	-0.10	-0.21	0.12
VF498032	676	0.82	0.33	-0.18	0.33	-0.12	-0.25
VF496892	676	0.35	0.33	-0.22	0.33	-0.07	-0.13
VF496900	676	0.81	0.45	-0.22	-0.26	0.45	-0.23
VF496949	676	0.81	0.31	0.31	-0.18	-0.17	-0.19
VF496944	676	0.63	0.35	-0.15	-0.21	0.35	-0.17
VF496937	676	0.72	0.34	0.34	-0.27	-0.15	-0.14
VF496906	676	0.82	0.48	-0.29	0.48	-0.18	-0.29
Form 3							
VF497744	667	0.87	0.24	-0.11	0.24	-0.10	-0.20
VF497727	667	0.78	0.31	-0.05	-0.20	-0.26	0.31
VF497750	667	0.85	0.42	-0.26	0.42	-0.19	-0.21
VF497745	667	0.37	0.14	-0.16	-0.01	0.14	0.01
VF497754	667	0.34	0.15	0.14	-0.21	-0.22	0.15
VF497721	667	0.79	0.36	0.36	-0.11	-0.27	-0.18
VF497961	667	0.63	0.45	0.45	-0.21	-0.30	-0.13
VF497974	667	0.54	0.33	-0.08	0.33	-0.22	-0.13
VF497975	667	0.52	0.30	-0.11	-0.09	0.30	-0.20
VF497958	667	0.68	0.39	-0.22	-0.16	-0.20	0.39
VF497955	667	0.76	0.48	0.48	-0.27	-0.25	-0.21
VF497959	667	0.49	0.14	0.14	-0.24	0.12	-0.28

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF497726	675	0.44	0.20	0.20	-0.12	-0.04	-0.09
VF497748	675	0.48	0.03	-0.02	0.04	0.03	-0.15
VF497724	675	0.83	0.37	-0.27	0.37	-0.17	-0.12
VF497757	675	0.37	0.14	-0.01	0.14	-0.17	-0.16
VF497715	675	0.57	0.35	-0.25	-0.16	0.35	-0.14
VF497730	675	0.32	0.14	0.04	-0.23	-0.20	0.14
VF497978	675	0.50	0.36	0.36	-0.20	-0.16	-0.12
VF497972	675	0.33	0.22	-0.22	-0.17	0.09	0.22
VF497951	675	0.67	0.35	-0.17	-0.30	0.35	-0.07
VF497969	675	0.52	0.44	-0.23	-0.23	-0.15	0.44
VF497956	675	0.72	0.27	0.02	0.27	-0.31	-0.11
VF497971	675	0.31	0.17	-0.22	-0.14	0.06	0.17
Form 5							
VF497930	670	0.72	0.43	0.43	-0.28	-0.26	-0.10
VF497938	670	0.71	0.39	-0.22	-0.26	0.39	-0.09
VF497950	670	0.75	0.46	-0.31	-0.27	0.46	-0.18
VF497946	670	0.81	0.47	0.47	-0.29	-0.27	-0.19
VF497935	670	0.75	0.28	-0.26	0.28	-0.20	-0.05
VF497942	670	0.35	0.19	-0.13	-0.20	0.19	0.07
VF497711	670	0.66	0.43	-0.19	0.43	-0.23	-0.23
VF497686	670	0.42	0.29	-0.23	0.29	-0.03	-0.13
VF497819	670	0.67	0.44	-0.34	-0.17	0.44	-0.21
VF497803	670	0.64	0.43	-0.20	-0.23	0.43	-0.21
VF497756	670	0.46	0.20	0.04	-0.28	0.20	-0.05
VF497735	670	0.21	0.09	0.04	0.09	-0.20	0.03
Form 6							
VF497931	668	0.77	0.38	-0.24	0.38	-0.20	-0.16
VF497933	668	0.72	0.26	-0.06	-0.24	-0.18	0.26
VF497936	668	0.63	0.10	-0.12	-0.02	0.10	-0.03
VF497941	668	0.64	0.39	0.39	-0.16	-0.19	-0.28
VF497943	668	0.54	0.39	-0.13	-0.20	-0.23	0.39
VF497934	668	0.70	0.38	-0.24	0.38	-0.18	-0.18
VF497701	668	0.72	0.47	-0.30	-0.25	0.47	-0.18
VF497814	668	0.78	0.31	-0.14	-0.21	0.31	-0.15
VF497794	668	0.80	0.48	0.48	-0.34	-0.23	-0.19
VF497772	668	0.42	0.21	0.06	-0.23	0.21	-0.25
VF497732	668	0.59	0.46	-0.21	-0.29	0.46	-0.18
VF497785	668	0.55	0.23	-0.26	0.00	0.23	-0.12

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF497873	657	0.71	0.40	0.40	-0.35	-0.23	-0.02
VF497894	657	0.87	0.39	-0.21	-0.18	0.39	-0.22
VF497893	657	0.75	0.32	0.32	-0.21	-0.22	-0.13
VF497890	657	0.78	0.41	-0.24	-0.20	0.41	-0.21
VF497882	657	0.67	0.32	-0.20	-0.10	-0.17	0.32
VF497885	657	0.64	0.42	0.42	-0.22	-0.18	-0.24
VF497179	657	0.88	0.40	-0.29	-0.13	0.40	-0.20
VF497219	657	0.29	-0.08	0.18	-0.24	0.03	-0.08
VF497198	657	0.43	0.32	-0.03	-0.26	0.32	-0.08
VF497186	657	0.83	0.48	-0.31	-0.25	-0.18	0.48
VF497224	657	0.60	0.39	-0.12	0.39	-0.29	-0.15
VF497245	657	0.84	0.48	0.48	-0.20	-0.22	-0.33
Form 8							
VF497868	664	0.45	0.31	-0.13	0.31	-0.25	0.00
VF497862	664	0.56	0.35	-0.22	-0.18	0.35	-0.09
VF497879	664	0.65	0.28	-0.13	0.28	-0.21	-0.06
VF497889	664	0.79	0.49	-0.20	-0.20	0.49	-0.36
VF497881	664	0.65	0.36	-0.22	0.36	-0.13	-0.21
VF497876	664	0.79	0.32	-0.26	-0.09	0.32	-0.28
VF497205	664	0.70	0.30	0.30	-0.18	-0.19	-0.10
VF497190	664	0.75	0.39	-0.27	-0.19	-0.14	0.39
VF497211	664	0.74	0.35	-0.15	-0.21	-0.18	0.35
VF497238	664	0.58	0.17	0.17	-0.19	-0.03	-0.03
VF497214	664	0.74	0.40	-0.26	0.40	-0.20	-0.17
VF497175	664	0.80	0.41	0.41	-0.21	-0.30	-0.14
Form 9							
VF497980	672	0.45	0.44	0.44	-0.24	-0.14	-0.21
VF498011	672	0.57	0.29	-0.24	-0.04	-0.16	0.29
VF497883	672	0.49	0.22	0.00	-0.25	0.22	-0.20
VF498018	672	0.83	0.43	-0.26	-0.24	0.43	-0.21
VF497995	672	0.72	0.54	-0.26	-0.34	-0.25	0.54
VF498030	672	0.79	0.35	0.35	-0.18	-0.29	-0.10
VF497260	672	0.82	0.47	-0.26	-0.27	-0.23	0.47
VF497294	672	0.49	0.18	-0.19	-0.16	0.18	0.02
VF497299	672	0.69	0.38	-0.31	-0.18	0.38	-0.09
VF497295	672	0.58	0.31	-0.11	-0.18	0.31	-0.14
VF497291	672	0.42	0.32	-0.04	-0.26	-0.15	0.32
VF497268	672	0.53	0.26	-0.17	0.26	-0.20	-0.04

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF497985	669	0.26	0.03	0.03	0.06	-0.20	-0.01
VF498023	669	0.86	0.40	0.40	-0.25	-0.24	-0.19
VF498062	669	0.62	0.33	-0.26	0.33	-0.11	-0.09
VF498058	669	0.64	0.31	-0.10	0.31	-0.20	-0.15
VF497877	669	0.50	0.22	-0.19	-0.07	-0.07	0.22
VF497870	669	0.54	0.24	0.24	-0.08	-0.23	-0.12
VF497263	669	0.75	0.40	-0.29	-0.23	0.40	-0.10
VF497292	669	0.77	0.38	-0.26	0.38	-0.16	-0.18
VF497293	669	0.71	0.37	0.37	-0.20	-0.19	-0.17
VF497275	669	0.63	0.26	-0.20	-0.02	0.26	-0.21
VF497301	669	0.58	0.34	-0.12	-0.09	-0.28	0.34
VF497281	669	0.84	0.33	0.33	-0.27	-0.17	-0.13

Table A6. Reading Grade 8 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF497094	760	0.78	0.39	-0.16	-0.31	0.39	-0.12
VF497096	760	0.60	0.43	-0.21	0.43	-0.23	-0.27
VF497101	760	0.62	0.41	-0.31	-0.07	0.41	-0.19
VF497099	760	0.40	0.34	-0.23	-0.29	0.34	-0.01
VF497103	760	0.75	0.44	-0.31	0.44	-0.18	-0.22
VF497114	760	0.72	0.32	-0.26	-0.16	0.32	-0.14
VF497121	760	0.77	0.17	-0.05	0.17	-0.12	-0.11
VF497118	760	0.56	0.25	-0.10	-0.13	0.25	-0.20
VF497122	760	0.24	-0.02	0.13	-0.02	0.02	-0.20
VF497130	760	0.75	0.41	-0.33	-0.18	0.41	-0.11
VF497132	760	0.55	0.23	-0.17	0.09	0.23	-0.33
VF497117	760	0.61	0.44	-0.10	-0.27	-0.28	0.44
Form2							
VF497115	664	0.80	0.37	-0.18	0.37	-0.20	-0.21
VF497093	664	0.34	0.08	0.08	-0.23	0.11	-0.27
VF497113	664	0.34	0.26	-0.23	-0.14	-0.02	0.26
VF497102	664	0.30	0.14	-0.13	-0.07	0.03	0.14
VF497098	664	0.54	0.27	-0.10	0.27	-0.14	-0.11
VF497095	664	0.87	0.44	0.44	-0.27	-0.24	-0.21
VF497123	664	0.71	0.31	0.31	-0.20	-0.21	-0.11
VF497120	664	0.48	0.43	0.43	-0.13	-0.27	-0.16
VF497128	664	0.61	0.28	-0.16	0.28	-0.04	-0.18
VF497127	664	0.84	0.46	-0.22	-0.30	0.46	-0.20
VF497119	664	0.28	0.07	0.26	-0.23	-0.17	0.07
VF497116	664	0.63	0.37	-0.24	-0.07	0.37	-0.27
Form 3							
VF497139	672	0.85	0.36	-0.19	-0.14	-0.25	0.36
VF497143	672	0.83	0.37	-0.17	0.37	-0.22	-0.24
VF497137	672	0.65	0.40	0.40	-0.17	-0.27	-0.15
VF497156	672	0.58	0.32	-0.22	0.32	-0.18	-0.15
VF497164	672	0.50	0.30	-0.07	-0.11	0.30	-0.22
VF497153	672	0.66	0.09	0.09	-0.04	-0.02	-0.14
VF497363	672	0.53	0.28	-0.11	0.28	-0.26	-0.08
VF497325	672	0.68	0.34	0.34	-0.09	-0.18	-0.22
VF497367	672	0.52	0.19	-0.14	0.19	0.02	-0.21
VF497319	672	0.71	0.48	-0.23	-0.29	-0.20	0.48
VF497336	672	0.62	0.38	0.38	-0.20	-0.23	-0.12
VF497331	672	0.21	0.14	0.14	0.11	-0.23	-0.22

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF497146	669	0.83	0.32	-0.14	-0.26	0.32	-0.11
VF497148	669	0.60	0.36	-0.15	0.36	-0.11	-0.27
VF497166	669	0.82	0.51	-0.24	-0.32	-0.28	0.51
VF497158	669	0.31	0.05	-0.14	0.05	-0.01	0.02
VF497161	669	0.55	0.20	-0.10	-0.04	0.20	-0.24
VF497165	669	0.42	0.17	0.17	0.07	-0.22	-0.17
VF497370	669	0.68	0.26	-0.22	-0.08	0.26	-0.12
VF497329	669	0.81	0.39	-0.20	-0.28	-0.16	0.39
VF497355	669	0.94	0.27	-0.18	-0.19	0.27	-0.10
VF497349	669	0.90	0.31	0.31	-0.15	-0.18	-0.19
VF497353	669	0.60	0.37	-0.25	0.37	-0.20	-0.12
VF497328	669	0.65	0.35	-0.07	-0.16	-0.30	0.35
Form 5							
VF497185	672	0.59	0.07	0.02	-0.02	0.07	-0.17
VF497178	672	0.59	0.34	-0.18	0.34	-0.30	-0.07
VF497193	672	0.59	0.34	-0.03	-0.32	-0.18	0.34
VF497203	672	0.64	0.40	0.40	-0.07	-0.28	-0.25
VF497207	672	0.86	0.44	-0.25	-0.17	-0.28	0.44
VF497209	672	0.67	0.32	-0.26	-0.07	0.32	-0.22
VF497787	672	0.60	0.21	0.21	-0.07	-0.18	-0.16
VF497796	672	0.49	0.13	0.05	0.13	-0.20	-0.13
VF497801	672	0.36	0.21	-0.07	0.21	-0.13	-0.08
VF497804	672	0.69	0.29	0.29	-0.21	-0.15	-0.08
VF497810	672	0.45	0.27	-0.13	-0.21	0.27	-0.08
VF497792	672	0.76	0.43	-0.22	-0.17	-0.27	0.43
Form 6							
VF497189	669	0.18	0.24	0.03	-0.03	0.24	-0.20
VF497173	669	0.49	0.25	-0.15	0.25	-0.18	-0.06
VF497180	669	0.91	0.34	0.34	-0.23	-0.18	-0.15
VF497196	669	0.77	0.35	-0.13	0.35	-0.18	-0.25
VF497199	669	0.53	0.21	-0.19	0.00	-0.13	0.21
VF497213	669	0.79	0.39	0.39	-0.16	-0.25	-0.24
VF497800	669	0.63	0.44	-0.31	-0.21	-0.20	0.44
VF497797	669	0.42	0.32	-0.21	0.32	-0.22	-0.01
VF497802	669	0.59	0.24	-0.08	-0.16	-0.19	0.24
VF497806	669	0.63	0.40	-0.25	-0.14	0.40	-0.20
VF497813	669	0.63	0.41	0.41	-0.26	-0.19	-0.14
VF497799	669	0.62	0.29	-0.10	-0.28	0.29	-0.08

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF497333	669	0.66	0.34	-0.14	-0.28	-0.05	0.34
VF497339	669	0.63	0.27	-0.12	-0.13	0.27	-0.19
VF497342	669	0.75	0.25	0.25	-0.20	-0.17	-0.08
VF497347	669	0.47	0.26	-0.17	-0.04	0.26	-0.15
VF497352	669	0.83	0.33	-0.13	0.33	-0.24	-0.19
VF497335	669	0.22	0.14	-0.21	0.14	0.03	-0.06
VF497306	669	0.93	0.26	-0.18	-0.21	0.26	-0.01
VF497323	669	0.35	-0.03	-0.04	-0.03	0.06	0.01
VF497296	669	0.84	0.27	0.27	-0.22	-0.19	-0.11
VF497316	669	0.82	0.39	-0.17	0.39	-0.28	-0.18
VF497313	669	0.57	0.44	-0.32	-0.12	-0.18	0.44
VF497312	669	0.95	0.31	-0.17	0.31	-0.15	-0.22
Form 8							
VF497341	649	0.79	0.36	-0.18	-0.27	-0.17	0.36
VF497350	649	0.73	0.22	-0.04	0.22	-0.28	-0.15
VF497348	649	0.67	0.37	-0.18	-0.21	0.37	-0.20
VF497346	649	0.83	0.40	-0.19	-0.24	0.40	-0.24
VF497351	649	0.39	0.19	-0.23	0.19	-0.17	0.06
VF497345	649	0.64	0.23	0.23	-0.20	-0.06	-0.11
VF497305	649	0.79	0.34	-0.29	-0.15	-0.08	0.34
VF497300	649	0.87	0.43	-0.29	-0.10	0.43	-0.28
VF497302	649	0.59	0.43	0.43	-0.25	-0.11	-0.29
VF497309	649	0.81	0.45	-0.31	-0.24	0.45	-0.19
VF497298	649	0.89	0.45	-0.29	-0.21	-0.26	0.45
VF497304	649	0.41	0.06	0.06	-0.07	-0.21	0.11
Form 9							
VF497427	675	0.92	0.42	0.42	-0.28	-0.24	-0.15
VF497441	675	0.85	0.37	-0.28	-0.15	0.37	-0.20
VF497443	675	0.87	0.35	-0.15	0.35	-0.23	-0.23
VF497436	675	0.84	0.35	0.35	-0.24	-0.24	-0.18
VF497445	675	0.76	0.43	-0.22	-0.13	-0.30	0.43
VF497444	675	0.92	0.43	-0.20	-0.24	-0.28	0.43
VF497235	675	0.90	0.45	-0.24	-0.30	-0.19	0.45
VF497259	675	0.55	0.31	0.31	-0.27	-0.08	-0.05
VF497269	675	0.37	0.17	-0.06	-0.25	0.17	0.04
VF497252	675	0.84	0.32	-0.23	0.32	-0.19	-0.10
VF497255	675	0.83	0.33	-0.25	-0.20	-0.09	0.33
VF497229	675	0.59	0.29	0.29	-0.22	-0.07	-0.24

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF497439	674	0.33	0.15	-0.09	0.15	-0.16	-0.01
VF497440	674	0.82	0.20	-0.07	-0.19	0.20	-0.11
VF497429	674	0.93	0.15	-0.09	-0.05	-0.14	0.15
VF497431	674	0.89	0.41	-0.24	0.41	-0.24	-0.21
VF497442	674	0.90	0.41	0.41	-0.25	-0.26	-0.15
VF497446	674	0.69	0.32	-0.17	-0.20	0.32	-0.11
VF497266	674	0.47	0.20	-0.20	-0.19	0.20	-0.07
VF497244	674	0.79	0.39	-0.20	-0.24	0.39	-0.16
VF497264	674	0.81	0.47	-0.24	0.47	-0.30	-0.21
VF497271	674	0.80	0.42	-0.27	-0.15	-0.25	0.42
VF497242	674	0.50	0.40	-0.12	-0.24	0.40	-0.15
VF497257	674	0.78	0.27	0.27	-0.11	-0.16	-0.21

Mathematics

Table A7. Mathematics Grade 3 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF394365	798	0.65	0.50	-0.22	-0.36	0.50	-0.12
VF393770	798	0.61	0.48	0.48	-0.07	-0.13	-0.42
VF394340	798	0.25	0.10	-0.27	0.10	0.09	-0.02
VF494690	798	0.24	0.19	0.02	-0.12	-0.19	0.19
VF406359	798	0.84	0.40	-0.17	0.40	-0.22	-0.25
VF393959	798	0.89	0.39	-0.21	0.39	-0.20	-0.24
VF394252	798	0.47	0.27	-0.14	-0.03	0.27	-0.20
VF393784	798	0.71	0.37	-0.22	-0.25	-0.18	0.37
VF394358	798	0.55	0.54	-0.29	0.54	-0.26	-0.23
VF493098	798	0.37	0.29	-0.19	0.29	-0.26	-0.13
VF494335	798	0.69	0.43	0.43	-0.31	-0.14	-0.19
VF394241	798	0.93	0.33	-0.12	-0.16	0.33	-0.25
Form 2							
VF493110	700	0.83	0.23	-0.26	0.23	-0.04	-0.07
VF394366	700	0.86	0.37	0.37	-0.28	-0.19	-0.09
VF394359	700	0.73	0.43	-0.36	-0.30	0.43	-0.05
VF406337	700	0.49	0.44	-0.23	-0.11	-0.29	0.44
VF494734	700	0.27	0.05	-0.09	-0.04	0.05	0.09
VF493314	700	0.84	0.40	-0.29	0.40	-0.07	-0.25
VF494670	700	0.23	0.18	-0.13	-0.01	0.00	0.18
VF394244	700	0.89	0.34	-0.24	-0.14	0.34	-0.17
VF387500	700	0.53	0.42	-0.15	0.42	-0.08	-0.35
VF394382	700	0.87	0.36	0.36	-0.19	-0.22	-0.23
VF394355	700	0.32	0.24	-0.32	-0.20	0.24	0.03
VF393783	700	0.92	0.39	-0.24	-0.27	0.39	-0.17
Form 3							
VF494873	696	0.83	0.33	-0.15	-0.25	0.33	-0.17
VF393751	696	0.83	0.45	-0.18	-0.34	0.45	-0.21
VF494665	696	0.60	0.32	-0.21	0.32	-0.09	-0.22
VF387507	696	0.54	0.40	-0.09	-0.10	-0.35	0.40
VF394224	696	0.69	0.45	-0.32	-0.20	0.45	-0.14
VF493094	696	0.78	0.38	0.38	-0.25	-0.17	-0.16
VF393748	696	0.51	0.48	-0.42	-0.06	-0.12	0.48
VF394254	696	0.37	0.24	-0.22	0.05	0.24	-0.20
VF406327	696	0.15	0.28	-0.31	0.28	-0.10	0.13
VF493445	696	0.63	0.45	-0.20	-0.21	-0.30	0.45
VF406306	696	0.83	0.29	0.29	-0.21	-0.24	-0.04
VF393824	696	0.43	0.28	0.28	-0.09	-0.12	-0.20

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF406331	699	0.72	0.42	-0.14	-0.32	0.42	-0.19
VF493287	699	0.59	0.40	-0.22	0.40	-0.15	-0.21
VF494756	699	0.57	0.25	0.25	-0.08	-0.16	-0.13
VF493124	699	0.42	0.27	-0.08	-0.10	0.27	-0.16
VF393823	699	0.54	0.48	-0.18	-0.28	-0.23	0.48
VF394248	699	0.68	0.41	0.41	-0.14	-0.27	-0.28
VF394229	699	0.56	0.27	-0.18	-0.14	-0.07	0.27
VF406297	699	0.45	0.28	-0.10	0.28	-0.20	-0.07
VF494802	699	0.68	0.30	-0.06	-0.13	-0.24	0.30
VF494855	699	0.66	0.25	-0.11	0.25	-0.06	-0.22
VF494674	699	0.29	0.20	-0.19	-0.17	0.20	0.17
VF493428	699	0.53	0.28	0.28	-0.11	-0.12	-0.17
Form 5							
VF494880	722	0.92	0.30	-0.19	-0.14	0.30	-0.19
VF394251	722	0.77	0.49	-0.41	-0.15	0.49	-0.18
VF493127	722	0.50	0.45	-0.15	-0.15	-0.31	0.45
VF394375	722	0.77	0.35	0.35	-0.20	-0.23	-0.10
VF393961	722	0.89	0.34	-0.15	0.34	-0.12	-0.28
VF394379	722	0.46	0.35	-0.20	0.35	-0.21	-0.07
VF494841	722	0.45	0.23	0.23	-0.10	-0.16	-0.02
VF394370	722	0.68	0.43	0.43	-0.16	-0.34	-0.17
VF394232	722	0.81	0.41	-0.27	-0.20	-0.19	0.41
VF387498	722	0.11	0.17	-0.09	-0.16	0.07	0.17
VF406339	722	0.78	0.39	-0.12	-0.37	0.39	-0.14
VF406324	722	0.74	0.29	-0.14	0.29	-0.14	-0.17
Form 6							
VF394368	705	0.64	0.39	-0.40	-0.02	-0.15	0.39
VF394378	705	0.43	0.19	-0.24	-0.07	0.01	0.19
VF493136	705	0.65	0.40	0.40	-0.20	-0.16	-0.25
VF393782	705	0.91	0.38	-0.29	-0.21	0.38	-0.20
VF394336	705	0.68	0.39	-0.33	-0.13	0.39	-0.13
VF393744	705	0.22	0.18	-0.35	-0.13	0.27	0.18
VF394371	705	0.54	0.42	0.42	-0.14	-0.35	-0.16
VF494820	705	0.28	0.27	-0.12	0.27	-0.04	-0.12
VF494886	705	0.89	0.38	-0.28	0.38	-0.19	-0.19
VF394208	705	0.63	0.40	0.40	-0.19	-0.38	-0.05
VF406295	705	0.63	0.14	-0.20	0.14	-0.16	0.03
VF393750	705	0.59	0.49	-0.34	-0.21	0.49	-0.16

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF394381	703	0.50	0.40	-0.11	-0.28	0.40	-0.14
VF493364	703	0.61	0.32	-0.16	-0.17	-0.16	0.32
VF494895	703	0.27	0.23	-0.08	-0.13	0.23	-0.15
VF394376	703	0.61	0.43	0.43	-0.16	-0.20	-0.27
VF394221	703	0.43	0.17	0.03	-0.19	-0.26	0.17
VF493146	703	0.62	0.42	-0.31	0.42	-0.14	-0.16
VF494693	703	0.28	0.14	-0.06	0.14	-0.12	0.02
VF394362	703	0.39	0.34	-0.26	-0.09	0.34	-0.08
VF387496	703	0.78	0.42	0.42	-0.27	-0.18	-0.23
VF393785	703	0.60	0.35	0.35	-0.15	-0.18	-0.21
VF393804	703	0.38	0.33	-0.38	-0.19	0.08	0.33
VF393752	703	0.48	0.42	-0.16	0.42	-0.12	-0.28
Form 8							
VF393786	698	0.76	0.39	0.39	-0.26	-0.23	-0.12
VF493161	698	0.51	0.19	-0.03	0.19	-0.16	-0.08
VF492342	698	0.75	0.47	-0.19	-0.43	0.47	-0.10
VF394361	698	0.79	0.47	-0.29	-0.34	-0.13	0.47
VF393749	698	0.56	0.48	-0.13	0.48	-0.12	-0.44
VF493461	698	0.48	0.22	-0.09	-0.03	0.22	-0.25
VF394239	698	0.58	0.60	0.60	-0.21	-0.17	-0.45
VF393746	698	0.13	0.16	-0.29	-0.13	0.28	0.16
VF394339	698	0.62	0.28	-0.18	-0.16	0.28	-0.04
VF493387	698	0.85	0.37	0.37	-0.10	-0.22	-0.24
VF494833	698	0.35	0.34	-0.04	0.34	-0.24	-0.16
VF394372	698	0.46	0.45	-0.30	-0.12	-0.17	0.45
Form 9							
VF393788	701	0.75	0.48	0.48	-0.38	-0.13	-0.23
VF394363	701	0.60	0.14	0.14	-0.21	-0.14	0.04
VF393815	701	0.89	0.44	-0.20	0.44	-0.29	-0.24
VF394250	701	0.62	0.31	-0.27	0.31	-0.03	-0.12
VF393772	701	0.62	0.55	-0.41	-0.25	-0.08	0.55
VF493978	701	0.92	0.34	-0.26	-0.19	-0.05	0.34
VF394235	701	0.99	0.12	-0.09	-0.01	0.12	.
VF493153	701	0.74	0.36	0.36	-0.13	-0.21	-0.20
VF387502	701	0.96	0.31	-0.19	-0.16	0.31	-0.13
VF494750	701	0.74	0.41	-0.26	0.41	-0.20	-0.16
VF493415	701	0.50	0.42	-0.21	0.42	-0.17	-0.18
VF394373	701	0.31	0.29	-0.07	-0.27	-0.13	0.29

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF494680	711	0.25	0.21	-0.10	-0.07	0.21	-0.09
VF393742	711	0.39	0.37	-0.23	-0.11	0.37	-0.16
VF494103	711	0.38	0.30	-0.11	0.30	-0.18	-0.11
VF394180	711	0.58	0.29	-0.08	0.29	-0.25	-0.18
VF393775	711	0.85	0.48	0.48	-0.14	-0.20	-0.38
VF394369	711	0.61	0.52	-0.12	-0.23	-0.36	0.52
VF406204	711	0.54	0.30	-0.22	-0.14	-0.04	0.30
VF494861	711	0.61	0.36	-0.15	-0.07	0.36	-0.27
VF406343	711	0.88	0.42	-0.36	0.42	-0.15	-0.08
VF393793	711	0.76	0.30	0.30	-0.15	-0.13	-0.17
VF393800	711	0.72	0.53	-0.20	-0.43	-0.09	0.53
VF394356	711	0.63	0.45	0.45	-0.20	-0.26	-0.22

Table A8. Mathematics Grade 4 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF492343	802	0.21	0.22	0.22	-0.08	-0.01	-0.18
VF497412	802	0.35	0.23	-0.15	0.08	0.23	-0.17
VF493128	802	0.30	0.25	-0.21	-0.03	-0.04	0.25
VF493366	802	0.38	0.17	-0.14	-0.11	0.17	0.03
VF492367	802	0.95	0.21	-0.12	0.21	-0.12	-0.10
VF493154	802	0.84	0.31	0.31	-0.14	-0.20	-0.15
VF492312	802	0.79	0.45	-0.23	-0.22	-0.29	0.45
VF492333	802	0.60	0.32	-0.19	0.32	-0.17	-0.16
VF492363	802	0.71	0.39	0.39	-0.20	-0.33	-0.18
VF493303	802	0.75	0.36	-0.14	-0.20	-0.22	0.36
VF393756	802	0.74	0.53	-0.29	-0.30	-0.21	0.53
VF493284	802	0.18	0.25	-0.11	0.00	-0.17	0.25
Form 2							
VF493305	707	0.33	0.22	-0.22	0.05	0.22	-0.10
VF492327	707	0.53	0.40	-0.21	0.40	-0.02	-0.33
VF493147	707	0.47	0.36	0.36	-0.28	-0.03	-0.17
VF492329	707	0.42	0.32	-0.15	0.32	-0.16	-0.10
VF492341	707	0.65	0.48	-0.26	-0.26	0.48	-0.19
VF393714	707	0.51	0.32	-0.26	-0.24	0.32	0.08
VF492361	707	0.64	0.42	-0.23	-0.20	-0.20	0.42
VF493219	707	0.72	0.42	-0.22	0.42	-0.22	-0.21
VF492346	707	0.97	0.29	0.29	-0.14	-0.23	-0.09
VF493356	707	0.88	0.26	-0.07	-0.18	-0.19	0.26
VF492381	707	0.35	-0.04	-0.04	-0.15	-0.09	0.23
VF493257	707	0.40	0.30	-0.20	-0.16	-0.02	0.30
Form 3							
VF493312	701	0.39	0.28	-0.24	-0.11	0.28	0.01
VF493226	701	0.75	0.34	-0.18	-0.28	-0.09	0.34
VF492331	701	0.62	0.43	-0.30	0.43	-0.12	-0.22
VF493142	701	0.58	0.49	-0.38	0.49	-0.16	-0.14
VF393726	701	0.74	0.23	-0.07	0.23	-0.20	-0.05
VF493133	701	0.50	0.41	-0.18	-0.13	-0.22	0.41
VF492375	701	0.71	0.22	0.22	-0.17	-0.12	-0.04
VF492348	701	0.97	0.27	0.27	-0.14	-0.19	-0.14
VF492336	701	0.51	0.44	-0.32	-0.14	-0.12	0.44
VF492380	701	0.42	0.02	0.02	-0.22	-0.08	0.22
VF493361	701	0.50	0.15	0.05	-0.19	0.15	-0.07
VF493249	701	0.37	0.20	-0.26	0.02	0.20	-0.09

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF493228	706	0.41	0.44	-0.01	0.44	-0.06	-0.40
VF492355	706	0.86	0.26	-0.10	-0.19	-0.15	0.26
VF493329	706	0.42	0.21	-0.16	0.21	-0.09	-0.03
VF497391	706	0.72	0.49	0.49	-0.25	-0.29	-0.21
VF493145	706	0.33	0.30	0.01	0.30	-0.05	-0.27
VF493354	706	0.32	0.10	0.01	0.02	-0.13	0.10
VF492332	706	0.62	0.34	-0.18	-0.11	0.34	-0.23
VF493371	706	0.60	0.38	-0.28	-0.12	-0.13	0.38
VF493134	706	0.72	0.39	-0.27	-0.22	0.39	-0.08
VF492334	706	0.83	0.46	-0.16	-0.27	0.46	-0.29
VF492364	706	0.76	0.43	0.43	-0.17	-0.33	-0.21
VF492349	706	0.98	0.17	-0.05	-0.15	0.17	-0.10
Form 5							
VF493140	710	0.53	0.57	-0.37	-0.18	-0.21	0.57
VF493260	710	0.48	0.38	-0.17	-0.26	-0.12	0.38
VF493318	710	0.35	0.35	-0.28	-0.05	0.35	-0.06
VF492351	710	0.50	0.35	-0.14	0.35	-0.14	-0.22
VF493373	710	0.82	0.30	0.30	-0.15	-0.18	-0.17
VF492326	710	0.36	0.21	-0.11	0.21	-0.09	-0.05
VF492376	710	0.72	0.35	-0.17	-0.33	-0.12	0.35
VF493152	710	0.26	0.14	-0.05	-0.12	0.14	-0.12
VF492389	710	0.69	0.41	0.41	-0.25	-0.35	-0.11
VF493295	710	0.25	0.25	-0.07	-0.18	0.25	0.04
VF492311	710	0.86	0.37	0.37	-0.15	-0.29	-0.16
VF493138	710	0.25	0.03	0.08	0.03	-0.18	0.00
Form 6							
VF493268	723	0.38	0.02	0.16	-0.15	-0.23	0.02
VF493310	723	0.35	0.03	-0.06	0.03	0.09	-0.08
VF497407	723	0.64	0.37	-0.35	0.37	-0.08	-0.02
VF492387	723	0.62	0.46	0.46	-0.15	-0.24	-0.30
VF493288	723	0.22	0.18	0.02	-0.08	0.18	-0.18
VF493143	723	0.65	0.10	0.10	-0.03	-0.10	-0.07
VF492321	723	0.54	0.40	-0.09	-0.32	0.40	-0.13
VF492350	723	0.92	0.34	-0.22	-0.24	0.34	-0.08
VF492371	723	0.90	0.25	-0.18	-0.13	-0.13	0.25
VF493130	723	0.48	0.10	0.10	-0.23	0.09	-0.05
VF493151	723	0.34	0.37	-0.21	-0.22	-0.05	0.37
VF493377	723	0.63	0.42	-0.27	0.42	-0.19	-0.14

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF492320	713	0.87	0.51	-0.14	-0.47	0.51	-0.02
VF493272	713	0.57	0.26	-0.13	-0.11	-0.12	0.26
VF492373	713	0.78	0.35	-0.25	0.35	-0.17	-0.10
VF492328	713	0.41	0.23	0.23	-0.21	-0.20	0.12
VF493381	713	0.93	0.35	-0.18	0.35	-0.20	-0.19
VF493242	713	0.65	0.31	-0.10	-0.11	0.31	-0.24
VF492392	713	0.69	0.37	-0.18	0.37	-0.17	-0.25
VF493126	713	0.88	0.37	-0.23	-0.22	0.37	-0.14
VF492353	713	0.57	0.36	-0.12	-0.28	-0.09	0.36
VF492306	713	0.94	0.24	0.24	-0.19	-0.16	-0.11
VF492338	713	0.74	0.32	-0.19	-0.19	-0.23	0.32
VF493344	713	0.81	0.39	0.39	-0.24	-0.25	-0.15
Form 8							
VF493236	695	0.74	0.33	-0.16	-0.25	-0.14	0.33
VF492315	695	0.90	0.50	-0.24	-0.43	0.50	-0.06
VF493280	695	0.36	0.39	-0.13	-0.09	0.39	-0.24
VF492386	695	0.58	0.38	-0.20	0.38	-0.21	-0.14
VF493131	695	0.89	0.26	0.26	-0.07	-0.13	-0.25
VF492339	695	0.85	0.49	-0.30	-0.31	0.49	-0.18
VF492323	695	0.67	0.43	-0.23	0.43	-0.26	-0.16
VF492354	695	0.73	0.43	-0.25	-0.26	0.43	-0.16
VF393648	695	0.27	0.31	0.04	0.31	-0.22	-0.17
VF492370	695	0.63	0.17	-0.08	-0.12	-0.05	0.17
VF493294	695	0.28	0.10	0.10	-0.21	0.13	-0.05
VF493327	695	0.31	0.18	-0.13	-0.04	-0.03	0.18
Form 9							
VF493223	693	0.43	0.38	-0.14	-0.22	0.38	-0.11
VF493337	693	0.44	0.36	-0.05	-0.16	0.36	-0.25
VF492369	693	0.75	0.39	-0.21	-0.22	-0.19	0.39
VF492330	693	0.49	0.28	-0.15	-0.19	-0.08	0.28
VF492352	693	0.97	0.17	-0.12	0.17	-0.07	-0.10
VF492337	693	0.61	0.38	0.38	-0.20	-0.20	-0.15
VF492359	693	0.55	0.36	-0.16	-0.26	-0.12	0.36
VF493149	693	0.87	0.28	0.28	-0.19	-0.13	-0.14
VF493262	693	0.46	0.14	0.14	-0.24	0.06	-0.12
VF493301	693	0.47	0.16	-0.06	0.16	-0.05	-0.14
VF497402	693	0.57	0.37	-0.29	0.37	-0.15	-0.08
VF393667	693	0.28	0.21	-0.09	-0.07	0.21	-0.08

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF493267	713	0.28	0.08	-0.23	0.20	-0.24	0.08
VF493238	713	0.93	0.14	-0.16	0.14	-0.05	-0.01
VF492358	713	0.85	0.31	-0.18	-0.13	-0.20	0.31
VF497395	713	0.61	0.47	0.47	-0.19	-0.25	-0.24
VF493349	713	0.37	0.40	-0.11	-0.09	-0.26	0.40
VF492390	713	0.94	0.23	-0.14	0.23	-0.12	-0.16
VF493144	713	0.60	0.52	-0.29	-0.12	-0.32	0.52
VF493334	713	0.33	0.35	-0.07	0.35	-0.22	-0.09
VF492344	713	0.21	0.19	0.09	-0.23	0.19	0.01
VF393675	713	0.54	0.46	0.46	-0.20	-0.27	-0.18
VF492310	713	0.64	0.55	-0.48	-0.20	0.55	-0.08
VF493135	713	0.48	0.39	-0.19	-0.19	0.39	-0.22

Table A9. Mathematics Grade 5 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF491627	772	0.82	0.36	-0.23	-0.14	0.36	-0.26
VF492281	772	0.45	0.50	-0.46	-0.08	0.50	0.02
VF491979	772	0.93	0.27	-0.16	0.27	-0.22	-0.06
VF491783	772	0.33	0.45	-0.19	-0.16	-0.16	0.45
VF491900	772	0.43	0.47	-0.36	-0.18	-0.15	0.47
VF491896	772	0.43	0.36	0.36	-0.26	-0.11	-0.12
VF492091	772	0.48	0.23	0.23	-0.12	-0.09	-0.10
VF492077	772	0.64	0.52	-0.23	-0.26	-0.28	0.52
VF492374	772	0.66	0.36	0.36	-0.17	-0.18	-0.18
VF491734	772	0.70	0.49	-0.18	0.49	-0.31	-0.25
VF492296	772	0.43	0.23	0.06	0.23	-0.11	-0.23
VF492521	772	0.38	0.21	-0.11	-0.08	0.21	-0.05
Form 2							
VF492001	656	0.68	0.29	-0.16	0.29	-0.29	0.00
VF492292	656	0.36	0.51	-0.42	-0.04	-0.08	0.51
VF491635	656	0.46	0.41	-0.13	0.41	0.00	-0.35
VF491895	656	0.75	0.48	-0.18	-0.36	-0.23	0.48
VF492095	656	0.57	0.57	-0.30	-0.34	0.57	-0.19
VF492048	656	0.53	0.47	-0.16	0.47	-0.28	-0.23
VF492366	656	0.48	0.27	-0.06	-0.20	0.27	-0.10
VF492532	656	0.56	0.23	-0.23	-0.04	0.23	-0.13
VF491937	656	0.49	0.33	0.33	-0.06	-0.25	-0.16
VF491911	656	0.92	0.23	0.23	-0.10	-0.14	-0.14
VF491817	656	0.77	0.50	-0.20	-0.29	-0.29	0.50
VF492113	656	0.48	0.38	-0.07	-0.16	-0.26	0.38
Form 3							
VF491973	667	0.83	0.39	0.39	-0.14	-0.20	-0.29
VF491926	667	0.61	0.40	-0.09	0.40	-0.22	-0.26
VF491905	667	0.67	0.44	-0.21	-0.20	0.44	-0.26
VF492120	667	0.40	0.26	-0.02	0.26	-0.21	-0.10
VF492423	667	0.70	0.36	-0.16	0.36	-0.19	-0.21
VF492010	667	0.51	0.38	-0.13	0.38	-0.13	-0.24
VF492235	667	0.53	0.36	-0.25	-0.10	0.36	-0.22
VF492246	667	0.42	0.36	-0.11	-0.19	-0.14	0.36
VF491946	667	0.71	0.39	0.39	-0.19	-0.18	-0.21
VF491963	667	0.35	0.33	-0.17	-0.18	-0.03	0.33
VF492083	667	0.36	0.39	-0.32	-0.01	0.39	-0.15
VF492421	667	0.66	0.34	0.34	-0.16	-0.24	-0.13

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF491951	668	0.83	0.39	-0.22	0.39	-0.17	-0.23
VF491922	668	0.71	0.54	-0.53	-0.16	-0.07	0.54
VF491987	668	0.84	0.38	-0.17	-0.21	0.38	-0.26
VF492251	668	0.26	0.36	-0.21	-0.11	-0.04	0.36
VF491727	668	0.52	0.42	0.42	-0.22	-0.25	-0.21
VF492401	668	0.61	0.34	-0.16	0.34	-0.19	-0.17
VF492088	668	0.54	0.43	-0.18	-0.19	-0.33	0.43
VF491969	668	0.88	0.24	-0.16	-0.15	-0.11	0.24
VF492099	668	0.55	0.53	0.53	-0.13	-0.14	-0.42
VF491630	668	0.50	0.36	-0.22	-0.13	0.36	-0.18
VF492128	668	0.70	0.08	0.08	0.00	-0.06	-0.07
VF491634	668	0.66	0.31	0.31	-0.14	-0.19	-0.14
Form 5							
VF492283	664	0.25	0.35	-0.24	-0.12	0.35	0.00
VF491924	664	0.61	0.38	-0.21	-0.20	-0.15	0.38
VF492000	664	0.66	0.27	0.27	-0.20	-0.12	-0.15
VF491804	664	0.56	0.58	0.58	-0.22	-0.33	-0.26
VF492288	664	0.73	0.27	-0.11	-0.18	-0.15	0.27
VF491962	664	0.70	0.41	-0.21	-0.22	0.41	-0.20
VF492304	664	0.55	0.45	-0.33	0.45	-0.16	-0.12
VF492427	664	0.39	0.44	-0.28	0.44	0.06	-0.31
VF492432	664	0.74	0.52	-0.23	-0.34	0.52	-0.26
VF491821	664	0.42	0.32	-0.08	-0.24	-0.10	0.32
VF492411	664	0.65	0.31	-0.19	-0.18	0.31	-0.14
VF492208	664	0.68	0.52	-0.37	-0.23	-0.18	0.52
Form 6							
VF491636	672	0.50	0.46	-0.41	-0.07	0.46	-0.07
VF491984	672	0.75	0.44	-0.26	-0.24	-0.23	0.44
VF492301	672	0.28	0.31	-0.25	-0.16	0.31	0.11
VF491939	672	0.77	0.49	-0.28	0.49	-0.28	-0.23
VF491952	672	0.85	0.35	-0.23	0.35	-0.19	-0.14
VF492391	672	0.67	0.32	-0.17	0.32	-0.23	-0.10
VF492174	672	0.56	0.52	0.52	-0.23	-0.42	-0.02
VF491885	672	0.45	0.20	-0.10	-0.16	0.20	0.02
VF491968	672	0.64	0.34	0.34	-0.21	-0.19	-0.10
VF492020	672	0.58	0.47	-0.22	-0.23	-0.22	0.47
VF492435	672	0.44	0.33	-0.03	0.33	-0.12	-0.26
VF491744	672	0.28	0.15	-0.02	-0.04	-0.10	0.15

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF492298	673	0.51	0.40	0.40	-0.11	-0.21	-0.22
VF491990	673	0.58	0.57	-0.28	-0.10	0.57	-0.38
VF491941	673	0.55	0.32	-0.03	0.32	-0.25	-0.20
VF492248	673	0.60	0.39	-0.27	-0.19	0.39	-0.14
VF491843	673	0.52	0.53	-0.21	-0.05	-0.41	0.53
VF492397	673	0.41	0.38	-0.25	-0.19	-0.05	0.38
VF491902	673	0.37	0.21	-0.12	-0.04	-0.09	0.21
VF492519	673	0.63	0.42	-0.26	0.42	-0.23	-0.13
VF491978	673	0.50	0.40	0.40	-0.12	-0.29	-0.14
VF491967	673	0.49	0.38	-0.18	0.38	-0.20	-0.10
VF492186	673	0.44	0.28	-0.17	-0.26	0.28	0.05
VF491753	673	0.60	0.20	0.20	-0.09	-0.06	-0.15
Form 8							
VF491998	657	0.59	0.46	-0.30	0.46	-0.20	-0.19
VF491903	657	0.31	0.45	0.45	-0.03	-0.23	-0.33
VF492214	657	0.35	0.51	-0.15	-0.30	-0.16	0.51
VF492303	657	0.37	0.28	-0.10	-0.21	-0.04	0.28
VF491945	657	0.79	0.45	-0.27	0.45	-0.21	-0.24
VF492524	657	0.90	0.21	-0.13	0.21	-0.13	-0.10
VF492403	657	0.60	0.30	-0.16	-0.17	0.30	-0.14
VF491914	657	0.55	0.29	0.29	-0.14	-0.18	-0.11
VF491992	657	0.59	0.29	0.29	-0.10	-0.17	-0.18
VF492027	657	0.46	0.18	-0.02	-0.15	0.18	-0.07
VF492203	657	0.85	0.52	-0.15	-0.35	-0.32	0.52
VF491761	657	0.69	0.53	-0.28	-0.26	-0.27	0.53
Form 9							
VF492313	663	0.48	0.37	-0.34	-0.10	0.00	0.37
VF491948	663	0.49	0.42	0.42	-0.20	-0.19	-0.18
VF492031	663	0.52	0.49	-0.25	-0.17	0.49	-0.25
VF492006	663	0.61	0.39	-0.16	-0.22	-0.21	0.39
VF491927	663	0.73	0.38	-0.23	-0.24	0.38	-0.13
VF491771	663	0.73	0.49	0.49	-0.24	-0.23	-0.29
VF492526	663	0.56	0.19	-0.08	-0.12	-0.07	0.19
VF492038	663	0.25	0.24	-0.24	0.24	-0.01	0.02
VF491794	663	0.69	0.46	0.46	-0.29	-0.21	-0.22
VF492007	663	0.70	0.34	-0.15	0.34	-0.23	-0.16
VF492406	663	0.81	0.36	0.36	-0.23	-0.17	-0.17
VF491997	663	0.87	0.33	-0.21	-0.22	0.33	-0.14

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF491888	673	0.79	0.45	-0.25	-0.23	-0.23	0.45
VF492255	673	0.47	0.32	-0.09	-0.14	-0.21	0.32
VF491626	673	0.69	0.38	-0.11	-0.16	-0.30	0.38
VF492015	673	0.35	0.26	-0.20	0.26	-0.08	-0.08
VF491916	673	0.56	0.55	-0.37	-0.20	-0.18	0.55
VF492003	673	0.47	0.30	0.30	-0.14	-0.19	-0.06
VF492528	673	0.67	0.26	-0.02	-0.18	0.26	-0.25
VF492416	673	0.74	0.26	-0.10	-0.19	0.26	-0.18
VF492211	673	0.15	0.22	-0.21	-0.18	0.22	0.07
VF492382	673	0.73	0.28	-0.13	-0.04	0.28	-0.22
VF491932	673	0.48	0.45	-0.27	0.45	-0.19	-0.22
VF491811	673	0.71	0.30	0.30	-0.22	-0.10	-0.17

Table A10. Mathematics Grade 6 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF491802	788	0.36	0.17	0.08	-0.10	0.17	-0.19
VF492134	788	0.26	-0.02	-0.02	0.03	-0.14	0.14
VF492025	788	0.37	0.45	-0.08	-0.31	-0.15	0.45
VF492388	788	0.62	0.36	-0.14	-0.18	-0.23	0.36
VF491930	788	0.59	0.48	-0.26	0.48	-0.26	-0.17
VF492210	788	0.44	0.34	-0.15	-0.16	-0.13	0.34
VF492588	788	0.84	0.31	-0.16	-0.22	0.31	-0.11
VF493006	788	0.52	0.39	-0.14	0.39	-0.24	-0.15
VF493002	788	0.45	0.28	-0.23	-0.10	0.28	-0.01
VF492790	788	0.17	0.25	0.25	-0.06	-0.14	-0.01
VF492716	788	0.45	0.45	-0.03	0.45	-0.33	-0.21
VF493058	788	0.63	0.48	-0.18	-0.27	0.48	-0.25
Form 2							
VF492533	661	0.67	0.49	-0.22	0.49	-0.38	-0.14
VF492093	661	0.39	0.43	-0.09	-0.15	-0.27	0.43
VF492287	661	0.50	0.23	-0.07	-0.09	-0.15	0.23
VF492220	661	0.47	0.20	-0.20	0.00	0.20	-0.07
VF491953	661	0.68	0.38	-0.24	0.38	-0.21	-0.14
VF492322	661	0.65	0.32	-0.17	-0.12	0.32	-0.18
VF491894	661	0.50	0.43	0.43	-0.27	-0.17	-0.13
VF492399	661	0.60	0.34	0.34	-0.18	-0.13	-0.18
VF492941	661	0.30	0.39	-0.35	-0.11	0.39	0.01
VF491981	661	0.94	0.29	-0.16	0.29	-0.17	-0.17
VF492747	661	0.20	0.24	-0.14	-0.25	0.10	0.24
VF491879	661	0.86	0.44	-0.19	-0.27	-0.25	0.44
Form 3							
VF492300	666	0.45	0.23	-0.19	0.03	0.23	-0.18
VF491960	666	0.69	0.33	-0.14	-0.17	-0.21	0.33
VF492542	666	0.61	0.42	-0.27	0.42	-0.20	-0.18
VF492660	666	0.28	0.40	0.40	-0.16	-0.14	-0.18
VF492280	666	0.46	0.35	-0.19	0.35	-0.26	-0.03
VF492233	666	0.84	0.35	-0.24	-0.17	-0.17	0.35
VF492229	666	0.19	0.16	-0.01	0.01	-0.14	0.16
VF491813	666	0.40	0.38	-0.14	0.38	-0.16	-0.18
VF492996	666	0.39	0.52	-0.27	-0.24	-0.11	0.52
VF491912	666	0.89	0.39	-0.17	-0.17	-0.31	0.39
VF492415	666	0.72	0.39	0.39	-0.21	-0.25	-0.15
VF492838	666	0.28	0.19	-0.13	-0.14	0.19	0.05

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF492593	660	0.66	0.47	-0.35	-0.08	-0.27	0.47
VF492634	660	0.24	0.33	-0.10	-0.33	0.33	0.11
VF492309	660	0.56	0.23	-0.17	-0.08	0.23	-0.09
VF492769	660	0.17	0.01	-0.14	0.01	0.02	0.12
VF492240	660	0.53	0.21	-0.20	0.21	-0.15	-0.01
VF492879	660	0.50	0.48	-0.25	-0.26	0.48	-0.12
VF492238	660	0.32	0.16	0.16	-0.04	-0.01	-0.14
VF491837	660	0.75	0.29	0.29	-0.19	-0.14	-0.12
VF491906	660	0.88	0.40	-0.13	0.40	-0.22	-0.28
VF492773	660	0.29	0.32	-0.13	-0.09	-0.15	0.32
VF493001	660	0.56	0.51	-0.31	-0.08	-0.27	0.51
VF491947	660	0.47	0.42	0.42	-0.17	-0.20	-0.17
Form 5							
VF492362	667	0.28	0.20	-0.05	-0.14	0.20	-0.03
VF492671	667	0.58	0.36	-0.22	-0.12	0.36	-0.19
VF493003	667	0.67	0.45	-0.25	-0.18	-0.26	0.45
VF491931	667	0.87	0.44	0.44	-0.22	-0.24	-0.28
VF492814	667	0.52	0.04	0.04	0.02	0.01	-0.12
VF491787	667	0.69	0.28	-0.15	-0.09	0.28	-0.18
VF492270	667	0.35	0.08	0.10	0.08	-0.21	0.00
VF492572	667	0.42	0.30	0.30	0.04	-0.23	-0.26
VF492308	667	0.26	0.19	-0.07	-0.05	-0.08	0.19
VF492543	667	0.35	0.27	-0.17	0.27	-0.10	-0.01
VF492890	667	0.48	0.25	-0.11	0.25	-0.14	-0.09
VF492562	667	0.81	0.41	-0.23	-0.25	0.41	-0.19
Form 6							
VF492585	674	0.60	0.44	0.44	-0.30	-0.23	-0.22
VF492383	674	0.67	0.50	0.50	-0.25	-0.35	-0.12
VF492732	674	0.37	0.17	-0.12	-0.09	0.17	0.06
VF491935	674	0.62	0.51	0.51	-0.29	-0.24	-0.23
VF492009	674	0.09	0.06	-0.09	0.06	-0.21	0.14
VF491864	674	0.67	0.23	-0.01	0.23	-0.22	-0.06
VF491996	674	0.52	0.31	-0.19	-0.22	0.31	-0.05
VF492610	674	0.58	0.15	-0.04	-0.12	0.15	-0.06
VF493013	674	0.30	0.19	-0.13	-0.13	0.19	0.08
VF492102	674	0.30	-0.04	-0.03	-0.02	-0.04	0.13
VF492605	674	0.75	0.45	0.45	-0.27	-0.28	-0.13
VF493008	674	0.71	0.29	-0.10	-0.10	-0.24	0.29

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF492408	667	0.42	0.22	-0.14	0.22	-0.08	-0.15
VF492577	667	0.65	0.51	-0.37	-0.21	-0.17	0.51
VF492030	667	0.49	0.24	-0.18	0.24	-0.01	-0.11
VF492273	667	0.81	0.31	-0.15	-0.18	-0.17	0.31
VF493062	667	0.63	0.39	0.39	-0.01	-0.29	-0.24
VF491966	667	0.64	0.36	0.36	-0.20	-0.21	-0.15
VF492759	667	0.30	0.12	0.07	-0.13	0.12	-0.05
VF492656	667	0.18	0.03	0.22	-0.14	0.03	-0.08
VF492652	667	0.86	0.35	0.35	-0.18	-0.33	-0.08
VF492160	667	0.29	0.12	-0.20	0.03	0.12	-0.11
VF492290	667	0.86	0.37	-0.24	-0.24	0.37	-0.14
VF493024	667	0.21	0.17	-0.16	-0.03	0.03	0.17
Form 8							
VF492066	658	0.43	0.19	0.19	-0.11	-0.13	0.00
VF492431	658	0.97	0.21	-0.12	0.21	-0.18	-0.01
VF492412	658	0.38	0.31	-0.18	-0.23	-0.05	0.31
VF492053	658	0.48	0.45	-0.30	0.45	-0.22	-0.12
VF492703	658	0.37	0.14	-0.02	-0.17	0.04	0.14
VF492284	658	0.80	0.35	-0.19	-0.19	0.35	-0.19
VF491860	658	0.65	0.38	0.38	-0.15	-0.10	-0.32
VF492181	658	0.67	0.30	-0.18	-0.25	0.30	-0.08
VF492676	658	0.55	0.15	0.15	-0.07	-0.07	-0.08
VF492926	658	0.29	0.18	-0.22	0.18	0.00	0.02
VF491970	658	0.65	0.31	-0.15	-0.18	-0.16	0.31
VF493068	658	0.41	0.44	-0.30	-0.17	-0.11	0.44
Form 9							
VF492422	677	0.40	0.30	-0.11	-0.20	0.30	-0.08
VF492594	677	0.41	0.53	-0.35	-0.27	-0.11	0.53
VF492078	677	0.47	0.33	-0.15	0.33	-0.17	-0.12
VF491940	677	0.58	0.45	-0.30	-0.15	-0.19	0.45
VF491976	677	0.50	0.21	-0.14	-0.07	0.21	-0.06
VF492582	677	0.70	0.38	-0.18	-0.15	-0.24	0.38
VF493092	677	0.53	0.51	-0.21	-0.30	-0.27	0.51
VF492831	677	0.21	0.09	-0.07	0.09	0.06	-0.14
VF492192	677	0.40	0.30	0.30	-0.13	-0.22	-0.02
VF492709	677	0.42	0.35	-0.14	0.35	-0.25	-0.16
VF493017	677	0.37	0.12	-0.17	0.12	0.01	0.07
VF493103	677	0.19	-0.13	-0.21	-0.14	-0.13	0.36

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF492260	676	0.41	0.39	-0.30	-0.15	-0.02	0.39
VF492931	676	0.40	0.50	-0.27	-0.31	0.50	-0.10
VF492202	676	0.45	0.27	-0.20	-0.10	0.27	-0.02
VF491874	676	0.57	0.49	0.49	-0.13	-0.26	-0.30
VF493089	676	0.58	0.51	-0.26	0.51	-0.13	-0.33
VF491986	676	0.66	0.50	-0.26	-0.27	-0.23	0.50
VF492209	676	0.51	0.17	-0.18	0.01	0.17	-0.13
VF492647	676	0.59	0.43	0.43	-0.26	-0.14	-0.22
VF492721	676	0.44	0.26	-0.09	0.26	-0.11	-0.17
VF492424	676	0.83	0.30	-0.06	0.30	-0.21	-0.21
VF492722	676	0.43	0.32	-0.14	0.32	-0.18	-0.09
VF492711	676	0.64	0.31	0.31	0.00	-0.37	-0.16

Table A11. Mathematics Grade 7 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF492677	837	0.55	0.41	-0.27	0.41	-0.18	-0.10
VF492878	837	0.28	0.36	-0.15	-0.16	0.36	-0.07
VF492999	837	0.54	0.33	0.33	0.00	-0.27	-0.24
VF492786	837	0.76	0.33	0.33	-0.19	-0.19	-0.17
VF492665	837	0.45	0.38	-0.25	0.38	-0.07	-0.15
VF492951	837	0.51	0.28	-0.16	0.28	-0.21	0.02
VF492236	837	0.18	0.26	-0.04	-0.18	-0.04	0.26
VF492780	837	0.54	0.41	-0.12	-0.21	0.41	-0.22
VF492649	837	0.08	0.06	-0.14	0.03	0.05	0.06
VF492583	837	0.32	0.13	0.06	0.13	-0.21	0.02
VF493043	837	0.30	0.16	0.01	-0.05	0.16	-0.14
VF492625	837	0.27	0.10	-0.06	0.03	-0.05	0.10
Form 2							
VF492915	670	0.79	0.31	0.31	-0.11	-0.24	-0.14
VF492560	670	0.86	0.37	-0.23	0.37	-0.23	-0.17
VF493038	670	0.67	0.57	-0.17	-0.48	-0.16	0.57
VF492407	670	0.48	0.35	0.35	-0.18	-0.21	-0.12
VF492645	670	0.41	0.41	0.41	-0.16	-0.12	-0.23
VF492980	670	0.39	0.12	-0.02	-0.18	0.12	0.04
VF493053	670	0.16	0.26	-0.10	-0.26	0.26	0.11
VF493012	670	0.18	0.07	0.01	0.08	0.07	-0.14
VF492578	670	0.23	0.29	-0.18	-0.08	-0.06	0.29
VF492402	670	0.62	0.43	-0.25	0.43	-0.33	-0.03
VF492673	670	0.19	0.27	-0.14	-0.27	0.14	0.27
VF492836	670	0.37	0.33	0.33	-0.09	-0.13	-0.17
Form 3							
VF493009	668	0.22	0.21	-0.12	-0.12	0.05	0.21
VF493057	668	0.92	0.24	0.24	-0.09	-0.16	-0.16
VF492340	668	0.17	0.27	0.09	-0.21	0.27	-0.09
VF492760	668	0.34	0.36	-0.13	0.36	-0.25	-0.05
VF492567	668	0.45	0.28	-0.05	0.28	-0.20	-0.12
VF492658	668	0.15	0.22	0.04	0.22	-0.23	0.03
VF492616	668	0.12	0.08	0.03	-0.07	-0.02	0.08
VF492765	668	0.28	0.15	-0.14	-0.01	-0.04	0.15
VF492247	668	0.41	0.35	0.35	0.06	-0.40	-0.07
VF492875	668	0.38	0.26	-0.10	0.26	-0.09	-0.17
VF492772	668	0.21	0.09	0.06	-0.13	0.09	-0.07
VF492973	668	0.34	0.25	-0.22	-0.04	0.25	-0.17

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF492835	663	0.72	0.41	0.41	-0.13	-0.23	-0.26
VF492307	663	0.75	0.50	0.50	-0.31	-0.32	-0.12
VF492394	663	0.79	0.39	-0.26	-0.23	0.39	-0.12
VF492888	663	0.52	0.44	-0.17	0.44	-0.18	-0.25
VF492654	663	0.21	0.13	-0.19	-0.13	0.13	0.13
VF493036	663	0.47	0.41	-0.06	0.41	-0.30	-0.17
VF492937	663	0.93	0.07	-0.03	-0.08	0.07	0.00
VF493061	663	0.44	0.51	-0.41	-0.14	-0.06	0.51
VF492959	663	0.31	0.11	0.11	-0.17	0.05	0.03
VF492756	663	0.55	0.12	-0.02	0.06	0.12	-0.16
VF492896	663	0.23	0.35	-0.37	-0.20	0.17	0.35
VF493071	663	0.37	0.31	0.31	-0.21	-0.15	-0.05
Form 5							
VF492663	668	0.73	0.46	-0.24	0.46	-0.21	-0.28
VF492871	668	0.78	0.44	-0.19	-0.31	0.44	-0.19
VF493021	668	0.38	0.33	-0.11	0.33	-0.19	-0.09
VF492859	668	0.65	0.28	0.28	-0.21	-0.15	0.00
VF492666	668	0.52	0.47	-0.20	-0.19	-0.27	0.47
VF492946	668	0.37	0.27	-0.24	0.27	-0.14	0.10
VF492546	668	0.29	0.22	0.08	-0.25	-0.23	0.22
VF492864	668	0.64	0.48	0.48	-0.32	-0.26	-0.09
VF493052	668	0.52	0.26	-0.16	-0.18	0.26	-0.02
VF492640	668	0.33	0.23	-0.17	0.23	-0.04	-0.04
VF492748	668	0.74	0.44	0.44	-0.24	-0.23	-0.21
VF493059	668	0.22	-0.05	-0.20	0.34	-0.05	-0.21
Form 6							
VF492413	667	0.50	0.20	-0.10	-0.09	-0.09	0.20
VF492966	667	0.69	0.49	0.49	-0.29	-0.13	-0.29
VF492853	667	0.59	0.36	-0.13	0.36	-0.24	-0.18
VF492576	667	0.39	0.25	-0.13	0.25	-0.12	-0.05
VF492910	667	0.56	0.14	-0.21	0.14	-0.11	0.00
VF493067	667	0.46	0.44	-0.22	-0.18	-0.21	0.44
VF492538	667	0.69	0.57	0.57	-0.27	-0.22	-0.37
VF492597	667	0.34	0.17	0.01	-0.13	0.17	-0.08
VF492846	667	0.57	0.37	-0.10	0.37	-0.29	-0.09
VF492653	667	0.10	0.17	-0.26	-0.17	0.23	0.17
VF492987	667	0.39	0.43	-0.35	-0.26	0.43	0.09
VF492834	667	0.63	0.41	-0.16	-0.26	0.41	-0.17

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF492824	654	0.75	0.42	-0.12	0.42	-0.26	-0.26
VF492828	654	0.35	0.44	0.44	-0.12	-0.12	-0.26
VF492302	654	0.84	0.36	0.36	-0.23	-0.17	-0.21
VF492708	654	0.44	0.45	-0.13	-0.37	-0.16	0.45
VF492967	654	0.53	0.25	-0.19	0.25	-0.11	-0.04
VF492898	654	0.43	0.34	0.34	-0.19	-0.10	-0.12
VF492644	654	0.30	0.14	0.20	-0.23	0.14	-0.17
VF492720	654	0.41	0.53	-0.26	-0.23	-0.17	0.53
VF492259	654	0.39	0.42	-0.16	-0.24	0.42	-0.12
VF493077	654	0.77	0.40	-0.24	-0.32	0.40	-0.10
VF492906	654	0.31	0.47	0.47	-0.02	-0.38	-0.17
VF492419	654	0.41	0.37	-0.18	0.37	-0.08	-0.24
Form 8							
VF492531	660	0.92	0.26	-0.21	-0.12	0.26	-0.11
VF492867	660	0.62	0.54	-0.31	0.54	-0.33	-0.14
VF492731	660	0.90	0.33	0.33	-0.22	-0.14	-0.18
VF493064	660	0.82	0.27	-0.20	-0.08	-0.17	0.27
VF493073	660	0.76	0.33	-0.07	0.33	-0.30	-0.10
VF492778	660	0.48	0.52	0.52	-0.15	-0.13	-0.38
VF492901	660	0.37	0.31	-0.31	0.31	-0.06	0.05
VF493046	660	0.42	0.31	0.00	-0.15	-0.24	0.31
VF493019	660	0.50	0.43	-0.42	-0.15	0.43	0.03
VF492995	660	0.62	0.33	-0.15	-0.19	-0.15	0.33
VF492295	660	0.68	0.36	-0.07	-0.18	0.36	-0.33
VF492629	660	0.29	0.09	0.09	0.04	-0.14	0.00
Form 9							
VF492929	671	0.93	0.21	-0.08	-0.11	-0.16	0.21
VF492425	671	0.55	0.42	0.42	-0.15	-0.27	-0.16
VF492908	671	0.23	0.23	0.23	0.16	-0.12	-0.28
VF492672	671	0.88	0.41	-0.27	-0.21	0.41	-0.22
VF492992	671	0.37	0.52	-0.30	-0.17	-0.18	0.52
VF492636	671	0.26	-0.03	-0.08	-0.12	0.13	-0.03
VF492701	671	0.37	0.24	-0.09	0.01	0.24	-0.22
VF492830	671	0.64	0.48	-0.17	0.48	-0.33	-0.21
VF492762	671	0.23	0.10	0.10	0.11	-0.22	-0.02
VF492986	671	0.49	0.31	-0.16	0.31	-0.12	-0.13
VF492599	671	0.47	0.36	-0.17	-0.21	0.36	-0.08
VF492955	671	0.56	0.38	-0.26	0.38	-0.20	-0.08

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF492771	667	0.69	0.38	0.38	-0.24	-0.18	-0.14
VF492921	667	0.46	-0.05	0.03	-0.05	-0.03	0.06
VF492357	667	0.32	0.25	0.25	-0.23	-0.03	-0.11
VF492975	667	0.56	0.52	-0.16	-0.37	0.52	-0.19
VF493015	667	0.48	0.33	0.33	-0.06	-0.26	-0.14
VF492595	667	0.13	0.14	0.03	-0.14	0.14	0.02
VF492714	667	0.19	0.05	-0.08	0.05	-0.12	0.10
VF493004	667	0.21	0.40	-0.15	-0.13	-0.13	0.40
VF493031	667	0.14	0.41	-0.34	-0.07	0.11	0.41
VF492589	667	0.20	0.32	0.02	-0.24	-0.05	0.32
VF492696	667	0.69	0.47	-0.31	-0.24	0.47	-0.17
VF492861	667	0.55	0.37	-0.19	0.37	-0.17	-0.19

Table A12. Mathematics Grade 8 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	<u>Option Discrimination (MC)</u>			
				A	B	C	D
Form 1							
VF492845	770	0.41	0.33	-0.17	0.33	-0.13	-0.14
VF492719	770	0.59	0.51	-0.04	-0.32	0.51	-0.37
VF492414	770	0.58	0.51	0.51	-0.23	-0.29	-0.20
VF494727	770	0.45	0.40	-0.22	0.40	-0.02	-0.29
VF491873	770	0.25	0.37	-0.10	-0.33	0.37	0.02
VF493157	770	0.50	0.29	-0.22	0.29	-0.25	0.06
VF493011	770	0.69	0.49	0.49	-0.26	-0.23	-0.23
VF494810	770	0.37	0.13	0.13	-0.04	-0.02	-0.10
VF492024	770	0.27	0.19	0.02	0.19	-0.15	-0.06
VF492409	770	0.18	0.08	0.08	-0.04	-0.10	0.08
VF492278	770	0.68	0.52	-0.22	-0.32	0.52	-0.22
VF494716	770	0.28	0.35	-0.11	-0.16	-0.08	0.35
Form 2							
VF492712	665	0.42	0.46	0.46	-0.39	-0.14	0.01
VF491918	665	0.13	0.21	0.35	-0.36	-0.23	0.21
VF492920	665	0.59	0.57	-0.33	-0.25	-0.25	0.57
VF493125	665	0.44	0.35	-0.04	-0.22	-0.21	0.35
VF491949	665	0.66	0.44	-0.20	0.44	-0.33	-0.11
VF492579	665	0.47	0.23	-0.01	-0.25	0.23	-0.07
VF492438	665	0.64	0.36	-0.17	-0.20	0.36	-0.16
VF492268	665	0.31	0.20	-0.19	-0.09	0.03	0.20
VF493113	665	0.61	0.35	-0.17	0.35	-0.19	-0.16
VF493083	665	0.30	0.44	0.44	-0.15	-0.14	-0.19
VF492178	665	0.46	0.42	0.42	-0.24	-0.21	-0.09
VF493160	665	0.29	0.17	-0.13	0.17	-0.10	0.05
Form 3							
VF492856	667	0.59	0.53	-0.24	-0.20	-0.32	0.53
VF492440	667	0.77	0.31	0.31	-0.18	-0.12	-0.18
VF494113	667	0.61	0.38	0.38	-0.22	-0.20	-0.12
VF492889	667	0.85	0.34	-0.20	0.34	-0.20	-0.17
VF492385	667	0.62	0.34	-0.16	-0.24	0.34	-0.09
VF491971	667	0.28	0.07	0.07	-0.04	0.12	-0.23
VF493115	667	0.80	0.20	-0.06	0.20	-0.12	-0.13
VF491920	667	0.77	0.39	-0.08	0.39	-0.25	-0.23
VF494968	667	0.37	0.17	-0.07	0.01	0.17	-0.13
VF492592	667	0.66	0.45	-0.24	-0.23	-0.21	0.45
VF492018	667	0.33	0.47	-0.17	-0.12	-0.24	0.47
VF494801	667	0.57	0.32	-0.19	-0.18	0.32	-0.06

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF492563	667	0.59	0.22	-0.19	-0.08	-0.08	0.22
VF493112	667	0.75	0.32	-0.12	-0.13	0.32	-0.24
VF493045	667	0.64	0.33	-0.07	0.33	-0.28	-0.14
VF492226	667	0.55	0.41	-0.11	-0.25	-0.23	0.41
VF492199	667	0.52	0.44	0.44	-0.18	-0.24	-0.17
VF491824	667	0.36	0.35	-0.07	-0.24	-0.08	0.35
VF492258	667	0.53	0.23	0.05	-0.24	-0.27	0.23
VF492393	667	0.48	0.35	0.35	-0.23	-0.14	-0.09
VF494709	667	0.62	0.16	0.16	-0.15	-0.16	0.07
VF491999	667	0.26	0.28	-0.15	0.28	-0.12	0.03
VF494819	667	0.29	0.28	-0.26	-0.03	0.28	0.04
VF491857	667	0.61	0.41	-0.20	0.41	-0.17	-0.24
Form 5							
VF492426	668	0.80	0.31	-0.15	-0.21	-0.18	0.31
VF492863	668	0.52	0.50	0.50	-0.14	-0.24	-0.32
VF492231	668	0.37	0.16	0.16	0.04	-0.14	-0.20
VF494751	668	0.47	0.31	-0.15	0.31	-0.12	-0.16
VF493605	668	0.16	-0.03	-0.03	0.02	0.10	-0.10
VF493034	668	0.46	0.42	-0.21	-0.18	-0.17	0.42
VF491980	668	0.49	0.51	-0.20	-0.25	-0.24	0.51
VF492410	668	0.58	0.45	-0.20	-0.27	0.45	-0.19
VF491938	668	0.83	0.39	-0.24	0.39	-0.16	-0.20
VF494769	668	0.69	0.38	-0.18	-0.24	0.38	-0.17
VF493107	668	0.40	0.30	-0.15	0.01	0.30	-0.22
VF494898	668	0.28	0.01	0.05	-0.13	0.01	0.10
Form 6							
VF492726	672	0.27	0.33	-0.40	0.03	0.09	0.33
VF491862	672	0.93	0.27	-0.18	-0.07	-0.16	0.27
VF493088	672	0.64	0.51	0.51	-0.35	-0.23	-0.16
VF494120	672	0.34	0.23	-0.04	0.23	-0.17	-0.05
VF494978	672	0.13	0.22	-0.14	-0.18	0.15	0.22
VF492360	672	0.35	0.12	0.05	0.12	-0.16	-0.02
VF492028	672	0.43	0.40	-0.10	-0.28	0.40	-0.11
VF492586	672	0.40	0.19	0.19	-0.04	-0.01	-0.20
VF494699	672	0.57	0.39	-0.21	-0.11	0.39	-0.20
VF491943	672	0.70	0.38	0.38	-0.26	-0.16	-0.12
VF493040	672	0.69	0.50	-0.19	-0.26	0.50	-0.30
VF493121	672	0.42	0.40	-0.23	-0.16	-0.13	0.40

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF492874	677	0.39	0.24	0.03	0.24	-0.24	-0.04
VF492212	677	0.48	0.29	0.29	-0.11	-0.06	-0.19
VF493150	677	0.42	0.43	0.43	-0.07	-0.19	-0.26
VF492216	677	0.66	0.28	-0.22	0.28	-0.08	-0.15
VF491915	677	0.31	0.22	0.22	-0.03	-0.09	-0.12
VF493097	677	0.34	0.14	0.00	-0.08	-0.06	0.14
VF494983	677	0.31	0.13	-0.21	0.13	0.15	-0.06
VF491965	677	0.62	0.38	-0.10	-0.18	-0.28	0.38
VF494099	677	0.35	0.10	-0.12	-0.16	0.10	0.13
VF492917	677	0.73	0.41	-0.21	-0.28	0.41	-0.13
VF494949	677	0.30	0.18	-0.02	-0.13	0.18	-0.03
VF492436	677	0.64	0.32	0.32	-0.22	-0.11	-0.16
Form 8							
VF492400	655	0.50	0.31	-0.15	0.31	-0.23	-0.08
VF492544	655	0.43	0.39	-0.19	0.39	-0.10	-0.21
VF491991	655	0.58	0.36	0.36	-0.23	-0.18	-0.09
VF494747	655	0.26	0.35	-0.25	-0.10	0.35	0.01
VF492430	655	0.56	0.36	-0.18	-0.26	0.36	-0.10
VF493025	655	0.30	0.24	0.01	-0.19	-0.11	0.24
VF493132	655	0.63	0.15	0.15	-0.13	-0.11	-0.03
VF491929	655	0.32	0.38	-0.08	-0.22	-0.14	0.38
VF494125	655	0.29	0.24	-0.12	0.24	-0.21	0.05
VF494963	655	0.25	0.10	-0.12	0.10	-0.03	0.12
VF492272	655	0.67	0.43	-0.18	-0.23	-0.23	0.43
VF492434	655	0.22	0.02	-0.05	0.04	0.02	-0.02
Form 9							
VF492880	670	0.19	0.30	0.08	0.11	0.30	-0.44
VF492550	670	0.73	0.39	0.39	-0.24	-0.17	-0.17
VF491907	670	0.59	0.28	-0.11	0.28	-0.20	-0.07
VF492063	670	0.42	0.37	-0.03	-0.26	0.37	-0.18
VF494972	670	0.35	-0.06	-0.06	0.07	0.01	0.01
VF492405	670	0.65	0.45	-0.14	0.45	-0.31	-0.16
VF492568	670	0.53	0.39	-0.18	0.39	-0.19	-0.16
VF491975	670	0.69	0.43	-0.16	-0.24	-0.20	0.43
VF494760	670	0.51	0.46	-0.33	0.46	-0.15	-0.09
VF492345	670	0.35	0.28	0.06	-0.15	-0.20	0.28
VF493000	670	0.76	0.45	0.45	-0.19	-0.32	-0.19
VF493117	670	0.38	0.40	-0.22	-0.10	-0.18	0.40

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF492848	667	0.20	0.23	0.12	0.23	-0.17	-0.20
VF492420	667	0.74	0.36	-0.22	-0.15	-0.22	0.36
VF492907	667	0.75	0.40	-0.29	0.40	-0.18	-0.16
VF494776	667	0.46	0.41	-0.27	-0.19	0.41	-0.08
VF492165	667	0.32	0.34	0.01	-0.07	-0.29	0.34
VF492289	667	0.60	0.49	0.49	-0.23	-0.26	-0.20
VF494928	667	0.30	0.19	-0.28	0.19	0.15	0.01
VF492008	667	0.45	0.42	-0.23	-0.21	0.42	-0.09
VF493102	667	0.33	0.29	0.29	-0.04	-0.10	-0.18
VF492439	667	0.43	0.35	-0.08	-0.20	-0.20	0.35
VF491923	667	0.84	0.37	0.37	-0.25	-0.15	-0.20
VF493159	667	0.60	0.36	-0.12	-0.23	-0.22	0.36

Science

Table A13. Science Grade 4 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF385248	798	0.80	0.19	-0.08	-0.13	-0.13	0.19
VF388627	1515	0.51	0.34	0.34	-0.19	-0.22	-0.03
VF386811	798	0.63	0.33	0.33	-0.15	-0.14	-0.19
VF386797	798	0.63	0.41	-0.12	0.41	-0.22	-0.29
VF386826	798	0.65	0.33	-0.18	-0.15	-0.21	0.33
VF386788	798	0.80	0.35	-0.21	-0.15	-0.26	0.35
VF386795	798	0.68	0.27	0.27	-0.13	-0.07	-0.21
VF386911	798	0.60	0.35	0.35	-0.17	-0.24	-0.10
VF386873	798	0.66	0.29	-0.23	0.29	-0.06	-0.15
VF417698	798	0.83	0.42	0.42	-0.27	-0.22	-0.23
VF417694	798	0.78	0.35	-0.23	-0.17	0.35	-0.18
VF386881	798	0.71	0.39	-0.14	0.39	-0.23	-0.22
Form 2							
VF385244	707	0.49	0.24	0.24	-0.02	-0.20	-0.07
VF388696	707	0.36	0.22	-0.09	-0.11	0.22	-0.06
VF386815	707	0.79	0.30	-0.18	0.30	-0.12	-0.18
VF386787	707	0.82	0.38	-0.20	-0.16	-0.27	0.38
VF386799	707	0.68	0.48	-0.26	-0.19	0.48	-0.29
VF386789	707	0.42	0.33	-0.17	0.33	-0.16	-0.11
VF386806	707	0.80	0.36	-0.22	0.36	-0.12	-0.23
VF386848	707	0.48	0.28	-0.15	-0.05	-0.16	0.28
VF386857	707	0.55	0.27	0.27	-0.19	-0.04	-0.26
VF386901	707	0.42	0.35	-0.30	0.02	0.35	-0.12
VF386847	707	0.92	0.30	-0.11	-0.22	0.30	-0.17
VF386861	707	0.27	0.19	-0.06	0.19	-0.03	-0.14
VF385250	698	0.93	0.34	-0.24	-0.06	0.34	-0.22
Form 3							
VF388708	698	0.80	0.42	-0.22	-0.14	-0.28	0.42
VF407131	698	0.84	0.33	0.33	-0.19	-0.22	-0.19
VF406427	698	0.61	0.38	-0.22	0.38	-0.13	-0.23
VF407137	698	0.95	0.21	-0.15	-0.16	0.21	-0.06
VF407152	698	0.43	0.28	-0.17	-0.17	-0.03	0.28
VF407128	698	0.77	0.43	-0.25	0.43	-0.23	-0.20
VF393924	698	0.70	0.30	-0.04	0.30	-0.18	-0.23
VF393826	698	0.64	0.40	-0.19	0.40	-0.23	-0.25
VF393954	698	0.77	0.38	-0.21	-0.16	0.38	-0.22
VF393813	698	0.91	0.25	-0.10	-0.13	0.25	-0.19
VF393854	698	0.96	0.18	0.18	-0.12	-0.11	-0.08

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF388716	707	0.76	0.36	-0.22	0.36	-0.14	-0.22
VF385245	707	0.85	0.20	-0.02	-0.17	-0.16	0.20
VF407139	707	0.61	0.27	-0.10	-0.20	0.27	-0.12
VF406430	707	0.65	0.26	-0.07	0.26	-0.20	-0.11
VF407138	707	0.41	0.16	0.16	-0.06	-0.18	-0.04
VF407141	707	0.87	0.14	-0.07	0.00	0.14	-0.12
VF406457	707	0.83	0.33	-0.21	0.33	-0.16	-0.19
VF393911	707	0.86	0.30	0.30	-0.16	-0.19	-0.14
VF393816	707	0.51	0.38	-0.20	0.38	-0.09	-0.24
VF393969	707	0.95	0.18	-0.12	-0.05	0.18	-0.13
VF393950	707	0.42	0.26	-0.23	0.26	0.02	-0.08
VF394018	707	0.38	0.06	-0.24	0.06	0.01	0.11
Form 5							
VF385249	706	0.83	0.32	-0.21	0.32	-0.14	-0.17
VF385241	706	0.42	0.27	-0.15	0.04	0.27	-0.20
VF416398	706	0.74	0.40	-0.18	0.40	-0.29	-0.16
VF311630	706	0.29	0.14	0.14	-0.13	0.02	0.01
VF311628	706	0.49	0.09	0.09	-0.18	-0.09	0.07
VF311629	706	0.55	0.33	-0.14	-0.15	-0.16	0.33
VF416433	706	0.50	0.11	0.08	0.11	-0.08	-0.23
VF416448	706	0.62	0.19	0.19	-0.04	-0.12	-0.19
VF387280	706	0.29	0.22	-0.04	-0.09	0.22	-0.10
VF387287	706	0.84	0.28	-0.13	-0.23	-0.13	0.28
VF387305	706	0.46	0.24	-0.01	-0.19	0.24	-0.11
VF387267	706	0.41	0.26	-0.05	0.26	-0.12	-0.16
VF311567	1413	0.67	0.35	0.35	-0.21	-0.14	-0.25
Form 6							
VF311632	718	0.89	0.30	-0.19	0.30	-0.21	-0.07
VF311640	718	0.79	0.22	0.22	-0.13	-0.08	-0.14
VF311572	718	0.70	0.38	-0.18	-0.26	-0.15	0.38
VF311535	718	0.85	0.29	-0.18	-0.17	0.29	-0.15
VF416378	718	0.81	0.28	0.28	-0.17	-0.18	-0.11
VF311559	718	0.53	0.32	0.32	-0.10	-0.20	-0.23
VF387252	718	0.64	0.41	-0.22	0.41	-0.26	-0.12
VF387402	718	0.36	0.19	0.19	0.09	-0.15	-0.19
VF387256	718	0.52	0.38	-0.18	-0.18	-0.15	0.38
VF387433	718	0.50	0.31	-0.20	-0.16	0.31	-0.05
VF387295	718	0.50	0.21	-0.11	-0.08	-0.11	0.21

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF386869	714	0.93	0.34	-0.14	-0.19	0.34	-0.26
VF386851	714	0.75	0.28	-0.16	-0.21	-0.03	0.28
VF311348	714	0.50	0.32	-0.30	0.32	-0.07	-0.09
VF311378	714	0.91	0.33	-0.07	-0.20	-0.24	0.33
VF311318	714	0.56	0.24	-0.04	-0.21	0.24	-0.15
VF416364	714	0.21	0.16	0.16	-0.21	0.10	-0.17
VF311330	714	0.63	0.29	-0.09	-0.27	0.29	-0.09
VF416382	714	0.41	0.29	-0.05	-0.25	-0.13	0.29
VF311584	714	0.88	0.39	-0.20	-0.27	0.39	-0.18
VF311586	714	0.69	0.43	-0.28	0.43	-0.13	-0.25
VF311548	714	0.78	0.33	-0.27	-0.14	0.33	-0.09
VF311576	714	0.83	0.27	-0.18	-0.17	-0.12	0.27
Form 8							
VF393818	697	0.74	0.40	0.40	-0.31	-0.20	-0.12
VF393964	697	0.42	0.19	-0.24	-0.05	0.06	0.19
VF393699	697	0.54	0.35	-0.25	-0.07	0.35	-0.14
VF393636	697	0.91	0.28	-0.21	0.28	-0.20	-0.11
VF393646	697	0.79	0.28	0.28	-0.22	-0.10	-0.14
VF417679	697	0.63	0.15	-0.13	0.15	0.00	-0.16
VF417684	697	0.77	0.31	-0.19	-0.23	0.31	-0.16
VF311363	697	0.88	0.41	-0.21	-0.26	-0.21	0.41
VF416359	697	0.33	0.26	-0.18	0.26	-0.09	-0.05
VF311367	697	0.96	0.17	-0.07	-0.13	0.17	-0.10
VF311343	697	0.71	0.40	0.40	-0.21	-0.27	-0.14
VF311353	697	0.72	0.40	0.40	-0.18	-0.24	-0.19
Form 9							
VF311580	695	0.72	0.31	-0.13	0.31	-0.16	-0.20
VF311567	1413	0.67	0.35	0.35	-0.21	-0.14	-0.25
VF386739	695	0.76	0.39	-0.20	0.39	-0.29	-0.09
VF417655	695	0.71	0.34	-0.24	-0.19	0.34	-0.11
VF386736	695	0.89	0.37	-0.14	0.37	-0.18	-0.28
VF386712	695	0.45	0.29	-0.13	-0.01	0.29	-0.23
VF386706	695	0.48	0.25	0.25	-0.02	-0.24	-0.13
VF393724	695	0.75	0.54	-0.24	-0.34	-0.26	0.54
VF393695	695	0.83	0.45	-0.14	0.45	-0.30	-0.25
VF393721	695	0.59	0.38	-0.12	-0.20	-0.23	0.38
VF393713	695	0.61	0.22	-0.08	0.22	-0.11	-0.13
VF417683	695	0.31	0.13	-0.08	-0.13	0.06	0.13

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF388627	1515	0.51	0.34	0.34	-0.19	-0.22	-0.03
VF385246	717	0.77	0.36	-0.20	-0.24	-0.15	0.36
VF386512	717	0.71	0.47	-0.24	-0.29	-0.20	0.47
VF386582	717	0.93	0.27	0.27	-0.09	-0.17	-0.18
VF386528	717	0.65	0.28	-0.19	0.28	-0.05	-0.17
VF386520	717	0.65	0.21	-0.10	-0.15	0.21	-0.10
VF386579	717	0.83	0.19	-0.13	0.01	-0.19	0.19
VF386702	717	0.67	0.33	-0.19	0.33	-0.20	-0.18
VF417662	717	0.84	0.36	0.36	-0.27	-0.18	-0.15
VF386718	717	0.79	0.38	-0.21	0.38	-0.18	-0.24
VF386732	717	0.29	0.22	-0.11	0.22	-0.07	-0.06
VF386697	717	0.27	0.11	-0.10	0.02	-0.07	0.11

Table A14. Science Grade 8 Classical Statistics for Field Test Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 1							
VF407480	782	0.55	0.41	-0.22	0.41	-0.22	-0.21
VF407468	782	0.43	0.20	0.20	-0.08	-0.13	-0.05
VF394814	782	0.41	0.29	-0.06	-0.10	-0.20	0.29
VF394815	782	0.29	0.10	-0.23	0.16	-0.14	0.10
VF394816	782	0.32	0.26	0.26	-0.13	-0.11	-0.06
VF394806	782	0.32	0.13	-0.19	0.13	-0.26	0.22
VF394811	782	0.29	0.13	0.13	-0.16	0.13	-0.22
VF407166	782	0.60	0.44	-0.24	0.44	-0.22	-0.16
VF407168	782	0.25	0.17	-0.15	-0.09	0.06	0.17
VF407160	782	0.41	0.33	0.33	-0.14	-0.26	-0.03
VF407167	782	0.57	0.43	-0.12	-0.30	-0.17	0.43
VF407159	782	0.44	0.43	-0.13	0.43	-0.26	-0.16
Form 2							
VF407471	658	0.73	0.44	-0.19	0.44	-0.30	-0.19
VF407482	658	0.61	0.33	-0.17	-0.25	0.33	-0.15
VF407165	658	0.45	0.29	-0.14	0.29	-0.10	-0.13
VF407162	658	0.37	0.02	-0.05	0.02	-0.23	0.15
VF407161	658	0.39	0.43	-0.18	-0.17	-0.17	0.43
VF407169	658	0.55	0.43	-0.17	-0.23	-0.21	0.43
VF407164	658	0.42	0.30	-0.05	-0.15	0.30	-0.17
VF394810	658	0.72	0.44	-0.23	-0.24	-0.24	0.44
VF394809	658	0.61	0.34	0.34	-0.19	-0.22	-0.09
VF394807	658	0.61	0.30	-0.08	0.30	-0.16	-0.19
VF394808	658	0.61	0.37	-0.25	0.37	-0.16	-0.13
VF394812	658	0.48	0.25	0.25	-0.13	-0.13	-0.08
Form 3							
VF407483	662	0.53	0.22	-0.08	0.22	-0.11	-0.12
VF407485	662	0.51	0.23	-0.13	-0.08	0.23	-0.12
VF407140	662	0.26	-0.03	-0.14	-0.03	-0.10	0.22
VF407213	662	0.48	-0.06	-0.12	-0.06	0.11	0.07
VF407155	662	0.70	0.39	-0.12	-0.28	0.39	-0.25
VF407254	662	0.40	0.29	-0.05	-0.25	-0.11	0.29
VF407156	662	0.81	0.31	-0.22	-0.19	0.31	-0.13
VF407356	662	0.88	0.31	-0.17	0.31	-0.18	-0.17
VF407352	662	0.43	0.43	-0.21	-0.18	0.43	-0.28
VF407323	662	0.44	0.06	-0.13	-0.05	0.06	0.13
VF407325	662	0.81	0.22	0.22	-0.07	-0.16	-0.20
VF407345	662	0.67	0.27	-0.16	-0.13	-0.12	0.27

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 4							
VF407470	664	0.44	-0.04	-0.09	-0.04	-0.03	0.10
VF407475	664	0.71	0.25	-0.18	-0.15	-0.14	0.25
VF407339	664	0.69	0.29	0.29	-0.13	-0.20	-0.17
VF407392	664	0.68	0.29	-0.11	-0.20	-0.12	0.29
VF407327	664	0.63	0.20	-0.09	-0.15	0.20	-0.07
VF407384	664	0.37	0.25	-0.11	-0.10	0.25	-0.08
VF407330	664	0.61	0.34	-0.12	0.34	-0.24	-0.13
VF407248	664	0.17	-0.01	0.22	-0.18	-0.07	-0.01
VF524715	664	0.30	-0.09	-0.03	-0.09	0.09	0.03
VF407154	664	0.44	0.21	-0.18	-0.17	0.21	0.03
VF407227	664	0.33	0.06	-0.06	0.06	-0.05	0.04
VF407242	664	0.72	0.26	0.26	-0.14	-0.14	-0.11
Form 5							
VF407478	665	0.68	0.20	0.20	-0.12	-0.14	-0.08
VF407473	665	0.43	0.23	-0.08	-0.10	0.23	-0.12
VF313256	665	0.48	0.23	-0.08	0.23	-0.20	-0.04
VF313265	665	0.13	0.04	0.04	-0.16	-0.10	0.16
VF313280	665	0.44	0.43	-0.21	-0.21	-0.16	0.43
VF313272	665	0.19	0.22	-0.15	-0.25	0.18	0.22
VF313274	665	0.30	0.26	0.26	-0.26	0.04	-0.11
VF394777	665	0.54	0.24	0.05	-0.18	-0.20	0.24
VF394786	665	0.17	0.08	0.14	0.08	-0.07	-0.16
VF394784	665	0.30	-0.01	-0.01	-0.09	0.18	-0.12
VF394773	665	0.33	0.11	0.05	0.11	-0.10	-0.16
VF524713	665	0.35	0.10	-0.20	0.10	-0.19	0.12
Form 6							
VF394571	667	0.67	0.37	-0.17	0.37	-0.20	-0.22
VF394557	667	0.55	0.28	0.28	-0.01	-0.26	-0.21
VF394775	667	0.64	0.23	-0.12	-0.18	-0.05	0.23
VF394780	667	0.56	0.31	0.31	-0.07	-0.18	-0.17
VF394782	667	0.18	-0.13	0.38	-0.20	-0.18	-0.13
VF394789	667	0.09	-0.16	-0.16	0.33	-0.20	-0.15
VF394787	667	0.33	0.24	-0.16	-0.19	0.02	0.24
VF313262	667	0.29	-0.07	-0.14	-0.08	-0.07	0.20
VF313276	667	0.27	0.11	-0.08	0.11	-0.04	0.06
VF313277	667	0.74	0.42	0.42	-0.22	-0.30	-0.14
VF313269	667	0.68	0.32	-0.24	-0.17	-0.12	0.32
VF313281	667	0.53	0.34	-0.17	0.34	-0.26	-0.10

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
Form 7							
VF394554	672	0.28	0.12	-0.11	-0.09	0.03	0.12
VF394574	672	0.63	0.45	-0.23	-0.17	-0.25	0.45
VF407454	672	0.74	0.39	0.39	-0.25	-0.19	-0.14
VF407456	672	0.61	0.28	-0.23	-0.20	0.28	-0.01
VF407466	672	0.49	0.27	0.27	-0.05	-0.28	-0.12
VF407460	672	0.55	0.12	-0.23	0.12	-0.18	0.09
VF407455	672	0.38	0.07	-0.13	0.07	-0.06	0.08
VF313294	672	0.45	0.19	0.19	-0.08	-0.08	-0.07
VF313297	672	0.51	0.22	-0.14	0.22	-0.15	0.00
VF313292	672	0.56	0.29	-0.13	-0.04	-0.24	0.29
VF313293	672	0.24	0.09	-0.14	0.17	-0.17	0.09
VF313300	672	0.53	0.34	-0.08	0.34	-0.15	-0.22
Form 8							
VF394567	653	0.65	0.30	-0.22	0.30	-0.18	-0.08
VF394569	653	0.43	0.44	-0.14	-0.18	-0.26	0.44
VF313287	653	0.26	0.08	-0.07	0.08	-0.11	0.08
VF313288	653	0.39	0.06	0.02	0.06	-0.07	-0.02
VF313289	653	0.48	0.31	-0.20	-0.06	0.31	-0.16
VF313291	653	0.65	0.40	0.40	-0.18	-0.24	-0.20
VF313295	653	0.38	0.18	0.03	-0.16	-0.09	0.18
VF407458	653	0.56	0.30	-0.17	0.30	-0.26	0.02
VF407461	653	0.81	0.32	0.32	-0.24	-0.16	-0.14
VF407465	653	0.36	0.17	-0.22	0.07	-0.12	0.17
VF407467	653	0.67	0.46	0.46	-0.32	-0.13	-0.21
VF407469	653	0.77	0.22	-0.17	0.22	-0.07	-0.17
Form 9							
VF394565	662	0.82	0.40	-0.28	0.40	-0.23	-0.11
VF394561	662	0.66	0.33	-0.12	-0.11	-0.29	0.33
VF388393	662	0.34	0.23	0.23	-0.12	-0.23	0.13
VF388404	662	0.36	0.11	0.16	0.11	-0.21	-0.13
VF388442	662	0.43	0.07	0.11	-0.16	0.07	-0.09
VF388418	662	0.12	-0.03	-0.03	0.33	-0.22	-0.22
VF388547	662	0.11	0.15	-0.09	-0.16	0.12	0.15
VF394486	662	0.72	0.23	-0.20	-0.16	0.23	-0.05
VF394482	662	0.22	0.25	0.25	-0.17	0.02	-0.09
VF394497	662	0.37	0.15	-0.06	0.15	-0.16	0.05
VF394477	662	0.42	0.48	-0.15	-0.20	-0.23	0.48
VF394491	662	0.66	0.33	-0.16	0.33	-0.28	-0.07

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
			Form 10				
VF394547	668	0.59	0.36	-0.13	0.36	-0.23	-0.22
VF394566	668	0.79	0.30	-0.22	0.30	-0.14	-0.11
VF394502	668	0.52	0.27	-0.26	-0.02	0.27	-0.07
VF394488	668	0.47	0.31	-0.18	-0.27	-0.06	0.31
VF394474	668	0.40	0.22	-0.01	0.22	-0.10	-0.14
VF394505	668	0.33	0.14	0.03	-0.18	-0.08	0.14
VF394514	668	0.15	-0.07	-0.07	0.14	-0.04	-0.07
VF388397	668	0.23	-0.12	0.17	0.00	-0.12	-0.06
VF388454	668	0.44	0.25	-0.16	-0.27	0.25	0.00
VF388425	668	0.29	0.05	0.13	0.05	-0.01	-0.21
VF388503	668	0.30	0.30	-0.15	-0.07	0.30	-0.12
VF388413	668	0.48	0.35	-0.16	0.35	-0.19	-0.11

SAWS

Table A15. Distributions of SAWS Grade 3 Field Test - 12-point Prompt

Score	Form											
	<u>1</u>		<u>2</u>		<u>3</u>		<u>4</u>		<u>5</u>		<u>6</u>	
	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N
0	49	4.0	44	3.6	33	2.7	41	3.4	55	4.5	48	3.9
1	22	1.8	36	3.0	25	2.1	9	0.7	15	1.2	1	0.1
2	24	2.0	21	1.7	11	0.9	13	1.1	20	1.6	13	1.1
3	13	1.1	12	1.0	8	0.7	13	1.1	15	1.2	7	0.6
4	136	11.1	172	14.3	193	16.0	147	12.2	181	14.8	217	17.8
5	120	9.8	135	11.2	127	10.5	106	8.8	123	10.1	136	11.2
6	147	12.0	153	12.7	151	12.5	165	13.7	153	12.5	141	11.6
7	173	14.1	173	14.3	167	13.8	175	14.5	187	15.3	164	13.5
8	251	20.4	179	14.8	273	22.6	262	21.7	224	18.3	266	21.8
9	76	6.2	91	7.5	67	5.5	85	7.0	73	6.0	70	5.7
10	74	6.0	66	5.5	45	3.7	59	4.9	62	5.1	58	4.8
11	54	4.4	55	4.6	52	4.3	53	4.4	49	4.0	47	3.9
12	89	7.2	70	5.8	57	4.7	78	6.5	65	5.3	51	4.2
N	1228		1207		1209		1206		1222		1219	
MEAN	6.87		6.60		6.66		6.95		6.59		6.61	
SD	2.88		2.88		2.63		2.69		2.79		2.62	

Table A16. Distributions of Rater 1 Scores for SAWS Grade 3 by Trait – 12-point Prompt

Form	Score	Idea Development		Organization		Voice		Conventions	
		N	PCT	N	PCT	N	PCT	N	PCT
1	0	50	4.1	91	7.4	82	6.7	110	9.0
	1	303	24.7	401	32.7	375	30.5	407	33.1
	2	636	51.8	542	44.1	599	48.8	548	44.6
	3	239	19.5	194	15.8	172	14	163	13.3
	N		1228		1228		1228		1228
	MEAN		1.87		1.68		1.70		1.62
	SD		0.77		0.83		0.79		0.82
2	0	50	4.1	98	8.1	87	7.2	106	8.8
	1	360	29.8	496	41.1	408	33.8	447	37.0
	2	582	48.2	444	36.8	544	45.1	507	42.0
	3	215	17.8	169	14.0	168	13.9	147	12.2
	N		1207		1207		1207		1207
	MEAN		1.80		1.57		1.66		1.58
	SD		0.78		0.83		0.81		0.81
3	0	33	2.7	70	5.8	59	4.9	75	6.2
	1	397	32.8	457	37.8	411	34.0	456	37.7
	2	620	51.3	550	45.5	602	49.8	537	44.4
	3	159	13.2	132	10.9	137	11.3	141	11.7
	N		1209		1209		1209		1209
	MEAN		1.75		1.62		1.68		1.62
	SD		0.71		0.76		0.74		0.77
4	0	50	4.1	58	4.8	58	4.8	72	6.0
	1	278	23.1	421	34.9	377	31.3	422	35.0
	2	658	54.6	565	46.8	601	49.8	561	46.5
	3	220	18.2	162	13.4	170	14.1	151	12.5
	N		1206		1206		1206		1206
	MEAN		1.87		1.69		1.73		1.66
	SD		0.75		0.76		0.76		0.77
Form	Score	Idea Development		Organization		Voice		Conventions	
		N	PCT	N	PCT	N	PCT	N	PCT
5	0	72	5.9	83	6.8	94	7.7	78	6.4
	1	377	30.9	457	37.4	415	34	455	37.2
	2	594	48.6	530	43.4	565	46.2	532	43.5
	3	179	14.6	152	12.4	148	12.1	157	12.8

	N	1222		1222		1222		1222	
	MEAN	1.72		1.61		1.63		1.63	
	SD	0.78		0.79		0.79		0.79	
6	0	48	3.9	61	5.0	57	4.7	70	5.7
	1	369	30.3	512	42.0	425	34.9	477	39.1
	2	624	51.2	525	43.1	604	49.5	544	44.6
	3	178	14.6	121	9.9	133	10.9	128	10.5
	N	1219		1219		1219		1219	
	MEAN	1.76		1.58		1.67		1.60	
	SD	0.74		0.74		0.73		0.75	

Table A17. Distributions of Scores for SAWS Grade 5 – 12-point Prompt

Score	<u>Form</u>											
	<u>1</u>		<u>2</u>		<u>3</u>		<u>4</u>		<u>5</u>		<u>6</u>	
	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N
0	33	2.8	38	3.3	44	3.8	43	3.7	21	1.8	35	3.0
1	2	0.2	7	0.6	4	0.3	5	0.4	4	0.3	6	0.5
2	4	0.3	9	0.8	2	0.2	4	0.3	4	0.3	17	1.5
3	7	0.6	6	0.5	8	0.7	8	0.7	7	0.6	8	0.7
4	128	11.0	162	14.1	115	10.0	106	9.2	99	8.6	182	15.8
5	89	7.7	127	11.1	73	6.3	66	5.7	74	6.5	150	13.1
6	120	10.3	145	12.7	111	9.6	145	12.5	121	10.6	143	12.4
7	169	14.5	143	12.5	162	14.1	149	12.9	163	14.2	146	12.7
8	271	23.3	244	21.3	267	23.2	307	26.6	291	25.4	251	21.8
9	91	7.8	66	5.8	81	7.0	70	6.1	74	6.5	63	5.5
10	68	5.9	66	5.8	69	6.0	55	4.8	66	5.8	46	4.0
11	78	6.7	56	4.9	73	6.3	68	5.9	85	7.4	48	4.2
12	102	8.8	76	6.6	143	12.4	130	11.2	136	11.9	54	4.7
N	1162		1145		1152		1156		1145		1149	
MEAN	7.44		6.93		7.58		7.48		7.76		6.64	
SD	2.68		2.72		2.87		2.80		2.64		2.60	

Table A18. Distributions of Scores for SAWS Grade 5 by Trait – 12-point Prompt

Form	Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>		
		N	PCT	N	PCT	N	PCT	N	PCT	
1	0	33	2.8	43	3.7	35	3.0	43	3.7	
	1	233	20.1	331	28.5	334	28.7	351	30.2	
	2	622	53.5	560	48.2	595	51.2	567	48.8	
	3	274	23.6	228	19.6	198	17.0	201	17.3	
	N		1162		1162		1162		1162	
	MEAN		1.98		1.84		1.82		1.80	
	SD		0.74		0.78		0.74		0.76	
2	0	41	3.6	59	5.2	51	4.5	50	4.4	
	1	349	30.5	379	33.1	396	34.6	398	34.8	
	2	546	47.7	526	45.9	531	46.4	559	48.8	
	3	209	18.3	181	15.8	167	14.6	138	12.1	
	N		1145		1145		1145		1145	
	MEAN		1.81		1.72		1.71		1.69	
	SD		0.77		0.79		0.77		0.74	
3	0	49	4.3	52	4.5	48	4.2	51	4.4	
	1	233	20.2	317	27.5	261	22.7	302	26.2	
	2	584	50.7	519	45.1	599	52.0	565	49.0	
	3	286	24.8	264	22.9	244	21.2	234	20.3	
	N		1152		1152		1152		1152	
	MEAN		1.96		1.86		1.90		1.85	
	SD		0.79		0.82		0.77		0.79	
4	0	43	3.7	54	4.7	48	4.2	59	5.1	
	1	236	20.4	316	27.3	269	23.3	291	25.2	
	2	618	53.5	550	47.6	620	53.6	600	51.9	
	3	259	22.4	236	20.4	219	18.9	206	17.8	
	N		1156		1156		1156		1156	
	MEAN		1.95		1.84		1.87		1.82	
	SD		0.76		0.80		0.76		0.78	
5	0	25	2.2	32	2.8	26	2.3	35	3.1	
	1	224	19.6	310	27.1	234	20.4	303	26.5	
	2	600	52.4	555	48.5	619	54.1	583	50.9	
	3	296	25.9	248	21.7	266	23.2	224	19.6	
	N		1145		1145		1145		1145	
	MEAN		2.02		1.89		1.98		1.87	
	SD		0.74		0.77		0.73		0.75	
6	0	40	3.5	61	5.3	49	4.3	52	4.5	
	1	413	35.9	398	34.6	457	39.8	423	36.8	
	2	560	48.7	540	47.0	531	46.2	536	46.6	
	3	136	11.8	150	13.1	112	9.7	138	12.0	
	N		1149		1149		1149		1149	
	MEAN		1.69		1.68		1.61		1.66	
	SD		0.72		0.77		0.72		0.75	

Table A19. Distributions of Scores for SAWS Grade 5 - 4-point Prompt¹⁰

Score	<u>1</u>		<u>2</u>		<u>Form</u> <u>3&6</u>		<u>4</u>		<u>5</u>	
	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N
0	39	3.4	43	3.8	79	3.4	48	4.2	29	2.5
1	75	6.5	59	5.2	67	2.9	103	8.9	33	2.9
2	244	21.0	285	24.9	349	15.2	241	20.8	186	16.2
3	304	26.2	313	27.3	659	28.6	286	24.7	289	25.2
4	500	43.0	445	38.9	1147	49.8	478	41.3	608	53.1
N	1162		1145		2301		1156		1145	
MEAN	2.99		2.92		3.19		2.90		3.23	
SD	1.10		1.08		1.02		1.16		0.99	

¹⁰ Distributions of rater 2 scores were not included since only 25% of the papers received second reads.

Table A20. Distributions of Scores for SAWS Grade 5 by Trait – 4-point Prompt

Form	Score	<u>Response to Text</u>		<u>Holistic</u>	
		N	PCT	N	PCT
1	0	142	12.2	39	3.4
	1	438	37.7	373	32.1
	2	582	50.1	750	64.5
	N	1162		1162	
	MEAN	1.38		1.61	
	SD	0.69		0.55	
2	0	116	10.1	43	3.8
	1	434	37.9	480	41.9
	2	595	52.0	622	54.3
	N	1145		1145	
	MEAN	1.42		1.51	
	SD	0.67		0.57	
3&6	0	169	7.3	79	3.4
	1	622	27.0	756	32.9
	2	1510	65.6	1466	63.7
	N	2301		2301	
	MEAN	1.58		1.60	
	SD	0.62		0.56	
4	0	180	15.6	48	4.2
	1	338	29.2	475	41.1
	2	638	55.2	633	54.8
	N	1156		1156	
	MEAN	1.40		1.51	
	SD	0.74		0.58	
5	0	73	6.4	29	2.5
	1	304	26.6	368	32.1
	2	768	67.1	748	65.3
	N	1145		1145	
	MEAN	1.61		1.63	
	SD	0.61		0.53	

Table A21. Distributions of Rater 1 Scores for SAWS Grade 5 - 8-point Prompt¹¹

Score	Form									
	<u>1</u>		<u>2</u>		<u>3&6</u>		<u>4</u>		<u>5</u>	
	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N
0	37	3.2	42	3.7	89	3.9	50	4.3	24	2.1
1	12	1.0	42	3.7	40	1.7	30	2.6	38	3.3
2	39	3.4	97	8.5	100	4.3	87	7.5	93	8.1
3	144	12.4	139	12.1	209	9.1	150	13.0	167	14.6
4	208	17.9	185	16.2	387	16.8	224	19.4	201	17.6
5	227	19.5	228	19.9	550	23.9	245	21.2	223	19.5
6	233	20.1	223	19.5	497	21.6	189	16.3	173	15.1
7	173	14.9	130	11.4	290	12.6	130	11.2	152	13.3
8	89	7.7	59	5.2	139	6.0	51	4.4	74	6.5
N	1162		1145		2301		1156		1145	
MEAN	5.00		4.59		4.91		4.52		4.66	
SD	1.85		1.96		1.85		1.92		1.92	

¹¹ Distributions of rater 2 scores were not included since only 25% of the papers received second reads.

Table B22. Distributions of Scores for SAWS Grade 5 by Trait – 8-point Prompt

Form	Score	<u>Response to Text</u>		<u>Holistic</u>		
		N	PCT	N	PCT	
1	0	130	11.2	37	3.2	
	1	407	35.0	46	4.0	
	2	625	53.8	178	15.3	
	3			286	24.6	
	4			285	24.5	
	5			227	19.5	
	6			103	8.9	
	N		1162		1162	
	MEAN		1.43		3.57	
	SD		0.68		1.45	
2	0	256	22.4	42	3.7	
	1	219	19.1	76	6.6	
	2	670	58.5	219	19.1	
	3			328	28.6	
	4			265	23.1	
	5			154	13.4	
	6			61	5.3	
	N		1145		1145	
	MEAN		1.36		3.23	
	SD		0.82		1.41	
3&6	0	319	13.9	89	3.9	
	1	216	9.4	121	5.3	
	2	1766	76.7	445	19.3	
	3			644	28.0	
	4			556	24.2	
	5			302	13.1	
	6			144	6.3	
	N		2301		2301	
	MEAN		1.63		3.28	
	SD		0.71		1.42	
4	0	360	31.1	50	4.3	
	1	358	31.0	41	3.5	
	2	438	37.9	167	14.4	
	3			301	26.0	
	4			330	28.5	
	5			207	17.9	
	6			60	5.2	
	N		1156		1156	
	MEAN		1.07		3.45	
	SD		0.83		1.39	
5	0	248	21.7	24	2.1	
	1	343	30.0	68	5.9	
	2	554	48.4	223	19.5	
	3			315	27.5	
	4			238	20.8	
	5			186	16.2	

Form	Score	<u>Response to Text</u>		<u>Holistic</u>	
		N	PCT	N	PCT
	6			91	7.9
	N	1145		1145	
	MEAN	1.27		3.39	
	SD	0.79		1.43	

Table A22. Distributions of Rater 1 Scores for SAWS Grade 7 - 12-point Prompt

Score	Form											
	<u>1</u>		<u>2</u>		<u>3</u>		<u>4</u>		<u>5</u>		<u>6</u>	
	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N
0	38	3.3	42	3.6	34	3.0	47	4.0	43	3.7	37	3.2
1	3	0.3	2	0.2	0	0.0	3	0.3	1	0.1	3	0.3
2	1	0.1	1	0.1	4	0.3	2	0.2	4	0.3	1	0.1
3	7	0.6	4	0.3	2	0.2	4	0.3	3	0.3	2	0.2
4	124	10.6	83	7.2	99	8.6	107	9.2	146	12.4	162	14.1
5	85	7.3	66	5.7	101	8.8	75	6.4	68	5.8	79	6.9
6	110	9.4	115	9.9	112	9.7	110	9.4	87	7.4	101	8.8
7	134	11.5	123	10.6	138	12.0	138	11.8	156	13.3	129	11.2
8	267	22.9	279	24.1	275	23.9	291	24.9	304	25.9	308	26.7
9	90	7.7	83	7.2	81	7.0	66	5.7	60	5.1	77	6.7
10	73	6.3	69	6.0	74	6.4	67	5.7	69	5.9	45	3.9
11	82	7.0	109	9.4	89	7.7	86	7.4	89	7.6	73	6.3
12	154	13.2	183	15.8	140	12.2	171	14.7	145	12.3	136	11.8
N	1168		1159		1149		1167		1175		1153	
MEAN	7.68		8.02		7.73		7.74		7.61		7.46	
SD	2.86		2.89		2.77		2.94		2.88		2.83	

Table A23. Distributions of Rater 1 Scores for SAWS Grade 7 by Trait – 12-point Prompt

Form	Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
		N	PCT	N	PCT	N	PCT	N	PCT
1	0	50	4.3	47	4.0	40	3.4	45	3.9
	1	275	23.5	295	25.3	280	24.0	279	23.9
	2	570	48.8	510	43.7	584	50.0	577	49.4
	3	273	23.4	316	27.1	264	22.6	267	22.9
	N		1168		1168		1168		1168
	MEAN		1.91		1.94		1.92		1.91
	SD		0.80		0.83		0.77		0.78
2	0	48	4.1	45	3.9	44	3.8	47	4.1
	1	194	16.7	248	21.4	210	18.1	261	22.5
	2	571	49.3	522	45.0	577	49.8	567	48.9
	3	346	29.9	344	29.7	328	28.3	284	24.5
	N		1159		1159		1159		1159
	MEAN		2.05		2.01		2.03		1.94
	SD		0.79		0.82		0.78		0.79
3	0	37	3.2	38	3.3	34	3.0	37	3.2
	1	277	24.1	294	25.6	254	22.1	268	23.3
	2	553	48.1	530	46.1	592	51.5	608	52.9
	3	282	24.5	287	25.0	269	23.4	236	20.5
	N		1149		1149		1149		1149
	MEAN		1.94		1.93		1.95		1.91
	SD		0.78		0.80		0.76		0.75
4	0	54	4.6	56	4.8	49	4.2	51	4.4
	1	214	18.3	269	23.1	256	21.9	298	25.5
	2	584	50.0	538	46.1	560	48.0	580	49.7
	3	315	27.0	304	26	302	25.9	238	20.4
	N		1167		1167		1167		1167
	MEAN		1.99		1.93		1.96		1.86
	SD		0.80		0.82		0.80		0.78
5	0	47	4.0	50	4.3	45	3.8	45	3.8
	1	232	19.7	310	26.4	264	22.5	340	28.9
	2	595	50.6	561	47.7	595	50.6	559	47.6
	3	301	25.6	254	21.6	271	23.1	231	19.7
	N		1175		1175		1175		1175
	MEAN		1.98		1.87		1.93		1.83
	SD		0.78		0.80		0.78		0.78
6	0	41	3.6	41	3.6	38	3.3	42	3.6
	1	297	25.8	309	26.8	276	23.9	357	31.0
	2	571	49.5	569	49.3	588	51.0	539	46.7
	3	244	21.2	234	20.3	251	21.8	215	18.6
	N		1153		1153		1153		1153
	MEAN		1.88		1.86		1.91		1.80
	SD		0.77		0.77		0.76		0.78

Table A24. Distributions of Scores for SAWS Grade 7 - 4-point Prompt¹²

Score	<u>Form</u>											
	<u>1</u>		<u>2</u>		<u>3</u>		<u>4</u>		<u>5</u>		<u>6</u>	
	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N
0	46	3.9	43	3.7	36	3.1	57	4.9	55	4.7	40	3.5
1	78	6.7	127	11.0	114	9.9	150	12.9	88	7.5	66	5.7
2	223	19.1	248	21.4	241	21.0	295	25.3	273	23.2	258	22.4
3	321	27.5	250	21.6	241	21.0	260	22.3	328	27.9	391	33.9
4	500	42.8	491	42.4	517	45.0	405	34.7	431	36.7	398	34.5
N	1168		1159		1149		1167		1175		1153	
MEAN	2.99		2.88		2.95		2.69		2.84		2.90	
SD	1.11		1.18		1.16		1.21		1.14		1.05	

¹² Distributions of rater 2 scores were not included since only 25% of the papers received second reads.

Table B26. Distributions of Scores for SAWS Grade 7 by Trait – 4-point Prompt

Form	Score	<u>Response to Text</u>		<u>Holistic</u>	
		N	PCT	N	PCT
1	0	164	14.0	46	3.9
	1	416	35.6	349	29.9
	2	588	50.3	773	66.2
	N	1168		1168	
	MEAN	1.36		1.62	
	SD	0.72		0.56	
2	0	187	16.1	43	3.7
	1	372	32.1	467	40.3
	2	600	51.8	649	56.0
	N	1159		1159	
	MEAN	1.36		1.52	
	SD	0.74		0.57	
3	0	180	15.7	36	3.1
	1	346	30.1	431	37.5
	2	623	54.2	682	59.4
	N	1149		1149	
	MEAN	1.39		1.56	
	SD	0.74		0.56	
4	0	230	19.7	57	4.9
	1	444	38.0	510	43.7
	2	493	42.2	600	51.4
	N	1167		1167	
	MEAN	1.23		1.47	
	SD	0.75		0.59	
5	0	183	15.6	55	4.7
	1	493	42.0	389	33.1
	2	499	42.5	731	62.2
	N	1175		1175	
	MEAN	1.27		1.58	
	SD	0.71		0.58	
6	0	121	10.5	40	3.5
	1	553	48.0	390	33.8
	2	479	41.5	723	62.7
	N	1153		1153	
	MEAN	1.31		1.59	
	SD	0.65		0.56	

Table A25. Distributions of Scores for SAWS Grade 7 - 8-point Prompts¹³

Score	Form											
	<u>1</u>		<u>2</u>		<u>3</u>		<u>4</u>		<u>5</u>		<u>6</u>	
	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N	N	Pct N
0	47	4.0	44	3.8	42	3.7	53	4.5	62	5.3	32	2.8
1	16	1.4	4	0.3	13	1.1	16	1.4	13	1.1	38	3.3
2	34	2.9	28	2.4	38	3.3	58	5.0	53	4.5	73	6.3
3	75	6.4	93	8.0	121	10.5	122	10.5	107	9.1	138	12.0
4	171	14.6	149	12.9	200	17.4	179	15.3	189	16.1	189	16.4
5	225	19.3	236	20.4	227	19.8	181	15.5	200	17.0	230	19.9
6	278	23.8	249	21.5	233	20.3	229	19.6	246	20.9	200	17.3
7	215	18.4	205	17.7	186	16.2	200	17.1	190	16.2	160	13.9
8	107	9.2	151	13.0	89	7.7	129	11.1	115	9.8	93	8.1
N	1168		1159		1149		1167		1175		1153	
MEAN	5.26		5.39		5.05		5.08		5.04		4.83	
SD	1.90		1.90		1.88		2.06		2.04		1.97	

¹³ Distributions of rater 2 scores were not included since only 25% of the papers received second reads.

Table B28. Distributions of Scores for SAWS Grade 7 by Trait – 8-point Prompt

Form	Score	Response to Text		Holistic		
		N	PCT	N	PCT	
1	0	110	9.4	47	4.0	
	1	308	26.4	31	2.7	
	2	750	64.2	104	8.9	
	3			302	25.9	
	4			335	28.7	
	5			240	20.5	
	6			109	9.3	
	N		1168		1168	
	MEAN		1.55		3.71	
	SD		0.66		1.41	
2	0	155	13.4	44	3.8	
	1	351	30.3	12	1.0	
	2	653	56.3	93	8.0	
	3			246	21.2	
	4			331	28.6	
	5			263	22.7	
	6			170	14.7	
	N		1159		1159	
	MEAN		1.43		3.96	
	SD		0.72		1.43	
3	0	168	14.6	42	3.7	
	1	418	36.4	19	1.7	
	2	563	49.0	136	11.8	
	3			279	24.3	
	4			340	29.6	
	5			231	20.1	
	6			102	8.9	
	N		1149		1149	
	MEAN		1.34		3.70	
	SD		0.72		1.39	
4	0	272	23.3	53	4.5	
	1	301	25.8	26	2.2	
	2	594	50.9	116	9.9	
	3			258	22.1	
	4			311	26.6	
	5			258	22.1	
	6			145	12.4	
	N		1167		1167	
	MEAN		1.28		3.80	
	SD		0.82		1.49	
5	0	143	12.2	62	5.3	
	1	364	31.0	34	2.9	
	2	668	56.9	148	12.6	
	3			281	23.9	
	4			320	27.2	
	5			211	18.0	

Form	Score	<u>Response to Text</u>		<u>Holistic</u>	
		N	PCT	N	PCT
	6			119	10.1
	N	1175		1175	
	MEAN	1.45		3.59	
	SD	0.70		1.51	
6	0	285	24.7	32	2.8
	1	367	31.8	50	4.3
	2	501	43.5	158	13.7
	3			259	22.5
	4			326	28.3
	5			216	18.7
	6			112	9.7
	N	1153		1153	
	MEAN	1.19		3.64	
	SD	0.80		1.43	

Appendix B: DIF Results for Field Test 2013 Items

Reading

Table B1. Grade 3 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	0	0	0	0	0	0	0	0	0	0
B-	5	4.3	0	0	0	0	5	4.3	0	0
A	104	89.7	0	0	0	0	52	44.8	0	0
B+	7	6.0	0	0	0	0	2	1.7	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	116	100	116	100	57	49.1	116	100
TOTAL	116	100	116	100	116	100	116	100	116	100

Table B2. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	0	0	0	0	0	0	0	0	0	0
B-	3	2.6	0	0	0	0	2	1.7	0	0
A	106	91.4	0	0	0	0	35	30.2	0	0
B+	7	6.0	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	116	100	116	100	79	68.1	116	100
TOTAL	116	100	116	100	116	100	116	100	116	100

Table B3. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	1	0.8	0	0	0	0	0	0	0	0
B-	5	4.2	0	0	0	0	0	0	0	0
A	105	89.0	0	0	0	0	12	10.2	0	0
B+	6	5.1	0	0	0	0	0	0	0	0
C+	1	0.8	0	0	0	0	0	0	0	0
SMALL N	0	0	118	100	118	100	106	89.8	118	100
TOTAL	118	100	118	100	118	100	118	100	118	100

Table B4. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	2	1.7	0	0	0	0	0	0	0	0
B-	8	6.7	0	0	0	0	0	0	0	0
A	109	90.8	0	0	0	0	12	10.0	0	0
B+	1	0.8	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	120	100	120	100	108	90.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table B5. Grade 7 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</u>		<u>White vs.</u> <u>Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	0	0	0	0	0	0	0	0	0	0
B-	8	6.8	0	0	0	0	2	1.7	0	0
A	103	88.0	0	0	0	0	10	8.5	0	0
B+	6	5.1	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	117	100	117	100	105	89.7	117	100
TOTAL	117	100	117	100	117	100	117	100	117	100

Table B6. 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</u>		<u>White vs.</u> <u>Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	1	0.8	0	0	0	0	0	0	0	0
B-	3	2.5	0	0	0	0	0	0	0	0
A	102	85.7	0	0	0	0	0	0	0	0
B+	12	10.1	0	0	0	0	0	0	0	0
C+	1	0.8	0	0	0	0	0	0	0	0
SMALL N	0	0	119	100	119	100	119	100.0	119	100
TOTAL	119	100	119	100	119	100	119	100	119	100

Mathematics

Table B7. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	0	0	0	0	0	0	0	0	0	0
B-	8	6.7	0	0	0	0	0	0	0	0
A	112	93.3	0	0	0	0	60	50.0	0	0
B+	0	0	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	120	100	120	100	60	50.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table B8. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	0	0	0	0	0	0	0	0	0	0
B-	9	7.5	0	0	0	0	0	0	0	0
A	108	90.0	0	0	0	0	24	20.0	0	0
B+	3	2.5	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	120	100	120	100	96	80.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table B9. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	0	0	0	0	0	0	0	0	0	0
B-	6	5.0	0	0	0	0	0	0	0	0
A	107	89.2	0	0	0	0	0	0	0	0
B+	7	5.8	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	120	100	120	100	120	100	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table B10. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	1	0.8	0	0	0	0	0	0	0	0
B-	7	5.8	0	0	0	0	0	0	0	0
A	109	90.8	0	0	0	0	12	10.0	0	0
B+	3	2.5	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	120	100	120	100	108	90.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table B11. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	0	0	0	0	0	0	0	0	0	0
B-	9	7.5	0	0	0	0	0	0	0	0
A	105	87.5	0	0	0	0	12	10.0	0	0
B+	5	4.2	0	0	0	0	0	0	0	0
C+	1	0.8	0	0	0	0	0	0	0	0
SMALL N	0	0	120	100	120	100	108	90.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table B12. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	0	0	0	0	0	0	0	0	0	0
B-	4	3.3	0	0	0	0	1	0.8	0	0
A	109	90.8	0	0	0	0	11	9.2	0	0
B+	7	5.8	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	120	100	120	100	108	90.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Science

Table B13. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	2	1.7	0	0	0	0	0	0	0	0
B-	14	11.7	0	0	0	0	3	2.5	0	0
A	102	85.0	0	0	0	0	22	18.3	0	0
B+	2	1.7	0	0	0	0	1	0.8	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	120	100	120	100	94	78.3	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table B14. 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</u>		<u>White vs. Native American</u>	
	N	%	N	%	N	%	N	%	N	%
C-	1	0.9	0	0	0	0	0	0	0	0
B-	6	5.1	0	0	0	0	1	0.9	0	0
A	107	91.5	0	0	0	0	11	9.4	0	0
B+	3	2.6	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL N	0	0	117	100	117	100	105	89.7	117	100
TOTAL	117	100	117	100	117	100	117	100	117	100

Appendix C: Rasch Difficulty, Standard Error, Fit Statistics, and N-Counts for
2013 Field Test Items

Reading

Table C1. Reading Grade 3 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF389165	792	-1.697	0.083	0.89	0.83
VF389470	792	-1.964	0.086	0.93	0.89
VF389474	792	-1.270	0.080	1.02	1.00
VF389951	792	0.629	0.092	1.06	1.35
VF389473	792	-1.179	0.080	1.09	1.08
VF389457	792	-0.557	0.080	1.14	1.21
VF493383	792	-2.387	0.093	0.89	0.85
VF494661	792	-1.810	0.085	0.91	0.84
VF494909	792	-0.849	0.080	1.06	1.05
VF494764	792	-2.062	0.088	0.91	0.84
VF494944	792	-0.231	0.082	1.14	1.33
VF494745	792	-0.665	0.080	0.99	1.03
Form 2					
VF389467	701	-1.758	0.091	1.06	1.03
VF389446	701	-1.352	0.087	1.10	1.16
VF389477	701	-1.667	0.090	0.96	0.91
VF389949	701	-1.895	0.093	1.00	0.98
VF389620	701	-1.611	0.090	1.03	1.00
VF389482	701	-0.040	0.087	1.27	1.46
VF493480	701	-1.817	0.092	0.95	0.93
VF494098	701	-1.651	0.090	0.94	0.90
VF494956	701	-0.986	0.086	0.99	1.00
VF494732	701	-2.686	0.109	0.81	0.62
VF494759	701	-2.325	0.101	0.82	0.66
VF494915	701	-1.965	0.094	0.95	0.87

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 3					
VF394046	693	-2.840	0.115	0.82	0.64
VF394049	693	-1.926	0.094	0.98	0.98
VF394057	693	-2.242	0.099	1.00	1.02
VF394054	693	-3.322	0.132	0.90	0.62
VF394052	693	-1.763	0.091	1.11	1.19
VF394045	693	-1.192	0.086	1.13	1.13
VF389093	693	-0.362	0.085	1.25	1.44
VF389113	693	-2.051	0.096	0.94	0.82
VF389095	693	-1.583	0.089	0.87	0.82
VF389115	693	-0.089	0.087	1.09	1.22
VF389100	693	-2.282	0.100	0.87	0.75
VF389116	693	-0.858	0.085	1.12	1.15
Form 4					
VF394050	701	-2.383	0.102	0.99	1.07
VF394051	701	-2.127	0.097	1.12	1.26
VF394055	701	0.405	0.091	1.28	1.69
VF394056	701	-1.035	0.084	1.13	1.13
VF394053	701	-2.242	0.099	0.95	0.87
VF394041	701	-1.521	0.088	1.13	1.16
VF389082	701	-1.662	0.089	0.87	0.76
VF389120	701	-0.391	0.084	1.08	1.16
VF389105	701	0.066	0.087	1.31	1.56
VF389074	701	-0.362	0.084	1.08	1.07
VF389123	701	-1.529	0.088	0.98	0.93
VF389119	701	-0.475	0.084	1.06	1.13
Form 5					
VF497812	726	-1.699	0.088	1.16	1.28
VF497773	726	-2.535	0.104	0.90	0.86
VF497775	726	-0.458	0.082	1.20	1.30
VF497822	726	-0.301	0.083	1.07	1.15
VF497827	726	-1.200	0.084	1.02	0.99
VF497820	726	-2.010	0.093	0.94	0.93
VF387075	726	-0.932	0.082	1.14	1.15
VF387072	726	-2.472	0.102	0.87	0.84
VF387055	726	-2.311	0.099	0.88	0.77
VF387084	726	-2.330	0.099	0.91	0.76
VF387096	726	-0.816	0.082	1.11	1.17
VF387089	726	-1.925	0.091	0.89	0.81

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 6					
VF497783	704	-2.787	0.112	0.82	0.67
VF497793	704	-1.831	0.092	0.93	0.85
VF497818	704	-1.814	0.091	0.92	0.84
VF497781	704	-1.306	0.086	0.92	0.89
VF497807	704	-1.659	0.089	0.97	0.99
VF497815	704	-1.748	0.091	1.03	1.05
VF387092	704	-0.953	0.084	1.06	1.04
VF387071	704	-1.314	0.086	1.18	1.24
VF387050	704	-0.736	0.084	1.13	1.16
VF387069	704	-2.494	0.104	0.95	0.98
VF387079	704	-1.942	0.093	0.87	0.74
VF387067	704	-0.792	0.084	1.33	1.43
Form 7					
VF497676	696	-2.105	0.097	0.97	0.93
VF497660	696	-2.579	0.107	0.92	0.82
VF497690	696	-1.268	0.087	1.09	1.14
VF497705	696	-0.209	0.087	1.15	1.23
VF497668	696	-2.031	0.096	0.97	0.95
VF497671	696	-1.215	0.087	0.98	0.94
VF497573	696	-3.568	0.143	0.80	0.38
VF497419	696	-1.185	0.087	1.09	1.08
VF497594	696	-1.437	0.088	0.93	0.87
VF497592	696	-1.571	0.090	0.85	0.78
VF497611	696	-0.527	0.085	1.01	1.05
VF497614	696	-0.519	0.085	1.03	1.08
Form 8					
VF497684	701	-0.731	0.084	1.04	1.05
VF497700	701	-1.218	0.086	1.01	1.02
VF497665	701	0.064	0.087	1.31	1.53
VF497695	701	-2.073	0.096	1.05	1.00
VF497696	701	-1.875	0.093	0.99	0.95
VF497683	701	-2.358	0.102	0.90	0.76
VF497424	701	-1.514	0.088	1.08	1.09
VF497577	701	-2.812	0.115	0.80	0.55
VF497599	701	-1.058	0.085	1.15	1.26
VF497585	701	-3.293	0.132	0.78	0.41
VF497606	701	-1.475	0.088	1.00	0.98
VF497566	701	-2.963	0.119	0.89	0.75

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 9					
VF394173	708	-1.390	0.087	1.14	1.28
VF394165	708	-1.776	0.092	0.87	0.77
VF394143	708	-0.857	0.084	1.07	1.11
VF394141	708	-2.512	0.107	0.86	0.68
VF394157	708	-1.181	0.086	1.08	1.11
VF394137	708	-1.352	0.087	0.99	0.96
VF497718	708	-2.088	0.097	1.15	1.26
VF497716	708	-1.151	0.086	1.09	1.11
VF497759	708	-0.596	0.084	0.98	0.97
VF497758	708	-1.093	0.085	0.99	1.00
VF497761	708	-1.810	0.093	0.88	0.79
VF497766	708	-0.265	0.084	1.18	1.34
Form 10					
VF524523	718	-2.523	0.105	0.92	0.82
VF394145	718	-1.634	0.089	0.96	0.91
VF394134	718	-0.698	0.083	1.10	1.14
VF394154	718	-2.624	0.107	0.92	0.80
VF394153	718	-2.105	0.096	0.98	0.91
VF394151	718	-0.856	0.083	0.94	0.93
VF497731	718	-2.895	0.115	0.82	0.67
VF497728	718	-0.318	0.083	1.16	1.23
VF522349	718	-0.912	0.083	0.99	0.98
VF497751	718	-1.200	0.085	1.06	1.06
VF497767	718	-0.842	0.083	1.02	1.06
VF497725	718	-0.995	0.084	0.94	0.93

Table C2. Reading Grade 4 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF495010	791	-2.839	0.176	0.84	0.40
VF495021	791	-0.276	0.085	1.13	1.27
VF495028	791	0.870	0.079	1.18	1.28
VF495001	791	-1.216	0.103	0.89	0.79
VF495647	791	0.788	0.079	1.33	1.47
VF494988	791	0.217	0.081	0.98	0.96
VF497327	791	-0.494	0.088	0.94	0.85
VF497297	791	-1.708	0.118	0.83	0.62
VF497314	791	-0.494	0.088	0.86	0.73
VF497322	791	0.171	0.081	0.95	0.92
VF497334	791	-0.298	0.085	1.04	1.02
VF497311	791	1.167	0.080	1.20	1.27
Form 2					
VF495015	705	-0.721	0.099	0.93	0.88
VF494997	705	-1.910	0.135	0.87	0.70
VF495644	705	-0.323	0.092	1.14	1.26
VF495003	705	0.131	0.087	0.94	0.94
VF495030	705	-0.427	0.093	1.27	1.73
VF494993	705	0.527	0.085	0.99	0.99
VF497330	705	-0.409	0.093	0.95	0.83
VF497326	705	0.213	0.086	0.96	0.92
VF497338	705	0.085	0.087	1.18	1.33
VF497307	705	0.146	0.087	1.02	0.99
VF497318	705	-0.861	0.102	1.00	1.05
VF497303	705	-0.861	0.102	0.91	0.76
Form 3					
VF497159	697	-1.686	0.127	0.94	0.82
VF497212	697	-0.419	0.094	1.03	0.98
VF497147	697	-1.563	0.122	0.90	0.73
VF497155	697	-0.274	0.092	1.00	0.92
VF497167	697	0.245	0.086	0.91	0.87
VF497215	697	1.442	0.087	1.05	1.21
VF497279	697	-0.062	0.089	1.15	1.17
VF497265	697	-1.891	0.135	0.93	0.73
VF497253	697	-0.118	0.090	1.23	1.33
VF497233	697	-0.183	0.090	1.09	1.10
VF497250	697	0.776	0.085	1.12	1.15
VF497256	697	-1.327	0.114	0.88	0.77

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF497188	707	-2.580	0.173	0.83	0.40
VF497206	707	-2.194	0.150	0.93	0.69
VF497195	707	0.606	0.084	1.12	1.13
VF497162	707	1.104	0.084	1.03	1.12
VF497151	707	0.970	0.084	1.15	1.24
VF497220	707	-0.378	0.092	1.01	0.97
VF497243	707	-0.327	0.092	1.12	1.24
VF497227	707	-1.821	0.132	0.93	0.74
VF497270	707	1.139	0.084	1.14	1.25
VF497261	707	-0.030	0.088	1.21	1.28
VF497286	707	-0.278	0.091	1.10	1.15
VF497247	707	-0.030	0.088	1.03	1.07
Form 5					
VF496940	713	-0.931	0.101	1.01	1.04
VF496899	713	-0.246	0.088	1.01	1.02
VF496934	713	0.938	0.082	1.29	1.41
VF496894	713	0.352	0.083	0.96	0.94
VF496883	713	0.551	0.083	1.07	1.10
VF496921	713	0.741	0.082	1.21	1.32
VF495798	713	-0.215	0.088	0.92	0.88
VF495794	713	0.170	0.084	1.06	1.05
VF495832	713	0.639	0.082	1.12	1.19
VF495856	713	0.978	0.083	1.00	1.04
VF495873	713	0.633	0.082	1.10	1.13
VF495789	713	1.751	0.089	1.24	1.55
Form 6					
VF496202	718	0.406	0.084	1.16	1.20
VF496219	718	-0.741	0.098	0.86	0.75
VF496270	718	-0.466	0.093	1.06	1.06
VF496902	718	1.026	0.083	1.35	1.54
VF496864	718	-0.573	0.095	0.95	0.88
VF496917	718	1.194	0.084	1.06	1.11
VF495786	718	-1.819	0.131	0.86	0.67
VF495812	718	0.832	0.083	1.01	1.06
VF495868	718	0.749	0.083	1.04	1.07
VF495823	718	1.237	0.084	1.16	1.26
VF495843	718	0.277	0.085	0.94	0.89
VF495775	718	0.617	0.084	1.10	1.14

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF495975	714	1.640	0.087	1.10	1.26
VF495970	714	-0.599	0.096	0.93	0.90
VF495987	714	0.695	0.083	1.15	1.16
VF496010	714	1.288	0.084	1.22	1.46
VF495955	714	-1.035	0.106	0.92	0.74
VF496016	714	1.224	0.084	1.19	1.41
VF407282	714	-1.608	0.124	0.89	0.61
VF407237	714	-0.842	0.101	1.05	1.14
VF407298	714	0.093	0.087	1.12	1.11
VF407232	714	-1.419	0.117	0.84	0.61
VF407247	714	-1.753	0.130	0.85	0.59
VF407239	714	0.449	0.084	1.09	1.18
Form 8					
VF495963	697	-0.688	0.097	1.09	1.26
VF495947	697	1.350	0.087	1.25	1.47
VF496019	697	-1.266	0.111	0.89	0.74
VF495994	697	-0.374	0.092	1.14	1.09
VF495982	697	-0.119	0.089	1.09	1.19
VF495944	697	-0.937	0.103	0.92	0.80
VF407250	697	-0.865	0.101	0.97	0.95
VF407243	697	-0.417	0.093	1.02	1.05
VF407287	697	-1.446	0.116	0.92	0.78
VF407295	697	-1.948	0.136	0.95	0.90
VF407297	697	0.068	0.087	1.09	1.10
VF407235	697	0.257	0.086	1.18	1.31
Form 9					
VF497387	696	-1.406	0.117	0.93	0.78
VF497361	696	-1.365	0.116	0.92	0.71
VF497390	696	-0.674	0.099	0.87	0.75
VF497365	696	0.924	0.085	1.18	1.40
VF497378	696	1.322	0.086	1.09	1.15
VF497356	696	1.664	0.089	1.07	1.18
VF494842	1412	-0.782	0.070	1.10	1.37
VF494863	696	-1.549	0.122	0.89	0.76
VF494852	696	-0.888	0.103	0.90	0.84
VF494927	696	0.724	0.085	1.14	1.23
VF494954	696	-1.260	0.113	0.85	0.73
VF494964	696	-0.986	0.106	0.98	0.96

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF497381	716	-1.251	0.110	0.86	0.66
VF497384	716	0.584	0.084	1.05	1.14
VF497359	716	0.208	0.085	0.99	0.99
VF497362	716	0.098	0.086	1.01	0.99
VF497396	716	1.382	0.086	1.25	1.44
VF497354	716	-1.469	0.117	0.92	1.10
VF494937	716	-1.717	0.125	0.87	0.71
VF494856	716	-0.939	0.102	1.03	1.13
VF494914	716	-0.255	0.090	0.96	1.02
VF494907	716	0.113	0.086	0.98	0.95
VF494969	716	0.280	0.085	1.03	1.06

Table C3. Reading Grade 5 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF495685	751	2.186	0.097	1.12	1.57
VF495826	751	1.295	0.083	1.21	1.34
VF495879	751	1.301	0.083	1.21	1.36
VF495811	751	-1.054	0.097	0.86	0.69
VF495906	751	-0.552	0.088	0.91	0.85
VF495818	751	-0.305	0.085	0.89	0.85
VF497016	751	-0.671	0.090	0.92	0.84
VF497025	751	-0.169	0.084	0.88	0.83
VF497001	751	0.853	0.081	1.19	1.25
VF497037	751	-1.231	0.101	0.84	0.64
VF497032	751	1.288	0.083	1.15	1.25
VF497039	751	0.269	0.081	0.96	0.94
Form 2					
VF495793	659	-0.573	0.097	0.93	0.84
VF495802	659	-1.535	0.122	0.83	0.71
VF495840	659	-0.746	0.100	0.85	0.69
VF495886	659	1.764	0.093	1.17	1.41
VF495831	659	-1.210	0.111	0.84	0.65
VF495909	659	-1.010	0.106	0.91	0.83
VF497012	659	-0.220	0.091	1.05	1.03
VF497020	659	-0.229	0.092	1.04	1.00
VF497027	659	1.738	0.093	1.15	1.54
VF497009	659	0.399	0.087	0.98	0.95
VF497028	659	0.711	0.086	0.90	0.88
VF497030	659	0.063	0.089	0.97	0.93
Form 3					
VF497274	668	0.788	0.085	0.99	1.00
VF497277	668	0.891	0.086	1.01	1.02
VF497287	668	-0.559	0.094	0.92	0.88
VF497280	668	-1.853	0.131	0.97	1.15
VF497282	668	0.832	0.086	1.23	1.29
VF497283	668	-0.328	0.091	0.93	0.88
VF496869	668	0.883	0.086	0.96	0.95
VF496889	668	-0.819	0.099	0.89	0.84
VF496876	668	0.262	0.086	1.23	1.32
VF496884	668	-0.622	0.095	0.88	0.78
VF496878	668	0.810	0.086	1.14	1.18
VF496875	668	-0.899	0.100	0.85	0.66

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF497272	670	1.056	0.085	1.22	1.31
VF497288	670	-0.540	0.096	0.99	0.99
VF497284	670	-0.614	0.097	0.92	0.88
VF497278	670	-0.938	0.104	0.83	0.69
VF497273	670	-1.460	0.120	0.94	0.97
VF497285	670	1.338	0.087	0.97	1.07
VF496871	670	-1.335	0.116	1.01	1.04
VF496886	670	-0.822	0.102	1.00	0.92
VF496887	670	-1.206	0.112	0.85	0.69
VF496882	670	-0.379	0.093	1.00	1.04
VF496872	670	-0.169	0.090	1.02	1.01
VF496868	670	-0.080	0.089	1.07	1.10
Form 5					
VF407329	677	-0.883	0.102	0.91	0.78
VF407357	677	-2.395	0.165	0.99	0.97
VF407332	677	-0.801	0.101	1.05	1.08
VF407338	677	0.243	0.086	1.01	0.97
VF407386	677	-0.625	0.097	0.94	0.91
VF407336	677	-0.409	0.093	1.06	1.11
VF497226	677	-1.692	0.128	0.90	0.83
VF497237	677	-0.070	0.089	0.97	0.95
VF497231	677	-0.190	0.090	0.99	1.03
VF497241	677	-0.588	0.096	0.92	0.82
VF497251	677	1.878	0.093	1.15	1.34
VF497240	677	0.448	0.085	1.15	1.18
Form 6					
VF407322	670	-0.105	0.090	1.06	1.12
VF407355	670	-1.153	0.110	0.96	0.93
VF407360	670	-1.848	0.136	0.98	0.87
VF407321	670	-1.531	0.123	0.96	0.87
VF407319	670	0.108	0.088	1.05	1.06
VF407388	670	-0.382	0.094	0.99	0.93
VF497228	670	0.824	0.085	1.04	1.06
VF497230	670	-0.818	0.102	1.07	1.22
VF497232	670	0.592	0.085	0.93	0.89
VF497248	670	0.253	0.087	1.09	1.12
VF497234	670	0.781	0.085	0.94	0.94
VF497254	670	0.534	0.085	1.05	1.06

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF497055	671	-0.080	0.088	0.99	0.98
VF497060	671	-1.690	0.126	0.87	0.62
VF497181	671	-0.554	0.095	0.95	0.96
VF497056	671	-0.347	0.092	0.99	0.97
VF497177	671	-0.264	0.090	0.89	0.80
VF497184	671	0.141	0.087	1.02	1.03
VF496048	671	-0.112	0.089	1.03	1.08
VF496115	671	-2.132	0.147	0.84	0.52
VF496193	671	1.603	0.091	1.19	1.40
VF496185	671	-0.167	0.089	1.06	1.11
VF496032	671	0.141	0.087	1.01	0.97
VF496195	671	0.702	0.085	1.06	1.07
Form 8					
VF497052	661	0.493	0.086	1.17	1.21
VF497170	661	-1.374	0.116	0.94	0.90
VF497174	661	-2.009	0.142	0.87	0.68
VF497182	661	-1.636	0.126	0.90	0.82
VF497172	661	-0.805	0.101	0.90	0.85
VF497176	661	0.817	0.086	1.09	1.15
VF496024	661	-0.553	0.096	0.86	0.75
VF496188	661	-0.271	0.092	1.03	0.99
VF496101	661	-0.427	0.094	0.95	0.91
VF496068	661	-0.628	0.097	0.96	0.92
VF496085	661	0.215	0.087	1.04	1.09
VF496194	661	0.758	0.086	1.08	1.14
Form 9					
VF495943	667	-1.766	0.128	0.98	0.79
VF495800	667	-1.850	0.131	0.95	0.78
VF495911	667	-0.912	0.101	0.94	0.87
VF495837	667	-1.782	0.129	0.97	0.93
VF495850	667	-0.254	0.089	1.11	1.22
VF495905	667	-0.327	0.090	0.97	0.97
VF496201	667	-2.746	0.185	0.87	0.43
VF496218	667	-0.468	0.092	1.06	1.02
VF496879	667	0.225	0.085	1.01	0.99
VF496209	667	-0.736	0.097	1.02	0.96
VF496865	667	-0.051	0.087	0.94	0.91
VF496220	667	-0.409	0.091	1.05	1.12

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF495780	676	0.429	0.085	0.98	0.96
VF495940	676	-1.756	0.132	0.91	0.90
VF495921	676	0.246	0.086	0.90	0.88
VF495924	676	-0.653	0.098	0.93	0.85
VF495859	676	-0.253	0.091	1.08	1.11
VF495935	676	1.680	0.090	1.36	1.82
VF496200	676	-1.306	0.115	0.98	1.12
VF496206	676	0.035	0.088	1.04	1.04
VF496213	676	-0.406	0.093	1.09	1.31
VF496212	676	-1.037	0.107	0.94	0.80
VF496211	676	-2.248	0.157	0.86	0.54
VF496221	676	-2.043	0.145	0.90	0.66

Table C4. Reading Grade 6 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF496164	769	-0.035	0.084	1.00	1.04
VF496187	769	1.579	0.081	1.05	1.17
VF496191	769	1.342	0.079	0.97	1.00
VF496863	769	0.521	0.079	0.94	0.91
VF496208	769	0.553	0.079	0.95	0.92
VF496873	769	0.271	0.081	0.95	0.94
VF497040	769	1.540	0.081	1.14	1.20
VF497034	769	-0.539	0.093	0.90	0.80
VF497042	769	-0.547	0.093	0.82	0.70
VF497046	769	0.671	0.079	1.16	1.21
VF497050	769	0.609	0.079	1.30	1.40
VF497031	769	1.076	0.078	1.12	1.14
Form 2					
VF496172	656	0.278	0.089	0.93	0.88
VF496180	656	1.930	0.089	0.97	1.07
VF496880	656	1.804	0.088	1.16	1.30
VF496204	656	0.801	0.085	0.98	0.97
VF496415	656	0.549	0.087	0.99	1.00
VF496867	656	0.270	0.089	0.90	0.86
VF497035	656	-0.128	0.095	0.92	0.90
VF497041	656	-0.688	0.108	0.83	0.65
VF497038	656	-1.719	0.150	0.86	0.55
VF497049	656	-0.336	0.099	1.03	1.00
VF497033	656	0.845	0.085	1.01	1.00
VF497047	656	-0.416	0.101	0.93	0.93
Form 3					
VF496916	666	1.331	0.084	1.02	1.04
VF496925	666	1.600	0.085	1.09	1.14
VF496938	666	-0.489	0.103	1.02	1.05
VF496935	666	0.988	0.084	0.96	0.95
VF496918	666	0.632	0.085	0.93	0.88
VF496926	666	0.632	0.085	0.97	0.96
VF523786	666	0.308	0.089	0.97	0.95
VF523801	666	0.566	0.086	0.96	0.95
VF523870	666	-0.326	0.099	0.99	0.98
VF523830	666	1.828	0.087	1.37	1.57
VF523863	666	-0.656	0.108	0.83	0.67
VF523813	666	1.184	0.084	1.06	1.10

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF496943	663	-0.687	0.109	0.82	0.69
VF496922	663	0.160	0.091	1.08	1.11
VF496920	663	-0.185	0.097	0.90	0.81
VF496933	663	0.534	0.087	0.98	0.94
VF496942	663	-0.584	0.106	0.94	0.83
VF496927	663	0.109	0.092	1.10	1.25
VF523852	663	0.075	0.093	0.99	1.07
VF523825	663	0.210	0.091	0.91	0.85
VF523818	663	0.347	0.089	0.93	0.90
VF523846	663	0.193	0.091	0.85	0.78
VF523861	663	0.135	0.092	0.92	0.88
VF523804	663	0.410	0.089	0.95	0.94
Form 5					
VF496043	679	-0.575	0.102	0.97	0.89
VF496083	679	-1.499	0.135	0.90	0.81
VF496029	679	-0.760	0.107	0.93	0.87
VF496097	679	-0.453	0.099	0.89	0.79
VF496047	679	1.171	0.083	1.14	1.19
VF496071	679	-0.627	0.103	0.85	0.73
VF497073	679	0.106	0.089	0.96	0.90
VF497053	679	1.612	0.085	1.03	1.12
VF497065	679	1.206	0.083	1.13	1.18
VF497064	679	1.826	0.087	1.07	1.18
VF497067	679	2.691	0.103	1.14	1.63
VF497062	679	0.403	0.086	1.12	1.18
Form 6					
VF496055	678	1.035	0.083	0.90	0.89
VF496087	678	-0.587	0.103	0.91	0.82
VF496036	678	0.708	0.084	1.00	0.99
VF496065	678	0.210	0.088	0.93	0.89
VF496100	678	-0.415	0.099	0.92	0.81
VF496051	678	0.680	0.084	0.98	0.98
VF497058	678	1.138	0.083	0.99	0.98
VF497059	678	0.187	0.088	1.00	1.02
VF497070	678	1.494	0.084	1.22	1.33
VF497069	678	0.581	0.085	1.04	1.04
VF497071	678	1.409	0.084	1.11	1.13
VF497068	678	0.722	0.084	1.03	1.04

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF497074	672	1.094	0.083	0.97	0.96
VF497078	672	0.493	0.086	1.04	1.02
VF497087	672	0.530	0.085	0.97	0.93
VF497077	672	1.450	0.084	1.02	1.05
VF497083	672	1.386	0.084	0.95	0.94
VF497076	672	1.506	0.084	0.96	0.99
VF496977	672	0.652	0.085	1.02	1.02
VF496958	672	-0.671	0.106	1.00	1.10
VF496968	672	0.149	0.089	1.02	1.04
VF496963	672	0.766	0.084	1.01	1.02
VF496975	672	0.275	0.088	0.99	1.01
VF496974	672	0.228	0.088	0.97	0.96
Form 8					
VF497085	664	2.470	0.097	0.99	1.06
VF497079	664	0.334	0.088	0.97	0.94
VF497082	664	0.983	0.084	0.99	1.02
VF497075	664	1.994	0.089	0.97	0.97
VF497084	664	1.068	0.084	1.01	1.01
VF497086	664	1.876	0.088	1.32	1.54
VF496954	664	-0.494	0.103	0.91	0.82
VF496964	664	1.216	0.084	1.00	1.01
VF496979	664	-0.473	0.102	0.93	0.90
VF496961	664	-0.865	0.113	0.90	0.74
VF496970	664	-0.188	0.096	0.96	0.96
VF496952	664	-0.216	0.097	0.98	1.07
Form 9					
VF388824	676	0.776	0.084	1.18	1.20
VF388868	676	-1.538	0.140	0.86	0.55
VF388909	676	1.543	0.084	1.10	1.22
VF388911	676	0.268	0.088	0.88	0.87
VF388907	676	1.805	0.086	1.04	1.14
VF388914	676	0.958	0.083	1.07	1.09
VF495938	676	1.131	0.083	1.06	1.07
VF495961	676	0.488	0.086	0.96	0.98
VF495954	676	-0.219	0.096	0.97	1.00
VF495945	676	0.804	0.084	1.02	1.03
VF495990	676	-0.247	0.096	0.84	0.75
VF495925	676	0.001	0.092	0.97	0.94

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF388848	676	-0.336	0.099	1.04	1.06
VF388851	676	0.920	0.083	0.96	0.94
VF388881	676	-1.115	0.123	0.95	0.89
VF388853	676	0.160	0.090	0.93	0.86
VF388905	676	2.450	0.095	1.07	1.29
VF388912	676	-0.184	0.096	0.95	0.90
VF495908	676	-0.655	0.107	0.94	0.88
VF495916	676	-1.345	0.133	0.90	0.65
VF495918	676	-0.147	0.095	0.99	0.98
VF495968	676	1.038	0.083	0.96	0.95
VF495980	676	1.648	0.084	1.27	1.37
VF495978	676	1.889	0.086	1.18	1.31

Table C5. Reading Grade 7 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF498064	809	1.361	0.077	1.01	0.99
VF498057	809	0.427	0.080	0.94	0.91
VF498052	809	1.349	0.077	1.03	1.07
VF498047	809	-0.682	0.097	1.05	1.27
VF498036	809	2.145	0.084	1.20	1.42
VF498034	809	-0.020	0.084	0.95	0.89
VF496895	809	1.409	0.078	0.95	0.96
VF496901	809	0.754	0.078	0.98	1.02
VF496908	809	-0.361	0.090	0.87	0.77
VF496913	809	-0.418	0.091	0.82	0.67
VF496932	809	-0.250	0.088	0.89	0.83
VF496903	809	-0.085	0.085	0.88	0.77
Form 2					
VF498063	676	0.869	0.085	1.09	1.11
VF498059	676	0.447	0.088	0.93	0.90
VF498054	676	0.621	0.086	0.95	0.91
VF498051	676	0.984	0.084	1.02	1.02
VF498060	676	2.212	0.091	1.13	1.34
VF498032	676	-0.558	0.106	0.96	0.90
VF496892	676	1.928	0.087	1.01	1.02
VF496900	676	-0.514	0.105	0.87	0.70
VF496949	676	-0.503	0.105	0.97	1.09
VF496944	676	0.598	0.086	1.00	1.01
VF496937	676	0.065	0.093	0.99	1.03
VF496906	676	-0.558	0.106	0.84	0.67
Form 3					
VF497744	667	-0.939	0.122	1.00	0.99
VF497727	667	-0.161	0.099	0.98	0.98
VF497750	667	-0.756	0.115	0.89	0.67
VF497745	667	1.873	0.086	1.17	1.24
VF497754	667	2.046	0.088	1.12	1.22
VF497721	667	-0.230	0.100	0.93	0.91
VF497961	667	0.673	0.086	0.89	0.84
VF497974	667	1.105	0.084	1.01	1.02
VF497975	667	1.182	0.084	1.02	1.03
VF497958	667	0.413	0.089	0.94	0.86
VF497955	667	-0.084	0.097	0.86	0.73
VF497959	667	1.336	0.084	1.18	1.23

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF497726	675	1.528	0.083	1.11	1.17
VF497748	675	1.334	0.083	1.27	1.33
VF497724	675	-0.540	0.107	0.92	0.87
VF497757	675	1.855	0.085	1.14	1.24
VF497715	675	0.920	0.084	0.98	0.98
VF497730	675	2.095	0.088	1.14	1.23
VF497978	675	1.238	0.083	0.97	0.96
VF497972	675	2.079	0.088	1.05	1.13
VF497951	675	0.440	0.088	0.97	0.97
VF497969	675	1.142	0.083	0.91	0.89
VF497956	675	0.169	0.091	1.02	1.09
VF497971	675	2.157	0.089	1.10	1.23
Form 5					
VF497930	670	0.194	0.093	0.92	0.91
VF497938	670	0.245	0.092	0.95	0.94
VF497950	670	-0.010	0.096	0.89	0.78
VF497946	670	-0.423	0.105	0.86	0.68
VF497935	670	-0.028	0.096	1.03	1.12
VF497942	670	2.011	0.087	1.10	1.22
VF497711	670	0.497	0.089	0.92	0.85
VF497686	670	1.650	0.085	1.03	1.08
VF497819	670	0.426	0.089	0.91	0.86
VF497803	670	0.613	0.087	0.93	0.88
VF497756	670	1.464	0.084	1.13	1.18
VF497735	670	2.810	0.100	1.07	1.59
Form 6					
VF497931	668	-0.121	0.099	0.94	0.91
VF497933	668	0.182	0.093	1.04	1.18
VF497936	668	0.702	0.087	1.22	1.34
VF497941	668	0.627	0.087	0.96	0.92
VF497943	668	1.124	0.084	0.96	0.95
VF497934	668	0.325	0.091	0.96	0.89
VF497701	668	0.207	0.093	0.88	0.78
VF497814	668	-0.161	0.100	0.99	1.07
VF497794	668	-0.283	0.103	0.85	0.72
VF497772	668	1.697	0.085	1.10	1.17
VF497732	668	0.858	0.085	0.90	0.85
VF497785	668	1.053	0.084	1.10	1.12

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF497873	657	0.176	0.093	0.94	0.94
VF497894	657	-0.983	0.123	0.89	0.77
VF497893	657	-0.077	0.097	1.01	1.02
VF497890	657	-0.233	0.100	0.92	0.83
VF497882	657	0.368	0.090	1.01	1.04
VF497885	657	0.559	0.088	0.94	0.90
VF497179	657	-1.108	0.127	0.87	0.66
VF497219	657	2.299	0.092	1.34	1.69
VF497198	657	1.552	0.086	1.02	1.05
VF497186	657	-0.644	0.111	0.85	0.66
VF497224	657	0.758	0.087	0.97	0.95
VF497245	657	-0.694	0.113	0.83	0.64
Form 8					
VF497868	664	1.488	0.085	1.01	1.05
VF497862	664	1.011	0.085	0.99	0.98
VF497879	664	0.543	0.088	1.05	1.03
VF497889	664	-0.263	0.101	0.85	0.79
VF497881	664	0.558	0.088	0.99	0.93
VF497876	664	-0.283	0.102	0.98	1.00
VF497205	664	0.262	0.092	1.02	1.04
VF497190	664	-0.038	0.097	0.94	0.89
VF497211	664	0.063	0.095	0.97	0.97
VF497238	664	0.895	0.085	1.15	1.20
VF497214	664	0.072	0.095	0.94	0.89
VF497175	664	-0.357	0.104	0.90	0.79
Form 9					
VF497980	672	1.540	0.084	0.90	0.91
VF498011	672	0.979	0.085	1.05	1.10
VF497883	672	1.355	0.084	1.12	1.17
VF498018	672	-0.534	0.108	0.87	0.71
VF497995	672	0.168	0.093	0.81	0.71
VF498030	672	-0.281	0.102	0.94	1.03
VF497260	672	-0.476	0.107	0.84	0.69
VF497294	672	1.327	0.084	1.16	1.20
VF497299	672	0.319	0.091	0.95	0.92
VF497295	672	0.892	0.085	1.04	1.03
VF497291	672	1.690	0.085	1.02	1.04
VF497268	672	1.143	0.084	1.07	1.09

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF497985	669	2.463	0.094	1.18	1.60
VF498023	669	-0.856	0.118	0.88	0.78
VF498062	669	0.684	0.086	1.01	1.00
VF498058	669	0.579	0.087	1.03	1.03
VF497877	669	1.248	0.084	1.10	1.12
VF497870	669	1.080	0.084	1.09	1.09
VF497263	669	-0.002	0.095	0.92	0.88
VF497292	669	-0.142	0.098	0.94	0.89
VF497293	669	0.181	0.092	0.96	0.92
VF497275	669	0.632	0.086	1.07	1.12
VF497301	669	0.880	0.085	1.00	0.99
VF497281	669	-0.623	0.110	0.94	0.87

Table C6. Reading Grade 8 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF497094	760	-0.141	0.093	0.93	0.87
VF497096	760	0.829	0.081	0.93	0.88
VF497101	760	0.757	0.081	0.94	0.97
VF497099	760	1.809	0.081	0.99	1.00
VF497103	760	0.010	0.090	0.89	0.91
VF497114	760	0.228	0.087	1.00	1.02
VF497121	760	-0.107	0.092	1.11	1.25
VF497118	760	1.014	0.080	1.09	1.09
VF497122	760	2.636	0.090	1.26	1.71
VF497130	760	0.026	0.090	0.93	0.83
VF497132	760	1.071	0.079	1.11	1.13
VF497117	760	0.790	0.081	0.93	0.89
Form 2					
VF497115	664	-0.245	0.104	0.94	0.89
VF497093	664	2.179	0.088	1.21	1.34
VF497113	664	2.171	0.088	1.03	1.13
VF497102	664	2.404	0.091	1.11	1.31
VF497098	664	1.225	0.084	1.08	1.11
VF497095	664	-0.778	0.120	0.84	0.77
VF497123	664	0.333	0.092	1.02	0.98
VF497120	664	1.488	0.084	0.91	0.92
VF497128	664	0.884	0.086	1.06	1.07
VF497127	664	-0.536	0.112	0.87	0.68
VF497119	664	2.488	0.092	1.19	1.41
VF497116	664	0.780	0.087	0.98	0.98
Form 3					
VF497139	672	-0.533	0.112	0.92	0.88
VF497143	672	-0.388	0.108	0.93	0.88
VF497137	672	0.670	0.087	0.94	0.90
VF497156	672	1.031	0.085	1.02	1.00
VF497164	672	1.439	0.084	1.03	1.05
VF497153	672	0.647	0.088	1.22	1.31
VF497363	672	1.271	0.084	1.05	1.07
VF497325	672	0.553	0.089	0.99	0.95
VF497367	672	1.334	0.084	1.14	1.17
VF497319	672	0.367	0.091	0.87	0.81
VF497336	672	0.835	0.086	0.96	0.98
VF497331	672	2.967	0.101	1.07	1.38

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF497146	669	-0.415	0.109	0.96	0.86
VF497148	669	0.970	0.085	0.97	0.94
VF497166	669	-0.299	0.106	0.82	0.63
VF497158	669	2.372	0.090	1.22	1.38
VF497161	669	1.177	0.084	1.12	1.15
VF497165	669	1.787	0.084	1.14	1.19
VF497370	669	0.557	0.089	1.06	1.10
VF497329	669	-0.201	0.103	0.91	0.96
VF497355	669	-1.726	0.173	0.93	0.64
VF497349	669	-1.012	0.132	0.93	0.84
VF497353	669	0.941	0.085	0.97	0.94
VF497328	669	0.704	0.087	0.98	1.00
Form 5					
VF497185	672	1.000	0.085	1.24	1.31
VF497178	672	1.015	0.085	1.00	0.98
VF497193	672	1.022	0.085	1.00	1.01
VF497203	672	0.765	0.087	0.95	0.93
VF497207	672	-0.622	0.116	0.87	0.67
VF497209	672	0.635	0.088	1.01	1.03
VF497787	672	0.957	0.085	1.10	1.15
VF497796	672	1.472	0.083	1.18	1.22
VF497801	672	2.113	0.086	1.08	1.15
VF497804	672	0.517	0.090	1.03	1.08
VF497810	672	1.696	0.084	1.05	1.09
VF497792	672	0.114	0.096	0.91	0.80
Form 6					
VF497189	669	3.198	0.107	0.99	1.16
VF497173	669	1.460	0.084	1.08	1.12
VF497180	669	-1.211	0.140	0.90	0.76
VF497196	669	0.027	0.098	0.97	0.92
VF497199	669	1.269	0.084	1.12	1.14
VF497213	669	-0.121	0.101	0.93	0.86
VF497800	669	0.781	0.087	0.92	0.90
VF497797	669	1.823	0.085	1.02	1.05
VF497802	669	0.996	0.085	1.09	1.07
VF497806	669	0.789	0.087	0.94	0.91
VF497813	669	0.819	0.087	0.94	0.91
VF497799	669	0.864	0.086	1.06	1.09

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF497333	669	0.727	0.088	0.98	0.96
VF497339	669	0.878	0.086	1.05	1.08
VF497342	669	0.266	0.094	1.03	1.06
VF497347	669	1.627	0.083	1.04	1.06
VF497352	669	-0.288	0.108	0.95	0.86
VF497335	669	2.970	0.099	1.05	1.29
VF497306	669	-1.426	0.158	0.94	0.81
VF497323	669	2.243	0.087	1.26	1.49
VF497296	669	-0.347	0.109	0.98	1.04
VF497316	669	-0.242	0.106	0.91	0.77
VF497313	669	1.167	0.084	0.91	0.88
VF497312	669	-1.674	0.175	0.91	0.57
Form 8					
VF497341	648	-0.075	0.103	0.95	0.96
VF497350	648	0.295	0.095	1.10	1.18
VF497348	648	0.606	0.091	0.99	0.95
VF497346	648	-0.381	0.111	0.90	0.79
VF497351	648	1.972	0.087	1.13	1.20
VF497345	648	0.776	0.089	1.11	1.14
VF497305	648	-0.096	0.103	0.96	0.98
VF497300	648	-0.732	0.122	0.86	0.65
VF497302	648	1.038	0.087	0.93	0.93
VF497309	648	-0.205	0.106	0.88	0.74
VF497298	648	-0.975	0.132	0.84	0.53
VF497304	648	1.889	0.087	1.28	1.35
Form 9					
VF497427	675	-1.269	0.145	0.87	0.50
VF497441	675	-0.566	0.115	0.90	0.81
VF497443	675	-0.719	0.120	0.92	0.84
VF497436	675	-0.439	0.111	0.94	0.89
VF497445	675	0.136	0.097	0.91	0.86
VF497444	675	-1.401	0.152	0.83	0.46
VF497235	675	-1.074	0.135	0.83	0.52
VF497259	675	1.268	0.085	1.05	1.08
VF497269	675	2.100	0.086	1.16	1.28
VF497252	675	-0.476	0.112	0.96	1.07
VF497255	675	-0.354	0.109	0.95	1.04
VF497229	675	1.044	0.086	1.06	1.06

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF497439	674	2.345	0.089	1.16	1.31
VF497440	674	-0.237	0.105	1.06	1.20
VF497429	674	-1.436	0.156	1.00	1.15
VF497431	674	-0.854	0.126	0.85	0.63
VF497442	674	-1.023	0.134	0.86	0.54
VF497446	674	0.543	0.090	1.01	1.02
VF497266	674	1.611	0.084	1.13	1.17
VF497244	674	-0.057	0.101	0.92	0.87
VF497264	674	-0.226	0.105	0.86	0.68
VF497271	674	-0.130	0.103	0.89	0.81
VF497242	674	1.499	0.084	0.96	0.95
VF497257	674	0.042	0.099	1.02	1.02

Mathematics

Table C7. Mathematics Grade 3 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF394365	797	-0.424	0.083	0.90	0.86
VF393770	797	-0.175	0.081	0.92	0.88
VF394340	797	1.777	0.091	1.25	1.79
VF494690	797	1.810	0.092	1.15	1.49
VF406359	797	-1.637	0.103	0.95	0.77
VF393959	797	-2.168	0.121	0.91	0.70
VF394252	797	0.520	0.080	1.15	1.22
VF393784	797	-0.738	0.086	1.02	0.99
VF394358	797	0.150	0.080	0.85	0.83
VF493098	797	1.036	0.082	1.11	1.19
VF494335	797	-0.628	0.085	0.97	0.86
VF394241	797	-2.727	0.146	0.91	0.80
Form 2					
VF493110	700	-1.441	0.109	1.12	1.28
VF394366	700	-1.694	0.116	0.97	0.80
VF394359	700	-0.770	0.095	0.96	1.04
VF406337	700	0.553	0.085	0.96	0.92
VF494734	700	1.778	0.095	1.37	1.76
VF493314	700	-1.551	0.112	0.95	0.84
VF494670	700	2.042	0.100	1.13	1.62
VF394244	700	-2.071	0.130	0.95	1.03
VF387500	700	0.342	0.085	0.98	0.98
VF394382	700	-1.849	0.122	0.96	0.82
VF394355	700	1.428	0.090	1.13	1.30
VF393783	700	-2.432	0.147	0.90	0.55
Form 3					
VF494873	695	-1.358	0.109	1.02	1.08
VF393751	695	-1.370	0.109	0.91	0.83
VF494665	695	0.077	0.087	1.09	1.07
VF387507	695	0.359	0.086	1.03	1.03
VF394224	695	-0.401	0.091	0.96	1.00
VF493094	695	-0.988	0.100	1.00	0.97
VF393748	695	0.519	0.085	0.90	0.90
VF394254	695	1.230	0.088	1.17	1.34
VF406327	695	2.752	0.116	1.00	1.22
VF493445	695	-0.113	0.088	0.96	0.94
VF406306	695	-1.394	0.110	1.06	1.19
VF393824	695	0.951	0.086	1.10	1.23

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF406331	697	-0.673	0.093	0.96	1.19
VF493287	697	0.075	0.085	0.99	0.96
VF494756	697	0.176	0.085	1.14	1.17
VF493124	697	0.905	0.085	1.10	1.21
VF393823	697	0.284	0.084	0.91	0.87
VF394248	697	-0.409	0.089	0.98	0.98
VF394229	697	0.190	0.085	1.12	1.12
VF406297	697	0.725	0.085	1.11	1.17
VF494802	697	-0.449	0.090	1.08	1.10
VF494855	697	-0.299	0.088	1.15	1.19
VF494674	697	1.604	0.092	1.16	1.35
VF493428	697	0.348	0.084	1.12	1.19
Form 5					
VF494880	720	-2.320	0.143	0.96	0.75
VF394251	720	-0.932	0.096	0.87	0.72
VF493127	720	0.538	0.083	0.92	0.94
VF394375	720	-0.941	0.096	0.99	0.99
VF393961	720	-1.915	0.124	0.94	0.82
VF394379	720	0.726	0.084	1.03	1.10
VF494841	720	0.775	0.084	1.17	1.23
VF394370	720	-0.380	0.088	0.95	0.91
VF394232	720	-1.195	0.102	0.94	0.90
VF387498	720	3.090	0.128	1.02	1.62
VF406339	720	-1.026	0.098	0.95	0.82
VF406324	720	-0.720	0.092	1.07	1.22
Form 6					
VF394368	703	-0.245	0.087	1.01	1.05
VF394378	703	0.843	0.085	1.25	1.34
VF493136	703	-0.291	0.088	0.98	0.96
VF393782	703	-2.299	0.139	0.90	0.93
VF394336	703	-0.431	0.089	1.01	1.03
VF393744	703	2.069	0.101	1.13	1.41
VF394371	703	0.274	0.084	0.97	0.95
VF494820	703	1.650	0.093	1.05	1.27
VF494886	703	-1.985	0.126	0.93	0.70
VF394208	703	-0.192	0.087	1.00	0.97
VF406295	703	-0.184	0.087	1.27	1.41
VF393750	703	0.015	0.085	0.90	0.87

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF394381	701	0.493	0.084	0.98	0.94
VF493364	701	-0.052	0.086	1.08	1.12
VF494895	701	1.747	0.095	1.07	1.28
VF394376	701	-0.081	0.086	0.96	0.94
VF394221	701	0.858	0.085	1.21	1.31
VF493146	701	-0.133	0.086	0.96	0.91
VF494693	701	1.676	0.094	1.15	1.58
VF394362	701	1.049	0.086	1.03	1.02
VF387496	701	-1.065	0.099	0.95	0.82
VF393785	701	-0.022	0.085	1.04	1.02
VF393804	701	1.109	0.087	1.04	1.14
VF393752	701	0.571	0.084	0.95	1.02
Form 8					
VF393786	698	-0.879	0.097	1.00	1.00
VF493161	698	0.497	0.085	1.24	1.30
VF492342	698	-0.804	0.096	0.91	0.90
VF394361	698	-1.106	0.102	0.90	0.79
VF393749	698	0.235	0.086	0.91	0.89
VF493461	698	0.648	0.085	1.21	1.29
VF394239	698	0.162	0.086	0.80	0.74
VF393746	698	2.847	0.120	1.09	1.73
VF394339	698	-0.093	0.087	1.14	1.24
VF493387	698	-1.578	0.114	0.96	0.83
VF494833	698	1.350	0.089	1.03	1.10
VF394372	698	0.743	0.085	0.94	0.92
Form 9					
VF393788	700	-0.764	0.096	0.91	0.86
VF394363	700	0.094	0.086	1.29	1.45
VF393815	700	-1.925	0.128	0.88	0.71
VF394250	700	-0.004	0.087	1.10	1.09
VF393772	700	-0.019	0.087	0.86	0.79
VF493978	700	-2.315	0.145	0.94	0.96
VF394235	700	-4.533	0.364	0.95	1.54
VF493153	700	-0.727	0.096	1.03	1.06
VF387502	700	-3.129	0.198	0.93	0.69
VF494750	700	-0.700	0.095	0.98	0.90
VF493415	700	0.584	0.085	0.96	0.96
VF394373	700	1.613	0.091	1.04	1.33

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF494680	711	1.863	0.095	1.14	1.29
VF393742	711	1.078	0.086	1.00	1.03
VF494103	711	1.130	0.086	1.10	1.22
VF394180	711	0.078	0.085	1.11	1.10
VF393775	711	-1.621	0.113	0.87	0.60
VF394369	711	-0.060	0.086	0.88	0.82
VF406204	711	0.314	0.084	1.12	1.16
VF494861	711	-0.052	0.086	1.06	1.05
VF406343	711	-1.928	0.124	0.89	0.78
VF393793	711	-0.924	0.096	1.07	1.17
VF393800	711	-0.677	0.092	0.85	0.77
VF394356	711	-0.141	0.086	0.94	0.87

Table C8. Mathematics Grade 4 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF492343	800	2.181	0.096	1.06	1.56
VF497412	800	1.273	0.083	1.18	1.29
VF493128	800	1.555	0.086	1.11	1.31
VF493366	800	1.106	0.081	1.24	1.42
VF492367	800	-2.941	0.169	1.00	0.87
VF493154	800	-1.488	0.103	0.99	1.14
VF492312	800	-1.112	0.094	0.91	0.80
VF492333	800	0.005	0.080	1.08	1.09
VF492363	800	-0.593	0.086	1.00	0.99
VF493303	800	-0.829	0.089	1.01	0.98
VF393756	800	-0.774	0.088	0.85	0.70
VF493284	800	2.354	0.100	1.01	1.42
Form 2					
VF493305	707	1.471	0.088	1.15	1.29
VF492327	707	0.424	0.084	0.99	0.99
VF493147	707	0.718	0.084	1.02	1.06
VF492329	707	0.987	0.085	1.05	1.10
VF492341	707	-0.183	0.087	0.91	0.86
VF393714	707	0.529	0.084	1.09	1.11
VF492361	707	-0.101	0.086	0.97	0.93
VF493219	707	-0.540	0.091	0.96	0.91
VF492346	707	-3.375	0.225	0.92	0.45
VF493356	707	-1.805	0.123	1.04	0.91
VF492381	707	1.370	0.087	1.49	1.80
VF493257	707	1.066	0.085	1.09	1.15
Form 3					
VF493312	701	1.175	0.086	1.09	1.17
VF493226	701	-0.722	0.095	1.04	1.00
VF492331	701	-0.016	0.086	0.95	0.90
VF493142	701	0.227	0.085	0.90	0.85
VF393726	701	-0.668	0.094	1.15	1.34
VF493133	701	0.633	0.084	0.98	0.98
VF492375	701	-0.496	0.092	1.16	1.39
VF492348	701	-3.320	0.221	0.94	0.49
VF492336	701	0.541	0.084	0.95	0.96
VF492380	701	0.998	0.085	1.43	1.64
VF493361	701	0.591	0.084	1.25	1.40
VF493249	701	1.287	0.087	1.21	1.34

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF493228	705	1.077	0.085	0.92	0.94
VF492355	705	-1.563	0.116	1.02	1.18
VF493329	705	1.004	0.085	1.17	1.28
VF497391	705	-0.540	0.092	0.89	0.78
VF493145	705	1.484	0.089	1.03	1.16
VF493354	705	1.580	0.090	1.22	1.62
VF492332	705	0.001	0.086	1.05	1.03
VF493371	705	0.118	0.085	1.02	1.03
VF493134	705	-0.548	0.092	0.98	0.95
VF492334	705	-1.301	0.108	0.89	0.70
VF492364	705	-0.777	0.096	0.93	0.85
VF492349	705	-3.626	0.257	0.98	1.03
Form 5					
VF493140	710	0.405	0.084	0.82	0.75
VF493260	710	0.636	0.084	1.00	1.01
VF493318	710	1.277	0.087	1.00	1.11
VF492351	710	0.524	0.084	1.05	1.09
VF493373	710	-1.330	0.106	1.04	0.99
VF492326	710	1.262	0.087	1.17	1.30
VF492376	710	-0.604	0.091	1.03	1.11
VF493152	710	1.794	0.094	1.17	1.53
VF492389	710	-0.473	0.090	0.99	0.97
VF493295	710	1.883	0.095	1.10	1.25
VF492311	710	-1.674	0.116	0.96	0.77
VF493138	710	1.856	0.095	1.30	1.87
Form 6					
VF493268	723	1.231	0.085	1.38	1.64
VF493310	723	1.356	0.087	1.38	1.66
VF497407	723	-0.113	0.086	1.03	1.01
VF492387	723	0.004	0.085	0.94	0.92
VF493288	723	2.179	0.099	1.11	1.41
VF493143	723	-0.164	0.086	1.31	1.39
VF492321	723	0.378	0.083	1.01	0.99
VF492350	723	-2.285	0.144	0.92	0.83
VF492371	723	-2.005	0.130	1.00	1.09
VF493130	723	0.716	0.083	1.32	1.49
VF493151	723	1.424	0.087	1.00	0.99
VF493377	723	-0.032	0.085	0.98	0.94

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF492320	712	-1.624	0.120	0.82	0.59
VF493272	712	0.353	0.085	1.16	1.19
VF492373	712	-0.895	0.100	1.02	1.11
VF492328	712	1.134	0.085	1.16	1.35
VF493381	712	-2.382	0.154	0.92	0.78
VF493242	712	-0.051	0.087	1.11	1.26
VF492392	712	-0.301	0.090	1.04	1.04
VF493126	712	-1.758	0.125	0.93	0.96
VF492353	712	0.324	0.085	1.06	1.07
VF492306	712	-2.532	0.163	1.01	1.11
VF492338	712	-0.596	0.094	1.08	1.13
VF493344	712	-1.092	0.104	0.97	0.91
Form 8					
VF493236	694	-0.669	0.096	1.07	1.04
VF492315	694	-1.961	0.133	0.81	0.59
VF493280	694	1.374	0.088	0.95	0.97
VF492386	694	0.231	0.086	1.03	1.02
VF493131	694	-1.843	0.128	1.02	1.03
VF492339	694	-1.503	0.116	0.85	0.66
VF492323	694	-0.251	0.090	0.97	0.92
VF492354	694	-0.588	0.094	0.97	0.90
VF393648	694	1.878	0.094	1.05	1.11
VF492370	694	-0.016	0.087	1.25	1.37
VF493294	694	1.825	0.094	1.27	1.79
VF493327	694	1.638	0.091	1.14	1.57
Form 9					
VF493223	692	0.939	0.085	0.99	0.99
VF493337	692	0.918	0.085	1.01	1.02
VF492369	692	-0.714	0.095	0.96	0.92
VF492330	692	0.659	0.084	1.10	1.14
VF492352	692	-3.370	0.231	0.96	1.20
VF492337	692	0.041	0.086	1.00	0.98
VF492359	692	0.360	0.085	1.03	1.00
VF493149	692	-1.620	0.118	1.01	1.00
VF493262	692	0.795	0.085	1.25	1.35
VF493301	692	0.766	0.085	1.24	1.38
VF497402	692	0.252	0.085	1.01	0.97
VF393667	692	1.763	0.094	1.11	1.34

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF493267	713	1.857	0.093	1.35	1.81
VF493238	713	-2.375	0.151	1.05	1.21
VF492358	713	-1.474	0.112	1.00	0.95
VF497395	713	0.101	0.086	0.95	0.89
VF493349	713	1.314	0.087	0.95	1.06
VF492390	713	-2.519	0.159	0.98	0.96
VF493144	713	0.137	0.086	0.88	0.81
VF493334	713	1.532	0.089	1.02	1.08
VF492344	713	2.319	0.102	1.10	1.50
VF393675	713	0.427	0.085	0.96	0.91
VF492310	713	-0.078	0.087	0.85	0.79
VF493135	713	0.749	0.085	1.04	1.06

Table C9. Mathematics Grade 5 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF491627	770	-0.843	0.100	0.97	0.98
VF492281	770	1.271	0.082	0.93	0.90
VF491979	770	-2.132	0.150	0.95	0.88
VF491783	770	1.929	0.087	0.95	0.95
VF491900	770	1.359	0.082	0.96	0.96
VF491896	770	1.366	0.083	1.06	1.12
VF492091	770	1.084	0.082	1.26	1.37
VF492077	770	0.248	0.084	0.89	0.84
VF492374	770	0.177	0.084	1.06	1.19
VF491734	770	-0.057	0.087	0.90	0.88
VF492296	770	1.373	0.083	1.23	1.33
VF492521	770	1.629	0.084	1.25	1.38
Form 2					
VF492001	656	0.215	0.092	1.11	1.17
VF492292	656	1.874	0.092	0.87	0.81
VF491635	656	1.356	0.089	1.00	1.00
VF491895	656	-0.227	0.098	0.89	0.83
VF492095	656	0.789	0.088	0.83	0.76
VF492048	656	0.983	0.088	0.94	0.91
VF492366	656	1.208	0.088	1.18	1.26
VF492532	656	0.836	0.088	1.21	1.27
VF491937	656	1.161	0.088	1.11	1.20
VF491911	656	-1.717	0.148	0.99	1.31
VF491817	656	-0.326	0.100	0.87	0.69
VF492113	656	1.216	0.088	1.06	1.06
Form 3					
VF491973	664	-0.809	0.110	0.94	0.92
VF491926	664	0.551	0.088	1.02	1.00
VF491905	664	0.182	0.091	0.95	0.88
VF492120	664	1.606	0.089	1.15	1.21
VF492423	664	0.039	0.093	1.03	0.99
VF492010	664	1.023	0.087	1.05	1.04
VF492235	664	0.948	0.087	1.07	1.11
VF492246	664	1.519	0.088	1.05	1.07
VF491946	664	-0.005	0.093	1.00	0.98
VF491963	664	1.856	0.091	1.11	1.16
VF492083	664	1.815	0.091	1.00	1.04
VF492421	664	0.240	0.091	1.06	1.11

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF491951	667	-0.752	0.110	0.95	0.79
VF491922	667	0.038	0.094	0.84	0.73
VF491987	667	-0.864	0.113	0.93	0.91
VF492251	667	2.453	0.098	0.99	1.05
VF491727	667	1.070	0.087	0.98	1.04
VF492401	667	0.582	0.088	1.08	1.06
VF492088	667	0.972	0.087	0.98	0.97
VF491969	667	-1.217	0.125	1.03	1.31
VF492099	667	0.904	0.087	0.87	0.82
VF491630	667	1.168	0.087	1.07	1.09
VF492128	667	0.125	0.093	1.32	1.63
VF491634	667	0.302	0.091	1.10	1.13
Form 5					
VF492283	661	2.490	0.100	1.01	1.13
VF491924	661	0.541	0.089	1.04	0.99
VF492000	661	0.266	0.091	1.15	1.35
VF491804	661	0.829	0.088	0.83	0.81
VF492288	661	-0.098	0.096	1.11	1.19
VF491962	661	0.028	0.094	0.98	0.90
VF492304	661	0.837	0.088	0.97	0.92
VF492427	661	1.707	0.090	0.99	0.95
VF492432	661	-0.181	0.097	0.85	0.76
VF491821	661	1.499	0.089	1.09	1.21
VF492411	661	0.316	0.091	1.11	1.14
VF492208	661	0.191	0.092	0.87	0.92
Form 6					
VF491636	671	1.170	0.087	0.96	0.95
VF491984	671	-0.205	0.098	0.92	0.98
VF492301	671	2.345	0.096	1.06	1.22
VF491939	671	-0.341	0.100	0.87	0.73
VF491952	671	-0.947	0.115	0.96	0.90
VF492391	671	0.273	0.091	1.10	1.08
VF492174	671	0.843	0.087	0.90	0.91
VF491885	671	1.414	0.088	1.28	1.38
VF491968	671	0.437	0.090	1.08	1.12
VF492020	671	0.751	0.088	0.96	0.98
VF492435	671	1.460	0.088	1.10	1.15
VF491744	671	2.363	0.096	1.27	1.57

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF492298	673	0.999	0.087	1.02	1.03
VF491990	673	0.660	0.087	0.83	0.78
VF491941	673	0.827	0.087	1.11	1.15
VF492248	673	0.546	0.088	1.02	0.96
VF491843	673	0.946	0.087	0.86	0.84
VF492397	673	1.507	0.088	1.03	1.09
VF491902	673	1.736	0.090	1.22	1.42
VF492519	673	0.367	0.089	0.98	0.92
VF491978	673	1.036	0.087	1.03	1.01
VF491967	673	1.111	0.087	1.06	1.07
VF492186	673	1.338	0.087	1.17	1.23
VF491753	673	0.561	0.088	1.24	1.40
Form 8					
VF491998	656	0.709	0.089	0.96	0.92
VF491903	656	2.156	0.095	0.91	0.93
VF492214	656	1.963	0.093	0.88	0.82
VF492303	656	1.803	0.091	1.13	1.26
VF491945	656	-0.511	0.104	0.90	0.76
VF492524	656	-1.423	0.134	1.03	1.20
VF492403	656	0.638	0.089	1.13	1.19
VF491914	656	0.881	0.088	1.15	1.18
VF491992	656	0.686	0.089	1.14	1.24
VF492027	656	1.346	0.088	1.26	1.40
VF492203	656	-0.917	0.115	0.81	0.54
VF491761	656	0.143	0.093	0.87	0.76
Form 9					
VF492313	661	1.138	0.087	1.03	1.04
VF491948	661	1.093	0.086	0.98	0.95
VF492031	661	0.921	0.086	0.92	0.90
VF492006	661	0.467	0.088	0.99	0.99
VF491927	661	-0.190	0.095	0.97	0.98
VF491771	661	-0.181	0.095	0.88	0.80
VF492526	661	0.741	0.087	1.22	1.32
VF492038	661	2.372	0.099	1.08	1.36
VF491794	661	0.039	0.092	0.91	0.97
VF492007	661	0.030	0.092	1.03	1.06
VF492406	661	-0.678	0.105	0.97	0.86
VF491997	661	-1.197	0.121	0.94	0.91

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF491888	673	-0.525	0.103	0.89	0.80
VF492255	673	1.272	0.087	1.08	1.10
VF491626	673	0.113	0.092	1.00	0.91
VF492015	673	1.909	0.091	1.14	1.27
VF491916	673	0.838	0.087	0.84	0.78
VF492003	673	1.295	0.087	1.14	1.26
VF492528	673	0.263	0.090	1.13	1.20
VF492416	673	-0.142	0.096	1.09	1.20
VF492211	673	3.246	0.117	1.07	1.43
VF492382	673	-0.105	0.095	1.08	1.09
VF491932	673	1.212	0.087	0.95	0.92
VF491811	673	-0.007	0.094	1.09	1.15

Table C10. Mathematics Grade 6 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF491802	788	2.052	0.084	1.30	1.48
VF492134	788	2.668	0.091	1.49	1.93
VF492025	788	1.969	0.083	0.98	0.95
VF492388	788	0.687	0.082	1.05	1.14
VF491930	788	0.871	0.081	0.93	0.92
VF492210	788	1.633	0.081	1.10	1.13
VF492588	788	-0.707	0.104	0.99	0.98
VF493006	788	1.226	0.080	1.03	1.02
VF493002	788	1.548	0.081	1.17	1.19
VF492790	788	3.335	0.105	1.04	1.49
VF492716	788	1.554	0.081	0.97	0.95
VF493058	788	0.647	0.082	0.92	0.86
Form 2					
VF492533	660	0.654	0.091	0.90	0.81
VF492093	660	2.077	0.089	0.94	0.98
VF492287	660	1.557	0.087	1.19	1.25
VF492220	660	1.670	0.087	1.21	1.28
VF491953	660	0.629	0.091	1.00	1.01
VF492322	660	0.784	0.090	1.07	1.02
VF491894	660	1.542	0.087	0.97	1.01
VF492399	660	1.058	0.088	1.05	1.08
VF492941	660	2.612	0.095	0.96	1.01
VF491981	660	-1.666	0.171	0.94	0.62
VF492747	660	3.242	0.107	1.07	1.26
VF491879	660	-0.616	0.118	0.88	0.62
Form 3					
VF492300	665	1.835	0.087	1.19	1.27
VF491960	665	0.629	0.091	1.05	1.06
VF492542	665	1.053	0.088	0.97	0.97
VF492660	665	2.757	0.096	0.93	1.01
VF492280	665	1.789	0.087	1.04	1.05
VF492233	665	-0.359	0.111	0.96	0.88
VF492229	665	3.373	0.108	1.21	1.46
VF491813	665	2.096	0.088	1.01	1.04
VF492996	665	2.166	0.089	0.84	0.85
VF491912	665	-0.874	0.130	0.90	0.60
VF492415	665	0.432	0.094	0.98	0.91
VF492838	665	2.748	0.096	1.15	1.38

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF492593	658	0.703	0.090	0.91	0.84
VF492634	658	2.975	0.102	1.04	1.04
VF492309	658	1.203	0.087	1.18	1.26
VF492769	658	3.472	0.114	1.30	2.12
VF492240	658	1.347	0.087	1.21	1.25
VF492879	658	1.490	0.087	0.92	0.90
VF492238	658	2.438	0.093	1.25	1.49
VF491837	658	0.171	0.097	1.07	1.10
VF491906	658	-0.837	0.126	0.89	0.63
VF492773	658	2.625	0.096	1.07	1.13
VF493001	658	1.203	0.087	0.89	0.85
VF491947	658	1.663	0.087	0.99	0.98
Form 5					
VF492362	667	2.687	0.095	1.15	1.30
VF492671	667	1.122	0.086	1.02	1.00
VF493003	667	0.627	0.090	0.93	0.86
VF491931	667	-0.666	0.119	0.86	0.62
VF492814	667	1.388	0.086	1.37	1.52
VF491787	667	0.537	0.091	1.08	1.13
VF492270	667	2.287	0.090	1.30	1.51
VF492572	667	1.929	0.087	1.07	1.12
VF492308	667	2.798	0.097	1.12	1.37
VF492543	667	2.263	0.090	1.08	1.15
VF492890	667	1.631	0.086	1.13	1.16
VF492562	667	-0.207	0.105	0.92	0.76
Form 6					
VF492585	674	1.040	0.086	0.95	0.96
VF492383	674	0.681	0.089	0.87	0.80
VF492732	674	2.143	0.088	1.20	1.30
VF491935	674	0.950	0.087	0.87	0.83
VF492009	674	4.220	0.141	1.08	1.92
VF491864	674	0.689	0.089	1.12	1.17
VF491996	674	1.425	0.085	1.07	1.07
VF492610	674	1.135	0.086	1.22	1.30
VF493013	674	2.561	0.092	1.12	1.32
VF492102	674	2.536	0.092	1.37	1.69
VF492605	674	0.199	0.096	0.90	0.87
VF493008	674	0.436	0.092	1.04	1.07

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF492408	667	1.907	0.087	1.15	1.25
VF492577	667	0.744	0.089	0.87	0.79
VF492030	667	1.582	0.086	1.14	1.16
VF492273	667	-0.226	0.106	1.00	1.00
VF493062	667	0.885	0.088	0.99	1.03
VF491966	667	0.815	0.088	1.01	0.98
VF492759	667	2.553	0.093	1.19	1.51
VF492656	667	3.307	0.108	1.25	1.81
VF492652	667	-0.637	0.118	0.93	0.96
VF492160	667	2.623	0.094	1.19	1.53
VF492290	667	-0.637	0.118	0.93	0.80
VF493024	667	3.138	0.104	1.14	1.38
Form 8					
VF492066	658	1.834	0.088	1.21	1.33
VF492431	658	-2.456	0.236	0.94	0.77
VF492412	658	2.093	0.090	1.10	1.17
VF492053	658	1.604	0.087	0.94	0.91
VF492703	658	2.166	0.090	1.27	1.39
VF492284	658	-0.136	0.104	0.98	0.95
VF491860	658	0.742	0.090	1.01	1.04
VF492181	658	0.635	0.091	1.08	1.14
VF492676	658	1.241	0.087	1.27	1.47
VF492926	658	2.640	0.096	1.19	1.38
VF491970	658	0.726	0.090	1.07	1.08
VF493068	658	1.966	0.089	0.97	0.96
Form 9					
VF492422	677	1.985	0.087	1.08	1.15
VF492594	677	1.962	0.087	0.85	0.79
VF492078	677	1.623	0.085	1.06	1.07
VF491940	677	1.123	0.086	0.94	0.88
VF491976	677	1.500	0.085	1.17	1.20
VF492582	677	0.459	0.091	0.98	0.93
VF493092	677	1.334	0.085	0.89	0.87
VF492831	677	3.146	0.104	1.22	1.67
VF492192	677	2.015	0.087	1.08	1.13
VF492709	677	1.872	0.086	1.02	1.02
VF493017	677	2.169	0.088	1.28	1.41
VF493103	677	3.246	0.106	1.48	2.22

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF492260	675	2.076	0.087	0.98	0.97
VF492931	675	2.121	0.087	0.88	0.84
VF492202	675	1.869	0.086	1.11	1.13
VF491874	675	1.309	0.086	0.90	0.86
VF493089	675	1.243	0.086	0.88	0.85
VF491986	675	0.849	0.089	0.89	0.80
VF492209	675	1.578	0.085	1.21	1.31
VF492647	675	1.199	0.086	0.94	0.89
VF492721	675	1.913	0.086	1.10	1.14
VF492424	675	-0.229	0.108	1.01	0.98
VF492722	675	1.964	0.086	1.05	1.06
VF492711	675	0.927	0.088	1.07	1.18

Table C11. Mathematics Grade 7 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF492677	837	1.361	0.076	0.97	0.93
VF492878	837	2.746	0.084	0.99	1.05
VF492999	837	1.407	0.076	1.03	1.04
VF492786	837	0.307	0.085	0.97	1.02
VF492665	837	1.845	0.076	1.01	0.99
VF492951	837	1.591	0.076	1.09	1.15
VF492236	837	3.453	0.098	1.02	1.28
VF492780	837	1.413	0.076	0.97	0.95
VF492649	837	4.512	0.135	1.08	2.17
VF492583	837	2.553	0.082	1.24	1.36
VF493043	837	2.676	0.083	1.19	1.33
VF492625	837	2.825	0.085	1.27	1.43
Form 2					
VF492915	670	0.233	0.102	1.00	1.01
VF492560	670	-0.269	0.116	0.91	0.77
VF493038	670	0.988	0.089	0.79	0.72
VF492407	670	1.902	0.085	1.04	1.06
VF492645	670	2.256	0.087	0.97	0.97
VF492980	670	2.346	0.087	1.28	1.38
VF493053	670	3.784	0.113	1.01	1.30
VF493012	670	3.672	0.110	1.22	1.60
VF492578	670	3.309	0.101	1.04	1.13
VF492402	670	1.235	0.087	0.94	0.91
VF492673	670	3.555	0.107	1.06	1.13
VF492836	670	2.446	0.088	1.05	1.05
Form 3					
VF493009	668	3.336	0.101	1.08	1.28
VF493057	668	-0.940	0.149	0.95	0.86
VF492340	668	3.739	0.111	0.97	1.19
VF492760	668	2.657	0.090	0.99	0.98
VF492567	668	2.099	0.085	1.07	1.09
VF492658	668	3.895	0.116	0.99	1.39
VF492616	668	4.238	0.129	1.17	1.65
VF492765	668	3.003	0.095	1.14	1.36
VF492247	668	2.259	0.086	1.03	1.03
VF492875	668	2.446	0.087	1.08	1.13
VF492772	668	3.440	0.103	1.19	1.52
VF492973	668	2.625	0.089	1.09	1.19

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF492835	663	0.733	0.093	0.92	0.93
VF492307	663	0.537	0.096	0.84	0.71
VF492394	663	0.305	0.101	0.92	0.82
VF492888	663	1.768	0.085	0.95	0.93
VF492654	663	3.461	0.104	1.14	1.59
VF493036	663	1.993	0.085	0.97	0.95
VF492937	663	-1.053	0.155	1.05	1.70
VF493061	663	2.155	0.086	0.88	0.85
VF492959	663	2.815	0.092	1.21	1.44
VF492756	663	1.608	0.085	1.25	1.32
VF492896	663	3.304	0.101	0.97	1.07
VF493071	663	2.498	0.089	1.07	1.06
Form 5					
VF492663	667	0.591	0.094	0.89	0.76
VF492871	667	0.302	0.100	0.89	0.80
VF493021	667	2.415	0.088	1.05	1.06
VF492859	667	1.077	0.088	1.08	1.16
VF492666	667	1.708	0.085	0.92	1.01
VF492946	667	2.438	0.088	1.11	1.21
VF492546	667	2.884	0.094	1.16	1.28
VF492864	667	1.092	0.088	0.89	0.83
VF493052	667	1.716	0.085	1.11	1.15
VF492640	667	2.646	0.091	1.14	1.24
VF492748	667	0.529	0.095	0.89	0.85
VF493059	667	3.364	0.103	1.34	1.95
Form 6					
VF492413	667	1.833	0.085	1.19	1.25
VF492966	667	0.905	0.090	0.88	0.78
VF492853	667	1.386	0.086	1.02	0.99
VF492576	667	2.366	0.087	1.12	1.25
VF492910	667	1.541	0.086	1.24	1.31
VF493067	667	2.008	0.086	0.94	0.93
VF492538	667	0.889	0.091	0.79	0.73
VF492597	667	2.657	0.090	1.20	1.35
VF492846	667	1.475	0.086	1.01	0.98
VF492653	667	4.453	0.137	1.00	1.56
VF492987	667	2.373	0.087	0.95	0.99
VF492834	667	1.190	0.088	0.96	0.92

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF492824	653	0.541	0.096	0.90	0.84
VF492828	653	2.585	0.091	0.93	0.95
VF492302	653	-0.131	0.113	0.91	0.80
VF492708	653	2.088	0.087	0.94	0.91
VF492967	653	1.654	0.086	1.13	1.13
VF492898	653	2.149	0.087	1.06	1.09
VF492644	653	2.852	0.095	1.26	1.47
VF492720	653	2.241	0.088	0.86	0.82
VF492259	653	2.359	0.089	0.97	0.99
VF493077	653	0.398	0.099	0.91	0.93
VF492906	653	2.799	0.094	0.91	0.89
VF492419	653	2.241	0.088	1.02	1.01
Form 8					
VF492531	659	-0.985	0.149	0.95	0.78
VF492867	659	1.210	0.088	0.84	0.76
VF492731	659	-0.670	0.133	0.93	0.70
VF493064	659	0.077	0.106	1.02	0.98
VF493073	659	0.430	0.098	0.98	0.94
VF492778	659	1.929	0.086	0.87	0.83
VF492901	659	2.488	0.089	1.08	1.10
VF493046	659	2.230	0.087	1.10	1.15
VF493019	659	1.847	0.086	0.97	0.95
VF492995	659	1.210	0.088	1.05	1.02
VF492295	659	0.940	0.090	0.99	1.03
VF492629	659	2.900	0.094	1.28	1.54
Form 9					
VF492929	670	-1.117	0.158	1.00	0.78
VF492425	670	1.639	0.085	0.96	0.97
VF492908	670	3.363	0.101	1.10	1.35
VF492672	670	-0.442	0.124	0.87	0.77
VF492992	670	2.528	0.089	0.86	0.84
VF492636	670	3.148	0.097	1.38	1.82
VF492701	670	2.497	0.088	1.15	1.19
VF492830	670	1.186	0.088	0.89	0.83
VF492762	670	3.312	0.100	1.23	1.54
VF492986	670	1.915	0.085	1.08	1.08
VF492599	670	2.010	0.085	1.03	1.02
VF492955	670	1.567	0.085	1.00	0.98

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF492771	667	0.844	0.090	0.96	0.91
VF492921	667	1.988	0.085	1.40	1.57
VF492357	667	2.676	0.091	1.09	1.18
VF492975	667	1.485	0.085	0.85	0.79
VF493015	667	1.894	0.085	1.03	1.02
VF492595	667	4.013	0.123	1.08	1.64
VF492714	667	3.494	0.107	1.25	1.63
VF493004	667	3.395	0.104	0.90	0.97
VF493031	667	3.911	0.119	0.86	0.81
VF492589	667	3.450	0.105	0.99	1.13
VF492696	667	0.844	0.090	0.88	0.81
VF492861	667	1.543	0.085	0.99	1.00

Table C12. Mathematics Grade 8 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF492845	770	2.354	0.081	1.07	1.10
VF492719	770	1.452	0.080	0.87	0.85
VF492414	770	1.516	0.080	0.88	0.86
VF494727	770	2.139	0.080	0.99	0.98
VF491873	770	3.216	0.091	0.98	1.05
VF493157	770	1.909	0.080	1.12	1.18
VF493011	770	0.957	0.084	0.87	0.83
VF494810	770	2.522	0.083	1.28	1.40
VF492024	770	3.109	0.090	1.17	1.41
VF492409	770	3.777	0.103	1.24	1.67
VF492278	770	0.971	0.084	0.84	0.73
VF494716	770	3.077	0.089	0.99	1.13
Form 2					
VF492712	665	2.358	0.087	0.93	0.91
VF491918	665	4.329	0.127	1.12	1.31
VF492920	665	1.517	0.086	0.81	0.74
VF493125	665	2.290	0.087	1.06	1.04
VF491949	665	1.188	0.089	0.91	0.85
VF492579	665	2.104	0.086	1.16	1.18
VF492438	665	1.274	0.088	1.01	1.00
VF492268	665	2.976	0.093	1.18	1.34
VF493113	665	1.419	0.087	1.02	1.12
VF493083	665	3.020	0.094	0.94	0.95
VF492178	665	2.149	0.086	0.97	0.97
VF493160	665	3.073	0.095	1.20	1.45
Form 3					
VF492856	667	1.558	0.086	0.85	0.79
VF492440	667	0.571	0.098	1.00	0.92
VF494113	667	1.446	0.087	0.99	1.01
VF492889	667	0.000	0.113	0.91	1.02
VF492385	667	1.408	0.087	1.03	0.98
VF491971	667	3.178	0.096	1.34	1.59
VF493115	667	0.381	0.102	1.06	1.32
VF491920	667	0.562	0.098	0.91	1.15
VF494968	667	2.645	0.089	1.25	1.32
VF492592	667	1.192	0.089	0.90	0.95
VF492018	667	2.847	0.091	0.91	0.92
VF494801	667	1.647	0.086	1.06	1.04

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF492563	667	1.502	0.086	1.14	1.26
VF493112	667	0.640	0.095	0.98	0.98
VF493045	667	1.271	0.087	1.02	1.03
VF492226	667	1.670	0.085	0.96	0.93
VF492199	667	1.836	0.085	0.94	0.92
VF491824	667	2.607	0.089	1.03	1.06
VF492258	667	1.778	0.085	1.14	1.24
VF492393	667	2.009	0.085	1.02	1.01
VF494709	667	1.362	0.086	1.18	1.33
VF491999	667	3.183	0.096	1.05	1.21
VF494819	667	3.012	0.094	1.08	1.16
VF491857	667	1.399	0.086	0.95	0.95
Form 5					
VF492426	668	0.371	0.102	0.99	0.94
VF492863	668	1.887	0.085	0.89	0.88
VF492231	668	2.620	0.088	1.22	1.29
VF494751	668	2.141	0.085	1.08	1.08
VF493605	668	3.931	0.113	1.33	1.84
VF493034	668	2.184	0.086	0.96	0.95
VF491980	668	2.010	0.085	0.88	0.84
VF492410	668	1.575	0.086	0.92	0.88
VF491938	668	0.105	0.109	0.90	0.74
VF494769	668	0.999	0.090	0.97	0.94
VF493107	668	2.444	0.087	1.09	1.11
VF494898	668	3.126	0.095	1.37	1.54
Form 6					
VF492726	671	3.159	0.095	1.06	1.05
VF491862	671	-0.912	0.152	0.94	0.90
VF493088	671	1.284	0.087	0.86	0.78
VF494120	671	2.792	0.090	1.17	1.27
VF494978	671	4.239	0.123	1.03	1.41
VF492360	671	2.743	0.090	1.28	1.37
VF492028	671	2.299	0.086	0.99	1.01
VF492586	671	2.456	0.087	1.21	1.28
VF494699	671	1.625	0.085	0.99	0.96
VF491943	671	0.977	0.090	0.96	0.99
VF493040	671	1.001	0.090	0.84	0.76
VF493121	671	2.336	0.086	0.99	1.00

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF492874	677	2.551	0.086	1.14	1.17
VF492212	677	2.095	0.084	1.08	1.12
VF493150	677	2.389	0.085	0.95	0.93
VF492216	677	1.217	0.087	1.05	1.05
VF491915	677	2.974	0.091	1.14	1.23
VF493097	677	2.788	0.089	1.22	1.30
VF494983	677	2.982	0.091	1.22	1.34
VF491965	677	1.419	0.086	0.98	0.97
VF494099	677	2.718	0.088	1.26	1.34
VF492917	677	0.857	0.092	0.92	0.85
VF494949	677	3.024	0.092	1.18	1.28
VF492436	677	1.323	0.086	1.02	1.03
Form 8					
VF492400	655	1.979	0.085	1.05	1.13
VF492544	655	2.318	0.087	0.99	1.01
VF491991	655	1.577	0.086	1.00	0.97
VF494747	655	3.218	0.097	0.97	1.11
VF492430	655	1.665	0.086	1.01	0.96
VF493025	655	2.973	0.093	1.11	1.19
VF493132	655	1.367	0.087	1.16	1.36
VF491929	655	2.870	0.092	0.97	1.04
VF494125	655	3.070	0.095	1.08	1.22
VF494963	655	3.285	0.099	1.25	1.46
VF492272	655	1.158	0.089	0.91	0.86
VF492434	655	3.467	0.102	1.30	1.66
Form 9					
VF492880	670	3.777	0.108	1.00	1.33
VF492550	670	0.777	0.094	0.95	0.95
VF491907	670	1.546	0.086	1.12	1.18
VF492063	670	2.380	0.087	1.03	1.04
VF494972	670	2.762	0.090	1.50	1.69
VF492405	670	1.225	0.089	0.92	0.90
VF492568	670	1.855	0.086	1.01	1.01
VF491975	670	1.024	0.091	0.92	0.98
VF494760	670	1.929	0.086	0.94	0.92
VF492345	670	2.770	0.090	1.13	1.19
VF493000	670	0.631	0.097	0.89	0.78
VF493117	670	2.611	0.089	1.00	1.04

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF492848	667	3.695	0.105	1.09	1.26
VF492420	667	0.799	0.095	0.96	1.02
VF492907	667	0.718	0.096	0.92	1.07
VF494776	667	2.235	0.086	0.98	0.95
VF492165	667	2.931	0.091	1.01	1.08
VF492289	667	1.525	0.087	0.89	0.87
VF494928	667	3.058	0.093	1.21	1.30
VF492008	667	2.257	0.086	0.97	0.97
VF493102	667	2.890	0.091	1.09	1.14
VF492439	667	2.346	0.086	1.03	1.04
VF491923	667	0.060	0.112	0.91	0.75
VF493159	667	1.540	0.086	1.01	1.06

Science

Table C13. Science Grade 4 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF385248	798	-0.916	0.094	1.07	1.16
VF386811	798	0.092	0.079	0.99	1.01
VF386797	798	0.060	0.079	0.93	0.90
VF386826	798	-0.042	0.080	1.00	0.95
VF386788	798	-0.907	0.094	0.93	0.86
VF386795	798	-0.153	0.082	1.02	1.10
VF386911	798	0.216	0.078	0.97	1.00
VF386873	798	-0.048	0.080	1.02	1.02
VF417698	798	-1.102	0.099	0.88	0.73
VF417694	798	-0.770	0.091	0.93	0.88
VF386881	798	-0.317	0.083	0.93	0.86
Form 2					
VF385244	707	0.798	0.081	1.07	1.10
VF388696	707	1.439	0.085	1.07	1.12
VF386815	707	-0.786	0.098	0.99	0.95
VF386787	707	-0.998	0.103	0.93	0.78
VF386799	707	-0.139	0.087	0.87	0.79
VF386789	707	1.106	0.082	0.99	1.04
VF386806	707	-0.874	0.100	0.94	0.87
VF386848	707	0.838	0.081	1.05	1.04
VF386857	707	0.513	0.082	1.05	1.08
VF386901	707	1.133	0.082	0.97	0.98
VF386847	707	-2.037	0.145	0.94	0.77
VF386861	707	1.897	0.091	1.06	1.23
Form 3					
VF385250	698	-2.130	0.149	0.91	0.61
VF388708	698	-0.897	0.101	0.90	0.77
VF407131	698	-1.149	0.108	0.96	0.88
VF406427	698	0.233	0.084	0.96	0.93
VF407137	698	-2.675	0.185	0.96	0.80
VF407152	698	1.094	0.083	1.06	1.10
VF407128	698	-0.673	0.096	0.90	0.80
VF393924	698	-0.238	0.089	1.03	1.03
VF393826	698	0.059	0.086	0.95	0.93
VF393954	698	-0.701	0.097	0.93	0.89
VF393813	698	-1.850	0.135	0.97	1.00
VF393854	698	-2.860	0.200	0.98	0.85

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF388716	707	-0.597	0.093	0.95	0.88
VF385245	707	-1.301	0.111	1.04	1.04
VF407139	707	0.203	0.083	1.05	1.03
VF406430	707	-0.002	0.085	1.05	1.14
VF407138	707	1.118	0.083	1.14	1.19
VF407141	707	-1.457	0.117	1.06	1.16
VF406457	707	-1.081	0.105	0.95	0.91
VF393911	707	-1.364	0.113	0.97	0.89
VF393816	707	0.675	0.082	0.95	0.94
VF393969	707	-2.516	0.172	0.98	0.90
VF393950	707	1.091	0.083	1.04	1.09
VF394018	707	1.291	0.084	1.25	1.33
Form 5					
VF385249	706	-1.085	0.104	0.95	0.85
VF385241	706	1.041	0.081	1.02	1.03
VF416398	706	-0.521	0.091	0.90	0.84
VF311630	706	1.691	0.088	1.09	1.19
VF311628	706	0.727	0.081	1.17	1.23
VF311629	706	0.460	0.081	0.98	0.98
VF416433	706	0.688	0.081	1.16	1.19
VF416448	706	0.113	0.083	1.08	1.13
VF387280	706	1.684	0.088	1.04	1.10
VF387287	706	-1.174	0.107	0.97	0.92
VF387305	706	0.870	0.081	1.04	1.06
VF387267	706	1.101	0.082	1.04	1.05
Form 6					
VF311632	718	-1.720	0.126	0.94	0.82
VF311640	718	-0.797	0.097	1.06	1.07
VF311572	718	-0.278	0.087	0.94	0.92
VF311535	718	-1.257	0.109	0.96	1.14
VF416378	718	-0.942	0.100	1.00	0.95
VF311559	718	0.584	0.081	1.02	1.02
VF387252	718	0.009	0.084	0.93	0.90
VF387402	718	1.397	0.084	1.14	1.18
VF387256	718	0.630	0.081	0.96	0.95
VF387433	718	0.689	0.081	1.02	1.03
VF387295	718	0.696	0.081	1.11	1.13

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF386869	714	-2.129	0.151	0.90	0.61
VF386851	714	-0.457	0.092	1.02	1.05
VF311348	714	0.815	0.082	1.02	1.03
VF311378	714	-1.806	0.134	0.93	0.69
VF311318	714	0.522	0.082	1.09	1.13
VF416364	714	2.381	0.099	1.08	1.32
VF311330	714	0.164	0.084	1.05	1.08
VF416382	714	1.218	0.083	1.04	1.05
VF311584	714	-1.485	0.120	0.89	0.68
VF311586	714	-0.121	0.087	0.92	0.83
VF311548	714	-0.688	0.097	0.98	0.95
VF311576	714	-1.054	0.106	1.00	0.99
Form 8					
VF393818	697	-0.488	0.092	0.92	0.82
VF393964	697	1.091	0.083	1.14	1.18
VF393699	697	0.540	0.082	0.99	1.00
VF393636	697	-1.920	0.139	0.93	0.79
VF393646	697	-0.806	0.099	1.00	0.97
VF417679	697	0.130	0.085	1.16	1.23
VF417684	697	-0.674	0.096	0.97	0.95
VF311363	697	-1.506	0.120	0.89	0.65
VF416359	697	1.575	0.087	1.02	1.10
VF311367	697	-2.793	0.196	0.97	0.96
VF311343	697	-0.307	0.089	0.94	0.85
VF311353	697	-0.371	0.090	0.93	0.87
Form 9					
VF311580	695	-0.329	0.090	1.02	1.02
VF386739	695	-0.587	0.095	0.94	0.85
VF311567	1413	-0.096	0.061	0.98	0.94
VF417655	695	-0.281	0.090	0.98	0.99
VF386736	695	-1.607	0.125	0.92	0.75
VF386712	695	0.983	0.083	1.03	1.05
VF386706	695	0.847	0.082	1.07	1.08
VF393724	695	-0.515	0.094	0.81	0.70
VF393695	695	-1.102	0.107	0.84	0.74
VF393721	695	0.328	0.084	0.96	0.97
VF393713	695	0.222	0.084	1.10	1.15
VF417683	695	1.709	0.088	1.12	1.34

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
		Form 10			
VF385246	717	-0.686	0.095	0.95	0.88
VF388627	1515	0.679	0.056	0.99	1.01
VF386512	717	-0.343	0.089	0.87	0.79
VF386582	717	-2.144	0.148	0.94	0.80
VF386528	717	0.021	0.084	1.04	1.06
VF386520	717	-0.015	0.085	1.11	1.12
VF386579	717	-1.078	0.104	1.05	1.21
VF386702	717	-0.101	0.085	0.99	0.97
VF417662	717	-1.167	0.107	0.92	0.83
VF386718	717	-0.815	0.097	0.93	0.83
VF386732	717	1.776	0.088	1.07	1.17
VF386697	717	1.855	0.090	1.16	1.29

Table C14. Science Grade 8 IRT Statistics for Field Test Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 1					
VF407480	782	-0.109	0.078	0.93	0.91
VF407468	782	0.476	0.078	1.12	1.13
VF394814	782	0.550	0.079	1.06	1.09
VF394815	782	1.178	0.085	1.20	1.33
VF394816	782	1.044	0.083	1.07	1.09
VF394806	782	1.010	0.083	1.19	1.24
VF394811	782	1.207	0.085	1.18	1.28
VF407166	782	-0.353	0.079	0.91	0.86
VF407168	782	1.396	0.088	1.12	1.28
VF407160	782	0.562	0.079	1.00	1.00
VF407167	782	-0.187	0.078	0.90	0.88
VF407159	782	0.408	0.078	0.92	0.91
Form 2					
VF407471	658	-0.848	0.093	0.89	0.78
VF407482	658	-0.226	0.086	1.01	1.02
VF407165	658	0.559	0.086	1.05	1.06
VF407162	658	0.912	0.088	1.33	1.44
VF407161	658	0.843	0.087	0.92	0.93
VF407169	658	0.053	0.085	0.92	0.88
VF407164	658	0.677	0.086	1.04	1.05
VF394810	658	-0.813	0.093	0.88	0.82
VF394809	658	-0.248	0.087	0.99	0.99
VF394807	658	-0.233	0.086	1.03	1.10
VF394808	658	-0.248	0.087	0.96	0.95
VF394812	658	0.400	0.085	1.09	1.11
Form 3					
VF407483	662	0.142	0.084	1.10	1.09
VF407485	662	0.213	0.084	1.09	1.10
VF407140	662	1.470	0.095	1.30	1.54
VF407213	662	0.368	0.084	1.36	1.51
VF407155	662	-0.690	0.090	0.92	0.87
VF407254	662	0.749	0.086	1.04	1.07
VF407156	662	-1.393	0.104	0.98	0.87
VF407356	662	-1.997	0.125	0.91	0.81
VF407352	662	0.589	0.085	0.92	0.92
VF407323	662	0.539	0.085	1.24	1.31
VF407325	662	-1.393	0.104	1.00	1.22
VF407345	662	-0.539	0.088	1.04	1.07

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 4					
VF407470	664	0.513	0.083	1.25	1.31
VF407475	664	-0.732	0.090	1.01	1.01
VF407339	664	-0.622	0.088	0.97	0.96
VF407392	664	-0.568	0.087	0.98	1.00
VF407327	664	-0.323	0.085	1.04	1.14
VF407384	664	0.859	0.085	1.01	1.06
VF407330	664	-0.251	0.084	0.95	0.92
VF407248	664	2.022	0.108	1.17	1.44
VF524715	664	1.218	0.090	1.27	1.43
VF407154	664	0.540	0.083	1.06	1.09
VF407227	664	1.038	0.087	1.16	1.22
VF407242	664	-0.806	0.091	0.99	1.02
Form 5					
VF407478	665	-0.559	0.088	1.06	1.12
VF407473	665	0.621	0.084	1.06	1.06
VF313256	665	0.397	0.083	1.05	1.07
VF313265	665	2.429	0.120	1.10	1.52
VF313280	665	0.557	0.084	0.91	0.89
VF313272	665	1.943	0.105	1.03	1.11
VF313274	665	1.286	0.091	1.03	1.05
VF394777	665	0.092	0.083	1.06	1.09
VF394786	665	2.079	0.108	1.10	1.37
VF394784	665	1.278	0.091	1.26	1.42
VF394773	665	1.109	0.088	1.14	1.26
VF524713	665	0.994	0.087	1.16	1.23
Form 6					
VF394571	667	-0.573	0.088	0.94	0.90
VF394557	667	0.016	0.084	1.03	1.05
VF394775	667	-0.429	0.086	1.06	1.08
VF394780	667	-0.019	0.084	1.00	1.00
VF394782	667	2.009	0.107	1.26	1.87
VF394789	667	2.877	0.141	1.21	2.18
VF394787	667	1.069	0.088	1.07	1.08
VF313262	667	1.262	0.091	1.31	1.55
VF313276	667	1.389	0.093	1.16	1.25
VF313277	667	-0.923	0.093	0.88	0.81
VF313269	667	-0.620	0.088	0.99	0.93
VF313281	667	0.099	0.083	0.97	0.96

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 7					
VF394554	672	1.373	0.092	1.14	1.32
VF394574	672	-0.336	0.085	0.90	0.86
VF407454	672	-0.959	0.093	0.91	0.87
VF407456	672	-0.264	0.085	1.02	1.04
VF407466	672	0.295	0.083	1.04	1.04
VF407460	672	0.018	0.083	1.17	1.22
VF407455	672	0.825	0.086	1.23	1.30
VF313294	672	0.476	0.084	1.11	1.12
VF313297	672	0.212	0.083	1.08	1.09
VF313292	672	-0.030	0.084	1.03	1.05
VF313293	672	1.614	0.097	1.17	1.32
VF313300	672	0.136	0.083	0.99	0.99
Form 8					
VF394567	653	-0.432	0.088	1.01	1.04
VF394569	653	0.639	0.086	0.91	0.91
VF313287	653	1.540	0.096	1.24	1.34
VF313288	653	0.842	0.087	1.26	1.37
VF313289	653	0.411	0.085	1.04	1.08
VF313291	653	-0.440	0.089	0.94	0.91
VF313295	653	0.895	0.088	1.16	1.22
VF407458	653	0.012	0.086	1.04	1.07
VF407461	653	-1.329	0.104	0.94	0.95
VF407465	653	0.973	0.089	1.15	1.22
VF407467	653	-0.519	0.089	0.87	0.81
VF407469	653	-1.065	0.098	1.03	1.18
Form 9					
VF394565	662	-1.424	0.106	0.90	0.71
VF394561	662	-0.447	0.088	0.98	1.02
VF388393	662	1.069	0.088	1.09	1.11
VF388404	662	0.991	0.088	1.19	1.29
VF388442	662	0.657	0.085	1.23	1.30
VF388418	662	2.633	0.127	1.19	1.85
VF388547	662	2.715	0.130	1.03	1.31
VF394486	662	-0.777	0.092	1.05	1.07
VF394482	662	1.766	0.100	1.03	1.11
VF394497	662	0.915	0.087	1.15	1.20
VF394477	662	0.693	0.085	0.87	0.84
VF394491	662	-0.439	0.088	0.99	0.97

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
Form 10					
VF394547	668	-0.096	0.084	0.98	0.97
VF394566	668	-1.231	0.101	0.97	0.98
VF394502	668	0.227	0.083	1.05	1.08
VF394488	668	0.471	0.084	1.03	1.02
VF394474	668	0.805	0.085	1.08	1.12
VF394505	668	1.151	0.089	1.16	1.20
VF394514	668	2.280	0.113	1.24	1.77
VF388397	668	1.700	0.098	1.31	1.71
VF388454	668	0.590	0.084	1.06	1.09
VF388425	668	1.353	0.091	1.22	1.37
VF388503	668	1.271	0.090	1.00	1.03
VF388413	668	0.387	0.083	0.98	0.97

Appendix D: Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Reading

Table D1. Reading Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	N	Mean Scale Score	Std Dev Scale Score
3	No	6118	609.3	49.9
	Yes	1023	564.0	51.9
4	No	6143	682.4	52.2
	Yes	1015	631.2	52.3
5	No	5838	677.3	45.5
	Yes	934	627.9	44.9
6	No	5944	695.3	40.6
	Yes	856	646.1	38.4
7	No	5958	703.4	41.6
	Yes	869	654.8	39.7
8	No	5950	712.9	42.4
	Yes	824	664.1	38.6

Mathematics

Table D2. Mathematics Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	N	Mean Scale Score	Std Dev Scale Score
3	No	6080	669.0	55.0
	Yes	1053	630.5	54.5
4	No	6182	673.6	54.0
	Yes	981	633.3	53.7
5	No	5843	698.6	54.9
	Yes	923	652.1	49.5
6	No	5957	718.4	52.6
	Yes	838	668.2	44.6
7	No	5969	731.7	49.0
	Yes	856	686.1	38.9
8	No	5952	740.1	49.2
	Yes	827	693.6	38.3

Science

Table D3. Science Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	N	Mean Scale Score	Std Dev Scale Score
4	No	6157	676.5	44.3
	Yes	1000	651.5	40.7
8	No	5918	656.4	44.6
	Yes	836	617.3	36.9

SAWS

Table D4. SAWS Mean Raw Scores, Counts, and Raw Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	N	Mean Scale Score	Std Dev Scale Score
3	No	6251	6.87	2.75
	Yes	1041	5.11	2.56
4	No	6291	6.91	2.52
	Yes	991	5.24	2.31
5	No	5964	7.61	2.65
	Yes	945	5.69	2.36
6	No	6054	7.02	2.62
	Yes	872	5.05	2.16
7	No	6077	7.85	2.71
	Yes	895	5.48	2.54
8	No	6075	8.12	2.73
	Yes	839	5.81	2.55

Appendix E: Frequency of Individual Accommodations for 2013 PAWS Tests

Students received the same accommodations for all subjects (Reading, Mathematics, and Science). The only exceptions are for those accommodations shaded in yellow which were not allowed for the Reading test.

Reading

Table E1. 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	1	0	1
2	Student uses a Large Print Special Test Form.	3	3	3	3	6	5
3	Student uses an Audio Special Test Form.	1	6	0	9	9	2
4	Student uses magnification devices.	2	7	1	3	2	1
5	Student uses color overlays to reduce glare or enhance text.	13	16	12	9	6	2
6	Student uses templates to reduce the amount of visible print.	14	23	11	8	10	6
7	Student uses tactile graphics.	1	0	0	2	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. Readers 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	0	0	0
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	18	12	11	7	12	14
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. Readers 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.	0	0	0	0	0	0
11	Student asks for clarification of directions (not test questions or answer choices).	361	434	488	397	479	441
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	8	7	6	16	14	12
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	0	0	0	0	0	0

Table E2. 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.	94	120	108	74	70	57
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.	12	28	40	36	55	41
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.	7	5	3	6	3	13
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	1	0	1
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	3	1	1	0	0
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	102	120	128	83	66	49
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.	113	155	190	179	167	131

Table E3. 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.	481	560	604	497	555	458

Table E4. 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.	425	530	589	503	537	485
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	385	443	500	348	363	309
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	172	187	187	140	99	79

Table E5. 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.	42	43	23	28	44	29
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.	111	81	55	59	72	50
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. Readers may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Readers 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.	152	99	72	57	66	48
28	Student uses a bilingual dictionary provided by the school.	93	54	35	31	54	32

Table E6. 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 .	159	101	75	52	64	43

Table E7. 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.	150	96	69	51	48	29
31	Student is allowed to complete the test over multiple days.	145	85	46	42	45	23

Mathematics

Table E8. 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	1	1	0	1
2	Student uses a Large Print Special Test Form.	3	2	45	2	5	4
3	Student uses an Audio Special Test Form.	33	37	4	50	79	49
4	Student uses magnification devices.	2	7	1	3	2	1
5	Student uses color overlays to reduce glare or enhance text.	14	15	13	9	6	2
6	Student uses templates to reduce the amount of visible print.	14	23	11	8	10	6
7	Student uses tactile graphics.	1	0	0	2	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. Readers 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.	4	0	2	0	1	2
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	18	12	12	8	11	14
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. Readers 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.	428	480	515	349	416	333
11	Student asks for clarification of directions (not test questions or answer choices).	377	417	486	389	471	439
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	8	7	6	15	14	12
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	7	6	2	7	11	20

Table E 9.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.	99	117	109	74	70	57
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. 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.	13	27	40	36	54	41
16	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.	7	5	3	6	3	13
17	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	0	0	1	0	1
18	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	0	3	1	1	0	0
19	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.	109	117	131	81	67	49
20		119	149	193	178	167	131

Table E10. 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.	498	541	598	486	545	455

Table E11. 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.	441	509	583	492	528	484
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	401	426	495	339	355	305
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	180	183	189	139	100	79

Table E12. 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.	43	44	24	31	44	30
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.	116	83	58	62	72	51
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. Readers may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Readers 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.	155	100	75	62	66	49
28	Student uses a bilingual dictionary provided by the school.	95	56	36	33	54	34

Table E13. 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 .	164	103	78	57	64	45

Table E14. 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.	155	99	71	54	49	31
31	Student is allowed to complete the test over multiple days.	148	88	48	44	46	25

Science

Table E15. Frequency of IEP Student's Standard Accommodations: Presentation Accommodations

Code	Accommodation	4	8
1	Student uses a Braille Special Test Form.	0	1
2	Student uses a Large Print Special Test Form.	3	3
3	Student uses an Audio Special Test Form.	38	57
4	Student uses magnification devices.	7	1
5	Student uses color overlays to reduce glare or enhance text.	16	2
6	Student uses templates to reduce the amount of visible print.	23	6
7	Student uses tactile graphics.	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. Readers 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	2
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	12	14
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. Readers 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.	489	335
11	Student asks for clarification of directions (not test questions or answer choices).	427	444
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	7	12
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	6	20

Table E 16.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.	119	58
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.	28	41
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.	5	13
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	1
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	0
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	118	51
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.	153	133

Table E17. 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.	552	461

Table E18. 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.	522	492
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	435	312
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	184	80

Table E19. 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.	45	32
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.	84	53
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. Readers may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Readers 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.	102	52
28	Student uses a bilingual dictionary provided by the school.	56	36

Table E20. 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 .	105	46

Table E21. 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.	100	32
31	Student is allowed to complete the test over multiple days.	88	25

Appendix F: Rasch Difficulty, Standard Error, and Fit Statistics for 2013
Operational Items

Reading

Table F1. Reading Grade 3 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF311162	7138	-1.472	0.028	1.06	1.08
VF311164	7138	0.184	0.028	1.14	1.41
VF311167	7138	-2.144	0.030	1.10	1.14
VF311170	7138	-0.690	0.026	1.13	1.17
VF311175	7138	-1.505	0.028	1.21	1.31
VF311176	7138	-2.576	0.033	1.05	0.87
VF311180	7138	-1.544	0.028	1.20	1.31
VF311185	7138	-1.149	0.027	1.11	1.14
VF311395	7138	-1.408	0.027	1.06	1.11
VF311399	7138	-0.472	0.027	1.07	1.11
VF311403	7138	-1.150	0.027	1.16	1.22
VF311461	7138	-1.511	0.028	0.92	0.86
VF311466	7138	-1.758	0.029	0.98	0.96
VF311469	7138	-2.077	0.030	1.04	1.03
VF311472	7138	-0.502	0.027	1.21	1.31
VF486860	7138	-1.840	0.029	0.96	0.88
VF486888	7138	-0.889	0.027	1.15	1.21
VF486838	7138	-1.202	0.027	0.95	0.91
VF486849	7138	-1.886	0.029	1.08	1.10
VF486867	7138	-0.829	0.027	0.96	0.95
VF486830	7138	-0.996	0.027	0.89	0.86
VF486875	7138	-1.518	0.028	0.89	0.82
VF486896	7138	-0.655	0.027	1.01	1.02
VF486908	7138	-1.388	0.027	0.87	0.80
VF311909	7138	-2.146	0.030	0.95	0.92
VF311911	7138	-2.263	0.031	0.98	0.93
VF311894	7138	-1.272	0.027	1.01	1.00
VF311897	7138	-0.141	0.027	1.22	1.30
VF311914	7138	-1.863	0.029	1.03	1.02
VF311904	7138	-2.219	0.031	0.95	0.91
VF311916	7138	-2.819	0.036	0.90	0.71
VF311906	7138	-2.012	0.030	0.90	0.81
VF311917	7138	-1.705	0.028	0.90	0.81
VF311805	7138	-1.361	0.027	0.94	0.91
VF311810	7138	-1.442	0.028	1.19	1.23
VF311812	7138	-1.874	0.029	0.78	0.67

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF311817	7138	-1.611	0.028	0.84	0.75
VF311818	7138	-1.216	0.027	1.21	1.28
VF311820	7138	-1.058	0.027	0.89	0.85
VF311815	7138	-2.350	0.032	0.88	0.77
VF486930	7138	-2.823	0.036	0.86	0.69
VF486974	7138	-1.226	0.027	0.92	0.88
VF486987	7138	-1.573	0.028	0.94	0.93
VF486997	7138	0.170	0.028	1.16	1.42
VF486984	7138	-1.343	0.027	1.02	1.01
VF486959	7138	-1.358	0.027	0.93	0.88
VF486969	7138	-1.577	0.028	0.98	0.97
VF486990	7138	-1.810	0.029	0.89	0.80
VF486977	7138	-1.050	0.027	1.08	1.13
VF486965	7138	-2.329	0.032	0.82	0.64

Table F2. Reading Grade 4 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF311990	7143	-1.189	0.034	1.04	1.13
VF311992	7143	-0.280	0.029	0.96	0.90
VF312007	7143	0.051	0.028	0.98	0.95
VF312008	7143	0.093	0.027	1.09	1.09
VF312017	7143	0.029	0.028	0.94	0.89
VF312013	7143	-2.243	0.047	0.88	0.57
VF312014	7143	-0.604	0.030	0.89	0.78
VF312011	7143	-0.014	0.028	0.89	0.84
VF312018	7143	0.540	0.027	1.04	1.06
VF311882	7143	-0.235	0.029	1.05	1.03
VF311886	7143	0.560	0.027	1.09	1.12
VF311888	7143	1.449	0.027	1.15	1.35
VF311889	7143	-1.508	0.037	0.91	0.71
VF311891	7143	-0.920	0.032	0.88	0.73
VF311895	7143	-0.029	0.028	1.04	1.05
VF311900	7143	-0.140	0.028	1.13	1.19
VF489900	7143	-0.781	0.031	0.84	0.70
VF489905	7143	-1.285	0.035	0.87	0.70
VF489906	7143	-1.757	0.040	0.92	0.92
VF489907	7143	-1.150	0.034	0.83	0.63
VF489909	7143	-0.925	0.032	0.85	0.72
VF489902	7143	-0.187	0.028	1.04	1.07
VF489901	7143	-1.870	0.042	0.84	0.59
VF489904	7143	-1.559	0.038	0.81	0.56
VF312142	7143	0.338	0.027	1.00	0.98
VF312148	7143	-0.546	0.030	1.02	0.96
VF312160	7143	-0.995	0.033	1.02	1.05
VF312169	7143	0.900	0.027	1.13	1.19
VF312188	7143	-0.419	0.029	0.74	0.65
VF312195	7143	-2.005	0.044	1.10	1.11
VF489874	7143	-0.127	0.028	1.26	1.43
VF489884	7143	-0.071	0.028	0.98	0.96
VF489879	7143	-0.316	0.029	1.04	1.13
VF489877	7143	0.006	0.028	1.13	1.18
VF489753	7143	0.181	0.027	1.06	1.13
VF311952	7143	0.866	0.027	1.20	1.29
VF311959	7143	0.190	0.027	1.18	1.24
VF311963	7143	0.577	0.027	1.14	1.20
VF311966	7143	0.881	0.027	1.01	1.04
VF311968	7143	0.841	0.027	1.06	1.12

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF311970	7143	0.577	0.027	1.05	1.08
VF311971	7143	0.333	0.027	0.90	0.86
VF311975	7143	-0.562	0.030	0.90	0.86
VF489917	7143	0.548	0.027	1.03	1.05
VF488634	7143	0.530	0.027	1.14	1.20
VF487559	7143	-0.231	0.029	0.89	0.80
VF488033	7143	0.189	0.027	0.93	0.90
VF487043	7143	-0.263	0.029	0.94	0.88
VF487554	7143	-1.152	0.034	0.82	0.60
VF488081	7143	0.304	0.027	1.01	1.00

Table F3. Reading Grade 5 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF312209	6769	-0.024	0.028	0.83	0.76
VF312215	6769	-0.045	0.028	0.90	0.84
VF312217	6769	-1.233	0.035	1.18	1.34
VF312220	6769	0.369	0.027	1.18	1.26
VF312221	6769	-0.946	0.032	1.05	1.03
VF313503	6769	-0.336	0.029	0.94	0.91
VF313494	6769	0.015	0.028	1.16	1.22
VF313978	6769	-0.267	0.029	0.96	0.92
VF313485	6769	-0.752	0.031	0.92	0.88
VF313463	6769	-0.287	0.029	0.94	0.90
VF313476	6769	0.253	0.027	1.15	1.21
VF312367	6769	-1.152	0.034	1.12	1.30
VF312372	6769	0.308	0.027	1.07	1.09
VF312382	6769	-0.106	0.028	1.02	1.02
VF312416	6769	-0.086	0.028	1.12	1.20
VF312419	6769	0.433	0.027	0.97	0.96
VF312458	6769	-0.377	0.029	0.83	0.75
VF312460	6769	-0.827	0.032	0.99	0.96
VF312461	6769	-0.850	0.032	0.97	0.87
VF312465	6769	0.278	0.027	1.15	1.17
VF312466	6769	0.753	0.027	1.01	1.03
VF312468	6769	-1.660	0.039	1.22	1.45
VF486978	6769	0.153	0.027	0.93	0.89
VF486982	6769	0.205	0.027	1.01	1.01
VF486986	6769	0.116	0.027	0.88	0.84
VF486971	6769	0.152	0.027	0.89	0.85
VF486976	6769	-0.499	0.030	0.98	1.00
VF312475	6769	0.056	0.028	1.03	1.04
VF312481	6769	0.227	0.027	1.16	1.29
VF312484	6769	-0.382	0.029	0.86	0.78
VF312487	6769	0.826	0.027	1.02	1.04
VF312489	6769	-0.534	0.030	1.03	1.13
VF488080	6769	0.124	0.027	1.02	0.99
VF487583	6769	-0.938	0.032	0.95	0.94
VF496891	6769	-0.015	0.028	1.06	1.09
VF487007	6769	0.668	0.027	0.96	0.96
VF487021	6769	0.198	0.027	0.99	0.98
VF487040	6769	0.673	0.027	1.12	1.16
VF487551	6769	-0.392	0.029	0.92	0.84
VF487047	6769	0.159	0.027	1.09	1.11

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF487003	6769	-0.626	0.030	0.94	0.98
VF487033	6769	-0.550	0.030	0.92	0.85
VF486694	6769	0.282	0.027	0.85	0.81
VF486693	6769	0.594	0.027	1.02	1.04
VF486696	6769	-0.812	0.031	0.87	0.76
VF486697	6769	-0.304	0.029	0.94	0.87
VF486695	6769	0.355	0.027	1.22	1.31
VF486683	6769	1.012	0.027	1.08	1.13
VF486686	6769	0.939	0.027	1.06	1.09
VF486953	6769	0.294	0.027	0.97	0.95
VF486879	6769	0.401	0.027	0.99	0.98
VF486952	6769	0.171	0.027	1.05	1.07
VF486856	6769	0.137	0.027	0.88	0.84
VF486886	6769	0.273	0.027	1.02	1.01

Table F4. Reading Grade 6 IRT Statistics for Operational Item

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF489857	6800	0.076	0.029	0.98	0.97
VF489868	6800	1.125	0.026	0.96	0.95
VF489854	6800	0.776	0.027	1.13	1.16
VF489865	6800	0.128	0.028	0.97	0.97
VF489869	6800	0.341	0.028	1.10	1.17
VF489860	6800	-0.535	0.032	0.91	0.82
VF489855	6800	-0.685	0.033	1.00	1.00
VF489863	6800	1.640	0.027	1.08	1.18
VF489867	6800	1.277	0.026	1.34	1.43
VF312738	6800	0.294	0.028	0.93	0.88
VF312739	6800	0.919	0.026	1.03	1.03
VF312760	6800	0.079	0.029	1.10	1.08
VF312767	6800	0.778	0.027	1.20	1.23
VF312769	6800	0.115	0.028	0.87	0.79
VF489895	6800	0.873	0.026	0.95	0.94
VF489889	6800	0.313	0.028	0.94	0.90
VF489888	6800	1.006	0.026	1.01	1.02
VF489891	6800	0.838	0.026	1.06	1.08
VF489892	6800	0.836	0.026	1.11	1.13
VF489896	6800	0.821	0.026	0.85	0.82
VF489894	6800	0.013	0.029	1.04	1.07
VF329860	6800	0.490	0.027	1.00	1.01
VF331459	6800	-0.239	0.030	0.95	0.93
VF332097	6800	-0.739	0.034	1.04	1.04
VF333313	6800	1.393	0.027	1.13	1.19
VF339080	6800	1.364	0.026	1.11	1.16
VF340193	6800	0.523	0.027	1.02	1.03
VF341575	6800	-0.024	0.029	0.89	0.83
VF312483	6800	-2.339	0.060	0.88	0.73
VF312485	6800	-0.087	0.029	0.79	0.73
VF312486	6800	-1.300	0.040	0.85	0.75
VF312490	6800	0.011	0.029	0.96	0.91
VF312491	6800	-0.588	0.033	0.89	0.77
VF312492	6800	-1.125	0.038	0.68	0.54
VF312493	6800	0.182	0.028	0.93	0.91
VF312786	6800	-0.068	0.029	0.86	0.77
VF313231	6800	0.182	0.028	0.91	0.85
VF312868	6800	0.705	0.027	1.08	1.08
VF312785	6800	1.391	0.027	1.06	1.10
VF312870	6800	-0.274	0.030	1.00	1.01

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF313235	6800	1.229	0.026	1.01	1.02
VF312920	6800	-0.769	0.034	0.95	0.88
VF489872	6800	0.420	0.027	1.02	1.04
VF489875	6800	0.341	0.028	0.90	0.86
VF489873	6800	0.683	0.027	1.04	1.05
VF489876	6800	0.229	0.028	0.95	0.91
VF489878	6800	-0.292	0.031	0.93	0.90
VF489880	6800	1.163	0.026	0.98	0.98
VF489882	6800	0.662	0.027	1.13	1.17
VF489835	6800	0.218	0.028	0.86	0.79
VF489828	6800	0.998	0.026	1.08	1.09
VF489158	6800	-0.284	0.030	0.92	0.87
VF489833	6800	-0.682	0.033	0.86	0.69
VF489837	6800	0.277	0.028	1.08	1.11
VF489829	6800	0.813	0.026	1.02	1.03
VF489152	6800	0.108	0.028	0.84	0.76

Table F5. Reading Grade 7 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF355076	6827	1.363	0.026	1.00	1.00
VF355080	6827	1.497	0.026	1.11	1.14
VF355087	6827	0.278	0.028	0.91	0.87
VF355089	6827	-0.103	0.030	0.96	0.95
VF355091	6827	-0.321	0.032	0.93	0.83
VF355660	6827	-1.009	0.038	0.97	0.81
VF486685	6827	-0.039	0.030	0.94	0.92
VF486690	6827	0.367	0.028	1.07	1.10
VF355682	6827	0.618	0.027	1.06	1.06
VF355686	6827	0.891	0.027	1.08	1.09
VF355689	6827	0.115	0.029	0.95	0.90
VF355691	6827	1.005	0.026	0.96	0.95
VF355700	6827	0.024	0.030	1.15	1.15
VF355701	6827	0.786	0.027	0.96	0.94
VF486676	6827	0.515	0.028	0.92	0.88
VF486168	6827	1.715	0.027	1.10	1.14
VF486680	6827	1.213	0.026	1.02	1.02
VF486679	6827	2.101	0.028	1.11	1.21
VF486482	6827	1.196	0.026	1.07	1.08
VF486674	6827	0.949	0.027	0.91	0.89
VF486682	6827	0.416	0.028	1.03	1.06
VF486684	6827	1.001	0.026	0.98	0.97
VF356105	6827	0.324	0.028	0.93	0.86
VF356106	6827	-1.631	0.048	1.30	0.94
VF356110	6827	1.211	0.026	0.92	0.91
VF497511	6827	-0.843	0.037	1.01	0.86
VF356116	6827	0.612	0.027	0.92	0.87
VF356129	6827	0.314	0.028	1.06	1.04
VF355106	6827	0.105	0.029	1.05	1.10
VF355116	6827	1.068	0.026	1.01	1.01
VF355119	6827	1.326	0.026	1.15	1.19
VF355127	6827	1.128	0.026	1.11	1.13
VF355132	6827	0.645	0.027	0.99	0.98
VF355137	6827	1.084	0.026	0.95	0.93
VF486129	6827	-0.311	0.032	0.89	0.82
VF486131	6827	2.463	0.030	1.16	1.35
VF486148	6827	1.004	0.026	1.05	1.07
VF486143	6827	0.140	0.029	1.09	1.31
VF486152	6827	-0.071	0.030	1.00	1.11
VF486150	6827	0.911	0.027	1.31	1.43

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF486142	6827	-0.054	0.030	0.89	0.84
VF486155	6827	0.579	0.027	1.01	1.03
VF355722	6827	0.090	0.029	0.95	0.92
VF355750	6827	0.961	0.027	0.93	0.91
VF355735	6827	0.848	0.027	0.97	0.96
VF355740	6827	2.222	0.028	1.00	1.10
VF355747	6827	1.062	0.026	0.92	0.91
VF355725	6827	0.986	0.027	0.92	0.90
VF485019	6827	0.842	0.027	1.22	1.30
VF485552	6827	-0.262	0.031	0.90	0.83
VF485087	6827	1.243	0.026	1.05	1.07
VF485016	6827	0.518	0.027	0.86	0.81
VF485010	6827	1.060	0.026	1.04	1.05
VF485021	6827	0.954	0.027	1.12	1.15
VF485049	6827	0.845	0.027	0.98	0.97
VF485550	6827	0.771	0.027	0.95	0.93

Table F6. Reading Grade 8 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF385503	6772	1.267	0.027	0.97	0.97
VF385505	6772	0.721	0.028	1.19	1.35
VF385507	6772	1.516	0.026	1.06	1.08
VF385509	6772	1.267	0.027	1.03	1.04
VF385510	6772	0.197	0.030	1.23	1.42
VF385512	6772	-0.751	0.038	0.95	0.93
VF486947	6772	1.715	0.027	1.20	1.27
VF486950	6772	0.064	0.031	0.98	0.99
VF385516	6772	1.443	0.026	1.20	1.26
VF385518	6772	1.494	0.026	0.99	0.99
VF385520	6772	-0.030	0.031	1.10	1.25
VF385521	6772	1.054	0.027	0.99	0.97
VF385523	6772	1.225	0.027	1.07	1.08
VF385525	6772	0.320	0.029	0.92	0.83
VF486918	6772	0.930	0.027	1.01	1.02
VF486942	6772	0.209	0.030	0.97	0.98
VF486924	6772	1.099	0.027	0.92	0.91
VF486901	6772	-0.054	0.031	0.90	0.77
VF486910	6772	0.818	0.027	1.05	1.04
VF486939	6772	0.591	0.028	0.89	0.83
VF486862	6772	0.123	0.030	0.94	0.90
VF486870	6772	0.705	0.028	0.90	0.85
VF486851	6772	0.762	0.027	0.87	0.81
VF486874	6772	1.318	0.027	0.95	0.94
VF486880	6772	0.889	0.027	0.89	0.85
VF486864	6772	1.200	0.027	1.01	0.99
VF486855	6772	0.315	0.029	0.88	0.82
VF497522	6772	0.171	0.030	0.94	0.86
VF497524	6772	0.931	0.027	1.00	0.99
VF385826	6772	0.764	0.027	1.08	1.07
VF385832	6772	0.453	0.029	0.94	0.90
VF385836	6772	0.499	0.028	0.96	0.93
VF385838	6772	1.524	0.026	1.02	1.03
VF497529	6772	0.321	0.029	0.95	0.92
VF497537	6772	0.851	0.027	1.01	0.99
VF385812	6772	1.013	0.027	1.03	1.04
VF385815	6772	1.284	0.027	0.92	0.90
VF385818	6772	0.913	0.027	1.06	1.09
VF497541	6772	0.382	0.029	1.03	1.04
VF385819	6772	0.648	0.028	0.94	0.89

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF385822	6772	0.885	0.027	1.22	1.28
VF385824	6772	1.818	0.027	1.02	1.05
VF486810	6772	0.898	0.027	1.05	1.07
VF486811	6772	1.118	0.027	1.02	1.03
VF486780	6772	0.610	0.028	1.01	1.03
VF486775	6772	1.452	0.026	1.02	1.02
VF486773	6772	1.247	0.027	0.90	0.88
VF486769	6772	1.893	0.027	1.04	1.08
VF496073	6772	0.768	0.027	1.03	1.03
VF496102	6772	0.302	0.029	0.90	0.85
VF496114	6772	0.839	0.027	1.10	1.16
VF496116	6772	1.020	0.027	1.08	1.11
VF496120	6772	1.124	0.027	1.03	1.03
VF496123	6772	0.173	0.030	0.91	0.83
VF496126	6772	-0.010	0.031	0.98	0.99
VF496129	6772	0.154	0.030	0.97	1.01

Mathematics

Table F7. Mathematics Grade 3 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF479807	7101	-2.144	0.042	0.98	0.83
VF384789	7101	-1.903	0.039	1.21	1.31
VF384855	7101	-1.318	0.033	0.92	0.92
VF384871	7101	-1.340	0.033	0.84	0.75
VF385277	7101	-0.066	0.027	1.10	1.11
VF384863	7101	-0.343	0.028	0.96	0.96
VF384884	7101	0.340	0.027	1.03	1.04
VF479864	7101	-0.911	0.030	1.07	1.17
VF385367	7101	-1.773	0.037	0.91	0.79
VF479799	7101	0.378	0.027	0.97	0.96
VF479852	7101	-0.898	0.030	1.10	1.27
VF385226	7101	-1.308	0.033	1.01	0.99
VF384901	7101	0.113	0.027	0.96	0.95
VF384937	7101	-0.812	0.030	0.86	0.76
VF479882	7101	0.752	0.027	1.29	1.47
VF384949	7101	0.520	0.027	1.07	1.11
VF479846	7101	-1.833	0.038	0.97	1.00
VF384954	7101	-1.444	0.034	1.14	1.62
VF384935	7101	-1.574	0.035	1.04	1.20
VF384958	7101	0.346	0.027	0.97	0.97
VF384957	7101	-1.610	0.036	0.92	1.00
VF384961	7101	-2.447	0.046	1.04	1.22
VF384952	7101	0.449	0.027	0.96	0.96
VF385203	7101	-2.862	0.054	0.96	0.96
VF384963	7101	-1.584	0.035	0.80	0.78
VF384967	7101	-0.940	0.031	0.96	0.96
VF385033	7101	-0.951	0.031	1.11	1.15
VF479804	7101	-1.598	0.036	1.01	1.19
VF479743	7101	-0.178	0.028	0.87	0.82
VF479802	7101	0.589	0.027	1.32	1.49
VF385064	7101	-0.428	0.028	1.11	1.15
VF385059	7101	-0.503	0.029	0.86	0.80
VF385206	7101	0.095	0.027	0.98	0.94
VF385051	7101	-0.194	0.028	0.96	0.94
VF385068	7101	0.034	0.027	1.00	0.98
VF479764	7101	-0.716	0.029	0.98	0.98
VF385432	7101	-1.339	0.033	0.88	0.72
VF479910	7101	-0.112	0.027	0.99	0.94
VF479746	7101	-0.356	0.028	1.17	1.25

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF385117	7101	-0.553	0.029	1.34	1.47
VF479868	7101	-1.451	0.034	1.00	0.91
VF479872	7101	-2.299	0.044	0.93	0.93
VF385124	7101	-0.811	0.030	0.96	0.94
VF385186	7101	0.120	0.027	0.96	0.94
VF385128	7101	-1.474	0.034	0.74	0.55
VF486889	7101	1.338	0.028	0.97	1.00
VF480086	7101	-1.289	0.033	0.96	0.98
VF385179	7101	-0.562	0.029	0.95	0.94
VF480094	7101	-0.032	0.027	0.93	0.91
VF479911	7101	-0.351	0.028	0.88	0.80
VF479791	7101	-0.988	0.031	0.90	0.81
VF479876	7101	0.329	0.027	0.87	0.85
VF479703	7101	-1.484	0.035	1.10	1.32
VF479899	7101	-2.511	0.048	0.97	1.07
VF479903	7101	-0.505	0.029	1.05	1.11
VF479908	7101	-0.647	0.029	0.94	0.90
VF385037	7101	-0.877	0.030	1.02	0.99
VF480056	7101	-1.533	0.035	0.89	0.74
VF479857	7101	-0.182	0.028	0.95	0.91
VF479891	7101	-0.210	0.028	1.10	1.12

Table F8. Mathematics Grade 4 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF385385	7144	-1.586	0.036	0.95	0.88
VF311036	7144	0.194	0.027	1.01	1.00
VF483381	7144	-0.790	0.030	0.91	0.90
VF384731	7144	-0.805	0.030	0.92	0.85
VF384718	7144	0.635	0.027	0.97	0.97
VF311041	7144	-1.070	0.032	0.91	0.83
VF384713	7144	-0.930	0.031	0.90	0.79
VF384703	7144	-0.553	0.029	0.86	0.78
VF384728	7144	0.750	0.027	0.93	0.92
VF385111	7144	-1.637	0.037	0.93	0.84
VF480680	7144	-0.246	0.028	0.87	0.79
VF483385	7144	0.174	0.027	0.93	0.91
VF385247	7144	-0.815	0.030	0.95	0.89
VF385218	7144	0.413	0.027	1.06	1.09
VF483430	7144	-1.982	0.041	0.89	0.73
VF384786	7144	-1.346	0.034	1.00	1.01
VF384744	7144	-0.760	0.030	0.97	1.00
VF384762	7144	-0.542	0.029	1.05	1.03
VF384824	7144	-1.029	0.032	0.92	0.93
VF385389	7144	-0.804	0.030	0.98	0.95
VF384777	7144	-0.137	0.028	1.05	1.06
VF483464	7144	-0.337	0.028	1.11	1.15
VF384845	7144	-0.786	0.030	1.07	1.08
VF384805	7144	-1.409	0.035	0.78	0.64
VF384815	7144	0.315	0.027	1.15	1.17
VF480281	7144	-0.042	0.027	1.16	1.24
VF311057	7144	0.534	0.027	0.91	0.88
VF385170	7144	-1.567	0.036	0.92	0.87
VF385229	7144	-0.828	0.030	0.93	0.86
VF384892	7144	0.693	0.027	1.08	1.13
VF480145	7144	0.367	0.027	0.93	0.91
VF384877	7144	-0.530	0.029	1.04	1.00
VF384923	7144	-0.626	0.029	1.24	1.42
VF384881	7144	-0.190	0.028	1.13	1.18
VF384943	7144	0.097	0.027	1.07	1.11
VF384942	7144	0.821	0.027	0.91	0.88
VF384933	7144	0.017	0.027	0.92	0.90
VF385234	7144	-0.322	0.028	0.95	0.92
VF480115	7144	-1.278	0.034	1.03	1.08
VF484939	7144	-0.016	0.027	1.00	0.96

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF483395	7144	-0.503	0.029	0.90	0.81
VF484394	7144	0.602	0.027	1.14	1.22
VF480129	7144	-0.483	0.029	1.01	0.97
VF385072	7144	-0.517	0.029	1.05	1.05
VF311061	7144	-0.525	0.029	0.93	0.91
VF385208	7144	-0.020	0.027	0.95	0.93
VF480679	7144	-2.390	0.047	0.97	0.81
VF483387	7144	-1.505	0.036	0.86	0.70
VF384953	7144	0.712	0.027	1.00	1.02
VF495781	7144	0.756	0.027	1.14	1.22
VF483426	7144	-0.334	0.028	1.02	1.00
VF385083	7144	-1.896	0.040	1.01	1.03
VF483422	7144	-1.278	0.034	0.90	0.80
VF385080	7144	-1.006	0.032	1.27	1.41
VF385088	7144	0.063	0.027	0.98	0.97
VF384951	7144	-0.141	0.028	1.04	1.01
VF480162	7144	-1.551	0.036	0.82	0.61
VF480133	7144	-0.401	0.028	0.91	0.83
VF385397	7144	-0.740	0.030	1.00	0.95
VF483932	7144	1.354	0.028	1.13	1.34
VF480107	7144	-0.263	0.028	1.10	1.21
VF483392	7144	-1.367	0.034	0.93	0.84
VF385405	7144	1.922	0.030	1.29	1.98
VF480124	7144	-1.256	0.033	0.95	0.87
VF483394	7144	-0.737	0.030	0.92	0.82

Table F9. Mathematics Grade 5 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF486165	6748	-0.230	0.031	1.00	0.96
VF385502	6748	0.075	0.029	0.97	0.95
VF385506	6748	-0.421	0.032	1.01	0.94
VF385585	6748	-0.214	0.030	0.81	0.72
VF387070	6748	-1.248	0.039	0.94	0.92
VF387090	6748	1.240	0.028	0.89	0.86
VF387039	6748	-0.059	0.030	0.87	0.77
VF385612	6748	-0.131	0.030	0.88	0.83
VF385630	6748	0.280	0.029	0.92	0.84
VF385616	6748	-0.121	0.030	0.91	0.83
VF387078	6748	0.572	0.028	1.09	1.12
VF486672	6748	0.243	0.029	1.19	1.39
VF387017	6748	-0.053	0.030	0.91	0.83
VF387041	6748	-1.607	0.044	0.95	0.83
VF385700	6748	-0.760	0.034	1.06	1.07
VF385698	6748	1.419	0.028	1.11	1.15
VF386701	6748	0.818	0.028	1.17	1.24
VF386751	6748	-0.111	0.030	1.16	1.21
VF386781	6748	0.407	0.028	0.81	0.73
VF386650	6748	1.224	0.028	1.07	1.14
VF386821	6748	1.825	0.029	0.96	1.02
VF386842	6748	0.783	0.028	1.01	1.01
VF386775	6748	0.024	0.029	0.92	0.89
VF386885	6748	0.532	0.028	1.15	1.29
VF386880	6748	0.920	0.027	0.95	0.92
VF386913	6748	0.842	0.028	1.01	0.99
VF386833	6748	0.227	0.029	1.20	1.31
VF386892	6748	-0.280	0.031	0.91	0.79
VF386983	6748	-1.006	0.036	0.91	0.86
VF486620	6748	-0.367	0.031	1.10	1.22
VF485873	6748	0.460	0.028	1.09	1.16
VF387074	6748	1.427	0.028	0.93	0.93
VF484945	6748	1.347	0.028	0.97	0.98
VF387091	6748	0.920	0.027	1.17	1.27
VF485046	6748	0.593	0.028	1.09	1.10
VF386920	6748	2.009	0.029	1.13	1.20
VF484996	6748	0.123	0.029	0.95	0.95
VF486162	6748	1.366	0.028	0.99	1.02
VF486134	6748	0.490	0.028	0.93	0.89
VF386940	6748	0.497	0.028	0.92	0.87

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF386946	6748	0.603	0.028	1.02	1.00
VF386980	6748	0.611	0.028	0.96	0.94
VF386956	6748	0.328	0.028	0.90	0.84
VF386971	6748	-0.452	0.032	1.04	1.11
VF486688	6748	0.776	0.028	1.03	1.03
VF486159	6748	0.834	0.028	1.12	1.15
VF485003	6748	0.777	0.028	0.97	0.99
VF387076	6748	0.339	0.028	1.00	0.97
VF486133	6748	-0.319	0.031	1.01	1.06
VF387004	6748	1.993	0.029	1.00	1.01
VF485000	6748	1.318	0.028	0.91	0.89
VF485007	6748	0.647	0.028	1.03	1.05
VF485017	6748	0.290	0.028	0.92	0.87
VF485053	6748	0.251	0.029	0.96	0.90
VF484986	6748	-0.178	0.030	0.93	0.87
VF486128	6748	0.004	0.029	1.06	1.25
VF486156	6748	0.090	0.029	0.98	0.94
VF484990	6748	1.055	0.027	1.09	1.12
VF387065	6748	0.669	0.028	1.01	1.01
VF387094	6748	1.223	0.028	1.16	1.21
VF387087	6748	0.601	0.028	1.15	1.15
VF387037	6748	-0.478	0.032	0.88	0.80
VF387068	6748	-0.077	0.030	0.97	0.90
VF485020	6748	0.367	0.028	1.12	1.16
VF386878	6748	0.056	0.029	0.98	0.98

Table F10. Mathematics Grade 6 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF388057	6780	0.479	0.029	1.07	1.16
VF387210	6780	-0.398	0.035	1.00	1.06
VF387199	6780	0.599	0.029	1.03	1.05
VF387232	6780	0.288	0.030	0.90	0.80
VF387205	6780	1.651	0.027	1.14	1.18
VF387878	6780	0.379	0.029	0.88	0.78
VF387920	6780	-0.023	0.032	0.94	0.84
VF479699	6780	1.643	0.027	1.09	1.15
VF387214	6780	1.492	0.027	0.98	0.96
VF387889	6780	0.092	0.031	0.88	0.76
VF387817	6780	1.235	0.027	0.95	0.91
VF479745	6780	0.371	0.029	0.89	0.81
VF479812	6780	-0.681	0.037	0.99	1.06
VF387238	6780	0.457	0.029	0.89	0.82
VF387250	6780	0.609	0.029	1.14	1.22
VF387254	6780	1.613	0.027	1.03	1.04
VF387247	6780	0.782	0.028	1.07	1.09
VF387360	6780	0.646	0.028	0.93	0.88
VF387261	6780	0.330	0.030	0.90	0.87
VF479756	6780	0.319	0.030	0.94	0.89
VF387349	6780	-0.167	0.033	1.10	1.03
VF387336	6780	1.503	0.027	1.00	0.99
VF387363	6780	0.307	0.030	1.17	1.38
VF479749	6780	1.106	0.027	1.15	1.26
VF387882	6780	1.267	0.027	1.12	1.16
VF387234	6780	0.958	0.028	1.04	1.03
VF479803	6780	1.636	0.027	1.10	1.12
VF387375	6780	2.122	0.028	1.10	1.20
VF479809	6780	2.461	0.029	1.03	1.12
VF387393	6780	0.751	0.028	1.02	1.03
VF479744	6780	0.935	0.028	1.06	1.10
VF387386	6780	1.065	0.027	0.95	0.93
VF479713	6780	0.095	0.031	0.94	0.89
VF387476	6780	-0.164	0.033	1.02	1.15
VF387505	6780	1.098	0.027	0.89	0.83
VF387480	6780	0.781	0.028	1.06	1.08
VF387483	6780	0.586	0.029	1.05	1.05
VF479796	6780	-0.065	0.032	1.05	1.15
VF387486	6780	1.146	0.027	1.07	1.08
VF387885	6780	-0.968	0.041	0.95	0.81

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF479814	6780	-0.066	0.032	0.92	0.85
VF387509	6780	0.571	0.029	0.91	0.85
VF387730	6780	1.169	0.027	1.07	1.13
VF479763	6780	1.504	0.027	1.04	1.05
VF387936	6780	1.207	0.027	1.05	1.05
VF479805	6780	0.621	0.029	0.91	0.86
VF387711	6780	0.968	0.028	0.97	0.93
VF387332	6780	0.776	0.028	0.95	0.93
VF479792	6780	1.036	0.028	1.13	1.20
VF387790	6780	0.614	0.029	0.91	0.87
VF479707	6780	0.067	0.031	0.91	0.84
VF479790	6780	0.643	0.028	0.96	0.95
VF479794	6780	0.724	0.028	0.95	0.94
VF387787	6780	0.033	0.031	0.93	0.87
VF387801	6780	0.912	0.028	1.01	1.06
VF479787	6780	1.531	0.027	1.05	1.06
VF479788	6780	0.466	0.029	0.88	0.80
VF479785	6780	2.203	0.028	0.97	1.03
VF479752	6780	1.719	0.027	1.02	1.04
VF479808	6780	1.771	0.027	1.22	1.32
VF388085	6780	2.498	0.029	1.01	1.08
VF479798	6780	0.608	0.029	1.00	0.97
VF388087	6780	0.996	0.028	1.01	1.00
VF388077	6780	-0.735	0.038	0.89	0.67
VF479806	6780	0.381	0.029	1.00	0.99

Table F11. Mathematics Grade 7 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF479865	6818	0.615	0.029	0.86	0.76
VF387798	6818	1.689	0.027	1.05	1.04
VF479853	6818	1.910	0.027	0.99	0.98
VF387755	6818	2.575	0.028	1.01	1.03
VF387914	6818	2.211	0.027	1.03	1.05
VF387673	6818	1.714	0.027	1.10	1.13
VF388037	6818	1.661	0.027	1.04	1.05
VF388095	6818	0.555	0.030	1.04	1.09
VF479890	6818	1.555	0.027	1.01	1.01
VF479858	6818	2.291	0.027	1.16	1.22
VF387746	6818	0.494	0.030	1.00	0.96
VF479860	6818	1.153	0.027	1.02	1.01
VF388115	6818	0.956	0.028	0.98	1.05
VF388118	6818	0.786	0.029	1.14	1.29
VF388105	6818	1.054	0.028	1.22	1.36
VF479862	6818	1.888	0.027	0.98	0.98
VF388121	6818	1.323	0.027	0.90	0.87
VF388127	6818	0.350	0.031	1.11	1.08
VF388130	6818	1.163	0.027	1.09	1.18
VF388523	6818	1.901	0.027	1.05	1.08
VF387744	6818	2.353	0.028	1.08	1.13
VF388337	6818	1.783	0.027	0.88	0.85
VF388357	6818	0.707	0.029	0.88	0.84
VF388528	6818	1.964	0.027	1.05	1.08
VF387723	6818	1.906	0.027	1.15	1.20
VF388485	6818	1.880	0.027	1.05	1.06
VF388414	6818	1.338	0.027	1.05	1.05
VF387716	6818	1.942	0.027	1.13	1.18
VF388432	6818	1.092	0.028	0.88	0.81
VF479895	6818	-0.046	0.034	0.97	0.94
VF479880	6818	2.012	0.027	1.03	1.03
VF479863	6818	1.986	0.027	0.89	0.86
VF479866	6818	1.932	0.027	1.07	1.08
VF479893	6818	0.873	0.028	0.94	0.89
VF479850	6818	2.466	0.028	0.98	0.96
VF388371	6818	2.087	0.027	1.00	1.01
VF388499	6818	1.863	0.027	1.13	1.18
VF388361	6818	-0.037	0.034	0.94	0.80
VF388223	6818	2.111	0.027	1.01	1.01
VF388380	6818	0.667	0.029	1.04	1.11

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF479856	6818	1.780	0.027	1.16	1.24
VF388375	6818	2.093	0.027	1.01	1.02
VF388378	6818	2.233	0.027	1.11	1.13
VF479896	6818	0.738	0.029	0.93	0.87
VF388353	6818	1.892	0.027	0.89	0.88
VF479888	6818	1.492	0.027	1.02	1.04
VF479875	6818	1.069	0.028	0.92	0.88
VF479884	6818	2.478	0.028	1.03	1.06
VF479887	6818	0.380	0.031	0.95	0.87
VF388340	6818	2.055	0.027	0.93	0.93
VF479894	6818	1.040	0.028	1.10	1.13
VF388219	6818	1.869	0.027	0.96	0.95
VF387788	6818	2.244	0.027	0.93	0.92
VF388132	6818	-0.389	0.037	0.81	0.62
VF479901	6818	1.243	0.027	0.87	0.83
VF479845	6818	1.014	0.028	1.01	1.01
VF479849	6818	2.349	0.027	0.98	0.98
VF479854	6818	2.097	0.027	1.01	1.03
VF479873	6818	2.570	0.028	0.95	0.94
VF479886	6818	0.892	0.028	0.93	0.87
VF387690	6818	1.589	0.027	0.95	0.92
VF479870	6818	0.676	0.029	1.04	1.09
VF479892	6818	1.120	0.027	0.94	0.90
VF479885	6818	1.227	0.027	0.82	0.76
VF479900	6818	2.770	0.029	0.95	0.93
VF479881	6818	0.395	0.031	1.08	1.19

Table F12. Mathematics Grade 8 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF479710	6778	0.258	0.033	0.91	0.80
VF388106	6778	1.116	0.028	1.06	1.11
VF393236	6778	1.876	0.027	1.22	1.33
VF393263	6778	1.835	0.027	1.04	1.05
VF388119	6778	2.002	0.027	1.11	1.16
VF388116	6778	1.964	0.027	1.12	1.17
VF479869	6778	1.633	0.027	0.99	0.98
VF388596	6778	1.927	0.027	0.91	0.89
VF393265	6778	2.976	0.029	0.94	0.96
VF479797	6778	1.550	0.027	0.99	1.01
VF388600	6778	1.673	0.027	1.02	1.05
VF388739	6778	1.226	0.028	1.05	1.08
VF393247	6778	1.946	0.027	1.07	1.09
VF479871	6778	2.235	0.027	1.16	1.20
VF390039	6778	1.540	0.027	1.06	1.12
VF388619	6778	1.512	0.027	1.01	1.02
VF388747	6778	1.339	0.027	1.03	1.03
VF479848	6778	2.879	0.029	1.03	1.08
VF479824	6778	2.533	0.028	1.09	1.15
VF479772	6778	0.443	0.031	0.94	0.94
VF390817	6778	2.444	0.027	1.04	1.05
VF390175	6778	1.736	0.027	1.01	1.01
VF390818	6778	0.959	0.029	0.97	0.91
VF479859	6778	2.500	0.028	1.03	1.06
VF479855	6778	0.659	0.030	0.96	0.95
VF390215	6778	2.924	0.029	1.11	1.19
VF390631	6778	1.422	0.027	0.87	0.81
VF390240	6778	1.988	0.027	1.04	1.05
VF388910	6778	1.559	0.027	1.07	1.15
VF388913	6778	1.549	0.027	1.05	1.08
VF479672	6778	0.636	0.030	1.02	1.15
VF479789	6778	2.754	0.028	0.95	0.96
VF391184	6778	1.568	0.027	0.86	0.79
VF388915	6778	2.540	0.028	1.08	1.11
VF388919	6778	3.194	0.030	1.02	1.03
VF479689	6778	2.230	0.027	1.02	1.03
VF388916	6778	2.005	0.027	1.06	1.09
VF479840	6778	0.537	0.031	0.95	0.94
VF393290	6778	2.815	0.029	1.05	1.08
VF388921	6778	1.837	0.027	0.94	0.92

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF391888	6778	1.052	0.028	1.01	1.00
VF479813	6778	1.339	0.027	1.01	1.04
VF479762	6778	2.222	0.027	0.93	0.93
VF392991	6778	0.800	0.029	0.81	0.71
VF393288	6778	2.854	0.029	0.96	0.96
VF391178	6778	2.378	0.027	0.94	0.95
VF393274	6778	2.558	0.028	1.18	1.24
VF479723	6778	1.470	0.027	1.02	1.03
VF479761	6778	0.975	0.029	0.90	0.87
VF479681	6778	2.209	0.027	0.89	0.86
VF393006	6778	1.922	0.027	1.09	1.09
VF479665	6778	2.130	0.027	0.99	0.97
VF393138	6778	0.602	0.030	0.88	0.81
VF393242	6778	1.888	0.027	0.99	1.02
VF393148	6778	1.430	0.027	0.91	0.87
VF479842	6778	1.187	0.028	0.93	0.93
VF393276	6778	2.101	0.027	0.93	0.91
VF393257	6778	2.184	0.027	1.10	1.14
VF393268	6778	1.921	0.027	0.99	0.99
VF393211	6778	2.057	0.027	0.88	0.84
VF479867	6778	0.802	0.029	0.84	0.77
VF479879	6778	2.419	0.027	1.03	1.03
VF479780	6778	1.467	0.027	1.01	1.01
VF479827	6778	0.625	0.030	0.88	0.80
VF479777	6778	1.453	0.027	0.94	0.95
VF392976	6778	2.707	0.028	1.14	1.23
VF393245	6778	1.493	0.027	0.95	0.92
VF393293	6778	2.494	0.028	1.07	1.12
VF479820	6778	1.721	0.027	0.96	0.95
VF388742	6778	1.338	0.027	0.89	0.81

Science

Table F13. Science Grade 4 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF483448	7155	-1.681	0.040	0.93	0.79
VF484396	7155	0.493	0.026	1.13	1.17
VF269828	7155	-1.409	0.036	0.96	0.97
VF269830	7155	0.400	0.026	1.11	1.15
VF269831	7155	0.568	0.026	1.04	1.04
VF484935	7155	1.201	0.026	1.19	1.29
VF290032	7155	-1.229	0.034	0.87	0.71
VF290777	7155	-0.996	0.032	0.92	0.84
VF269869	7155	-0.165	0.027	0.99	0.96
VF269871	7155	0.740	0.026	0.99	1.00
VF269873	7155	-0.205	0.028	0.94	0.91
VF282620	7155	0.990	0.026	1.03	1.06
VF282623	7155	0.423	0.026	1.05	1.05
VF282626	7155	-0.707	0.030	0.87	0.75
VF282637	7155	0.033	0.027	0.97	0.97
VF282639	7155	-0.117	0.027	0.85	0.78
VF282642	7155	-0.588	0.029	0.97	1.02
VF282661	7155	-0.032	0.027	0.99	0.99
VF282669	7155	0.320	0.026	1.00	0.98
VF282670	7155	-1.240	0.035	0.90	0.82
VF282688	7155	0.860	0.026	1.00	1.01
VF283022	7155	-0.497	0.029	1.02	0.96
VF283606	7155	-1.556	0.038	1.12	0.98
VF284002	7155	0.237	0.026	1.01	1.01
VF284006	7155	0.355	0.026	0.92	0.90
VF284665	7155	0.872	0.026	1.01	1.01
VF284979	7155	0.706	0.026	1.23	1.30
VF293507	7155	-0.916	0.032	0.91	0.77
VF292879	7155	-0.332	0.028	1.00	0.97
VF294472	7155	-0.010	0.027	0.97	0.94
VF293853	7155	1.154	0.026	1.13	1.19
VF483424	7155	-0.183	0.027	0.95	0.91
VF483437	7155	-2.379	0.051	0.98	0.95
VF483431	7155	-0.465	0.029	0.88	0.80
VF287740	7155	0.843	0.026	1.02	1.03
VF287742	7155	0.762	0.026	1.01	1.03
VF287745	7155	0.753	0.026	1.09	1.13
VF287765	7155	1.092	0.026	1.07	1.09
VF287766	7155	-0.014	0.027	0.88	0.84

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF287770	7155	-2.067	0.046	1.14	1.30
VF269857	7155	1.079	0.026	1.06	1.10
VF269846	7155	0.586	0.026	1.04	1.06
VF269841	7155	0.916	0.026	1.14	1.20
VF296839	7155	0.832	0.026	1.03	1.05
VF296833	7155	0.432	0.026	0.99	0.96
VF287864	7155	-0.205	0.028	0.98	0.96
VF287849	7155	-0.387	0.028	0.92	0.88
VF287846	7155	0.103	0.027	0.98	0.97
VF287870	7155	0.722	0.026	1.03	1.04
VF287856	7155	0.584	0.026	1.01	1.00

Table F14. Science Grade 8 IRT Statistics for Operational Items

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF484958	6752	-0.002	0.027	0.96	0.95
VF484974	6752	0.472	0.027	1.12	1.18
VF486678	6752	1.360	0.029	1.07	1.12
VF486675	6752	0.195	0.027	1.07	1.09
VF308274	6752	0.766	0.027	1.13	1.20
VF308369	6752	0.438	0.027	0.97	0.96
VF308379	6752	-0.293	0.027	0.95	0.93
VF308708	6752	-0.072	0.027	1.08	1.12
VF308712	6752	0.534	0.027	0.90	0.88
VF308713	6752	0.278	0.027	1.07	1.09
VF484993	6752	0.464	0.027	0.98	0.97
VF484999	6752	-0.030	0.027	1.09	1.14
VF308868	6752	-1.865	0.037	0.98	0.81
VF308869	6752	-0.166	0.027	1.10	1.18
VF308871	6752	-0.202	0.027	1.02	1.04
VF486699	6752	-0.922	0.029	1.02	1.14
VF486698	6752	-0.739	0.029	0.96	0.98
VF486135	6752	-0.091	0.027	0.93	0.91
VF486126	6752	0.177	0.027	1.02	1.01
VF308893	6752	-0.363	0.027	0.98	0.97
VF308894	6752	0.686	0.027	1.12	1.16
VF308895	6752	-0.967	0.030	0.98	0.95
VF308899	6752	-0.285	0.027	1.09	1.16
VF308900	6752	-0.053	0.027	0.94	0.92
VF308901	6752	0.388	0.027	1.10	1.13
VF308906	6752	-0.106	0.027	0.87	0.82
VF308909	6752	1.513	0.030	1.10	1.26
VF308910	6752	0.181	0.027	1.03	1.02
VF308928	6752	1.363	0.029	1.12	1.22
VF308929	6752	-0.134	0.027	0.90	0.87
VF308930	6752	0.027	0.027	0.95	0.93
VF486847	6752	-0.523	0.028	0.98	0.97
VF486858	6752	0.948	0.028	1.00	1.03
VF486815	6752	-0.329	0.027	0.92	0.89
VF486821	6752	-0.302	0.027	0.94	0.89
VF486149	6752	0.651	0.027	0.97	0.98
VF486151	6752	-0.401	0.027	1.02	1.03
VF486146	6752	0.063	0.027	1.04	1.03
VF486166	6752	-0.275	0.027	1.08	1.14
VF486163	6752	-0.227	0.027	0.93	0.90

Accession Number	N	Rasch Difficulty	Rasch SE	Infit	Outfit
VF486771	6752	-0.098	0.027	1.02	1.02
VF486782	6752	-0.126	0.027	0.99	0.98
VF486765	6752	-0.182	0.027	0.91	0.86
VF486687	6752	-0.274	0.027	0.93	0.90
VF486692	6752	-1.038	0.030	0.92	0.92
VF486914	6752	0.961	0.028	1.07	1.11
VF486941	6752	0.413	0.027	1.00	1.00
VF486948	6752	0.884	0.027	1.00	1.02
VF485018	6752	-0.177	0.027	1.05	1.08
VF485023	6752	0.300	0.027	1.05	1.08

Appendix G: Classical Item Statistics for 2013 Operational Items

Reading

Table G1. Reading Grade 3 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF311162	7141	0.68	0.35	0.35	-0.18	-0.19	-0.20
VF311164	7141	0.30	0.21	0.21	-0.17	-0.08	-0.06
VF311167	7141	0.73	0.40	-0.26	-0.21	-0.16	0.40
VF311170	7141	0.51	0.33	-0.25	-0.10	-0.11	0.33
VF311175	7141	0.56	0.38	-0.21	0.38	-0.14	-0.20
VF311176	7141	0.77	0.54	-0.26	-0.30	-0.30	0.54
VF311180	7141	0.62	0.32	0.32	-0.15	-0.20	-0.21
VF311185	7141	0.59	0.35	-0.16	0.35	-0.21	-0.14
VF311395	7141	0.60	0.42	-0.17	-0.27	0.42	-0.19
VF311399	7141	0.50	0.40	-0.24	0.40	-0.06	-0.22
VF311403	7141	0.59	0.31	-0.12	-0.23	-0.15	0.31
VF311461	7141	0.71	0.45	-0.20	0.45	-0.25	-0.25
VF311466	7141	0.69	0.45	0.45	-0.28	-0.32	-0.11
VF311469	7141	0.71	0.48	-0.30	-0.23	0.48	-0.20
VF311472	7141	0.55	0.29	0.29	-0.17	-0.13	-0.13
VF486860	7141	0.71	0.47	0.47	-0.21	-0.28	-0.22
VF486888	7141	0.53	0.31	-0.16	-0.13	-0.18	0.31
VF486838	7141	0.60	0.49	-0.28	0.49	-0.24	-0.19
VF486849	7141	0.72	0.36	0.36	-0.21	-0.17	-0.21
VF486867	7141	0.52	0.48	-0.11	-0.22	0.48	-0.31
VF486830	7141	0.56	0.54	-0.39	0.54	-0.16	-0.19
VF486875	7141	0.66	0.54	-0.32	-0.23	0.54	-0.26
VF486896	7141	0.49	0.43	-0.22	-0.14	-0.27	0.43
VF486908	7141	0.63	0.56	-0.27	-0.30	0.56	-0.25
VF311909	7141	0.76	0.46	-0.26	-0.21	-0.28	0.46
VF311911	7141	0.78	0.42	0.42	-0.23	-0.23	-0.21
VF311894	7141	0.61	0.44	-0.25	0.44	-0.17	-0.21
VF311897	7141	0.39	0.24	0.11	-0.39	0.24	-0.26
VF311914	7141	0.72	0.40	-0.12	0.40	-0.29	-0.18
VF311904	7141	0.77	0.45	-0.26	-0.26	-0.22	0.45
VF311916	7141	0.85	0.47	-0.26	-0.24	-0.26	0.47
VF311906	7141	0.74	0.51	-0.28	-0.21	-0.32	0.51
VF311917	7141	0.69	0.53	-0.28	0.53	-0.29	-0.26
VF311805	7141	0.70	0.43	-0.21	0.43	-0.21	-0.24
VF311810	7141	0.59	0.33	0.33	-0.18	-0.15	-0.14
VF311812	7141	0.76	0.56	-0.33	-0.27	0.56	-0.31
VF311817	7141	0.71	0.54	-0.32	-0.26	0.54	-0.24
VF311818	7141	0.52	0.33	-0.33	-0.23	-0.06	0.33
VF311820	7141	0.62	0.53	-0.27	-0.26	-0.24	0.53

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF311815	7141	0.78	0.54	-0.32	-0.30	-0.25	0.54
VF486930	7141	0.85	0.49	-0.26	-0.28	0.49	-0.27
VF486974	7141	0.60	0.52	-0.23	-0.34	-0.21	0.52
VF486987	7141	0.67	0.49	-0.23	0.49	-0.32	-0.24
VF486997	7141	0.33	0.23	-0.25	-0.19	0.23	0.10
VF486984	7141	0.62	0.43	0.43	-0.25	-0.17	-0.20
VF486959	7141	0.63	0.51	-0.33	0.51	-0.23	-0.19
VF486969	7141	0.67	0.46	-0.22	-0.30	0.46	-0.22
VF486990	7141	0.71	0.53	-0.26	-0.29	0.53	-0.28
VF486977	7141	0.57	0.38	-0.26	-0.20	-0.10	0.38
VF486965	7141	0.79	0.58	-0.32	-0.30	0.58	-0.30

Table G2. Reading Grade 4 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF311990	7158	0.83	0.34	-0.27	-0.18	0.34	-0.17
VF311992	7158	0.70	0.48	-0.32	0.48	-0.19	-0.26
VF312007	7158	0.64	0.46	-0.20	0.46	-0.22	-0.26
VF312008	7158	0.63	0.37	-0.22	-0.14	-0.23	0.37
VF312017	7158	0.65	0.50	-0.31	-0.25	0.50	-0.23
VF312013	7158	0.93	0.42	-0.24	0.42	-0.20	-0.27
VF312014	7158	0.75	0.53	0.53	-0.35	-0.25	-0.25
VF312011	7158	0.65	0.54	-0.25	0.54	-0.30	-0.31
VF312018	7158	0.55	0.42	0.42	-0.14	-0.20	-0.32
VF311882	7158	0.69	0.41	-0.21	-0.14	0.41	-0.26
VF311886	7158	0.45	0.40	-0.16	-0.22	-0.18	0.40
VF311888	7158	0.36	0.26	0.26	-0.20	-0.31	0.04
VF311889	7158	0.86	0.47	0.47	-0.24	-0.26	-0.28
VF311891	7158	0.78	0.57	-0.30	-0.26	-0.33	0.57
VF311895	7158	0.63	0.45	-0.16	-0.26	-0.25	0.45
VF311900	7158	0.62	0.41	-0.23	0.41	-0.17	-0.27
VF489900	7158	0.78	0.56	-0.33	-0.34	0.56	-0.27
VF489905	7158	0.84	0.50	-0.28	-0.29	-0.25	0.50
VF489906	7158	0.89	0.41	-0.27	0.41	-0.23	-0.18
VF489907	7158	0.83	0.56	-0.29	-0.31	-0.31	0.56
VF489909	7158	0.80	0.55	0.55	-0.27	-0.29	-0.31
VF489902	7158	0.68	0.41	-0.21	0.41	-0.21	-0.21
VF489901	7158	0.90	0.48	0.48	-0.26	-0.27	-0.27
VF489904	7158	0.87	0.54	0.54	-0.34	-0.27	-0.28
VF312142	7158	0.62	0.43	-0.28	-0.17	-0.19	0.43
VF312148	7158	0.72	0.46	0.46	-0.21	-0.32	-0.18
VF312160	7158	0.81	0.38	-0.17	-0.22	0.38	-0.26
VF312169	7158	0.49	0.33	0.33	-0.10	-0.18	-0.23
VF312188	7158	0.85	0.43	-0.30	-0.23	-0.20	0.43
VF312195	7158	0.89	0.38	-0.23	0.38	-0.20	-0.25
VF489874	7158	0.67	0.21	0.21	-0.07	-0.26	-0.08
VF489884	7158	0.66	0.46	-0.29	-0.14	-0.24	0.46
VF489879	7158	0.71	0.40	-0.22	-0.21	-0.21	0.40
VF489877	7158	0.65	0.33	0.33	-0.12	-0.20	-0.22
VF489753	7158	0.62	0.40	-0.09	0.40	-0.25	-0.29
VF311952	7158	0.43	0.26	-0.25	0.26	-0.09	-0.04
VF311959	7158	0.59	0.32	-0.14	-0.14	-0.17	0.32
VF311963	7158	0.52	0.32	-0.11	-0.16	0.32	-0.17
VF311966	7158	0.51	0.44	-0.22	-0.24	-0.18	0.44
VF311968	7158	0.54	0.40	-0.31	-0.23	0.40	-0.07
VF311970	7158	0.57	0.39	-0.20	-0.18	0.39	-0.19
VF311971	7158	0.61	0.52	-0.37	0.52	-0.22	-0.17

Accession Number	N	Mean	Item-Test Corr	<u>Option Discrimination (MC)</u>			
				A	B	C	D
VF311975	7158	0.77	0.46	-0.23	-0.26	0.46	-0.25
VF489917	7158	0.55	0.42	0.42	-0.26	-0.18	-0.16
VF488634	7158	0.55	0.32	-0.11	-0.23	0.32	-0.12
VF487559	7158	0.69	0.54	-0.32	0.54	-0.27	-0.21
VF488033	7158	0.62	0.51	-0.23	-0.27	0.51	-0.24
VF487043	7158	0.70	0.50	-0.24	0.50	-0.28	-0.27
VF487554	7158	0.83	0.56	-0.30	-0.34	-0.27	0.56
VF488081	7158	0.59	0.44	-0.20	-0.31	-0.17	0.44

Table G3. Reading Grade 5 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF312209	6772	0.72	0.50	-0.18	-0.25	-0.33	0.50
VF312215	6772	0.67	0.49	-0.31	-0.25	0.49	-0.18
VF312217	6772	0.81	0.29	-0.21	-0.15	-0.10	0.29
VF312220	6772	0.52	0.27	-0.14	0.27	-0.12	-0.16
VF312221	6772	0.78	0.37	-0.24	-0.15	0.37	-0.21
VF313503	6772	0.70	0.46	-0.22	-0.16	-0.31	0.46
VF313494	6772	0.63	0.25	-0.07	0.25	-0.22	-0.18
VF313978	6772	0.69	0.45	-0.21	-0.30	0.45	-0.19
VF313485	6772	0.77	0.46	-0.16	-0.36	0.46	-0.23
VF313463	6772	0.69	0.47	-0.23	-0.24	-0.24	0.47
VF313476	6772	0.59	0.27	-0.04	-0.11	0.27	-0.26
VF312367	6772	0.80	0.31	-0.05	-0.19	-0.27	0.31
VF312372	6772	0.58	0.34	0.34	-0.17	-0.18	-0.16
VF312382	6772	0.64	0.42	-0.26	0.42	-0.19	-0.16
VF312416	6772	0.63	0.32	-0.16	-0.20	0.32	-0.12
VF312419	6772	0.58	0.44	-0.16	-0.25	-0.26	0.44
VF312458	6772	0.75	0.49	-0.33	0.49	-0.18	-0.25
VF312460	6772	0.77	0.40	0.40	-0.24	-0.25	-0.21
VF312461	6772	0.76	0.49	0.49	-0.23	-0.24	-0.31
VF312465	6772	0.52	0.32	-0.24	-0.16	0.32	-0.18
VF312466	6772	0.50	0.40	-0.15	0.40	-0.20	-0.18
VF312468	6772	0.85	0.31	-0.16	-0.22	-0.16	0.31
VF486978	6772	0.61	0.49	-0.31	0.49	-0.21	-0.19
VF486982	6772	0.60	0.40	-0.30	-0.25	0.40	-0.09
VF486986	6772	0.61	0.53	-0.31	-0.28	0.53	-0.19
VF486971	6772	0.61	0.53	0.53	-0.24	-0.27	-0.25
VF486976	6772	0.73	0.41	-0.21	0.41	-0.24	-0.21
VF312475	6772	0.67	0.33	-0.19	0.33	-0.17	-0.22
VF312481	6772	0.60	0.26	0.26	-0.07	-0.12	-0.18
VF312484	6772	0.72	0.51	-0.32	-0.22	0.51	-0.26
VF312487	6772	0.52	0.40	0.40	-0.22	-0.15	-0.24
VF312489	6772	0.72	0.38	-0.26	0.38	-0.18	-0.18
VF488080	6772	0.61	0.39	-0.23	-0.18	-0.21	0.39
VF487583	6772	0.80	0.41	-0.19	0.41	-0.21	-0.26
VF496891	6772	0.64	0.36	-0.18	0.36	-0.12	-0.22
VF487007	6772	0.50	0.45	-0.31	0.45	-0.18	-0.17
VF487021	6772	0.60	0.42	-0.22	-0.28	0.42	-0.12
VF487040	6772	0.50	0.30	-0.09	-0.12	-0.19	0.30
VF487551	6772	0.71	0.48	-0.22	-0.26	0.48	-0.27
VF487047	6772	0.61	0.33	-0.22	0.33	-0.25	-0.11
VF487003	6772	0.75	0.44	0.44	-0.27	-0.21	-0.21
VF487033	6772	0.74	0.47	-0.23	-0.23	0.47	-0.28

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF486694	6772	0.58	0.57	-0.25	-0.31	-0.24	0.57
VF486693	6772	0.52	0.39	-0.19	0.39	-0.20	-0.21
VF486696	6772	0.78	0.51	-0.33	-0.27	-0.21	0.51
VF486697	6772	0.69	0.46	-0.21	-0.30	0.46	-0.17
VF486695	6772	0.57	0.21	0.21	-0.20	-0.07	-0.04
VF486683	6772	0.43	0.32	-0.13	-0.12	-0.16	0.32
VF486686	6772	0.45	0.34	-0.23	0.34	-0.14	-0.09
VF486953	6772	0.58	0.45	-0.20	0.45	-0.22	-0.21
VF486879	6772	0.56	0.43	0.43	-0.20	-0.14	-0.27
VF486952	6772	0.60	0.36	-0.25	-0.17	0.36	-0.20
VF486856	6772	0.61	0.53	-0.30	-0.20	-0.29	0.53
VF486886	6772	0.58	0.40	0.40	-0.26	-0.20	-0.12

Table G4. Reading Grade 6 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF489857	6800	0.70	0.40	-0.14	-0.20	-0.28	0.40
VF489868	6800	0.49	0.42	-0.19	-0.21	-0.17	0.42
VF489854	6800	0.56	0.25	0.25	-0.20	-0.08	-0.14
VF489865	6800	0.69	0.40	-0.21	0.40	-0.28	-0.14
VF489869	6800	0.65	0.27	-0.21	-0.26	-0.07	0.27
VF489860	6800	0.80	0.45	0.45	-0.18	-0.27	-0.26
VF489855	6800	0.82	0.33	-0.23	0.33	-0.18	-0.17
VF489863	6800	0.38	0.26	-0.26	0.04	0.26	-0.12
VF489867	6800	0.46	0.03	-0.18	0.02	0.01	0.03
VF312738	6800	0.70	0.41	0.41	-0.19	-0.22	-0.23
VF312739	6800	0.55	0.35	-0.09	-0.20	-0.18	0.35
VF312760	6800	0.67	0.34	0.34	-0.20	-0.20	-0.14
VF312767	6800	0.53	0.19	0.19	-0.06	-0.06	-0.16
VF312769	6800	0.71	0.50	-0.28	0.50	-0.25	-0.23
VF489895	6800	0.54	0.43	0.43	-0.23	-0.24	-0.17
VF489889	6800	0.66	0.44	-0.17	-0.22	-0.26	0.44
VF489888	6800	0.51	0.37	0.37	-0.20	-0.12	-0.24
VF489891	6800	0.55	0.32	-0.19	0.32	-0.26	-0.10
VF489892	6800	0.55	0.26	-0.04	-0.16	0.26	-0.19
VF489896	6800	0.55	0.55	-0.29	0.55	-0.26	-0.21
VF489894	6800	0.71	0.32	-0.21	-0.24	-0.13	0.32
VF329860	6800	0.62	0.39	-0.32	0.39	-0.18	-0.08
VF331459	6800	0.75	0.44	0.44	-0.31	-0.23	-0.19
VF332097	6800	0.80	0.44	-0.26	-0.18	-0.31	0.44
VF333313	6800	0.39	0.20	0.03	0.20	-0.14	-0.23
VF339080	6800	0.38	0.22	-0.03	0.22	-0.16	-0.07
VF340193	6800	0.59	0.39	-0.20	-0.18	0.39	-0.24
VF341575	6800	0.72	0.49	-0.24	-0.27	-0.26	0.49
VF312483	6800	0.96	0.25	-0.17	-0.14	0.25	-0.11
VF312485	6800	0.81	0.39	-0.12	-0.25	0.39	-0.25
VF312486	6800	0.90	0.33	0.33	-0.27	-0.12	-0.13
VF312490	6800	0.70	0.44	0.44	-0.23	-0.14	-0.30
VF312491	6800	0.81	0.46	-0.22	-0.23	0.46	-0.28
VF312492	6800	0.91	0.35	-0.19	0.35	-0.20	-0.17
VF312493	6800	0.73	0.36	-0.19	-0.19	-0.23	0.36
VF312786	6800	0.73	0.53	-0.30	-0.32	-0.22	0.53
VF313231	6800	0.68	0.48	0.48	-0.26	-0.24	-0.22
VF312868	6800	0.58	0.30	0.30	-0.18	-0.10	-0.20
VF312785	6800	0.43	0.30	-0.16	-0.20	0.30	-0.17
VF312870	6800	0.76	0.35	-0.13	0.35	-0.11	-0.28
VF313235	6800	0.47	0.36	0.36	0.02	-0.24	-0.28
VF312920	6800	0.83	0.38	0.38	-0.21	-0.26	-0.15

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF489872	6800	0.64	0.36	-0.16	-0.15	0.36	-0.21
VF489875	6800	0.65	0.49	-0.25	-0.18	-0.29	0.49
VF489873	6800	0.58	0.34	-0.28	0.34	-0.16	-0.10
VF489876	6800	0.67	0.44	0.44	-0.27	-0.19	-0.19
VF489878	6800	0.76	0.44	-0.23	-0.25	-0.22	0.44
VF489880	6800	0.48	0.40	-0.18	0.40	-0.21	-0.13
VF489882	6800	0.59	0.25	-0.17	-0.07	-0.22	0.25
VF489835	6800	0.67	0.54	0.54	-0.35	-0.26	-0.20
VF489828	6800	0.52	0.30	-0.02	0.30	-0.30	-0.22
VF489158	6800	0.76	0.45	-0.30	-0.16	0.45	-0.23
VF489833	6800	0.82	0.51	-0.32	-0.27	-0.23	0.51
VF489837	6800	0.66	0.29	0.29	-0.20	-0.09	-0.20
VF489829	6800	0.56	0.36	-0.30	-0.09	0.36	-0.22
VF489152	6800	0.70	0.56	-0.27	0.56	-0.30	-0.29

Table G5. Reading Grade 7 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF355076	6827	0.59	0.43	-0.23	-0.29	0.43	-0.12
VF355080	6827	0.39	0.24	-0.03	0.24	-0.18	-0.19
VF355087	6827	0.73	0.40	-0.20	-0.27	0.40	-0.23
VF355089	6827	0.75	0.43	0.43	-0.19	-0.20	-0.29
VF355091	6827	0.78	0.48	-0.28	-0.26	0.48	-0.22
VF355660	6827	0.86	0.43	-0.23	-0.23	-0.24	0.43
VF486685	6827	0.75	0.42	-0.26	-0.20	-0.19	0.42
VF486690	6827	0.68	0.31	-0.17	-0.18	0.31	-0.11
VF355682	6827	0.59	0.38	0.38	-0.28	-0.20	-0.12
VF355686	6827	0.54	0.33	0.33	-0.25	-0.03	-0.28
VF355689	6827	0.71	0.47	-0.23	0.47	-0.27	-0.24
VF355691	6827	0.56	0.43	-0.16	-0.21	-0.25	0.43
VF355700	6827	0.67	0.39	-0.21	0.39	-0.21	-0.21
VF355701	6827	0.58	0.45	-0.28	-0.20	0.45	-0.14
VF486676	6827	0.65	0.47	0.47	-0.28	-0.24	-0.20
VF486168	6827	0.40	0.28	-0.09	-0.07	-0.21	0.28
VF486680	6827	0.51	0.36	-0.20	-0.14	-0.15	0.36
VF486679	6827	0.33	0.23	0.23	-0.23	0.04	-0.20
VF486482	6827	0.51	0.32	-0.23	-0.15	0.32	-0.10
VF486674	6827	0.56	0.48	-0.20	-0.32	0.48	-0.18
VF486682	6827	0.67	0.35	-0.23	-0.27	0.35	-0.11
VF486684	6827	0.55	0.41	-0.22	0.41	-0.25	-0.05
VF356105	6827	0.65	0.53	-0.32	-0.26	-0.19	0.53
VF356106	6827	0.89	0.44	0.44	-0.26	-0.22	-0.23
VF356110	6827	0.57	0.48	-0.20	-0.25	0.48	-0.25
VF497511	6827	0.84	0.45	-0.32	-0.14	-0.24	0.45
VF356116	6827	0.67	0.43	-0.16	-0.17	0.43	-0.30
VF356129	6827	0.65	0.40	-0.15	-0.26	-0.20	0.40
VF355106	6827	0.73	0.29	-0.16	0.29	-0.23	-0.08
VF355116	6827	0.62	0.37	-0.22	-0.12	-0.23	0.37
VF355119	6827	0.42	0.23	0.23	-0.21	-0.21	0.01
VF355127	6827	0.58	0.27	-0.20	-0.12	0.27	-0.08
VF355132	6827	0.61	0.41	0.41	-0.17	-0.26	-0.15
VF355137	6827	0.61	0.43	-0.20	0.43	-0.22	-0.22
VF486129	6827	0.79	0.46	-0.25	-0.31	-0.20	0.46
VF486131	6827	0.26	0.15	-0.08	0.15	-0.17	0.01
VF486148	6827	0.55	0.35	0.35	-0.18	-0.19	-0.18
VF486143	6827	0.72	0.25	-0.31	-0.26	0.04	0.25
VF486152	6827	0.75	0.34	0.34	-0.26	-0.12	-0.21
VF486150	6827	0.57	0.07	-0.21	-0.11	0.07	0.07
VF486142	6827	0.75	0.48	-0.24	-0.25	-0.25	0.48
VF486155	6827	0.64	0.37	-0.20	0.37	-0.30	-0.09

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF355722	6827	0.73	0.42	-0.20	-0.25	-0.21	0.42
VF355750	6827	0.56	0.46	-0.20	-0.20	-0.23	0.46
VF355735	6827	0.58	0.42	-0.20	0.42	-0.21	-0.18
VF355740	6827	0.30	0.32	0.32	-0.08	-0.13	-0.15
VF355747	6827	0.54	0.47	-0.22	-0.24	0.47	-0.19
VF355725	6827	0.55	0.47	-0.21	-0.19	-0.25	0.47
VF485019	6827	0.58	0.16	0.16	-0.23	0.03	-0.13
VF485552	6827	0.78	0.47	0.47	-0.28	-0.25	-0.21
VF485087	6827	0.50	0.34	-0.13	0.34	-0.18	-0.19
VF485016	6827	0.65	0.54	-0.22	-0.33	0.54	-0.28
VF485010	6827	0.54	0.35	-0.21	-0.20	0.35	-0.10
VF485021	6827	0.56	0.26	-0.24	0.02	-0.21	0.26
VF485049	6827	0.58	0.42	-0.21	-0.30	0.42	-0.08
VF485550	6827	0.60	0.44	-0.19	0.44	-0.24	-0.24

Table G6. Reading Grade 8 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF385503	6774	0.60	0.41	-0.31	-0.12	0.41	-0.15
VF385505	6774	0.63	0.20	-0.03	0.20	-0.13	-0.21
VF385507	6774	0.50	0.33	-0.18	-0.11	0.33	-0.21
VF385509	6774	0.60	0.34	-0.22	0.34	-0.19	-0.18
VF385510	6774	0.70	0.22	-0.04	-0.18	0.22	-0.13
VF385512	6774	0.87	0.31	-0.15	-0.20	-0.16	0.31
VF486947	6774	0.44	0.17	-0.22	-0.09	0.01	0.17
VF486950	6774	0.76	0.37	0.37	-0.22	-0.19	-0.20
VF385516	6774	0.51	0.18	-0.11	-0.20	0.18	-0.03
VF385518	6774	0.48	0.39	-0.16	-0.14	0.39	-0.21
VF385520	6774	0.78	0.21	0.21	-0.19	-0.16	-0.03
VF385521	6774	0.59	0.40	-0.18	-0.21	-0.18	0.40
VF385523	6774	0.52	0.33	0.33	-0.16	-0.21	-0.16
VF385525	6774	0.72	0.46	-0.29	0.46	-0.25	-0.19
VF486918	6774	0.61	0.38	-0.16	-0.25	0.38	-0.17
VF486942	6774	0.74	0.40	-0.31	0.40	-0.28	-0.09
VF486924	6774	0.57	0.47	-0.21	-0.19	-0.28	0.47
VF486901	6774	0.78	0.47	-0.24	0.47	-0.26	-0.24
VF486910	6774	0.63	0.34	-0.20	-0.28	0.34	-0.09
VF486939	6774	0.67	0.51	-0.30	-0.25	-0.22	0.51
VF486862	6774	0.75	0.43	0.43	-0.25	-0.21	-0.20
VF486870	6774	0.65	0.50	-0.26	-0.21	0.50	-0.26
VF486851	6774	0.64	0.53	0.53	-0.28	-0.27	-0.23
VF486874	6774	0.53	0.44	0.44	-0.20	-0.27	-0.13
VF486880	6774	0.61	0.51	-0.19	0.51	-0.31	-0.23
VF486864	6774	0.55	0.38	-0.18	-0.28	0.38	-0.21
VF486855	6774	0.72	0.50	-0.27	-0.30	-0.22	0.50
VF497522	6774	0.75	0.43	-0.23	-0.22	-0.23	0.43
VF497524	6774	0.61	0.39	-0.29	-0.09	-0.16	0.39
VF385826	6774	0.57	0.40	-0.28	-0.15	0.40	-0.11
VF385832	6774	0.73	0.38	-0.19	0.38	-0.26	-0.17
VF385836	6774	0.67	0.45	-0.25	-0.26	0.45	-0.20
VF385838	6774	0.48	0.37	-0.21	-0.12	-0.18	0.37
VF497529	6774	0.72	0.43	-0.23	-0.29	-0.15	0.43
VF497537	6774	0.62	0.38	-0.23	0.38	-0.20	-0.15
VF385812	6774	0.58	0.36	-0.23	-0.12	0.36	-0.18
VF385815	6774	0.55	0.48	-0.15	-0.27	-0.28	0.48
VF385818	6774	0.62	0.32	0.32	-0.06	-0.11	-0.29
VF497541	6774	0.71	0.34	-0.11	-0.24	0.34	-0.17
VF385819	6774	0.64	0.48	-0.29	-0.27	-0.22	0.48
VF385822	6774	0.53	0.25	-0.11	0.25	-0.07	-0.17
VF385824	6774	0.49	0.41	0.41	-0.23	-0.19	-0.16

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF486810	6774	0.61	0.33	0.33	-0.33	-0.10	-0.06
VF486811	6774	0.57	0.37	-0.25	-0.07	0.37	-0.22
VF486780	6774	0.67	0.37	0.37	-0.15	-0.26	-0.18
VF486775	6774	0.50	0.38	0.38	-0.15	-0.18	-0.22
VF486773	6774	0.54	0.50	-0.26	-0.17	0.50	-0.27
VF486769	6774	0.41	0.33	-0.26	-0.23	-0.04	0.33
VF496073	6774	0.64	0.36	0.36	-0.32	-0.23	-0.06
VF496102	6774	0.72	0.48	-0.31	0.48	-0.27	-0.15
VF496114	6774	0.62	0.28	-0.14	-0.14	0.28	-0.21
VF496116	6774	0.59	0.31	-0.16	-0.19	-0.12	0.31
VF496120	6774	0.57	0.36	-0.23	0.36	-0.12	-0.15
VF496123	6774	0.75	0.47	-0.26	-0.28	0.47	-0.18
VF496126	6774	0.78	0.37	0.37	-0.22	-0.21	-0.16
VF496129	6774	0.75	0.38	-0.20	-0.24	0.38	-0.20

Mathematics

Table G7. Mathematics Grade 3 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF479807	7133	0.90	0.34	0.34	-0.20	-0.20	-0.18
VF384789	7133	0.84	0.39	-0.25	0.39	-0.20	-0.20
VF384855	7133	0.83	0.40	-0.16	-0.31	0.40	-0.14
VF384871	7133	0.85	0.41	0.41	-0.22	-0.22	-0.23
VF385277	7133	0.61	0.36	-0.13	-0.22	0.36	-0.22
VF384863	7133	0.69	0.45	-0.19	-0.18	-0.32	0.45
VF384884	7133	0.49	0.44	-0.08	-0.10	0.44	-0.36
VF479864	7133	0.76	0.34	-0.30	-0.15	0.34	-0.14
VF385367	7133	0.87	0.43	-0.23	0.43	-0.16	-0.31
VF479799	7133	0.53	0.48	-0.22	-0.26	0.48	-0.20
VF479852	7133	0.76	0.32	0.32	-0.17	-0.21	-0.12
VF385226	7133	0.81	0.38	-0.18	0.38	-0.13	-0.28
VF384901	7133	0.55	0.50	0.50	-0.13	-0.21	-0.36
VF384937	7133	0.77	0.50	-0.27	0.50	-0.22	-0.30
VF479882	7133	0.45	0.19	0.04	0.19	-0.17	-0.10
VF384949	7133	0.50	0.38	-0.23	-0.05	0.38	-0.22
VF479846	7133	0.87	0.37	-0.25	0.37	-0.18	-0.17
VF384954	7133	0.81	0.30	-0.25	-0.13	0.30	-0.10
VF384935	7133	0.84	0.35	0.35	-0.17	-0.18	-0.22
VF384958	7133	0.59	0.47	0.47	-0.22	-0.27	-0.17
VF384957	7133	0.86	0.37	-0.17	0.37	-0.25	-0.17
VF384961	7133	0.92	0.21	-0.08	0.21	-0.16	-0.08
VF384952	7133	0.50	0.47	0.47	-0.06	-0.38	-0.24
VF385203	7133	0.95	0.28	0.28	-0.19	-0.13	-0.14
VF384963	7133	0.88	0.38	-0.24	-0.19	-0.19	0.38
VF384967	7133	0.76	0.44	-0.28	-0.20	0.44	-0.24
VF385033	7133	0.74	0.37	0.37	-0.26	-0.10	-0.21
VF479804	7133	0.85	0.34	-0.25	0.34	-0.11	-0.20
VF479743	7133	0.63	0.56	0.56	-0.20	-0.35	-0.32
VF479802	7133	0.49	0.17	-0.14	-0.15	0.00	0.17
VF385064	7133	0.65	0.39	-0.20	-0.22	-0.19	0.39
VF385059	7133	0.72	0.51	-0.37	0.51	-0.18	-0.23
VF385206	7133	0.58	0.47	-0.17	-0.23	-0.27	0.47
VF385051	7133	0.68	0.44	-0.28	-0.18	0.44	-0.19
VF385068	7133	0.62	0.43	0.43	-0.19	-0.30	-0.16
VF479764	7133	0.73	0.44	-0.15	-0.13	0.44	-0.38
VF385432	7133	0.82	0.51	0.51	-0.26	-0.28	-0.31
VF479910	7133	0.62	0.46	-0.31	-0.17	-0.17	0.46
VF479746	7133	0.67	0.29	0.29	-0.15	-0.14	-0.16
VF385117	7133	0.56	0.40	-0.30	-0.17	0.40	-0.09
VF479868	7133	0.83	0.38	0.38	-0.18	-0.25	-0.18

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF479872	7133	0.91	0.37	-0.17	-0.25	-0.19	0.37
VF385124	7133	0.73	0.50	-0.45	0.50	-0.09	-0.17
VF385186	7133	0.60	0.47	-0.24	-0.23	-0.22	0.47
VF385128	7133	0.86	0.52	-0.32	-0.24	-0.30	0.52
VF486889	7133	0.34	0.43	0.43	0.02	-0.15	-0.32
VF480086	7133	0.81	0.42	-0.14	0.42	-0.18	-0.34
VF385179	7133	0.72	0.45	-0.23	-0.24	0.45	-0.26
VF480094	7133	0.61	0.51	-0.07	0.51	-0.31	-0.35
VF479911	7133	0.67	0.55	0.55	-0.18	-0.27	-0.38
VF479791	7133	0.77	0.51	-0.32	-0.26	0.51	-0.23
VF479876	7133	0.54	0.55	-0.43	-0.23	-0.10	0.55
VF479703	7133	0.84	0.26	-0.17	-0.16	0.26	-0.14
VF479899	7133	0.93	0.30	-0.19	-0.19	-0.13	0.30
VF479903	7133	0.69	0.40	-0.15	-0.29	0.40	-0.17
VF479908	7133	0.72	0.49	-0.25	-0.29	0.49	-0.22
VF385037	7133	0.75	0.40	0.40	-0.31	-0.19	-0.08
VF480056	7133	0.84	0.48	-0.34	-0.23	0.48	-0.18
VF479857	7133	0.63	0.49	-0.19	0.49	-0.23	-0.31
VF479891	7133	0.64	0.36	-0.21	-0.19	-0.12	0.36

Table G8. Mathematics Grade 4 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF385385	7163	0.86	0.40	-0.27	-0.23	0.40	-0.15
VF311036	7163	0.60	0.42	0.42	-0.29	-0.19	-0.18
VF483381	7163	0.76	0.49	-0.18	0.49	-0.36	-0.22
VF384731	7163	0.76	0.48	0.48	-0.23	-0.20	-0.33
VF384718	7163	0.49	0.47	-0.47	0.47	-0.01	-0.02
VF311041	7163	0.81	0.44	-0.35	0.44	-0.15	-0.19
VF384713	7163	0.78	0.49	-0.38	-0.21	-0.19	0.49
VF384703	7163	0.73	0.53	0.53	-0.25	-0.26	-0.35
VF384728	7163	0.38	0.48	-0.44	0.48	0.03	-0.10
VF385111	7163	0.86	0.41	-0.29	-0.20	0.41	-0.17
VF480680	7163	0.67	0.56	-0.29	-0.32	-0.24	0.56
VF483385	7163	0.59	0.51	0.51	-0.26	-0.25	-0.22
VF385247	7163	0.76	0.46	-0.30	-0.16	0.46	-0.26
VF385218	7163	0.54	0.40	-0.26	-0.24	0.40	-0.04
VF483430	7163	0.90	0.42	-0.20	-0.27	-0.24	0.42
VF384786	7163	0.84	0.31	0.31	-0.21	-0.16	-0.13
VF384744	7163	0.76	0.41	-0.23	-0.20	-0.23	0.41
VF384762	7163	0.71	0.39	-0.30	0.39	-0.20	-0.12
VF384824	7163	0.82	0.37	-0.27	-0.17	0.37	-0.21
VF385389	7163	0.76	0.43	0.43	-0.28	-0.23	-0.18
VF384777	7163	0.65	0.40	0.40	-0.21	-0.21	-0.20
VF483464	7163	0.68	0.33	-0.25	0.33	-0.17	-0.14
VF384845	7163	0.76	0.33	0.33	-0.20	-0.17	-0.18
VF384805	7163	0.88	0.40	-0.20	-0.19	-0.28	0.40
VF384815	7163	0.56	0.31	0.31	-0.19	-0.14	-0.16
VF480281	7163	0.63	0.30	-0.17	-0.20	0.30	-0.07
VF311057	7163	0.58	0.53	0.53	-0.39	-0.18	-0.15
VF385170	7163	0.86	0.43	-0.27	-0.21	-0.21	0.43
VF385229	7163	0.76	0.47	-0.32	-0.14	0.47	-0.25
VF384892	7163	0.44	0.38	-0.25	-0.18	-0.11	0.38
VF480145	7163	0.55	0.51	-0.39	-0.25	-0.10	0.51
VF384877	7163	0.69	0.43	-0.26	-0.22	0.43	-0.21
VF384923	7163	0.69	0.30	-0.12	0.30	-0.19	-0.16
VF384881	7163	0.59	0.41	-0.21	-0.25	0.41	-0.16
VF384943	7163	0.62	0.37	-0.20	-0.17	-0.17	0.37
VF384942	7163	0.45	0.52	-0.32	-0.35	0.52	0.06
VF384933	7163	0.66	0.49	-0.32	-0.21	-0.20	0.49
VF385234	7163	0.68	0.48	0.48	-0.27	-0.26	-0.20
VF480115	7163	0.82	0.34	0.34	-0.21	-0.23	-0.16
VF484939	7163	0.62	0.44	-0.16	-0.18	0.44	-0.34
VF483395	7163	0.71	0.53	-0.23	0.53	-0.32	-0.26
VF484394	7163	0.50	0.33	-0.07	0.33	-0.19	-0.19

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF480129	7163	0.71	0.42	-0.19	-0.23	-0.24	0.42
VF385072	7163	0.72	0.38	-0.20	-0.25	0.38	-0.14
VF311061	7163	0.71	0.49	-0.32	-0.26	-0.18	0.49
VF385208	7163	0.62	0.49	-0.36	-0.17	0.49	-0.20
VF480679	7163	0.93	0.31	-0.10	-0.25	0.31	-0.14
VF483387	7163	0.85	0.50	0.50	-0.29	-0.24	-0.29
VF384953	7163	0.52	0.45	-0.20	0.45	-0.27	-0.14
VF495781	7163	0.47	0.32	-0.12	0.32	-0.16	-0.14
VF483426	7163	0.68	0.42	0.42	-0.25	-0.31	-0.15
VF385083	7163	0.88	0.34	0.34	-0.25	-0.18	-0.14
VF483422	7163	0.82	0.47	-0.18	-0.25	-0.33	0.47
VF385080	7163	0.74	0.32	0.32	-0.17	-0.22	-0.12
VF385088	7163	0.65	0.43	-0.12	-0.33	0.43	-0.16
VF384951	7163	0.65	0.41	-0.27	-0.21	0.41	-0.16
VF480162	7163	0.86	0.53	-0.34	-0.32	-0.22	0.53
VF480133	7163	0.69	0.51	-0.27	0.51	-0.30	-0.23
VF385397	7163	0.75	0.41	-0.34	-0.12	-0.15	0.41
VF483932	7163	0.36	0.29	-0.12	-0.13	0.29	-0.11
VF480107	7163	0.67	0.34	-0.15	0.34	-0.15	-0.24
VF483392	7163	0.83	0.43	-0.24	-0.29	-0.20	0.43
VF385405	7163	0.26	0.09	-0.02	-0.07	0.01	0.09
VF480124	7163	0.82	0.43	-0.30	0.43	-0.24	-0.12
VF483394	7163	0.75	0.49	0.49	-0.25	-0.25	-0.27

Table G9. Mathematics Grade 5 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF486165	6766	0.75	0.41	-0.20	-0.27	0.41	-0.18
VF385502	6766	0.67	0.50	-0.18	0.50	-0.23	-0.34
VF385506	6766	0.74	0.48	0.48	-0.14	-0.33	-0.29
VF385585	6766	0.75	0.55	-0.49	-0.15	-0.16	0.55
VF387070	6766	0.88	0.36	-0.17	0.36	-0.15	-0.27
VF387090	6766	0.47	0.55	-0.37	-0.18	-0.15	0.55
VF387039	6766	0.72	0.54	-0.14	0.54	-0.41	-0.26
VF385612	6766	0.76	0.46	-0.32	-0.25	0.46	-0.16
VF385630	6766	0.66	0.52	-0.23	-0.29	-0.28	0.52
VF385616	6766	0.69	0.58	0.58	-0.30	-0.40	-0.19
VF387078	6766	0.60	0.37	-0.18	-0.20	0.37	-0.20
VF486672	6766	0.66	0.26	0.26	-0.24	-0.09	-0.20
VF387017	6766	0.72	0.50	0.50	-0.22	-0.24	-0.32
VF387041	6766	0.91	0.33	-0.21	0.33	-0.20	-0.14
VF385700	6766	0.80	0.39	-0.21	-0.27	0.39	-0.12
VF385698	6766	0.48	0.38	0.38	-0.10	-0.26	-0.16
VF386701	6766	0.48	0.34	-0.16	-0.24	0.34	-0.12
VF386751	6766	0.69	0.34	-0.24	0.34	-0.16	-0.18
VF386781	6766	0.66	0.59	-0.18	-0.46	-0.20	0.59
VF386650	6766	0.47	0.38	-0.21	-0.27	0.38	-0.10
VF386821	6766	0.32	0.41	-0.12	-0.22	-0.13	0.41
VF386842	6766	0.57	0.44	-0.19	-0.34	0.44	-0.04
VF386775	6766	0.74	0.42	-0.26	-0.15	0.42	-0.24
VF386885	6766	0.64	0.29	-0.16	-0.22	-0.17	0.29
VF386880	6766	0.54	0.50	-0.39	0.50	-0.12	-0.18
VF386913	6766	0.59	0.44	0.44	-0.10	-0.38	-0.10
VF386833	6766	0.63	0.31	-0.16	-0.16	-0.14	0.31
VF386892	6766	0.74	0.53	-0.29	0.53	-0.22	-0.32
VF386983	6766	0.85	0.42	0.42	-0.23	-0.24	-0.26
VF486620	6766	0.77	0.29	-0.15	-0.16	0.29	-0.15
VF485873	6766	0.62	0.37	0.37	-0.16	-0.24	-0.15
VF387074	6766	0.44	0.51	-0.29	-0.25	0.51	-0.06
VF484945	6766	0.45	0.48	-0.27	-0.28	0.48	-0.07
VF387091	6766	0.54	0.31	-0.18	-0.13	-0.14	0.31
VF485046	6766	0.60	0.37	-0.16	0.37	-0.28	-0.18
VF386920	6766	0.38	0.41	-0.24	-0.20	0.41	-0.21
VF484996	6766	0.69	0.47	-0.24	-0.23	0.47	-0.27
VF486162	6766	0.45	0.46	0.46	-0.14	-0.28	-0.16
VF486134	6766	0.62	0.51	-0.29	-0.26	-0.22	0.51
VF386940	6766	0.63	0.50	-0.32	-0.14	-0.28	0.50
VF386946	6766	0.56	0.46	-0.25	0.46	-0.17	-0.23
VF386980	6766	0.62	0.47	-0.39	-0.17	0.47	-0.05

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF386956	6766	0.68	0.49	-0.32	0.49	-0.16	-0.25
VF386971	6766	0.77	0.36	-0.36	-0.05	0.36	-0.08
VF486688	6766	0.56	0.43	-0.21	0.43	-0.26	-0.16
VF486159	6766	0.55	0.35	-0.26	0.35	-0.12	-0.12
VF485003	6766	0.56	0.48	-0.18	-0.25	-0.26	0.48
VF387076	6766	0.65	0.45	-0.15	-0.26	-0.25	0.45
VF486133	6766	0.76	0.38	-0.20	0.38	-0.22	-0.21
VF387004	6766	0.36	0.48	-0.28	-0.23	-0.14	0.48
VF485000	6766	0.46	0.54	-0.37	-0.16	-0.15	0.54
VF485007	6766	0.59	0.42	0.42	-0.27	-0.18	-0.19
VF485017	6766	0.66	0.51	-0.45	-0.13	0.51	-0.12
VF485053	6766	0.66	0.47	-0.26	0.47	-0.23	-0.21
VF484986	6766	0.74	0.47	0.47	-0.23	-0.30	-0.25
VF486128	6766	0.71	0.36	-0.32	-0.17	0.36	-0.07
VF486156	6766	0.69	0.45	-0.33	0.45	-0.21	-0.11
VF484990	6766	0.51	0.39	-0.23	-0.31	0.39	0.05
VF387065	6766	0.58	0.45	-0.15	-0.33	0.45	-0.18
VF387094	6766	0.48	0.32	0.32	-0.10	-0.16	-0.17
VF387087	6766	0.60	0.32	-0.13	0.32	-0.16	-0.19
VF387037	6766	0.78	0.49	-0.17	-0.31	-0.28	0.49
VF387068	6766	0.72	0.45	0.45	-0.24	-0.24	-0.20
VF485020	6766	0.64	0.33	-0.20	-0.16	0.33	-0.14
VF386878	6766	0.70	0.45	-0.18	-0.22	-0.31	0.45

Table G10. Mathematics Grade 6 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF388057	6795	0.70	0.34	-0.16	0.34	-0.26	-0.19
VF387210	6795	0.82	0.41	0.41	-0.11	-0.19	-0.34
VF387199	6795	0.69	0.37	-0.23	-0.19	0.37	-0.16
VF387232	6795	0.76	0.44	-0.25	0.44	-0.27	-0.17
VF387205	6795	0.52	0.33	-0.15	-0.23	-0.05	0.33
VF387878	6795	0.72	0.53	-0.37	-0.27	0.53	-0.17
VF387920	6795	0.78	0.44	-0.18	0.44	-0.26	-0.26
VF479699	6795	0.48	0.36	-0.22	-0.24	-0.15	0.36
VF387214	6795	0.50	0.45	-0.31	-0.15	0.45	-0.15
VF387889	6795	0.76	0.50	0.50	-0.32	-0.22	-0.26
VF387817	6795	0.56	0.49	-0.16	-0.38	-0.15	0.49
VF479745	6795	0.72	0.52	0.52	-0.24	-0.28	-0.28
VF479812	6795	0.86	0.31	-0.13	-0.17	-0.22	0.31
VF387238	6795	0.69	0.53	-0.26	-0.25	0.53	-0.30
VF387250	6795	0.68	0.28	-0.19	0.28	-0.17	-0.07
VF387254	6795	0.54	0.43	-0.21	-0.25	0.43	-0.15
VF387247	6795	0.68	0.33	-0.15	-0.22	0.33	-0.15
VF387360	6795	0.67	0.48	-0.33	0.48	-0.19	-0.21
VF387261	6795	0.72	0.51	0.51	-0.37	-0.20	-0.19
VF479756	6795	0.73	0.46	-0.24	0.46	-0.24	-0.23
VF387349	6795	0.77	0.40	0.40	-0.32	-0.18	-0.10
VF387336	6795	0.52	0.45	-0.24	0.45	-0.14	-0.25
VF387363	6795	0.70	0.29	-0.15	-0.25	0.29	-0.10
VF479749	6795	0.58	0.31	-0.04	0.31	-0.25	-0.21
VF387882	6795	0.55	0.34	-0.12	-0.23	-0.14	0.34
VF387234	6795	0.61	0.40	0.40	-0.20	-0.26	-0.18
VF479803	6795	0.48	0.35	-0.08	0.35	-0.22	-0.16
VF387375	6795	0.38	0.34	0.34	-0.16	-0.20	-0.15
VF479809	6795	0.32	0.38	-0.31	-0.22	-0.01	0.38
VF387393	6795	0.63	0.44	-0.34	0.44	-0.20	-0.08
VF479744	6795	0.61	0.37	-0.14	-0.27	0.37	-0.17
VF387386	6795	0.59	0.49	-0.26	-0.33	0.49	-0.07
VF479713	6795	0.76	0.44	0.44	-0.31	-0.19	-0.19
VF387476	6795	0.81	0.31	-0.12	-0.14	-0.22	0.31
VF387505	6795	0.55	0.55	-0.35	0.55	-0.31	-0.10
VF387480	6795	0.69	0.31	-0.20	-0.11	-0.20	0.31
VF387483	6795	0.60	0.50	-0.29	-0.27	0.50	-0.15
VF479796	6795	0.79	0.32	-0.24	0.32	-0.16	-0.11
VF387486	6795	0.58	0.37	-0.21	-0.21	0.37	-0.15
VF387885	6795	0.89	0.34	-0.20	-0.23	0.34	-0.15
VF479814	6795	0.79	0.45	-0.24	-0.27	-0.21	0.45
VF387509	6795	0.70	0.48	-0.31	-0.28	-0.18	0.48

VF387730	6795	0.59	0.37	-0.37	0.37	-0.05	-0.05
VF479763	6795	0.50	0.40	-0.21	-0.17	0.40	-0.17
VF387936	6795	0.56	0.39	-0.32	-0.22	0.39	-0.05
VF479805	6795	0.67	0.50	-0.25	0.50	-0.32	-0.20
VF387711	6795	0.63	0.44	-0.19	-0.27	0.44	-0.18
VF387332	6795	0.65	0.48	-0.41	0.48	-0.12	-0.13
VF479792	6795	0.60	0.32	-0.22	-0.14	0.32	-0.11
VF387790	6795	0.69	0.48	0.48	-0.17	-0.36	-0.22
VF479707	6795	0.77	0.47	-0.29	-0.22	-0.23	0.47
VF479790	6795	0.67	0.46	-0.19	-0.24	-0.26	0.46
VF479794	6795	0.65	0.47	-0.28	-0.17	0.47	-0.26
VF387787	6795	0.75	0.50	-0.30	-0.14	-0.32	0.50
VF387801	6795	0.55	0.50	0.50	-0.16	-0.37	-0.16
VF479787	6795	0.50	0.39	-0.19	0.39	-0.23	-0.13
VF479788	6795	0.70	0.53	-0.41	-0.22	-0.13	0.53
VF479785	6795	0.37	0.44	0.44	-0.29	0.00	-0.24
VF479752	6795	0.46	0.42	-0.20	-0.16	-0.18	0.42
VF479808	6795	0.45	0.26	-0.16	-0.09	0.26	-0.10
VF388085	6795	0.31	0.40	0.40	-0.13	-0.22	-0.14
VF479798	6795	0.68	0.42	-0.20	-0.16	-0.28	0.42
VF388087	6795	0.60	0.42	0.42	-0.24	-0.18	-0.31
VF388077	6795	0.87	0.43	-0.27	-0.26	0.43	-0.18
VF479806	6795	0.72	0.41	-0.23	-0.22	-0.21	0.41

Table G11. Mathematics Grade 7 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF479865	6825	0.73	0.51	-0.23	-0.34	-0.23	0.51
VF387798	6825	0.54	0.37	-0.11	0.37	-0.28	-0.11
VF479853	6825	0.48	0.43	-0.22	-0.18	-0.26	0.43
VF387755	6825	0.35	0.41	0.41	-0.21	-0.26	0.02
VF387914	6825	0.43	0.41	-0.40	0.41	-0.12	0.04
VF387673	6825	0.52	0.32	-0.06	-0.27	0.32	-0.07
VF388037	6825	0.52	0.39	-0.14	0.39	-0.21	-0.23
VF388095	6825	0.74	0.31	0.31	-0.24	-0.17	-0.12
VF479890	6825	0.55	0.41	-0.17	-0.31	0.41	-0.10
VF479858	6825	0.40	0.28	-0.19	0.28	-0.11	-0.04
VF387746	6825	0.75	0.36	-0.19	-0.19	0.36	-0.23
VF479860	6825	0.63	0.38	-0.20	-0.19	-0.19	0.38
VF388115	6825	0.68	0.38	-0.18	-0.16	-0.25	0.38
VF388118	6825	0.65	0.34	0.34	-0.07	-0.26	-0.23
VF388105	6825	0.53	0.35	-0.28	-0.10	0.35	-0.06
VF479862	6825	0.48	0.44	-0.13	-0.23	-0.24	0.44
VF388121	6825	0.67	0.45	-0.38	-0.13	0.45	-0.14
VF388127	6825	0.71	0.45	0.45	-0.27	-0.25	-0.18
VF388130	6825	0.60	0.34	0.34	-0.16	-0.19	-0.17
VF388523	6825	0.47	0.38	-0.16	-0.15	-0.19	0.38
VF387744	6825	0.39	0.34	-0.22	-0.04	0.34	-0.16
VF388337	6825	0.57	0.54	-0.21	-0.35	-0.21	0.54
VF388357	6825	0.77	0.37	0.37	-0.19	-0.23	-0.16
VF388528	6825	0.44	0.36	-0.08	-0.05	-0.32	0.36
VF387723	6825	0.48	0.28	-0.21	-0.17	0.28	0.01
VF388485	6825	0.45	0.38	-0.27	-0.08	-0.15	0.38
VF388414	6825	0.59	0.37	-0.16	0.37	-0.21	-0.16
VF387716	6825	0.47	0.30	-0.12	0.30	-0.28	-0.05
VF388432	6825	0.65	0.51	-0.25	-0.27	0.51	-0.26
VF479895	6825	0.83	0.34	-0.22	0.34	-0.18	-0.15
VF479880	6825	0.46	0.40	-0.06	0.40	-0.19	-0.25
VF479863	6825	0.46	0.54	-0.31	-0.29	-0.11	0.54
VF479866	6825	0.47	0.36	-0.22	-0.16	0.36	-0.11
VF479893	6825	0.68	0.45	-0.25	-0.20	-0.22	0.45
VF479850	6825	0.37	0.45	0.00	-0.25	-0.35	0.45
VF388371	6825	0.44	0.42	-0.07	0.42	-0.31	-0.23
VF388499	6825	0.59	0.32	-0.13	-0.11	-0.22	0.32
VF388361	6825	0.81	0.44	-0.28	-0.22	0.44	-0.20
VF388223	6825	0.44	0.41	-0.17	0.41	-0.23	-0.16
VF388380	6825	0.72	0.32	-0.21	0.32	-0.17	-0.11
VF479856	6825	0.50	0.27	-0.23	-0.04	-0.27	0.27
VF388375	6825	0.45	0.42	0.42	-0.17	-0.17	-0.20

VF388378	6825	0.46	0.36	0.36	-0.22	-0.22	-0.05
VF479896	6825	0.71	0.44	-0.31	0.44	-0.21	-0.17
VF388353	6825	0.47	0.52	-0.39	-0.21	0.52	-0.10
VF479888	6825	0.56	0.40	0.40	-0.23	-0.21	-0.18
VF479875	6825	0.65	0.48	-0.23	-0.31	0.48	-0.14
VF479884	6825	0.37	0.39	-0.16	-0.10	-0.19	0.39
VF479887	6825	0.77	0.41	0.41	-0.30	-0.19	-0.13
VF388340	6825	0.38	0.47	-0.23	-0.23	0.47	-0.14
VF479894	6825	0.65	0.30	-0.22	0.30	-0.15	-0.10
VF388219	6825	0.51	0.46	-0.29	-0.31	0.46	0.05
VF387788	6825	0.41	0.49	-0.34	-0.22	-0.19	0.49
VF388132	6825	0.88	0.40	0.40	-0.27	-0.22	-0.18
VF479901	6825	0.61	0.54	-0.40	-0.21	-0.18	0.54
VF479845	6825	0.66	0.38	-0.23	0.38	-0.18	-0.16
VF479849	6825	0.39	0.44	-0.27	-0.18	0.44	-0.10
VF479854	6825	0.44	0.41	-0.20	0.41	-0.21	-0.13
VF479873	6825	0.35	0.47	-0.15	-0.22	-0.17	0.47
VF479886	6825	0.68	0.46	-0.16	-0.27	0.46	-0.25
VF387690	6825	0.54	0.47	-0.23	0.47	-0.25	-0.17
VF479870	6825	0.72	0.32	-0.10	-0.25	-0.12	0.32
VF479892	6825	0.64	0.46	-0.24	0.46	-0.24	-0.22
VF479885	6825	0.62	0.58	-0.30	-0.27	-0.31	0.58
VF479900	6825	0.31	0.46	0.46	-0.27	-0.26	-0.16
VF479881	6825	0.77	0.26	-0.14	0.26	-0.11	-0.16

Table G12. Mathematics Grade 8 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF479710	6779	0.81	0.41	-0.18	-0.27	-0.21	0.41
VF388106	6779	0.68	0.31	0.31	-0.22	-0.10	-0.16
VF393236	6779	0.52	0.21	-0.18	-0.09	0.21	-0.13
VF393263	6779	0.53	0.39	-0.23	-0.14	0.39	-0.17
VF388119	6779	0.47	0.32	0.32	-0.15	-0.13	-0.18
VF388116	6779	0.51	0.31	-0.14	0.31	-0.19	-0.15
VF479869	6779	0.57	0.42	-0.19	-0.22	0.42	-0.19
VF388596	6779	0.42	0.52	-0.15	-0.20	-0.30	0.52
VF393265	6779	0.31	0.46	-0.30	-0.16	0.46	-0.02
VF479797	6779	0.59	0.42	-0.29	0.42	-0.22	-0.11
VF388600	6779	0.56	0.39	0.39	-0.17	-0.20	-0.20
VF388739	6779	0.66	0.32	-0.10	-0.20	0.32	-0.18
VF393247	6779	0.51	0.36	0.36	-0.19	-0.18	-0.13
VF479871	6779	0.45	0.28	-0.18	0.28	-0.23	0.06
VF390039	6779	0.57	0.36	-0.22	-0.18	0.36	-0.12
VF388619	6779	0.58	0.40	-0.15	-0.25	0.40	-0.19
VF388747	6779	0.59	0.41	0.41	-0.25	-0.12	-0.22
VF479848	6779	0.32	0.39	0.39	-0.22	-0.20	-0.01
VF479824	6779	0.39	0.33	-0.12	-0.13	0.33	-0.18
VF479772	6779	0.79	0.38	0.38	-0.24	-0.20	-0.16
VF390817	6779	0.43	0.41	-0.21	0.41	-0.10	-0.22
VF390175	6779	0.51	0.43	-0.16	-0.21	-0.21	0.43
VF390818	6779	0.68	0.45	-0.21	0.45	-0.27	-0.19
VF479859	6779	0.39	0.40	0.40	-0.27	-0.13	-0.06
VF479855	6779	0.75	0.38	-0.22	-0.24	0.38	-0.16
VF390215	6779	0.32	0.32	-0.16	0.32	-0.14	-0.02
VF390631	6779	0.65	0.50	-0.20	-0.26	-0.27	0.50
VF390240	6779	0.49	0.39	-0.12	-0.25	0.39	-0.13
VF388910	6779	0.60	0.32	-0.18	0.32	-0.13	-0.22
VF388913	6779	0.61	0.33	-0.18	-0.13	-0.19	0.33
VF479672	6779	0.76	0.31	-0.20	0.31	-0.13	-0.15
VF479789	6779	0.35	0.47	-0.35	-0.13	0.47	0.00
VF391184	6779	0.57	0.56	0.56	-0.41	-0.20	-0.14
VF388915	6779	0.39	0.35	-0.14	-0.12	-0.16	0.35
VF388919	6779	0.33	0.52	-0.27	-0.05	-0.24	0.52
VF479689	6779	0.45	0.41	-0.12	-0.20	0.41	-0.20
VF388916	6779	0.50	0.36	0.36	-0.21	-0.22	-0.12
VF479840	6779	0.77	0.39	-0.18	-0.24	0.39	-0.19
VF393290	6779	0.34	0.38	-0.02	-0.17	-0.28	0.38
VF388921	6779	0.56	0.47	-0.19	-0.26	-0.23	0.47
VF391888	6779	0.69	0.36	-0.24	0.36	-0.21	-0.06
VF479813	6779	0.63	0.39	0.39	-0.14	-0.29	-0.12

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF479762	6779	0.45	0.50	-0.18	-0.29	-0.17	0.50
VF392991	6779	0.77	0.46	-0.28	0.46	-0.28	-0.13
VF393288	6779	0.33	0.46	-0.28	-0.19	-0.07	0.46
VF391178	6779	0.42	0.48	0.48	-0.26	-0.27	-0.03
VF393274	6779	0.38	0.26	-0.02	-0.27	-0.08	0.26
VF479723	6779	0.60	0.39	-0.15	-0.25	0.39	-0.16
VF479761	6779	0.70	0.48	-0.29	0.48	-0.23	-0.18
VF479681	6779	0.45	0.54	-0.23	-0.30	-0.15	0.54
VF393006	6779	0.50	0.34	-0.28	-0.19	0.34	0.00
VF479665	6779	0.47	0.44	-0.09	0.44	-0.33	-0.12
VF393138	6779	0.77	0.45	-0.13	-0.18	-0.35	0.45
VF393242	6779	0.52	0.43	0.43	-0.11	-0.24	-0.23
VF393148	6779	0.61	0.49	0.49	-0.26	-0.26	-0.18
VF479842	6779	0.66	0.46	-0.20	-0.21	0.46	-0.28
VF393276	6779	0.47	0.49	0.49	-0.18	-0.24	-0.21
VF393257	6779	0.46	0.33	-0.10	0.33	-0.12	-0.20
VF393268	6779	0.51	0.43	-0.25	0.43	-0.19	-0.13
VF393211	6779	0.48	0.54	0.54	-0.24	-0.28	-0.18
VF479867	6779	0.73	0.52	-0.39	-0.22	-0.18	0.52
VF479879	6779	0.41	0.40	0.40	-0.22	-0.06	-0.22
VF479780	6779	0.60	0.40	-0.27	0.40	-0.14	-0.17
VF479827	6779	0.76	0.47	-0.24	-0.25	-0.23	0.47
VF479777	6779	0.61	0.46	-0.15	0.46	-0.23	-0.28
VF392976	6779	0.36	0.28	-0.16	-0.01	0.28	-0.16
VF393245	6779	0.60	0.46	-0.17	-0.25	0.46	-0.24
VF393293	6779	0.40	0.36	0.36	-0.13	-0.18	-0.11
VF479820	6779	0.55	0.45	-0.20	0.45	-0.22	-0.21
VF388742	6779	0.63	0.51	-0.25	0.51	-0.30	-0.22

Science

Table G13. Science Grade 4 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF483448	7157	0.89	0.36	-0.19	0.36	-0.22	-0.18
VF484396	7157	0.55	0.24	-0.11	-0.14	0.24	-0.11
VF269828	7157	0.86	0.35	-0.15	-0.25	0.35	-0.21
VF269830	7157	0.56	0.27	-0.16	-0.10	-0.13	0.27
VF269831	7157	0.55	0.35	0.35	-0.13	-0.25	-0.19
VF484935	7157	0.40	0.17	0.17	-0.02	-0.09	-0.13
VF290032	7157	0.85	0.47	-0.31	-0.22	-0.22	0.47
VF290777	7157	0.82	0.41	-0.27	0.41	-0.18	-0.25
VF269869	7157	0.69	0.37	-0.21	0.37	-0.18	-0.18
VF269871	7157	0.47	0.39	-0.15	-0.21	0.39	-0.15
VF269873	7157	0.69	0.43	-0.15	-0.28	-0.29	0.43
VF282620	7157	0.51	0.39	0.39	-0.19	-0.21	-0.13
VF282623	7157	0.56	0.34	-0.17	0.34	-0.12	-0.21
VF282626	7157	0.78	0.50	-0.27	-0.30	0.50	-0.23
VF282637	7157	0.66	0.38	-0.31	-0.10	-0.24	0.38
VF282639	7157	0.71	0.49	-0.27	0.49	-0.29	-0.20
VF282642	7157	0.77	0.33	-0.13	-0.20	0.33	-0.19
VF282661	7157	0.64	0.41	-0.23	-0.23	-0.18	0.41
VF282669	7157	0.54	0.42	-0.17	0.42	-0.24	-0.16
VF282670	7157	0.85	0.41	-0.20	-0.26	0.41	-0.23
VF282688	7157	0.44	0.38	-0.11	0.38	-0.15	-0.22
VF283022	7157	0.69	0.49	-0.22	0.49	-0.29	-0.24
VF283606	7157	0.85	0.38	-0.25	0.38	-0.17	-0.20
VF284002	7157	0.60	0.37	-0.24	-0.24	-0.14	0.37
VF284006	7157	0.58	0.47	-0.24	-0.24	0.47	-0.20
VF284665	7157	0.47	0.38	-0.15	-0.19	-0.16	0.38
VF284979	7157	0.51	0.14	-0.04	-0.07	0.14	-0.10
VF293507	7157	0.81	0.45	-0.19	0.45	-0.20	-0.31
VF292879	7157	0.71	0.37	-0.12	-0.22	0.37	-0.22
VF294472	7157	0.65	0.42	0.42	-0.18	-0.20	-0.25
VF293853	7157	0.41	0.24	-0.17	0.02	-0.17	0.24
VF483424	7157	0.69	0.44	0.44	-0.25	-0.21	-0.20
VF483437	7157	0.94	0.23	-0.11	0.23	-0.17	-0.10
VF483431	7157	0.74	0.49	-0.27	-0.23	0.49	-0.29
VF287740	7157	0.48	0.36	-0.13	-0.20	0.36	-0.15
VF287742	7157	0.49	0.37	0.37	-0.18	-0.15	-0.16
VF287745	7157	0.50	0.29	-0.15	-0.17	0.29	-0.09
VF287765	7157	0.46	0.34	-0.24	-0.18	-0.08	0.34
VF287766	7157	0.69	0.45	0.45	-0.24	-0.19	-0.25
VF287770	7157	0.91	0.23	-0.14	-0.13	-0.13	0.23
VF269857	7157	0.43	0.31	-0.23	0.31	-0.21	0.02

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF269846	7157	0.53	0.34	-0.17	-0.10	0.34	-0.21
VF269841	7157	0.46	0.23	-0.18	-0.11	-0.04	0.23
VF296839	7157	0.48	0.35	-0.14	-0.20	0.35	-0.11
VF296833	7157	0.56	0.40	-0.25	-0.19	0.40	-0.11
VF287864	7157	0.69	0.40	-0.18	-0.26	0.40	-0.17
VF287849	7157	0.72	0.45	-0.18	-0.28	0.45	-0.28
VF287846	7157	0.63	0.40	-0.15	0.40	-0.26	-0.20
VF287870	7157	0.50	0.36	-0.21	-0.20	-0.08	0.36
VF287856	7157	0.53	0.38	0.38	-0.17	-0.21	-0.14

Table G14. Science Grade 8 Classical Statistics for Operational Items

Accession Number	N	Mean	Item-Test Corr	Option Discrimination (MC)			
				A	B	C	D
VF484958	6754	0.56	0.43	-0.16	-0.20	-0.26	0.43
VF484974	6754	0.46	0.27	0.27	-0.14	-0.16	-0.06
VF486678	6754	0.29	0.31	-0.25	0.00	-0.08	0.31
VF486675	6754	0.52	0.33	-0.19	0.33	-0.23	-0.05
VF308274	6754	0.47	0.32	-0.13	-0.10	0.32	-0.23
VF308369	6754	0.48	0.43	-0.13	0.43	-0.31	-0.16
VF308379	6754	0.64	0.41	-0.17	-0.22	0.41	-0.24
VF308708	6754	0.59	0.28	0.28	-0.26	-0.16	-0.05
VF308712	6754	0.44	0.50	-0.10	-0.32	-0.25	0.50
VF308713	6754	0.48	0.32	-0.21	-0.04	0.32	-0.22
VF484993	6754	0.46	0.42	0.42	-0.20	-0.11	-0.23
VF484999	6754	0.56	0.29	-0.14	-0.08	-0.22	0.29
VF308868	6754	0.85	0.40	-0.23	-0.20	-0.22	0.40
VF308869	6754	0.58	0.28	-0.03	0.28	-0.21	-0.18
VF308871	6754	0.54	0.41	-0.18	-0.20	-0.19	0.41
VF486699	6754	0.74	0.31	-0.13	-0.21	0.31	-0.14
VF486698	6754	0.70	0.39	0.39	-0.18	-0.22	-0.19
VF486135	6754	0.58	0.46	0.46	-0.25	-0.24	-0.18
VF486126	6754	0.52	0.37	-0.19	0.37	-0.18	-0.18
VF308893	6754	0.60	0.45	-0.20	0.45	-0.22	-0.26
VF308894	6754	0.43	0.29	-0.25	-0.16	0.29	-0.08
VF308895	6754	0.72	0.44	-0.24	0.44	-0.23	-0.21
VF308899	6754	0.53	0.38	0.38	-0.16	-0.18	-0.17
VF308900	6754	0.53	0.48	0.48	-0.21	-0.25	-0.19
VF308901	6754	0.44	0.29	0.29	-0.21	-0.02	-0.16
VF308906	6754	0.62	0.50	-0.22	-0.20	-0.30	0.50
VF308909	6754	0.28	0.29	-0.02	0.29	-0.16	-0.12
VF308910	6754	0.54	0.37	0.37	-0.19	-0.24	-0.10
VF308928	6754	0.32	0.33	-0.10	-0.14	-0.13	0.33
VF308929	6754	0.63	0.46	-0.28	-0.24	0.46	-0.20
VF308930	6754	0.56	0.44	0.44	-0.22	-0.21	-0.19
VF486847	6754	0.66	0.39	-0.16	-0.27	0.39	-0.16
VF486858	6754	0.36	0.38	0.38	-0.23	-0.17	-0.04
VF486815	6754	0.62	0.46	-0.16	0.46	-0.27	-0.24
VF486821	6754	0.62	0.45	-0.20	0.45	-0.22	-0.23
VF486149	6754	0.42	0.42	-0.24	-0.24	0.42	-0.10
VF486151	6754	0.64	0.35	-0.13	-0.14	0.35	-0.23
VF486146	6754	0.54	0.35	0.35	-0.20	-0.16	-0.14
VF486166	6754	0.61	0.29	-0.22	-0.11	0.29	-0.13
VF486163	6754	0.60	0.46	-0.24	-0.23	-0.20	0.46
VF486771	6754	0.58	0.37	-0.21	-0.23	0.37	-0.14
VF486782	6754	0.58	0.40	-0.25	-0.16	0.40	-0.21

Accession Number	N	Mean	Item-Test Corr	<u>Option Discrimination (MC)</u>			
				A	B	C	D
VF486765	6754	0.59	0.49	-0.21	-0.21	-0.28	0.49
VF486687	6754	0.61	0.46	0.46	-0.23	-0.24	-0.19
VF486692	6754	0.75	0.41	-0.26	-0.20	0.41	-0.19
VF486914	6754	0.36	0.32	-0.02	-0.18	-0.22	0.32
VF486941	6754	0.47	0.40	-0.15	0.40	-0.20	-0.16
VF486948	6754	0.37	0.38	-0.09	-0.24	-0.14	0.38
VF485018	6754	0.59	0.33	0.33	-0.17	-0.20	-0.10
VF485023	6754	0.49	0.34	-0.03	-0.24	-0.22	0.34

SAWS

Table G15. Distributions of Rating 1 Scores for SAWS

Score	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	268	3.68	202	2.77	193	2.79	198	2.86	235	3.37	218	3.15
1	91	1.25	29	0.40	12	0.17	30	0.43	23	0.33	19	0.27
2	77	1.06	35	0.48	24	0.35	57	0.82	21	0.30	11	0.16
3	175	2.40	144	1.98	41	0.59	72	1.04	31	0.44	13	0.19
4	1186	16.26	1112	15.27	779	11.28	1253	18.09	691	9.91	621	8.98
5	780	10.70	888	12.19	612	8.86	675	9.75	489	7.01	418	6.05
6	878	12.04	1016	13.95	813	11.77	772	11.15	744	10.67	613	8.87
7	996	13.66	1202	16.51	1057	15.30	915	13.21	932	13.37	817	11.82
8	1261	17.29	1104	15.16	1427	20.65	1621	23.40	1724	24.73	1846	26.70
9	443	6.08	550	7.55	479	6.93	331	4.78	415	5.95	430	6.22
10	380	5.21	382	5.25	409	5.92	275	3.97	408	5.85	442	6.39
11	370	5.07	274	3.76	458	6.63	291	4.20	449	6.44	486	7.03
12	387	5.31	344	4.72	605	8.76	436	6.30	810	11.62	980	14.17
N	7292		7282		6909		6926		6972		6914	
MEAN	6.62		6.68		7.35		6.77		7.55		7.84	
STD	2.79		2.56		2.69		2.65		2.8		2.81	

Table G16. Distributions of Trait Scores for SAWS - Grade 3

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	300	4.11	404	5.54	448	6.14	598	8.20
1	2324	31.87	2539	34.82	2812	38.56	2991	41.02
2	3483	47.76	3290	45.12	3092	42.40	2814	38.59
3	1185	16.25	1059	14.52	940	12.89	889	12.19
N	7292		7292		7292		7292	
MEAN	1.76		1.69		1.62		1.55	
SD	0.77		0.79		0.79		0.81	

Table G17. Distributions of Trait Scores for SAWS - Grade 4

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	215	2.95	537	7.37	252	3.46	268	3.68
1	2444	33.56	3523	48.38	2629	36.10	2591	35.58
2	3581	49.18	2111	28.99	3463	47.56	3366	46.22
3	1042	14.31	1111	15.26	938	12.88	1057	14.52
N	7282		7282		7282		7282	
MEAN	1.75		1.52		1.70		1.72	
SD	0.73		0.84		0.73		0.75	

Table G18. Distributions of Trait Scores for SAWS - Grade 5

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	215	3.11	221	3.20	231	3.34	262	3.79
1	1399	20.25	2055	29.74	2095	30.32	2436	35.26
2	3790	54.86	3208	46.43	3247	47.00	3155	45.67
3	1505	21.78	1425	20.63	1336	19.34	1056	15.28
N	6909		6909		6909		6909	
MEAN	1.95		1.84		1.82		1.72	
SD	0.74		0.78		0.77		0.76	

Table G19. Distributions of Trait Scores for SAWS - Grade 6

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	223	3.22	258	3.73	293	4.23	316	4.56
1	2422	34.97	2359	34.06	2387	34.46	2748	39.68
2	3337	48.18	3376	48.74	3287	47.46	3096	44.70
3	944	13.63	933	13.47	959	13.85	766	11.06
N	6926		6926		6926		6926	
MEAN	1.72		1.72		1.71		1.62	
SD	0.73		0.74		0.75		0.74	

Table G20. Distributions of Trait Scores for SAWS - Grade 7

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	249	3.57	268	3.84	278	3.99	306	4.39
1	1484	21.29	1829	26.23	1585	22.73	2060	29.55
2	3578	51.32	3393	48.67	3484	49.97	3369	48.32
3	1661	23.82	1482	21.26	1625	23.31	1237	17.74
N	6972		6972		6972		6972	
MEAN	1.95		1.87		1.93		1.79	
SD	0.77		0.78		0.78		0.78	

Table G21. Distributions of Trait Scores for SAWS - Grade 8

Score	<u>Idea Development</u>		<u>Organization</u>		<u>Voice</u>		<u>Conventions</u>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	233	3.37	239	3.46	244	3.53	257	3.72
1	1210	17.50	1494	21.61	1419	20.52	1857	26.86
2	3532	51.08	3387	48.99	3616	52.30	3339	48.29
3	1939	28.04	1794	25.95	1635	23.65	1461	21.13
N	6914		6914		6914		6914	
MEAN	2.04		1.97		1.96		1.87	
SD	0.77		0.78		0.76		0.78	

Appendix H: Scale Score Descriptive Statistics by Demographic Subgroup

Table H1. Summary Statistics of Reading, Mathematics, and Science Scale Score by Grade

Grade	N	Mean	SD
Reading			
3	7141	602.8	52.7
4	7158	675.2	55.2
5	6772	670.5	48.5
6	6800	689.1	43.5
7	6827	697.2	44.4
8	6774	706.9	44.9
Mathematics			
3	7133	663.4	56.6
4	7163	668.1	55.7
5	6766	692.2	56.5
6	6795	712.2	54.2
7	6825	726.0	50.2
8	6779	734.4	50.4
Science			
4	7157	673.1	44.6
8	6754	651.6	45.6

Table H2. Summary Statistics of Reading Grade 3 Scale Score

Group	N	Mean	SD
Total	7141	602.8	52.7
Male	3703	597.9	52.6
Female	3428	608.1	52.2
Unknown	10	590.9	56.4
American Indian	274	562.9	47.8
Asian	67	621.6	47.0
Black	90	583.9	46.0
Hawaiian/Pacific Islander	15	589.5	64.9
Hispanic	989	579.8	48.3
White	5557	608.9	51.7
Multiracial	128	609.1	50.8
Unknown	21	589.3	52.7
Free/Reduced Lunch	4472	611.9	52.3
Not Free/Reduced Lunch	2669	587.5	49.8
Special Education	6118	609.3	49.9
Not Special Education	1023	564.0	51.9
English Language Learner	6792	605.0	52.1
Not English Language Learner	349	558.7	43.5

Table H3. Summary Statistics of Reading Grade 4 Scale Score

Group	N	Mean	SD
Total	7158	675.2	55.2
Male	3687	669.4	55.3
Female	3461	681.3	54.3
Unknown	10	648.3	58.6
American Indian	278	637.0	49.0
Asian	67	685.0	62.8
Black	81	658.8	62.4
Hawaiian/Pacific Islander	11	659.7	47.7
Hispanic	936	656.5	51.9
White	5651	680.4	54.4
Multiracial	110	669.5	52.2
Unknown	24	680.0	53.0
Free/Reduced Lunch	4474	683.9	54.9
Not Free/Reduced Lunch	2684	660.6	52.5
Special Education	6143	682.4	52.2
Not Special Education	1015	631.2	52.3
English Language Learner	6963	676.7	54.6
Not English Language Learner	195	621.9	46.4

Table H4. Summary Statistics of Reading Grade 5 Scale Score

Group	N	Mean	SD
Total	6772	670.5	48.5
Male	3498	667.6	48.9
Female	3266	673.8	47.8
Unknown	8	640.4	47.7
American Indian	264	642.3	47.3
Asian	46	685.7	42.2
Black	79	643.3	48.6
Hawaiian/Pacific Islander	17	640.8	44.6
Hispanic	893	654.2	45.2
White	5358	675.2	47.8
Multiracial	90	660.0	52.8
Unknown	25	664.7	49.0
Free/Reduced Lunch	4301	677.7	48.6
Not Free/Reduced Lunch	2471	658.0	45.6
Special Education	5838	677.3	45.5
Not Special Education	934	627.9	44.9
English Language Learner	6608	672.0	47.9
Not English Language Learner	164	611.4	32.7

Table H5. Summary Statistics of Reading Grade 6 Scale Score

Group	N	Mean	SD
Total	6800	689.1	43.5
Male	3483	682.9	44.5
Female	3313	695.5	41.5
Unknown	4	683.0	33.6
American Indian	241	656.3	35.3
Asian	53	703.5	48.5
Black	85	669.0	43.9
Hawaiian/Pacific Islander	17	662.2	41.7
Hispanic	889	677.3	41.2
White	5391	692.8	42.9
Multiracial	103	691.9	47.6
Unknown	21	662.5	48.7
Free/Reduced Lunch	4370	695.2	43.4
Not Free/Reduced Lunch	2430	678.1	41.5
Special Education	5944	695.3	40.6
Not Special Education	856	646.1	38.4
English Language Learner	6651	690.1	43.1
Not English Language Learner	149	643.6	34.1

Table H6. Summary Statistics of Reading Grade 7 Scale Score

Group	N	Mean	SD
Total	6827	697.2	44.4
Male	3577	693.0	45.5
Female	3242	701.8	42.7
Unknown	8	697.0	53.6
American Indian	221	672.0	40.6
Asian	51	707.4	46.7
Black	106	675.6	45.2
Hawaiian/Pacific Islander	15	685.9	61.9
Hispanic	873	681.4	40.7
White	5435	701.3	43.9
Multiracial	117	690.7	46.3
Unknown	9	646.0	22.6
Free/Reduced Lunch	4517	704.2	44.3
Not Free/Reduced Lunch	2310	683.4	41.4
Special Education	5958	703.4	41.6
Not Special Education	869	654.8	39.7
English Language Learner	6674	698.3	44.1
Not English Language Learner	153	650.0	30.7

Table H7. Summary Statistics of Reading Grade 8 Scale Score

Group	N	Mean	SD
Total	6774	706.9	44.9
Male	3528	700.6	44.4
Female	3240	713.8	44.4
Unknown	6	696.0	49.3
American Indian	219	674.3	42.3
Asian	59	719.5	41.7
Black	77	696.1	39.3
Hawaiian/Pacific Islander	10	667.9	43.7
Hispanic	831	691.9	42.4
White	5469	710.7	44.4
Multiracial	94	704.9	46.5
Unknown	15	690.4	29.6
Free/Reduced Lunch	4509	713.3	44.3
Not Free/Reduced Lunch	2265	694.2	43.4
Special Education	5950	712.9	42.4
Not Special Education	824	664.1	38.6
English Language Learner	6647	707.8	44.5
Not English Language Learner	127	659.1	41.3

Table H8. Summary Statistics of Mathematics Grade 3 Scale Score

Group	N	Mean	SD
Total	7133	663.4	56.6
Male	3696	663.9	57.5
Female	3431	662.9	55.6
Unknown	6	585.8	83.0
American Indian	271	624.5	51.2
Asian	66	679.5	63.0
Black	89	643.5	51.9
Hawaiian/Pacific Islander	15	644.4	56.5
Hispanic	990	640.4	51.0
White	5562	669.6	56.0
Multiracial	132	662.2	54.3
Unknown	8	626.3	51.7
Free/Reduced Lunch	4368	672.5	57.0
Not Free/Reduced Lunch	2765	649.0	52.9
Special Education	6080	669.0	55.0
Not Special Education	1053	630.5	54.5
English Language Learner	6777	665.5	56.3
Not English Language Learner	356	622.5	47.8

Table H9. Summary Statistics of Mathematics Grade 4 Scale Score

Group	N	Mean	SD
Total	7163	668.1	55.7
Male	3686	669.4	56.2
Female	3461	666.7	55.1
Unknown	16	642.6	57.6
American Indian	280	633.6	56.2
Asian	73	679.2	58.7
Black	81	645.6	58.9
Hawaiian/Pacific Islander	13	659.8	69.4
Hispanic	928	648.8	52.7
White	5661	673.2	54.8
Multiracial	109	666.3	52.0
Unknown	18	659.1	49.7
Free/Reduced Lunch	4553	675.4	56.0
Not Free/Reduced Lunch	2610	655.2	52.8
Special Education	6182	673.6	54.0
Not Special Education	981	633.3	53.7
English Language Learner	6969	669.5	55.3
Not English Language Learner	194	615.8	45.9

Table H10. Summary Statistics of Mathematics Grade 5 Scale Score

Group	N	Mean	SD
Total	6766	692.2	56.5
Male	3494	693.2	55.9
Female	3265	691.3	57.0
Unknown	7	640.1	44.5
American Indian	263	668.2	58.8
Asian	47	721.0	61.9
Black	79	667.6	55.9
Hawaiian/Pacific Islander	16	657.9	52.8
Hispanic	890	673.2	52.6
White	5360	697.2	55.7
Multiracial	93	677.9	54.1
Unknown	18	645.5	64.0
Free/Reduced Lunch	4277	699.4	57.2
Not Free/Reduced Lunch	2489	679.9	52.9
Special Education	5843	698.6	54.9
Not Special Education	923	652.1	49.5
English Language Learner	6598	693.5	56.2
Not English Language Learner	168	642.1	41.2

Table H11. Summary Statistics of Mathematics Grade 6 Scale Score

Group	N	Mean	SD
Total	6795	712.2	54.2
Male	3475	712.0	55.4
Female	3314	712.4	53.0
Unknown	6	711.3	61.2
American Indian	243	670.4	41.8
Asian	54	741.0	62.3
Black	84	684.9	48.3
Hawaiian/Pacific Islander	17	681.8	39.5
Hispanic	891	695.6	49.3
White	5385	717.1	53.8
Multiracial	103	713.5	63.8
Unknown	18	694.9	51.9
Free/Reduced Lunch	4373	719.4	55.2
Not Free/Reduced Lunch	2422	699.3	49.9
Special Education	5957	718.4	52.6
Not Special Education	838	668.2	44.6
English Language Learner	6640	713.4	54.0
Not English Language Learner	155	663.0	36.6

Table H12. Summary Statistics of Mathematics Grade 7 Scale Score

Group	N	Mean	SD
Total	6825	726.0	50.2
Male	3579	727.7	51.5
Female	3242	724.1	48.5
Unknown	4	674.8	43.1
American Indian	219	696.2	37.4
Asian	52	746.7	47.1
Black	105	699.0	45.2
Hawaiian/Pacific Islander	15	716.2	59.6
Hispanic	874	708.5	43.3
White	5429	730.7	50.6
Multiracial	119	717.0	46.0
Unknown	12	685.4	35.4
Free/Reduced Lunch	4507	733.8	51.4
Not Free/Reduced Lunch	2318	710.9	44.0
Special Education	5969	731.7	49.0
Not Special Education	856	686.1	38.9
English Language Learner	6672	727.0	50.1
Not English Language Learner	153	682.9	31.8

Table H13. Summary Statistics of Mathematics Grade 8 Scale Score

Group	N	Mean	SD
Total	6779	734.4	50.4
Male	3532	734.2	50.6
Female	3241	734.8	50.1
Unknown	6	694.7	28.6
American Indian	217	700.8	38.1
Asian	59	759.1	56.7
Black	78	709.9	40.1
Hawaiian/Pacific Islander	11	693.8	29.7
Hispanic	833	719.5	44.2
White	5471	738.4	50.7
Multiracial	96	730.2	49.0
Unknown	14	696.1	30.5
Free/Reduced Lunch	4494	741.0	51.4
Not Free/Reduced Lunch	2285	721.4	45.4
Special Education	5952	740.1	49.2
Not Special Education	827	693.6	38.3
English Language Learner	6647	735.2	50.3
Not English Language Learner	132	695.5	37.0

Table H14. Summary Statistics of Science Grade 4 Scale Score

Group	N	Mean	SD
Total	7157	673.1	44.6
Male	3683	671.9	44.8
Female	3459	674.3	44.4
Unknown	15	665.0	43.7
American Indian	275	644.0	38.5
Asian	71	677.5	45.4
Black	82	652.3	55.1
Hawaiian/Pacific Islander	14	656.2	42.6
Hispanic	935	657.4	39.6
White	5652	677.5	44.3
Multiracial	109	672.0	45.3
Unknown	19	648.6	49.1
Free/Reduced Lunch	4479	679.8	44.8
Not Free/Reduced Lunch	2678	661.8	42.1
Special Education	6157	676.5	44.3
Not Special Education	1000	651.5	40.7
English Language Learner	6958	674.1	44.4
Not English Language Learner	199	636.3	34.5

Table H15. Summary Statistics of Science Grade 8 Scale Score

Group	N	Mean	SD
Total	6754	651.6	45.6
Male	3523	651.2	46.3
Female	3225	652.2	44.8
Unknown	6	607.8	26.9
American Indian	213	618.0	36.2
Asian	58	669.6	48.1
Black	77	626.3	34.1
Hawaiian/Pacific Islander	11	608.8	34.1
Hispanic	825	634.6	39.9
White	5456	655.9	45.5
Multiracial	102	645.5	50.7
Unknown	12	612.5	27.7
Free/Reduced Lunch	4439	658.4	46.4
Not Free/Reduced Lunch	2315	638.5	41.1
Special Education	5918	656.4	44.6
Not Special Education	836	617.3	36.9
English Language Learner	6621	652.4	45.5
Not English Language Learner	133	613.5	34.6

Table H16. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 3

Group	N	Percent	Prompt Total		Idea Development		Organization		Voice		Conventions	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	7292	100	6.62	2.79	1.76	0.77	1.69	0.79	1.62	0.79	1.55	0.81
Male	3786	52	6.04	2.65	1.62	0.74	1.54	0.76	1.47	0.75	1.41	0.77
Female	3480	48	7.25	2.79	1.92	0.77	1.85	0.78	1.79	0.78	1.70	0.82
Unknown	26	0	4.46	3.98	1.12	1.03	1.19	1.06	1.04	1.00	1.12	0.99
American Indian	283	4	5.30	2.63	1.47	0.72	1.37	0.76	1.26	0.70	1.20	0.76
Asian	65	1	7.82	2.81	2.03	0.79	1.95	0.74	1.92	0.82	1.91	0.79
Black	93	1	6.56	2.95	1.73	0.75	1.70	0.83	1.61	0.82	1.52	0.87
Hawaiian/Pacific Islander	16	0	6.06	3.34	1.50	0.82	1.69	0.87	1.50	1.10	1.38	0.89
Hispanic	1011	14	5.99	2.65	1.63	0.74	1.53	0.76	1.46	0.74	1.37	0.77
White	5646	77	6.80	2.77	1.80	0.76	1.73	0.78	1.67	0.78	1.60	0.81
Multiracial	135	2	6.28	3.00	1.61	0.86	1.65	0.81	1.53	0.84	1.49	0.85
Unknown	43	1	5.26	3.83	1.37	1.02	1.33	0.97	1.26	0.98	1.30	1.01
Free/Reduced Lunch	2745	38	6.01	2.72	1.61	0.75	1.54	0.77	1.47	0.77	1.39	0.80
Not Free or Reduced Lunch	4547	62	6.98	2.77	1.85	0.76	1.78	0.78	1.71	0.78	1.64	0.80
Special Education	1041	14	5.11	2.56	1.44	0.71	1.32	0.75	1.23	0.73	1.12	0.74
Not Special Education	6251	86	6.87	2.75	1.82	0.76	1.75	0.77	1.69	0.77	1.62	0.80
English Language Learner	353	5	5.33	2.39	1.51	0.67	1.37	0.72	1.29	0.68	1.15	0.67
Not English Language Learner	6939	95	6.68	2.79	1.77	0.77	1.70	0.78	1.64	0.79	1.57	0.81

Table H17. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 4

Group	N	Percent	Prompt Total		Idea Development		Organization		Voice		Conventions	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	7282	100	6.68	2.56	1.75	0.73	1.52	0.84	1.70	0.73	1.72	0.75
Male	3742	51	6.20	2.43	1.63	0.71	1.40	0.80	1.58	0.70	1.59	0.73
Female	3515	48	7.21	2.58	1.88	0.73	1.65	0.86	1.83	0.74	1.85	0.75
Unknown	25	0	4.72	3.22	1.16	0.85	1.00	0.87	1.28	0.94	1.28	0.94
American Indian	285	4	5.29	2.27	1.38	0.63	1.20	0.73	1.34	0.66	1.38	0.70
Asian	73	1	7.10	2.83	1.79	0.74	1.60	0.92	1.81	0.79	1.89	0.83
Black	83	1	6.04	2.47	1.57	0.70	1.37	0.71	1.51	0.69	1.59	0.72
Hawaiian/Pacific Islander	12	0	7.00	2.49	1.75	0.75	1.58	0.79	1.75	0.75	1.92	0.67
Hispanic	952	13	6.23	2.44	1.65	0.70	1.42	0.82	1.57	0.70	1.60	0.72
White	5726	79	6.84	2.55	1.79	0.73	1.56	0.84	1.74	0.73	1.75	0.75
Multiracial	109	1	6.68	2.53	1.70	0.71	1.59	0.80	1.67	0.71	1.72	0.74
Unknown	42	1	5.31	3.25	1.29	0.83	1.14	0.87	1.36	0.93	1.52	0.99
Free/Reduced Lunch	2649	36	6.17	2.49	1.62	0.72	1.42	0.81	1.55	0.71	1.57	0.73
Not Free or Reduced Lunch	4633	64	6.98	2.55	1.82	0.73	1.58	0.85	1.78	0.73	1.80	0.75
Special Education	991	14	5.24	2.31	1.40	0.67	1.24	0.77	1.30	0.65	1.29	0.66
Not Special Education	6291	86	6.91	2.52	1.80	0.73	1.57	0.84	1.76	0.73	1.78	0.74
English Language Learner	185	3	5.23	2.24	1.35	0.61	1.22	0.72	1.30	0.63	1.36	0.69
Not English Language Learner	7097	97	6.72	2.56	1.76	0.73	1.53	0.84	1.71	0.73	1.72	0.75

Table H18. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 5

Group	N	Percent	Prompt Total		Idea Development		Organization		Voice		Conventions	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	6909	100	7.35	2.69	1.95	0.74	1.84	0.78	1.82	0.77	1.72	0.76
Male	3575	52	6.81	2.58	1.82	0.72	1.71	0.75	1.68	0.75	1.59	0.73
Female	3316	48	7.94	2.68	2.10	0.72	1.99	0.78	1.98	0.77	1.87	0.77
Unknown	18	0	4.89	3.53	1.39	0.98	1.33	0.97	1.11	0.90	1.06	0.87
American Indian	274	4	6.23	2.21	1.72	0.64	1.53	0.69	1.52	0.68	1.46	0.63
Asian	50	1	8.38	2.56	2.20	0.64	2.10	0.79	2.08	0.72	2.00	0.78
Black	81	1	6.33	2.64	1.68	0.70	1.59	0.74	1.49	0.74	1.57	0.79
Hawaiian/Pacific Islander	17	0	8.00	2.29	2.12	0.60	1.88	0.70	1.94	0.66	2.06	0.75
Hispanic	908	13	6.99	2.55	1.90	0.72	1.74	0.75	1.72	0.76	1.62	0.71
White	5451	79	7.48	2.71	1.98	0.74	1.88	0.78	1.86	0.77	1.76	0.77
Multiracial	92	1	7.15	2.72	1.97	0.72	1.80	0.76	1.75	0.79	1.63	0.75
Unknown	36	1	5.83	3.46	1.58	0.87	1.53	0.94	1.39	0.93	1.33	0.86
Free/Reduced Lunch	2503	36	6.94	2.61	1.86	0.72	1.75	0.76	1.72	0.76	1.62	0.74
Not Free or Reduced Lunch	4406	64	7.58	2.71	2.01	0.74	1.90	0.79	1.88	0.78	1.79	0.77
Special Education	945	14	5.69	2.36	1.59	0.70	1.44	0.70	1.39	0.68	1.27	0.67
Not Special Education	5964	86	7.61	2.65	2.01	0.73	1.91	0.77	1.89	0.77	1.80	0.75
English Language Learner	161	2	5.63	1.88	1.64	0.57	1.41	0.59	1.35	0.58	1.23	0.57
Not English Language Learner	6748	98	7.39	2.70	1.96	0.74	1.86	0.78	1.83	0.77	1.74	0.76

Table H19. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 6

Group	N	Percent	Prompt Total		Idea Development		Organization		Voice		Conventions	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	6926	100	6.77	2.65	1.72	0.73	1.72	0.74	1.71	0.75	1.62	0.74
Male	3557	51	6.14	2.50	1.55	0.70	1.56	0.70	1.54	0.73	1.48	0.71
Female	3353	48	7.46	2.63	1.91	0.72	1.89	0.74	1.89	0.74	1.78	0.74
Unknown	16	0	3.75	3.40	0.94	0.93	1.00	0.89	1.00	0.97	0.81	0.75
American Indian	247	4	5.39	2.45	1.38	0.69	1.36	0.68	1.35	0.73	1.30	0.68
Asian	57	1	7.25	2.86	1.79	0.77	1.81	0.81	1.82	0.76	1.82	0.78
Black	93	1	5.62	2.79	1.46	0.73	1.44	0.81	1.41	0.77	1.31	0.74
Hawaiian/Pacific Islander	16	0	6.19	2.34	1.69	0.70	1.50	0.73	1.50	0.73	1.50	0.63
Hispanic	912	13	6.32	2.53	1.62	0.71	1.62	0.71	1.58	0.71	1.50	0.71
White	5453	79	6.94	2.63	1.76	0.73	1.76	0.73	1.76	0.75	1.66	0.74
Multiracial	105	2	6.71	2.88	1.74	0.80	1.70	0.81	1.69	0.80	1.58	0.77
Unknown	43	1	5.70	3.43	1.47	0.93	1.49	0.91	1.42	0.98	1.33	0.87
Free/Reduced Lunch	2476	36	6.30	2.56	1.61	0.71	1.60	0.72	1.60	0.73	1.49	0.72
Not Free or Reduced Lunch	4450	64	7.04	2.67	1.78	0.74	1.79	0.74	1.77	0.76	1.69	0.74
Special Education	872	13	5.05	2.16	1.33	0.61	1.29	0.63	1.27	0.66	1.16	0.63
Not Special Education	6054	87	7.02	2.62	1.78	0.73	1.78	0.73	1.77	0.75	1.69	0.73
English Language Learner	153	2	5.35	2.05	1.43	0.59	1.34	0.59	1.36	0.64	1.22	0.63
Not English Language Learner	6773	98	6.81	2.66	1.73	0.73	1.73	0.74	1.72	0.75	1.63	0.74

Table H20. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 7

Group	N	Percent	Prompt Total		Idea Development		Organization		Voice		Conventions	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	6972	100	7.55	2.80	1.95	0.77	1.87	0.78	1.93	0.78	1.79	0.78
Male	3664	53	6.96	2.73	1.81	0.76	1.73	0.77	1.78	0.78	1.64	0.76
Female	3302	47	8.20	2.74	2.11	0.75	2.03	0.77	2.09	0.76	1.97	0.76
Unknown	6	0	6.50	1.97	1.67	0.52	1.50	0.55	1.67	0.52	1.67	0.52
American Indian	234	3	6.51	2.76	1.71	0.76	1.63	0.76	1.63	0.80	1.54	0.75
Asian	52	1	7.88	2.46	2.02	0.70	1.96	0.71	2.02	0.67	1.88	0.70
Black	110	2	6.71	2.97	1.78	0.84	1.61	0.83	1.80	0.81	1.52	0.80
Hawaiian/Pacific Islander	14	0	6.07	4.32	1.71	1.07	1.43	1.09	1.50	1.16	1.43	1.16
Hispanic	893	13	7.04	2.68	1.84	0.75	1.73	0.76	1.80	0.76	1.67	0.75
White	5532	79	7.70	2.80	1.99	0.77	1.91	0.78	1.96	0.78	1.83	0.78
Multiracial	117	2	7.48	2.74	1.91	0.73	1.79	0.76	1.93	0.75	1.84	0.77
Unknown	20	0	6.70	2.58	1.70	0.80	1.70	0.66	1.70	0.86	1.60	0.75
Free/Reduced Lunch	2344	34	6.94	2.73	1.80	0.76	1.72	0.77	1.78	0.77	1.64	0.76
Not Free or Reduced Lunch	4628	66	7.86	2.79	2.03	0.76	1.95	0.78	2.00	0.78	1.87	0.78
Special Education	895	13	5.48	2.54	1.48	0.74	1.38	0.72	1.38	0.73	1.24	0.70
Not Special Education	6077	87	7.85	2.71	2.02	0.75	1.95	0.76	2.01	0.76	1.88	0.76
English Language Learner	161	2	5.89	2.47	1.62	0.75	1.42	0.71	1.55	0.72	1.30	0.64
Not English Language Learner	6811	98	7.59	2.80	1.96	0.77	1.88	0.78	1.93	0.78	1.81	0.78

Table H21. Gender and Ethnicity Performance by SAWS Prompt and Trait –Grade 8

Group	N	Percent	Prompt Total		Idea Development		Organization		Voice		Conventions	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	6914	100	7.84	2.81	2.04	0.77	1.97	0.78	1.96	0.76	1.87	0.78
Male	3593	52	7.28	2.75	1.91	0.76	1.83	0.77	1.82	0.76	1.72	0.76
Female	3300	48	8.47	2.73	2.18	0.75	2.14	0.76	2.12	0.74	2.03	0.77
Unknown	21	0	5.05	4.08	1.29	1.06	1.24	1.09	1.24	1.04	1.29	1.10
American Indian	232	3	6.24	3.04	1.64	0.83	1.55	0.84	1.54	0.80	1.51	0.83
Asian	58	1	8.83	2.69	2.24	0.73	2.17	0.70	2.19	0.74	2.22	0.73
Black	79	1	7.10	2.98	1.84	0.84	1.78	0.78	1.77	0.78	1.71	0.77
Hawaiian/Pacific Islander	13	0	6.46	1.45	1.69	0.48	1.62	0.51	1.69	0.48	1.46	0.52
Hispanic	838	12	7.26	2.69	1.90	0.75	1.81	0.76	1.82	0.74	1.73	0.76
White	5569	81	8.01	2.77	2.08	0.76	2.02	0.77	2.00	0.75	1.90	0.78
Multiracial	94	1	8.02	2.81	2.06	0.76	2.01	0.77	1.99	0.75	1.96	0.76
Unknown	31	0	5.87	4.02	1.45	1.03	1.52	1.09	1.45	1.06	1.45	0.99
Free/Reduced Lunch	2302	33	7.20	2.86	1.88	0.79	1.81	0.79	1.81	0.78	1.71	0.78
Not Free or Reduced Lunch	4612	67	8.16	2.73	2.12	0.74	2.05	0.77	2.04	0.74	1.95	0.77
Special Education	839	12	5.81	2.55	1.56	0.75	1.47	0.73	1.47	0.73	1.30	0.68
Not Special Education	6075	88	8.12	2.73	2.10	0.75	2.04	0.76	2.03	0.74	1.95	0.76
English Language Learner	134	2	5.55	2.96	1.49	0.85	1.38	0.80	1.41	0.82	1.27	0.75
Not English Language Learner	6780	98	7.89	2.79	2.05	0.76	1.99	0.78	1.97	0.76	1.88	0.78

Appendix I: Raw Score to Scale Score Tables

Table II. Reading Grade 3 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-6.264	336	69	1
1	-5.556	370	49	1
2	-4.834	404	35	1
3	-4.399	425	29	1
4	-4.081	441	26	1
5	-3.828	453	23	1
6	-3.615	463	21	1
7	-3.429	472	20	1
8	-3.264	480	19	1
9	-3.114	487	18	1
10	-2.976	494	18	1
11	-2.847	500	17	1
12	-2.726	506	17	1
13	-2.612	512	16	1
14	-2.502	517	16	1
15	-2.397	522	15	2
16	-2.296	527	15	2
17	-2.198	532	15	2
18	-2.102	536	15	2
19	-2.008	541	15	2
20	-1.916	545	15	2
21	-1.826	549	14	2
22	-1.736	554	14	2
23	-1.648	558	14	2
24	-1.559	562	14	2
25	-1.471	567	14	2
26	-1.383	571	14	2
27	-1.295	575	14	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
28	-1.206	579	14	2
29	-1.116	584	15	3
30	-1.025	588	15	3
31	-0.933	593	15	3
32	-0.838	597	15	3
33	-0.742	602	15	3
34	-0.643	607	15	3
35	-0.540	611	16	3
36	-0.434	617	16	3
37	-0.323	622	16	3
38	-0.207	628	17	3
39	-0.085	633	17	3
40	0.046	640	18	3
41	0.186	646	18	3
42	0.338	654	19	3
43	0.506	662	20	4
44	0.694	671	22	4
45	0.910	681	23	4
46	1.166	694	26	4
47	1.488	709	29	4
48	1.926	730	35	4
49	2.653	765	49	4
50	3.363	800	69	4

Table I2. Reading Grade 4 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-5.230	385	69	1
1	-4.517	420	49	1
2	-3.783	455	35	1
3	-3.337	477	29	1
4	-3.009	492	26	1
5	-2.745	505	24	1
6	-2.523	516	22	1
7	-2.328	525	21	1
8	-2.153	534	20	1
9	-1.994	541	19	1
10	-1.848	548	18	1
11	-1.711	555	18	1
12	-1.582	561	17	1
13	-1.460	567	17	1
14	-1.343	573	16	2
15	-1.231	578	16	2
16	-1.123	583	16	2
17	-1.018	588	15	2
18	-0.916	593	15	2
19	-0.816	598	15	2
20	-0.719	603	15	2
21	-0.623	607	15	2
22	-0.528	612	15	2
23	-0.434	617	15	2
24	-0.341	621	15	2
25	-0.249	625	15	2
26	-0.156	630	15	2
27	-0.064	634	15	3
28	0.029	639	15	3
29	0.122	643	15	3
30	0.217	648	15	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.313	653	15	3
32	0.410	657	15	3
33	0.509	662	15	3
34	0.611	667	15	3
35	0.716	672	16	3
36	0.824	677	16	3
37	0.936	683	16	3
38	1.054	688	17	3
39	1.178	694	17	3
40	1.310	701	18	4
41	1.451	707	18	4
42	1.603	715	19	4
43	1.771	723	20	4
44	1.960	732	22	4
45	2.175	742	23	4
46	2.431	755	26	4
47	2.752	770	29	4
48	3.189	791	35	4
49	3.914	826	49	4
50	4.623	860	69	4

Table I3. Reading Grade 5 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.912	401	69	1
1	-4.206	435	49	1
2	-3.485	469	35	1
3	-3.052	490	29	1
4	-2.736	506	25	1
5	-2.485	518	23	1
6	-2.274	528	21	1
7	-2.091	537	20	1
8	-1.929	545	19	1
9	-1.782	552	18	1
10	-1.647	558	17	1
11	-1.522	564	17	1
12	-1.404	570	16	1
13	-1.293	575	16	1
14	-1.188	580	15	1
15	-1.087	585	15	1
16	-0.990	590	15	2
17	-0.896	594	15	2
18	-0.806	599	14	2
19	-0.717	603	14	2
20	-0.631	607	14	2
21	-0.546	611	14	2
22	-0.463	615	14	2
23	-0.381	619	14	2
24	-0.300	623	14	2
25	-0.219	627	14	2
26	-0.139	631	14	2
27	-0.059	635	14	2
28	0.020	638	14	2
29	0.100	642	14	3
30	0.181	646	14	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.261	650	14	3
32	0.343	654	14	3
33	0.426	658	14	3
34	0.510	662	14	3
35	0.595	666	14	3
36	0.683	670	14	3
37	0.773	675	15	3
38	0.865	679	15	3
39	0.961	684	15	3
40	1.060	689	15	3
41	1.164	694	16	3
42	1.273	699	16	3
43	1.389	704	17	3
44	1.512	710	17	4
45	1.645	717	18	4
46	1.790	724	19	4
47	1.950	732	20	4
48	2.131	740	21	4
49	2.339	750	23	4
50	2.587	762	25	4
51	2.899	777	29	4
52	3.329	798	35	4
53	4.047	833	49	4
54	4.752	867	69	4

Table I4. Reading Grade 6 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.771	407	69	1
1	-4.055	442	49	1
2	-3.318	478	36	1
3	-2.871	499	29	1
4	-2.544	515	26	1
5	-2.283	527	23	1
6	-2.063	538	22	1
7	-1.873	547	20	1
8	-1.704	555	19	1
9	-1.551	563	18	1
10	-1.411	569	18	1
11	-1.281	576	17	1
12	-1.159	582	17	1
13	-1.045	587	16	1
14	-0.936	592	16	1
15	-0.832	597	15	2
16	-0.732	602	15	2
17	-0.635	607	15	2
18	-0.542	611	15	2
19	-0.452	616	14	2
20	-0.363	620	14	2
21	-0.277	624	14	2
22	-0.192	628	14	2
23	-0.108	632	14	2
24	-0.026	636	14	2
25	0.055	640	14	2
26	0.136	644	14	2
27	0.217	648	14	2
28	0.297	652	14	3
29	0.377	656	14	3
30	0.457	660	14	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.537	663	14	3
32	0.618	667	14	3
33	0.699	671	14	3
34	0.782	675	14	3
35	0.865	679	14	3
36	0.950	683	14	3
37	1.037	687	14	3
38	1.126	692	14	3
39	1.217	696	15	3
40	1.311	701	15	3
41	1.408	705	15	3
42	1.508	710	15	3
43	1.614	715	16	3
44	1.725	721	16	4
45	1.842	726	17	4
46	1.967	732	17	4
47	2.102	739	18	4
48	2.248	746	19	4
49	2.410	754	20	4
50	2.592	762	21	4
51	2.802	773	23	4
52	3.052	785	25	4
53	3.366	800	29	4
54	3.798	821	35	4
55	4.517	855	49	4
56	5.223	889	69	4

Table I5. Reading Grade 7 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.356	427	69	1
1	-3.644	462	49	1
2	-2.914	497	35	1
3	-2.473	518	29	1
4	-2.150	534	26	1
5	-1.893	546	23	1
6	-1.676	557	22	1
7	-1.488	566	20	1
8	-1.321	574	19	1
9	-1.170	581	18	1
10	-1.031	588	18	1
11	-0.903	594	17	1
12	-0.782	600	16	1
13	-0.668	605	16	1
14	-0.560	611	16	2
15	-0.457	615	15	2
16	-0.358	620	15	2
17	-0.262	625	15	2
18	-0.170	629	15	2
19	-0.080	634	14	2
20	0.008	638	14	2
21	0.094	642	14	2
22	0.179	646	14	2
23	0.262	650	14	2
24	0.344	654	14	2
25	0.425	658	14	2
26	0.505	662	14	2
27	0.585	666	14	2
28	0.665	670	14	3
29	0.745	673	14	3
30	0.824	677	14	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.905	681	14	3
32	0.985	685	14	3
33	1.067	689	14	3
34	1.149	693	14	3
35	1.233	697	14	3
36	1.318	701	14	3
37	1.405	705	14	3
38	1.494	710	14	3
39	1.585	714	15	3
40	1.679	718	15	3
41	1.777	723	15	3
42	1.879	728	16	3
43	1.985	733	16	3
44	2.096	739	16	3
45	2.215	744	17	3
46	2.341	750	17	4
47	2.477	757	18	4
48	2.625	764	19	4
49	2.789	772	20	4
50	2.973	781	21	4
51	3.185	791	23	4
52	3.438	803	26	4
53	3.755	819	29	4
54	4.191	840	35	4
55	4.913	874	49	4
56	5.621	908	69	4

Table I6. Reading Grade 8 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.063	442	69	1
1	-3.358	476	49	1
2	-2.640	510	35	1
3	-2.209	531	29	1
4	-1.896	546	25	1
5	-1.648	558	23	1
6	-1.440	568	21	1
7	-1.260	577	20	1
8	-1.100	584	19	1
9	-0.955	591	18	1
10	-0.823	598	17	1
11	-0.700	604	17	1
12	-0.585	609	16	1
13	-0.476	615	16	1
14	-0.373	620	15	1
15	-0.275	624	15	2
16	-0.181	629	15	2
17	-0.089	633	14	2
18	-0.001	637	14	2
19	0.085	642	14	2
20	0.169	646	14	2
21	0.251	650	14	2
22	0.331	653	14	2
23	0.411	657	14	2
24	0.489	661	13	2
25	0.566	665	13	2
26	0.643	668	13	2
27	0.720	672	13	2
28	0.796	676	13	3
29	0.873	680	13	3
30	0.949	683	13	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	1.026	687	13	3
32	1.104	691	13	3
33	1.182	694	13	3
34	1.261	698	14	3
35	1.341	702	14	3
36	1.423	706	14	3
37	1.507	710	14	3
38	1.592	714	14	3
39	1.680	718	14	3
40	1.771	723	15	3
41	1.865	727	15	3
42	1.963	732	15	3
43	2.065	737	16	3
44	2.173	742	16	3
45	2.287	748	17	3
46	2.409	754	17	4
47	2.541	760	18	4
48	2.685	767	19	4
49	2.844	775	20	4
50	3.023	783	21	4
51	3.230	793	23	4
52	3.478	805	25	4
53	3.789	820	29	4
54	4.218	841	35	4
55	4.935	875	49	4
56	5.639	909	69	4

Table I7. Mathematics Grade 3 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-5.935	351	69	1
1	-5.225	386	49	1
2	-4.499	421	35	1
3	-4.060	442	29	1
4	-3.739	457	26	1
5	-3.483	470	23	1
6	-3.268	480	22	1
7	-3.080	489	20	1
8	-2.913	497	19	1
9	-2.762	504	18	1
10	-2.622	511	18	1
11	-2.493	517	17	1
12	-2.372	523	17	1
13	-2.257	529	16	1
14	-2.148	534	16	1
15	-2.044	539	15	1
16	-1.944	544	15	1
17	-1.848	548	15	1
18	-1.754	553	15	1
19	-1.663	557	14	1
20	-1.575	562	14	2
21	-1.488	566	14	2
22	-1.403	570	14	2
23	-1.319	574	14	2
24	-1.237	578	14	2
25	-1.156	582	14	2
26	-1.075	586	14	2
27	-0.995	590	14	2
28	-0.916	593	14	2
29	-0.837	597	14	2
30	-0.758	601	14	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	-0.680	605	14	3
32	-0.601	609	14	3
33	-0.522	612	14	3
34	-0.442	616	14	3
35	-0.362	620	14	3
36	-0.281	624	14	3
37	-0.199	628	14	3
38	-0.116	632	14	3
39	-0.032	636	14	3
40	0.054	640	14	3
41	0.142	644	14	3
42	0.232	649	15	3
43	0.325	653	15	3
44	0.420	658	15	3
45	0.519	663	15	3
46	0.622	667	16	3
47	0.730	673	16	3
48	0.843	678	16	3
49	0.963	684	17	4
50	1.091	690	18	4
51	1.228	697	18	4
52	1.378	704	19	4
53	1.543	712	20	4
54	1.729	721	21	4
55	1.943	731	23	4
56	2.197	743	26	4
57	2.516	759	29	4
58	2.952	780	35	4
59	3.676	815	49	4
60	4.385	849	69	4

Table I8. Mathematics Grade 4 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-5.633	366	69	1
1	-4.927	400	49	1
2	-4.206	435	35	1
3	-3.774	456	29	1
4	-3.459	471	25	1
5	-3.209	483	23	1
6	-2.999	493	21	1
7	-2.818	502	20	1
8	-2.657	509	19	1
9	-2.511	516	18	1
10	-2.378	523	17	1
11	-2.254	529	17	1
12	-2.139	534	16	1
13	-2.030	540	16	1
14	-1.927	545	15	1
15	-1.828	549	15	1
16	-1.734	554	15	1
17	-1.643	558	14	1
18	-1.555	563	14	1
19	-1.470	567	14	1
20	-1.388	571	14	1
21	-1.307	574	14	1
22	-1.228	578	13	1
23	-1.151	582	13	1
24	-1.075	586	13	2
25	-1.000	589	13	2
26	-0.926	593	13	2
27	-0.853	596	13	2
28	-0.781	600	13	2
29	-0.709	603	13	2
30	-0.638	607	13	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	-0.567	610	13	2
32	-0.496	614	13	2
33	-0.425	617	13	2
34	-0.354	620	13	3
35	-0.283	624	13	3
36	-0.211	627	13	3
37	-0.139	631	13	3
38	-0.067	634	13	3
39	0.007	638	13	3
40	0.081	641	13	3
41	0.156	645	13	3
42	0.233	649	13	3
43	0.311	652	14	3
44	0.390	656	14	3
45	0.472	660	14	3
46	0.555	664	14	3
47	0.641	668	14	3
48	0.730	673	14	3
49	0.822	677	15	3
50	0.917	682	15	3
51	1.017	687	15	3
52	1.122	692	16	3
53	1.232	697	16	3
54	1.350	703	17	4
55	1.475	709	17	4
56	1.611	715	18	4
57	1.759	722	19	4
58	1.923	730	20	4
59	2.107	739	21	4
60	2.321	749	23	4
61	2.575	762	26	4
62	2.894	777	29	4

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
63	3.331	798	35	4
64	4.057	833	49	4
65	4.766	867	69	4

Table I9. Mathematics Grade 5 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.730	409	69	1
1	-4.023	444	49	1
2	-3.303	478	35	1
3	-2.871	499	29	1
4	-2.556	514	25	1
5	-2.307	526	23	1
6	-2.098	536	21	1
7	-1.917	545	20	1
8	-1.757	553	19	1
9	-1.613	560	18	1
10	-1.481	566	17	1
11	-1.359	572	17	1
12	-1.245	577	16	1
13	-1.137	583	16	1
14	-1.036	588	15	1
15	-0.939	592	15	1
16	-0.847	597	15	1
17	-0.758	601	14	1
18	-0.672	605	14	1
19	-0.589	609	14	2
20	-0.508	613	14	2
21	-0.430	617	13	2
22	-0.353	620	13	2
23	-0.278	624	13	2
24	-0.204	628	13	2
25	-0.131	631	13	2
26	-0.059	635	13	2
27	0.011	638	13	2
28	0.081	641	13	2
29	0.151	645	13	3
30	0.220	648	13	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.289	651	13	3
32	0.357	655	13	3
33	0.426	658	13	3
34	0.495	661	13	3
35	0.563	665	13	3
36	0.632	668	13	3
37	0.702	671	13	3
38	0.772	675	13	3
39	0.843	678	13	3
40	0.914	682	13	3
41	0.987	685	13	3
42	1.061	689	13	3
43	1.136	692	13	3
44	1.213	696	13	3
45	1.291	700	14	3
46	1.372	704	14	3
47	1.455	708	14	3
48	1.541	712	14	3
49	1.630	716	15	3
50	1.722	721	15	4
51	1.818	725	15	4
52	1.920	730	16	4
53	2.027	735	16	4
54	2.140	741	17	4
55	2.262	747	17	4
56	2.393	753	18	4
57	2.537	760	19	4
58	2.696	767	20	4
59	2.876	776	21	4
60	3.084	786	23	4
61	3.332	798	25	4
62	3.645	813	29	4

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
63	4.076	834	35	4
64	4.794	869	49	4
65	5.499	903	69	4

Table I10. Mathematics Grade 6 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.337	428	69	1
1	-3.631	462	49	1
2	-2.913	497	35	1
3	-2.482	518	29	1
4	-2.169	533	25	1
5	-1.921	545	23	1
6	-1.713	555	21	1
7	-1.533	564	20	1
8	-1.374	571	19	1
9	-1.230	578	18	1
10	-1.099	585	17	1
11	-0.977	590	17	1
12	-0.863	596	16	1
13	-0.756	601	16	1
14	-0.655	606	15	1
15	-0.558	611	15	1
16	-0.466	615	15	1
17	-0.377	619	14	1
18	-0.291	623	14	1
19	-0.208	627	14	1
20	-0.128	631	14	1
21	-0.049	635	13	2
22	0.028	639	13	2
23	0.103	642	13	2
24	0.177	646	13	2
25	0.250	650	13	2
26	0.322	653	13	2
27	0.393	656	13	2
28	0.463	660	13	2
29	0.533	663	13	3
30	0.602	667	13	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.671	670	13	3
32	0.740	673	13	3
33	0.809	677	13	3
34	0.878	680	13	3
35	0.947	683	13	3
36	1.017	687	13	3
37	1.087	690	13	3
38	1.157	693	13	3
39	1.228	697	13	3
40	1.301	700	13	3
41	1.374	704	13	3
42	1.448	707	13	3
43	1.524	711	13	3
44	1.601	715	13	3
45	1.680	718	14	3
46	1.761	722	14	3
47	1.845	726	14	3
48	1.931	731	14	3
49	2.020	735	15	3
50	2.113	739	15	3
51	2.210	744	15	4
52	2.312	749	16	4
53	2.420	754	16	4
54	2.534	760	17	4
55	2.657	766	17	4
56	2.789	772	18	4
57	2.934	779	19	4
58	3.094	787	20	4
59	3.274	795	21	4
60	3.483	805	23	4
61	3.733	817	25	4
62	4.046	833	29	4

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
63	4.478	853	35	4
64	5.197	888	49	4
65	5.903	922	69	4

Table I11. Mathematics Grade 7 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-3.669	461	69	1
1	-2.963	495	49	1
2	-2.244	529	35	1
3	-1.812	550	29	1
4	-1.498	565	25	1
5	-1.248	577	23	1
6	-1.039	587	21	1
7	-0.858	596	20	1
8	-0.698	604	19	1
9	-0.553	611	18	1
10	-0.421	617	17	1
11	-0.298	623	17	1
12	-0.184	629	16	1
13	-0.076	634	16	1
14	0.026	639	15	1
15	0.123	643	15	1
16	0.216	648	15	1
17	0.306	652	14	1
18	0.392	656	14	2
19	0.476	660	14	2
20	0.557	664	14	2
21	0.636	668	13	2
22	0.713	672	13	2
23	0.789	676	13	2
24	0.863	679	13	2
25	0.936	683	13	2
26	1.007	686	13	2
27	1.078	689	13	3
28	1.148	693	13	3
29	1.218	696	13	3
30	1.287	700	13	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	1.355	703	13	3
32	1.423	706	13	3
33	1.491	709	13	3
34	1.559	713	13	3
35	1.627	716	13	3
36	1.695	719	13	3
37	1.764	723	13	3
38	1.833	726	13	3
39	1.902	729	13	3
40	1.972	733	13	3
41	2.043	736	13	3
42	2.115	739	13	3
43	2.188	743	13	3
44	2.262	747	13	3
45	2.338	750	13	3
46	2.415	754	13	3
47	2.494	758	14	4
48	2.576	762	14	4
49	2.661	766	14	4
50	2.748	770	14	4
51	2.839	774	15	4
52	2.933	779	15	4
53	3.033	784	15	4
54	3.138	789	16	4
55	3.250	794	16	4
56	3.369	800	17	4
57	3.498	806	18	4
58	3.640	813	19	4
59	3.797	821	20	4
60	3.974	829	21	4
61	4.179	839	23	4
62	4.424	851	25	4

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
63	4.734	866	29	4
64	5.161	886	35	4
65	5.876	921	49	4
66	6.579	955	69	4

Table I12. Mathematics Grade 8 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-3.394	474	69	1
1	-2.689	508	49	1
2	-1.974	542	35	1
3	-1.546	563	29	1
4	-1.236	578	25	1
5	-0.990	590	23	1
6	-0.785	600	21	1
7	-0.608	608	20	1
8	-0.451	616	19	1
9	-0.310	623	18	1
10	-0.181	629	17	1
11	-0.062	635	16	1
12	0.049	640	16	1
13	0.154	645	15	1
14	0.252	650	15	1
15	0.346	654	15	1
16	0.436	659	14	1
17	0.522	663	14	1
18	0.605	667	14	1
19	0.685	671	14	1
20	0.763	674	13	1
21	0.839	678	13	2
22	0.913	682	13	2
23	0.985	685	13	2
24	1.055	688	13	2
25	1.125	692	13	2
26	1.193	695	13	2
27	1.260	698	12	2
28	1.326	701	12	2
29	1.392	705	12	2
30	1.457	708	12	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	1.521	711	12	3
32	1.585	714	12	3
33	1.649	717	12	3
34	1.713	720	12	3
35	1.776	723	12	3
36	1.840	726	12	3
37	1.903	729	12	3
38	1.967	732	12	3
39	2.031	735	12	3
40	2.095	738	12	3
41	2.160	742	12	3
42	2.225	745	12	3
43	2.291	748	12	3
44	2.358	751	13	3
45	2.426	754	13	3
46	2.495	758	13	3
47	2.566	761	13	3
48	2.637	765	13	3
49	2.711	768	13	3
50	2.786	772	13	3
51	2.863	776	13	3
52	2.943	779	14	4
53	3.025	783	14	4
54	3.111	787	14	4
55	3.200	792	15	4
56	3.293	796	15	4
57	3.391	801	15	4
58	3.495	806	16	4
59	3.606	811	16	4
60	3.724	817	17	4
61	3.852	823	18	4
62	3.993	830	19	4

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
63	4.148	837	20	4
64	4.325	846	21	4
65	4.529	856	23	4
66	4.774	868	25	4
67	5.083	883	29	4
68	5.510	903	35	4
69	6.224	938	49	4
70	6.928	971	69	4

Table I13. Science Grade 4 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-5.030	395	69	1
1	-4.311	430	49	1
2	-3.567	466	36	1
3	-3.113	487	30	1
4	-2.777	504	26	1
5	-2.506	517	24	1
6	-2.276	528	22	1
7	-2.075	537	21	1
8	-1.896	546	20	1
9	-1.732	554	19	1
10	-1.581	561	18	1
11	-1.440	568	18	1
12	-1.308	574	17	1
13	-1.182	581	17	1
14	-1.062	586	16	1
15	-0.948	592	16	1
16	-0.837	597	16	1
17	-0.730	602	16	1
18	-0.626	607	15	1
19	-0.525	612	15	2
20	-0.425	617	15	2
21	-0.328	622	15	2
22	-0.232	626	15	2
23	-0.137	631	15	2
24	-0.043	635	15	2
25	0.050	640	15	2
26	0.143	644	15	2
27	0.236	649	15	2
28	0.330	653	15	2
29	0.423	658	15	2
30	0.518	662	15	2

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.614	667	15	3
32	0.711	672	15	3
33	0.810	677	15	3
34	0.912	681	15	3
35	1.016	686	16	3
36	1.124	692	16	3
37	1.236	697	16	3
38	1.353	703	17	3
39	1.476	709	17	3
40	1.606	715	18	3
41	1.746	722	18	3
42	1.898	729	19	4
43	2.064	737	20	4
44	2.251	746	22	4
45	2.464	756	23	4
46	2.719	769	26	4
47	3.037	784	29	4
48	3.472	805	35	4
49	4.194	840	49	4
50	4.902	874	69	4

Table I14. Science Grade 8 Raw Score to Scale Score

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
0	-4.739	409	69	1
1	-4.030	443	49	1
2	-3.304	478	35	1
3	-2.867	499	29	1
4	-2.547	515	26	1
5	-2.292	527	23	1
6	-2.078	537	22	1
7	-1.892	546	20	1
8	-1.726	554	19	1
9	-1.576	562	18	1
10	-1.438	568	18	1
11	-1.309	574	17	1
12	-1.188	580	16	1
13	-1.074	586	16	1
14	-0.965	591	16	1
15	-0.861	596	15	1
16	-0.760	601	15	1
17	-0.662	606	15	2
18	-0.567	610	15	2
19	-0.475	615	15	2
20	-0.384	619	14	2
21	-0.294	623	14	2
22	-0.205	628	14	2
23	-0.118	632	14	2
24	-0.031	636	14	2
25	0.056	640	14	2
26	0.143	644	14	2
27	0.230	649	14	2
28	0.318	653	14	2
29	0.407	657	14	3
30	0.497	661	15	3

Raw Score	Theta	Rounded Scaled Score	Standard Error	Performance Level
31	0.588	666	15	3
32	0.681	670	15	3
33	0.776	675	15	3
34	0.874	680	15	3
35	0.975	685	15	3
36	1.079	690	16	3
37	1.188	695	16	3
38	1.303	700	17	3
39	1.423	706	17	3
40	1.552	712	18	3
41	1.690	719	18	4
42	1.840	726	19	4
43	2.006	734	20	4
44	2.191	743	21	4
45	2.405	753	23	4
46	2.659	766	26	4
47	2.977	781	29	4
48	3.413	802	35	4
49	4.136	837	49	4
50	4.844	871	69	4

Table I15. SAWS State Level Normative Percentage Distributions

Grade	State N count	0-3 Points	4-7 Points	8-10 Points	11-12-points
3	7292	8	53	29	10
4	7282	6	58	28	8
5	6909	4	47	34	15
6	6926	5	52	32	10
7	6972	4	41	37	18
8	6914	4	36	39	21

Appendix J: Sample PAWS Reports



2013 Proficiency Assessments for Wyoming Students (PAWS)

For the parent or guardian of:
Claire Carroll
 Student ID: 89100054 SCHOOL: Urie Elementary
 Grade: 04 DISTRICT: Uinta County School District #

Claire's Performance on the PAWS Test



Dear Parent or Guardian,

I am pleased to provide you with your child's results from the 2013 Proficiency Assessment of Wyoming Students (PAWS). Two 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 (reading, mathematics and science). The other is the scale score your child earned for each content area. These results provide your child's school with useful information about how well your child is learning the Wyoming Content and Performance Standards.

Each of the charts in this report contains your child's actual score, which is shown by a star. Charts on pages 2 – 4 show your child's scores in each of the content and skill areas, as compared to scores of other students in the same grade across the state. This information can help both you and your child's teacher identify areas of strength as well as areas to work on during the coming school year.

A more detailed explanation of this score report can be found on the Wyoming Department of Education website at www.edu.wyoming.gov.

I would also like to point out the Lexile® measure and Lexile® website (bottom of page 2) and the Quantile® measure and Quantile® website (bottom of page 3). Entering your child's scores on these websites will provide you with access to valuable tools and resources which can support your child's academic growth.

The Wyoming Department of Education is dedicated to working together with districts, teachers and parents to ensure your child's continued educational progress. I encourage you to be involved in your child's school and classroom as much as possible to further support his or her success in school.

Sincerely,
 Jim Rose
 Interim Director, Wyoming Department of Education

YOUR CHILD'S SCALE SCORES AND PERFORMANCE LEVELS

ADVANCED	975	975	975	
	700	726	698	731 *
PROFICIENT	699	725	697	
	634	666	620	
BASIC	633	665	619	
	570	612	584	
BELOW BASIC	569	611	583	
	300	300	300	
* = Student's Scale Score		324 *	326 *	
	Reading	Science	Mathematics	

Score Definitions

Performance Level - Performance Level (Below Basic, Basic, Proficient, or Advanced) is derived from the student's scale score. This reporting information enables teachers to identify a student's overall performance within a subject (Reading, Mathematics, and Science).

Scale Scores - Scale Scores are derived from the raw scores earned on the PAWS. For Grades 3-8, the scale scores range from 300-975. While these scale ranges are identical, scale scores across subjects are not comparable.

Not Tested (NT) - Your child did not participate in this content area.

Unfair Advantage (UA) - If your child received an UA, please contact your school for further information.

Not Applicable (NA) - This test was not given at your child's grade level.

Additional Resources and Information

Visit the Wyoming Department of Education online

Go to the Wyoming Department of Education's website at edu.wyoming.gov for more information about the Proficiency Assessments for Wyoming Students (PAWS).

2013 Proficiency Assessments for Wyoming Students (PAWS)

Claire's Performance on the PAWS Reading Test

YOUR CHILD'S READING SCALE SCORE AND PERFORMANCE LEVEL			
Performance Level: <i>Basic</i>			
Scale Score: 324			
ADVANCED	975	<input type="checkbox"/> ADVANCED: Scale Score range: 700-975 Students performing at an advanced level demonstrate complete mastery of understanding literary and informational texts and use a creative approach in unfamiliar settings. Students can consistently and independently demonstrate an accurate understanding of literary text; determine information related to and among informational texts; determine relevant information; relate story elements; respond to a variety of literary genres by using details; apply sufficient information to complete a task effectively.	
	700		
PROFICIENT	699		
	634		
BASIC	633		
	570	<input type="checkbox"/> PROFICIENT: Scale Score range: 634-699 Students performing at a proficient level meet the standard of understanding literary and informational texts and perform in several familiar settings. Students can consistently show a literal understanding of the text; identify information related to and among informational texts; identify main idea and supporting details; identify story elements; respond to a variety of literary genres by identifying main ideas; provide some information to complete a task satisfactorily.	
BELOW BASIC	569		
	300		
	324 *	329.4 ‡	<input checked="" type="checkbox"/> BASIC: Scale Score range: 570-633 Students performing at a basic level are working toward the standard for the grade. Understanding of literary and informational texts is emerging given external support and multiple prompts in limited familiar settings. Students can recognize important information in literary and informational texts given teacher-provided choices; match story elements; respond to or actively attend to familiar genres.
Your Child		State Average	
Scale Score		Scale Score	
<small>* = Student's Scale Score ‡ = State Average Scale Score</small>			

YOUR CHILD'S READING SCORE ANALYSIS BY TEXT TYPE		
Content Standards/Skills Description	Student's Scale Score	State Average Scale Score
<p>Functional Texts: These texts include reading materials such as directions, schedules, maps, diagrams, internet websites, electronic databases, and explanations for doing something or getting somewhere. They provide basic information readers need to accomplish day-to-day tasks.</p> <p>Relevance and Importance - The reader identifies and locates information from the text and understands how the information is relevant and important for accomplishing a specified task.</p>	908	908.2
<p>Expository Texts: These texts include such things as textbooks, encyclopedias, documentaries, speeches, public documents, print news media, the internet, websites, electronic databases, microfiche, almanacs, news, biographies, scientific explanations, and historical and political analyses.</p> <p>Major Points and Supporting Details - The reader identifies main ideas from the text and recognizes relevant details which support those main ideas.</p> <p>Information Relationships - The reader understands how information from the text relates to broader topics and issues.</p>	905	911.9
<p>Narrative Texts: These texts include stories, poems, novels, plays, and essays about America and various world cultures that are read to learn about people, to vicariously experience the characters and settings, to escape to imaginary places and times, and to become absorbed in adventure and fictional events, and various problems and solutions that structure the plots of these texts.</p> <p>Story Elements - The reader recognizes and understands an author's development of story element and how the story element contributes to the development of plot.</p> <p>Plot - The reader identifies and understands the development of a story's plot, central problem, and resolution.</p>	511	509.3

Lexile® measure = 824L

A Lexile® measure helps readers select materials at their reading level. The Lexile website helps build a customized list of books that match your child's Lexile level and interests. This list can serve as a guide in selecting books at your school or public library and/or local bookstore. For more information, and to search for books by Lexile measure, visit www.Lexile.com.

2013 Proficiency Assessments for Wyoming Students (PAWS)

Claire's Performance on the PAWS Mathematics Test

YOUR CHILD'S MATHEMATICS SCALE SCORE AND PERFORMANCE LEVEL			
Performance Level: <i>Basic</i> Scale Score: 731		<input type="checkbox"/> ADVANCED: Scale Score range: 698-975 Students performing at an advanced level demonstrate exemplary performance or understanding; use a creative mathematical approach or multiple complex methods of problem solving including interpreting complex information; completing complex tasks involving several processing steps; developing a suitable mathematical representation to produce a solution; applying knowledge in an unfamiliar problem context; displaying generalization, reasoning, and argumentation in communication of results.	
ADVANCED	975 698 731 * 740.0 ‡	<input type="checkbox"/> PROFICIENT: Scale Score range: 620-697 Students performing at a proficient level meet the acceptable standard for the grade by demonstrating solid performance or understanding of problem solving including completing tasks involving more than a single processing step; combining different pieces of information; interpreting different representations; recognizing which elements are important and how they relate to one another; working with given mathematical representations; carrying out sequence of processing steps to produce a solution.	
PROFICIENT	697 620	<input checked="" type="checkbox"/> BASIC: Scale Score range: 584-619 Students performing at a basic level have not yet met the acceptable standard for the grade. Although errors are being made, performance and understanding are emerging as demonstrated by the use of routine computations and problem solving including solving or using single step processes; recognizing familiar contexts and mathematically well-defined problems; reproducing of facts or processes; applying simple computational skills.	
BASIC	619 584	<input type="checkbox"/> BELOW BASIC: Scale Score range: 300-583 The student is not yet meeting the acceptable standard for the grade; errors are being made and there is no evidence of understanding problem solving or computation skills. Students can sometimes use simple computational skills to solve simple problems; perform some basic procedures with inconsistent accuracy; communicate with little use of mathematical language and often includes errors.	
BELOW BASIC	583 300		
Your Child State Average Scale Score Scale Score			
* = Student's Scale Score ‡ = State Average Scale Score			

YOUR CHILD'S MATHEMATICS SCORE ANALYSIS BY STANDARD		
Content Standards/Skills Examples	Student's Scale Score	State Average Scale Score
Number Operations and Concepts - Students use numbers, number sense, and number relationships in a problem-solving situation. Number Representations - Replace the □ with the correct symbol {<, >, =}: 0.80 □ 70/100. Number Operations and Concepts - What is 4012 - 309?	313	310.9
Geometry - Students apply geometric concepts, properties, and relationships in a problem-solving situation. Spatial Relationships - When are intersecting lines perpendicular? Two-Dimensional & Three-Dimensional Shapes - How many diagonals can be drawn in a regular pentagon? Transformations & Symmetry - Draw the reflection of the letter <u>W</u> across the given line.	407	409.0
Measurement - Students use a variety of tools and techniques of measurement in a problem-solving situation. Measurement Systems - How many minutes are there between 3:37 pm and 4:12 pm? Perimeter, Area & Volume - A rectangular swimming pool has a perimeter of 190 ft. The length of the pool is 60 ft. What is the width of the pool?	904	907.2
Algebra - Students use algebraic methods to investigate, model, and interpret patterns and functions involving numbers, shapes, data, and graphs in a problem-solving situation. Patterns, Relations & Functions - What is the rule that produces the following pattern? 96, 89, 82, 75	705	707.5
Data Analysis and Probability - Students use data analysis and probability to analyze given situations and the results of experiments. Collect and Analyze Data - 6 people have a dog for a pet, 8 people have a cat and 4 people have both. Draw a Venn diagram to show this situation. Inferences/Predictions - Numbered balls 1 thru 8 are put in a bag. You draw 2 balls from the bag. Is it (certain, likely, unlikely, impossible) the sum of the numbers is greater than 16?	502	505.5

Quantile® measure = 931Q

A Quantile® measure is similar to the Lexile and can help you identify math activities to do at home. These activities will help your child practice mathematical skills at an appropriate level, leading to increased mathematical understanding. For more information, visit the "Math at Home" section at www.Quantiles.com.

2013 Proficiency Assessments for Wyoming Students (PAWS)

Claire's Performance on the PAWS Science Test

YOUR CHILD'S SCIENCE SCALE SCORE AND PERFORMANCE LEVEL			
Performance Level: <i>Basic</i>			
Scale Score: 326			
ADVANCED	975		
	726		
PROFICIENT	725		
	666		
BASIC	665		
	612		
BELOW BASIC	611		
	300	326 *	329.5 ‡
	Your Child	State Average	Scale Score
* = Student's Scale Score ‡ = State Average Scale Score			
<input type="checkbox"/> ADVANCED: Scale Score range: 726-975 Grade 4 students performing at an advanced level demonstrate exemplary understanding and application of science content to identify, describe, compare, and classify objects and living things to explain the natural world. Students demonstrate a thorough understanding of the appropriate use of scientific inquiry procedures to cite evidence, represent data, draw conclusions, communicate results, make connections to daily life, suggest solutions to science-related issues and generate new questions. Students describe and compare properties of objects and living things, and describe cycles, patterns and interrelationships among objects and living things. Students demonstrate a thorough understanding of the appropriate use of a variety of scientific procedures and tools.			
<input type="checkbox"/> PROFICIENT: Scale Score range: 666-725 Grade 4 students performing at a proficient level demonstrate solid understanding and application of science content to identify, describe, compare, and classify objects and living things to explain the natural world. Students demonstrate an understanding of the use of the scientific inquiry process to communicate results and suggest solutions to science-related issues. Students give examples of observable properties and cycles, such as changes in objects in the sky, states of matter, and life cycles. Students demonstrate an understanding of the appropriate use of a variety of scientific procedures and tools.			
<input checked="" type="checkbox"/> BASIC: Scale Score range: 612-665 Grade 4 students performing at a basic level demonstrate limited understanding and application of science content to identify, describe, and recall science content. Students performing at this level demonstrate a limited ability to use the inquiry process to apply data, draw appropriate conclusions, and suggest solutions to science related issues, and to generate or validate information to communicate explanations. Students describe and record some characteristics of objects and living things. They demonstrate a limited understanding of the appropriate use of scientific procedures and tools.			
<input type="checkbox"/> BELOW BASIC: Scale Score range: 300-611 Grade 4 students performing at the below basic level are not yet meeting the acceptable standard for the grade; errors are being made and there is little or no evidence of understanding science content or inquiry procedures. Students can sometimes recognize, use, identify, describe, or recall scientific information/content. These students demonstrate little or no ability to use the inquiry process to generate or validate scientific information to communicate appropriate conclusions and explanations. They demonstrate little or no ability to identify or describe characteristics of objects and living things. They demonstrate little or no understanding of the appropriate use of scientific procedures and tools.			

YOUR CHILD'S SCIENCE SCORE ANALYSIS BY CONTENT STANDARD		
Content Standard - Concepts and Processes	Student's Scale Score	State Average Scale Score
Life Science - Cells and Cellular Processes: Key concepts include living organisms and distinguishing distinct structures and body systems that serve specific functions in growth, reproduction, and survival. Heredity and Biological Evolution: Key concepts include the process of how plants and animals progress through life cycles of birth, growth and development, reproduction and death; and the comparison of fossils to one another and to living organisms to observe their similarities and differences. Interactions and Energy Flow: A key concept is how unique features of plants and animals help them live in different environments.	509	509.6
Physical Science - Properties and Changes: Key concepts include the different states of matter and that each state has distinct physical properties; that some common materials, such as water, can be changed from one state to another by heating or cooling, and the classification of substances by their physical and chemical properties. Energy Types: A key concept is the Sun and how it supplies heat and light to Earth. Forces and Motion: Key concepts include sound produced by vibrating objects, which can be characterized by pitch and volume; and the change in the position and motion of an object resulting from pushing or pulling.	707	710.3
Earth and Space Science - Earth's Processes and Features: Key concepts are the composition and properties of earth materials. Solar System and Universe: Key concepts are the basic patterns of movement by objects in the sky.	610	609.6

SKILL - PROCESS AS INQUIRY DESCRIPTIONS
Observe and Question - Recognize and note facts, occurrences, or phenomena in the natural world. Understand the types of questions science can and cannot answer. Design and Conduct a Scientific Investigation - Identify appropriate and safe methods for collecting fair and adequate data to answer a scientific question. Understand the application of a variety of technologies, including hand tools, measuring instruments, calculators, and computers. Organize and Represent Data - Demonstrate attention to detail, recognize the need for reproducible results. Choose an appropriate method of data representation. Evaluate the quality and relevance of data to answer a given scientific problem. Draw Conclusions and Make Connections - Analyze patterns and trends in data, and use them to formulate explanations. Relate observations to knowledge of scientific concepts. Explain, communicate, and support conclusions.

Appendix K: Sample SAWS Reports

STUDENT REPORT

2013 Student Assessment of Writing Skills (SAWS)

For the parent or guardian of:

CAND10 W. RITING

Student ID: 11111120

SCHOOL: Saint Laurence Catholic School

Grade: 06

DISTRICT: AlbanyCountySchoolDistrict#1

CAND10's Performance on the SAWS Test



Dear Parent or Guardian,

I am pleased to provide you with your child's results from the 2013 Student Assessment of Writing Skills (SAWS). Two types of information are provided in this report. One is a description of your child's writing performance in each skill area. The other is a bar graph providing you with information about how other Wyoming students in your child's grade level performed on the test. These results provide your child's school with useful information about how well your child is learning the Wyoming Content and Performance Standards.

The first section of the SAWS score report identifies how your child scored in each of the four skill areas. Students can score 0 – 3 points for each of the four skills. The overall score is the sum of all four skill scores. A description of the skills and the score points can be found on the back of this report. The second section of the SAWS score report indicates how others across the state scored on the SAWS assessment. You can use this information to compare your student's performance with others in his or her grade.

A more detailed explanation of this score report can be found on the Wyoming Department of Education website at www.edu.wyoming.gov.

The Wyoming Department of Education is dedicated to working together with districts, teachers and parents to ensure your child's continued educational progress. I encourage you to be involved in your child's school and classroom as much as possible to further support his or her success in school.

Sincerely,

Jim Rose
Interim Director, Wyoming Department of Education

This report shows your student's essay score based on each content standard* and his or her overall score.

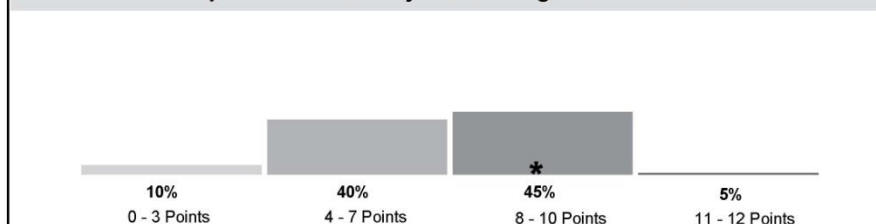
Idea Development:	3 out of 3
Organization:	3 out of 3
Voice:	2 out of 3
Convention:	2 out of 3
Overall Score:	9 out of 12

* Skill description for each content standard is on the back of this report.

Not Tested (NT) - Your child did not participate in this content area.

Unfair Advantage (UA) - If your child received an UA, please contact your school for further information.

The chart below shows how other Wyoming students in this grade level performed on this year's writing assessment.



* Indicates your student's score.

Additional Resources and Information

Visit the Wyoming Department of Education online

Go to the Wyoming Department of Education's website at edu.wyoming.gov for more information about the Student Assessment of Writing Skills (SAWS).

2013 Student Assessment of Writing Skills (SAWS)

Content Standards / Skills Descriptions

Expository Writing – Persuasive Essay – Students support a definite position to impact the opinion, attitude, or belief of others regarding a specific topic.

Idea Development - The writer develops the content of the message through the use of details.

Organization - The writer builds the structure to support the purpose and effectiveness of the writing.

Voice - The writer uses precise, appropriate language to communicate clear directions or procedures to an audience in a way that is informative, compelling, and engaging.

Conventions - The writer develops the mechanical correctness of the piece including spelling, capitalization, punctuation, and grammar.

SAWS Writing Scoring Guide - Grade 06 Targets				
<i>Expository: Write a Set of Directions or Procedures — Develops a set of directions or procedures to inform an audience</i>				
SKILLS	3	2	1	0
Idea Development The writer develops the content of the message through the use of details.	<ul style="list-style-type: none"> Develops clear and focused directions or procedures in response to the topic Uses descriptive details to enrich idea development 	<ul style="list-style-type: none"> Presents directions or procedures in response to the topic Uses relevant details 	<ul style="list-style-type: none"> Attempts to present directions or procedures in response to the topic Limited use of relevant details 	<ul style="list-style-type: none"> No response to the topic Details are consistently irrelevant
Organization The writer builds the structure to support the purpose and effectiveness of the writing.	<ul style="list-style-type: none"> Develops an effective introduction, body, and conclusion Sequencing of directions or procedures demonstrates effective logic and coherence Indicates paragraphs using appropriate spacing or indentation consistently 	<ul style="list-style-type: none"> Presents an introduction, body, and conclusion Sequencing of directions or procedures demonstrates overall logic and coherence Indicates paragraphs using appropriate spacing or indentation 	<ul style="list-style-type: none"> Presents an introduction or conclusion Sequencing of directions or procedures demonstrates some evidence of logic and coherence Similar ideas are grouped together without appropriate spacing or indentation 	<ul style="list-style-type: none"> Introduction and conclusion are unidentifiable Organization of directions or procedures lacks a logical sequence Similar ideas are not grouped together; no evidence of appropriate spacing or indentation
Voice The writer uses precise, appropriate language to communicate directly to the audience in a way that is informative, compelling, and engaging.	<ul style="list-style-type: none"> Directions or procedures consistently reveal voice or style appropriate to the purpose Uses a variety of precise and appropriate words or phrases 	<ul style="list-style-type: none"> Directions or procedures reveal voice or style appropriate to the purpose Uses precise and appropriate words or phrases 	<ul style="list-style-type: none"> Directions or procedures reveal limited voice or style appropriate to the purpose Demonstrates little variation in word choice and/or repetitious use of simple words or phrases 	<ul style="list-style-type: none"> Directions or procedures lack voice or style appropriate to the purpose Uses an extremely limited range of words or phrases or consistently uses words incorrectly
Conventions The writer develops the mechanical correctness of the piece including spelling, capitalization, punctuation, and grammar.	<ul style="list-style-type: none"> Uses grade-appropriate spelling consistently Uses grade-appropriate capitalization and punctuation consistently Uses grade-appropriate grammar and usage consistently Uses varied and correct sentences 	<ul style="list-style-type: none"> Uses grade-appropriate spelling Uses grade-appropriate capitalization and punctuation Uses grade-appropriate grammar and usage Uses varied and mostly correct sentences 	<ul style="list-style-type: none"> Spells common words correctly, but other grade-appropriate words incorrectly Uses limited grade-appropriate capitalization or punctuation Demonstrates limited control over grade-appropriate grammar and usage Attempts to use varied sentences; inconsistently uses correct sentences 	<ul style="list-style-type: none"> Misspells common words Demonstrates incorrect use of grade-appropriate capitalization or punctuation throughout Demonstrates incorrect use of grade-appropriate grammar and usage throughout Sentences are run-on, incomplete, or fragmented

Appendix L: Performance Level Percentages by Demographic Subgroup

Table L1. Performance Levels of Reading, Mathematics, and Science by Grade

Grade	Below Basic	Basic	Proficient	Advanced
Reading				
3	7.1	27.3	50.2	15.4
4	2.9	18.7	43.5	34.9
5	4.3	23.3	49.5	23.0
6	1.4	17.9	54.2	26.5
7	2.2	23.7	61.0	13.1
8	2.9	21.1	59.8	16.2
Mathematics				
3	2.7	10.0	51.1	36.2
4	6.2	12.9	53.6	27.4
5	5.0	15.5	48.5	31.0
6	5.7	12.7	53.5	28.1
7	5.2	19.0	50.2	25.6
8	10.3	22.4	47.7	19.6
Science				
4	8.0	34.6	45.6	11.8
8	14.1	42.2	34.7	8.9

Table L2. Performance Levels of Reading by Grade 3 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	7.1	27.3	50.2	15.4
Male	8.5	29.3	49.2	13.0
Female	5.4	25.2	51.3	18.0
Unknown	20.0	20.0	50.0	10.0
American Indian	22.6	41.6	31.8	4.0
Asian	1.5	23.9	47.8	26.9
Black	11.1	38.9	44.4	5.6
Hawaiian/Pacific Islander	20.0	20.0	46.7	13.3
Hispanic	11.7	40.0	41.7	6.6
White	5.5	24.2	52.8	17.5
Multiracial	3.9	27.3	51.6	17.2
Unknown	9.5	33.3	47.6	9.5
Free/Reduced Lunch	10.5	34.7	47.1	7.7
Not Free/Reduced Lunch	5.0	22.9	52.1	20.0
Special Education	23.7	41.8	29.7	4.8
Not Special Education	4.3	24.9	53.7	17.2
English Language Learner	22.1	49.3	26.4	2.3
Not English Language Learner	6.3	26.2	51.5	16.1

Table L3. Performance Levels of Reading by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	2.9	18.7	43.5	34.9
Male	3.8	21.0	44.3	30.9
Female	1.9	16.2	42.7	39.2
Unknown	10.0	30.0	50.0	10.0
American Indian	7.6	38.5	43.9	10.1
Asian	3.0	19.4	34.3	43.3
Black	8.6	23.5	38.3	29.6
Hawaiian/Pacific Islander	9.1	18.2	54.5	18.2
Hispanic	4.0	28.3	45.5	22.2
White	2.4	16.0	43.4	38.2
Multiracial	5.5	18.2	42.7	33.6
Unknown	0.0	25.0	41.7	33.3
Free/Reduced Lunch	4.4	24.6	46.6	24.4
Not Free/Reduced Lunch	2.0	15.2	41.7	41.1
Special Education	11.4	41.8	35.0	11.8
Not Special Education	1.5	14.9	44.9	38.7
English Language Learner	10.8	49.2	32.3	7.7
Not English Language Learner	2.7	17.8	43.8	35.6

Table L4. Performance Levels of Reading by Grade 5 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	4.3	23.3	49.5	23.0
Male	5.1	24.5	48.8	21.6
Female	3.4	21.9	50.3	24.5
Unknown	12.5	37.5	50.0	0.0
American Indian	9.5	41.7	38.6	10.2
Asian	0.0	13.0	60.9	26.1
Black	15.2	31.6	45.6	7.6
Hawaiian/Pacific Islander	11.8	35.3	47.1	5.9
Hispanic	6.4	34.0	47.1	12.4
White	3.4	20.4	50.4	25.7
Multiracial	10.0	26.7	45.6	17.8
Unknown	8.0	20.0	56.0	16.0
Free/Reduced Lunch	6.0	30.2	49.7	14.1
Not Free/Reduced Lunch	3.3	19.3	49.4	28.0
Special Education	18.4	46.1	28.9	6.5
Not Special Education	2.0	19.6	52.8	25.6
English Language Learner	22.0	60.4	16.5	1.2
Not English Language Learner	3.8	22.3	50.3	23.5

Table L5. Performance Levels of Reading by Grade 6 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	1.4	17.9	54.2	26.5
Male	2.0	21.8	53.5	22.7
Female	0.8	13.7	55.0	30.5
Unknown	0.0	25.0	50.0	25.0
American Indian	2.1	44.0	49.4	4.6
Asian	0.0	13.2	41.5	45.3
Black	3.5	30.6	51.8	14.1
Hawaiian/Pacific Islander	5.9	35.3	52.9	5.9
Hispanic	1.9	24.7	55.3	18.0
White	1.2	15.4	54.4	29.0
Multiracial	1.9	16.5	54.4	27.2
Unknown	9.5	23.8	57.1	9.5
Free/Reduced Lunch	1.7	24.4	56.1	17.8
Not Free/Reduced Lunch	1.3	14.3	53.2	31.3
Special Education	6.9	49.6	39.4	4.1
Not Special Education	0.6	13.3	56.3	29.7
English Language Learner	6.7	55.0	35.6	2.7
Not English Language Learner	1.3	17.0	54.6	27.0

Table L6. Performance Levels of Reading by Grade 7 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	2.2	23.7	61.0	13.1
Male	3.3	25.6	59.6	11.5
Female	1.1	21.6	62.5	14.9
Unknown	0.0	37.5	37.5	25.0
American Indian	6.8	37.1	53.4	2.7
Asian	0.0	23.5	51.0	25.5
Black	7.5	35.8	51.9	4.7
Hawaiian/Pacific Islander	13.3	26.7	40.0	20.0
Hispanic	3.7	34.6	56.2	5.5
White	1.7	21.0	62.5	14.8
Multiracial	1.7	28.2	58.1	12.0
Unknown	11.1	77.8	11.1	0.0
Free/Reduced Lunch	3.2	32.9	57.7	6.1
Not Free/Reduced Lunch	1.7	19.0	62.6	16.7
Special Education	10.2	54.4	33.1	2.2
Not Special Education	1.1	19.2	65.0	14.7
English Language Learner	9.2	65.4	25.5	0.0
Not English Language Learner	2.1	22.7	61.8	13.4

Table L7. Performance Levels of Reading by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	2.9	21.1	59.8	16.2
Male	4.0	24.9	57.8	13.2
Female	1.7	16.8	62.0	19.5
Unknown	0.0	33.3	50.0	16.7
American Indian	8.7	43.4	44.7	3.2
Asian	1.7	11.9	61.0	25.4
Black	2.6	32.5	58.4	6.5
Hawaiian/Pacific Islander	10.0	50.0	40.0	0.0
Hispanic	6.3	28.0	57.2	8.5
White	2.2	18.9	60.9	18.0
Multiracial	1.1	24.5	57.4	17.0
Unknown	0.0	33.3	66.7	0.0
Free/Reduced Lunch	4.7	28.9	56.4	10.1
Not Free/Reduced Lunch	2.0	17.1	61.5	19.3
Special Education	13.7	50.0	33.7	2.5
Not Special Education	1.4	17.1	63.4	18.1
English Language Learner	19.7	48.8	28.3	3.1
Not English Language Learner	2.6	20.5	60.4	16.5

Table L8. Performance Levels of Mathematics by Grade 3 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	2.7	10.0	51.1	36.2
Male	2.7	10.3	50.2	36.8
Female	2.6	9.8	52.0	35.7
Unknown	33.3	0.0	66.7	0.0
American Indian	8.5	26.6	50.9	14.0
Asian	1.5	6.1	47.0	45.5
Black	7.9	9.0	62.9	20.2
Hawaiian/Pacific Islander	6.7	13.3	46.7	33.3
Hispanic	3.7	17.2	59.3	19.8
White	2.1	8.0	49.4	40.4
Multiracial	2.3	8.3	53.8	35.6
Unknown	12.5	12.5	62.5	12.5
Free/Reduced Lunch	3.5	14.2	56.7	25.6
Not Free/Reduced Lunch	2.1	7.4	47.6	42.9
Special Education	8.0	23.3	50.4	18.3
Not Special Education	1.7	7.7	51.2	39.3
English Language Learner	7.9	25.8	55.6	10.7
Not English Language Learner	2.4	9.2	50.8	37.6

Table L9. Performance Levels of Mathematics by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	6.2	12.9	53.6	27.4
Male	6.1	12.1	54.0	27.8
Female	6.2	13.6	53.2	27.0
Unknown	6.3	31.3	43.8	18.8
American Indian	18.2	26.1	41.8	13.9
Asian	4.1	11.0	53.4	31.5
Black	13.6	18.5	50.6	17.3
Hawaiian/Pacific Islander	15.4	0.0	69.2	15.4
Hispanic	10.5	18.5	55.3	15.7
White	4.8	11.4	53.7	30.2
Multiracial	6.4	7.3	60.6	25.7
Unknown	5.6	11.1	66.7	16.7
Free/Reduced Lunch	8.2	17.0	55.6	19.2
Not Free/Reduced Lunch	5.0	10.5	52.4	32.2
Special Education	18.6	22.0	49.0	10.4
Not Special Education	4.2	11.4	54.3	30.1
English Language Learner	27.3	28.9	40.2	3.6
Not English Language Learner	5.6	12.4	53.9	28.1

Table L10. Performance Levels of Mathematics by Grade 5 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	5.0	15.5	48.5	31.0
Male	5.0	14.7	48.7	31.6
Female	4.9	16.3	48.4	30.4
Unknown	28.6	42.9	28.6	0.0
American Indian	12.2	24.3	45.6	17.9
Asian	2.1	2.1	38.3	57.4
Black	19.0	15.2	48.1	17.7
Hawaiian/Pacific Islander	25.0	18.8	43.8	12.5
Hispanic	8.8	22.0	49.7	19.6
White	3.7	13.8	48.8	33.8
Multiracial	7.5	28.0	40.9	23.7
Unknown	27.8	22.2	33.3	16.7
Free/Reduced Lunch	6.7	20.0	50.7	22.6
Not Free/Reduced Lunch	4.0	12.8	47.2	35.9
Special Education	18.1	31.1	39.8	11.1
Not Special Education	2.9	13.0	49.9	34.2
English Language Learner	20.2	34.5	39.9	5.4
Not English Language Learner	4.6	15.0	48.7	31.7

Table L11. Performance Levels of Mathematics by Grade 6 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	5.7	12.7	53.5	28.1
Male	6.2	13.5	51.3	29.1
Female	5.3	11.9	55.9	27.0
Unknown	16.7	16.7	16.7	50.0
American Indian	17.3	30.9	44.9	7.0
Asian	1.9	1.9	51.9	44.4
Black	16.7	15.5	54.8	13.1
Hawaiian/Pacific Islander	17.6	11.8	64.7	5.9
Hispanic	7.9	19.6	55.0	17.5
White	4.6	10.9	53.5	31.1
Multiracial	8.7	9.7	60.2	21.4
Unknown	22.2	5.6	50.0	22.2
Free/Reduced Lunch	7.4	16.8	55.9	19.9
Not Free/Reduced Lunch	4.8	10.4	52.2	32.6
Special Education	21.0	31.0	41.2	6.8
Not Special Education	3.6	10.1	55.2	31.1
English Language Learner	20.0	34.2	42.6	3.2
Not English Language Learner	5.4	12.2	53.7	28.7

Table L12. Performance Levels of Mathematics by Grade 7 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	5.2	19.0	50.2	25.6
Male	5.4	18.4	49.0	27.2
Female	4.9	19.6	51.6	23.8
Unknown	50.0	25.0	25.0	0.0
American Indian	14.2	30.1	49.8	5.9
Asian	0.0	7.7	51.9	40.4
Black	11.4	35.2	41.9	11.4
Hawaiian/Pacific Islander	6.7	40.0	26.7	26.7
Hispanic	7.9	28.1	50.3	13.6
White	4.2	16.8	50.4	28.6
Multiracial	8.4	21.0	51.3	19.3
Unknown	25.0	25.0	50.0	0.0
Free/Reduced Lunch	7.4	26.2	51.5	14.9
Not Free/Reduced Lunch	4.1	15.3	49.6	31.1
Special Education	20.4	39.6	34.1	5.8
Not Special Education	3.0	16.0	52.5	28.4
English Language Learner	17.6	47.1	34.0	1.3
Not English Language Learner	4.9	18.4	50.6	26.1

Table L13. Performance Levels of Mathematics by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	10.3	22.4	47.7	19.6
Male	10.4	22.7	46.9	19.9
Female	10.2	22.0	48.6	19.2
Unknown	16.7	50.0	33.3	0.0
American Indian	27.6	34.6	32.7	5.1
Asian	6.8	11.9	47.5	33.9
Black	20.5	33.3	39.7	6.4
Hawaiian/Pacific Islander	27.3	45.5	27.3	0.0
Hispanic	15.4	27.5	46.5	10.7
White	8.7	20.9	48.8	21.6
Multiracial	12.5	25.0	40.6	21.9
Unknown	21.4	50.0	21.4	7.1
Free/Reduced Lunch	14.5	27.4	46.5	11.6
Not Free/Reduced Lunch	8.2	19.9	48.3	23.6
Special Education	36.3	35.8	24.3	3.6
Not Special Education	6.7	20.5	51.0	21.8
English Language Learner	30.3	37.9	28.8	3.0
Not English Language Learner	9.9	22.1	48.1	19.9

Table L14. Performance Levels of Science by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	8.0	34.6	45.6	11.8
Male	8.5	34.9	45.3	11.4
Female	7.4	34.2	46.0	12.3
Unknown	13.3	46.7	33.3	6.7
American Indian	17.5	52.0	27.6	2.9
Asian	5.6	31.0	49.3	14.1
Black	22.0	40.2	26.8	11.0
Hawaiian/Pacific Islander	21.4	42.9	28.6	7.1
Hispanic	11.1	47.8	35.9	5.1
White	6.7	31.5	48.4	13.4
Multiracial	8.3	33.9	44.0	13.8
Unknown	21.1	47.4	26.3	5.3
Free/Reduced Lunch	11.1	42.0	40.0	6.8
Not Free/Reduced Lunch	6.1	30.1	48.9	14.8
Special Education	14.7	50.0	31.1	4.2
Not Special Education	6.9	32.1	48.0	13.1
English Language Learner	20.6	60.8	18.1	0.5
Not English Language Learner	7.6	33.8	46.4	12.2

Table L15. Performance Levels of Science by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	14.1	42.2	34.7	8.9
Male	15.1	40.9	34.7	9.3
Female	12.9	43.7	34.9	8.5
Unknown	50.0	50.0	0.0	0.0
American Indian	38.0	47.9	12.7	1.4
Asian	8.6	32.8	39.7	19.0
Black	32.5	48.1	18.2	1.3
Hawaiian/Pacific Islander	54.5	36.4	9.1	0.0
Hispanic	22.8	51.0	22.2	4.0
White	11.3	40.8	37.8	10.0
Multiracial	23.5	34.3	31.4	10.8
Unknown	41.7	50.0	8.3	0.0
Free/Reduced Lunch	20.1	48.3	27.2	4.4
Not Free/Reduced Lunch	11.0	39.1	38.7	11.3
Special Education	38.2	47.6	12.9	1.3
Not Special Education	10.7	41.5	37.8	10.0
English Language Learner	42.9	45.1	12.0	0.0
Not English Language Learner	13.5	42.2	35.2	9.1

Appendix M: SAWS Field Test Demographic Performance

Table M1. SAWS Field Test Gender and Ethnicity Frequency Representation – Grade 3

	<u>Form 1</u>		<u>Form 2</u>		<u>Form 3</u>		<u>Form 4</u>		<u>Form 5</u>		<u>Form 6</u>	
	N	%	N	%	N	%	N	%	N	%	N	%
Total	1228	100.0	1207	100.0	1209	100.0	1206	100.0	1222	100.0	1219	100.0
Male	647	52.7	621	51.5	619	51.2	626	51.9	617	50.5	655	53.7
Female	574	46.7	583	48.3	586	48.5	575	47.7	601	49.2	561	46.0
Unknown	7	0.6	3	0.2	4	0.3	5	0.4	4	0.3	3	0.2
American Indian	44	3.6	49	4.1	42	3.5	41	3.4	51	4.2	56	4.6
Asian	8	0.7	19	1.6	15	1.2	8	0.7	8	0.7	7	0.6
Black	15	1.2	14	1.2	19	1.6	12	1.0	17	1.4	16	1.3
Haw. Pac. Islander	2	0.2	3	0.2	0	0.0	7	0.6	1	0.1	3	0.2
Hispanic	161	13.1	164	13.6	159	13.2	177	14.7	172	14.1	177	14.5
White	971	79.1	926	76.7	931	77.0	937	77.7	944	77.3	937	76.9
Multiracial	16	1.3	26	2.2	35	2.9	19	1.6	22	1.8	17	1.4
Unknown	11	0.9	6	0.5	8	0.7	5	0.4	7	0.6	6	0.5
Free Lunch	445	36.2	474	39.3	430	35.6	453	37.6	472	38.6	470	38.6
Not Free Lunch	783	63.8	733	60.7	779	64.4	753	62.4	750	61.4	749	61.4
Special Education	179	14.6	172	14.3	165	13.6	167	13.8	178	14.6	179	14.7
Not Special Education	1049	85.4	1035	85.8	1044	86.4	1039	86.2	1044	85.4	1040	85.3
English Language Learner	58	4.7	48	4.0	56	4.6	69	5.7	56	4.6	66	5.4
Not English Language Learner	1170	95.3	1159	96.0	1153	95.4	1137	94.3	1166	95.4	1153	94.6

Table M2. SAWS Field Test Gender and Ethnicity Frequency Representation – Grade 5

	Form 1		Form 2		Form 3		Form 4		Form 5		Form 6	
	N	%	N	%	N	%	N	%	N	%	N	%
Total	1162	100.0	1145	100.0	1152	100.0	1156	100.0	1145	100.0	1149	100.0
Male	580	49.9	580	50.7	591	51.3	629	54.4	608	53.1	587	51.1
Female	578	49.7	564	49.3	559	48.5	524	45.3	535	46.7	556	48.4
Unknown	4	0.3	1	0.1	2	0.2	3	0.3	2	0.2	6	0.5
American Indian	47	4.0	38	3.3	58	5.0	47	4.1	46	4.0	38	3.3
Asian	5	0.4	8	0.7	10	0.9	12	1.0	8	0.7	7	0.6
Black	12	1.0	17	1.5	13	1.1	12	1.0	16	1.4	11	1.0
Haw. Pac. Islander	4	0.3	1	0.1	2	0.2	3	0.3	4	0.3	3	0.3
Hispanic	161	13.9	156	13.6	143	12.4	144	12.5	143	12.5	161	14.0
White	915	78.7	905	79.0	902	78.3	919	79.5	910	79.5	900	78.3
Multiracial	9	0.8	14	1.2	21	1.8	14	1.2	14	1.2	20	1.7
Unknown	9	0.8	6	0.5	3	0.3	5	0.4	4	0.3	9	0.8
Free Lunch	401	34.5	430	37.6	426	37.0	423	36.6	416	36.3	407	35.4
Not Free Lunch	761	65.5	715	62.4	726	63.0	733	63.4	729	63.7	742	64.6
Special Education	160	13.8	133	11.6	160	13.9	176	15.2	155	13.5	161	14.0
Not Special Education	1002	86.2	1012	88.4	992	86.1	980	84.8	990	86.5	988	86.0
English Language Learner	22	1.9	34	3.0	20	1.7	23	2.0	30	2.6	32	2.8
Not English Language Learner	1140	98.1	1111	97.0	1132	98.3	1133	98.0	1115	97.4	1117	97.2

Table M3. SAWS Field Test Gender and Ethnicity Frequency Representation – Grade 7

	<u>Form 1</u>		<u>Form 2</u>		<u>Form 3</u>		<u>Form 4</u>		<u>Form 5</u>		<u>Form 6</u>	
	N	%	N	%	N	%	N	%	N	%	N	%
Total	1168	100.0	1159	100.0	1149	100.0	1167	100.0	1175	100.0	1153	100.0
Male	634	54.3	597	51.5	629	54.7	601	51.5	587	50.0	616	53.4
Female	532	45.5	562	48.5	520	45.3	564	48.3	586	49.9	537	46.6
Unknown	2	0.2	0	0.0	0	0.0	2	0.2	2	0.2	0	0.0
American Indian	42	3.6	40	3.5	35	3.0	39	3.3	37	3.1	41	3.6
Asian	6	0.5	10	0.9	12	1.0	6	0.5	14	1.2	4	0.3
Black	12	1.0	18	1.6	15	1.3	24	2.1	20	1.7	21	1.8
Haw. Pac. Islander	2	0.2	2	0.2	2	0.2	2	0.2	2	0.2	4	0.3
Hispanic	155	13.3	160	13.8	147	12.8	149	12.8	141	12.0	141	12.2
White	928	79.5	912	78.7	915	79.6	925	79.3	931	79.2	920	79.8
Multiracial	20	1.7	16	1.4	19	1.7	18	1.5	25	2.1	19	1.6
Unknown	3	0.3	1	0.1	4	0.3	4	0.3	5	0.4	3	0.3
Free Lunch	404	34.6	381	32.9	372	32.4	395	33.8	397	33.8	395	34.3
Not Free Lunch	764	65.4	778	67.1	777	67.6	772	66.2	778	66.2	758	65.7
Special Education	165	14.1	145	12.5	144	12.5	156	13.4	141	12.0	143	12.4
Not Special Education	1003	85.9	1014	87.5	1005	87.5	1011	86.6	1034	88.0	1010	87.6
English Language Learner	32	2.7	30	2.6	20	1.7	30	2.6	21	1.8	28	2.4
Not English Language Learner	1136	97.3	1129	97.4	1129	98.3	1137	97.4	1154	98.2	1125	97.6

Table M4. SAWS Field Test Gender and Ethnicity Performance by 12-point Prompt - Grade 3

Group	<u>Form 1</u>			<u>Form 2</u>			<u>Form 3</u>			<u>Form 4</u>			<u>Form 5</u>			<u>Form 6</u>		
	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD
Total Group	1228	6.87	2.88	1207	6.60	2.88	1209	6.66	2.63	1206	6.95	2.69	1222	6.59	2.79	1219	6.61	2.62
Male	647	6.36	2.77	621	6.09	2.73	619	6.30	2.62	626	6.63	2.57	617	6.05	2.77	655	6.20	2.51
Female	574	7.47	2.84	583	7.16	2.91	586	7.04	2.59	575	7.31	2.76	601	7.16	2.70	561	7.11	2.64
Unknown	7	5.00	5.23	3	1.33	2.31	4	5.50	3.79	5	5.40	3.58	4	4.25	3.10	3	2.00	3.46
American Indian	44	5.50	2.66	49	4.69	2.97	42	5.52	2.74	41	5.95	2.53	51	5.14	2.92	56	6.11	2.36
Asian	8	7.75	1.98	19	8.53	2.09	15	7.00	2.95	8	7.88	1.36	8	8.13	1.55	7	8.71	2.21
Black	15	4.27	2.66	14	7.71	1.59	19	7.68	2.85	12	5.92	2.35	17	4.53	2.37	16	6.44	2.25
Haw. Pac. Islander	2	4.50	3.54	3	9.33	2.08				7	5.86	3.08	1	1.00	.	3	4.67	5.03
Hispanic	161	5.86	2.86	164	5.74	2.69	159	6.16	2.61	177	6.61	2.57	172	6.21	2.76	177	6.13	2.53
White	971	7.14	2.81	926	6.83	2.83	931	6.77	2.58	937	7.09	2.71	944	6.78	2.75	937	6.74	2.63
Multiracial	16	7.63	2.09	26	5.88	2.85	35	6.43	3.15	19	6.11	2.47	22	6.23	3.15	17	6.82	2.38
Unknown	11	5.64	4.63	6	1.67	1.97	8	6.50	2.73	5	5.60	3.51	7	6.86	3.63	6	3.00	2.45
Not Free Lunch	783	7.22	2.81	733	6.90	2.90	779	6.98	2.53	753	7.20	2.67	750	7.01	2.68	749	6.96	2.56
Free Lunch	445	6.26	2.88	474	6.12	2.77	430	6.07	2.71	453	6.52	2.67	472	5.92	2.84	470	6.04	2.62
Not Special Education	1049	7.11	2.83	1035	6.88	2.79	1044	6.87	2.60	1039	7.12	2.67	1044	6.85	2.72	1040	6.81	2.57
Special Education	179	5.46	2.74	172	4.90	2.80	165	5.30	2.46	167	5.86	2.55	178	5.08	2.75	179	5.44	2.62
Not English Lang. Learner	1170	6.95	2.87	1159	6.62	2.89	1153	6.71	2.65	1137	6.99	2.72	1166	6.64	2.79	1153	6.66	2.62
English Lang. Learner	58	5.31	2.59	48	5.92	2.45	56	5.59	2.01	69	6.22	2.07	56	5.66	2.74	66	5.80	2.49

Table M5. Gender and Ethnicity Performance by 12-point Prompt - Grade 5

Group	Form 1			Form 2			Form 3			Form 4			Form 5			Form 6		
	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD
Total Group	1162	7.44	2.68	1145	6.93	2.72	1152	7.58	2.87	1156	7.48	2.80	1145	7.76	2.64	1149	6.64	2.60
Male	580	6.92	2.48	580	6.31	2.54	591	7.13	2.89	629	6.94	2.73	608	7.25	2.60	587	6.20	2.48
Female	578	7.95	2.77	564	7.57	2.73	559	8.07	2.77	524	8.14	2.73	535	8.37	2.53	556	7.12	2.63
Unknown	4	7.00	4.83	1	0.00	.	2	4.50	0.71	3	5.00	4.36	2	2.00	2.83	6	5.33	3.93
American Indian	47	6.62	2.35	38	5.79	2.34	58	6.29	2.64	47	6.17	2.78	46	6.22	2.15	38	5.18	1.66
Asian	5	9.60	2.30	8	8.63	3.50	10	8.80	2.97	12	7.83	3.07	8	9.38	4.17	7	8.86	2.79
Black	12	6.17	2.69	17	6.35	3.18	13	6.08	4.17	12	5.92	2.19	16	6.50	2.00	11	4.00	2.19
Haw. Pac. Islander	4	9.75	2.06	1	7.00	.	2	5.00	7.07	3	5.33	2.31	4	7.50	1.73	3	6.00	2.00
Hispanic	161	7.11	2.37	156	6.59	2.39	143	7.11	2.76	144	7.07	2.87	143	7.32	2.37	161	6.12	2.47
White	915	7.53	2.74	905	7.05	2.75	902	7.75	2.85	919	7.66	2.75	910	7.92	2.65	900	6.85	2.58
Multiracial	9	7.44	1.88	14	6.71	2.55	21	7.76	2.72	14	6.14	2.82	14	9.00	2.39	20	5.40	3.23
Unknown	9	7.00	3.24	6	4.17	3.19	3	7.33	2.89	5	6.80	4.32	4	4.00	3.27	9	6.11	3.52
Not Free Lunch	761	7.65	2.71	715	7.19	2.74	726	7.84	2.93	733	7.74	2.83	729	7.96	2.69	742	6.97	2.62
Free Lunch	401	7.03	2.58	430	6.48	2.61	426	7.14	2.71	423	7.03	2.68	416	7.41	2.52	407	6.04	2.45
Not Special Education	1002	7.69	2.64	1012	7.14	2.66	992	7.85	2.81	980	7.79	2.71	990	7.96	2.60	988	6.90	2.56
Special Education	160	5.85	2.42	133	5.29	2.55	160	5.89	2.68	176	5.74	2.61	155	6.48	2.55	161	5.09	2.27
Not English Lang. Learner	1140	7.47	2.67	1111	6.98	2.72	1132	7.62	2.85	1133	7.51	2.80	1115	7.81	2.64	1117	6.70	2.59
English Lang. Learner	22	5.68	2.53	34	5.12	2.00	20	5.10	2.79	23	5.91	2.23	30	5.90	2.04	32	4.56	2.09

Table M6. Gender and Ethnicity Performance by 4-point Prompt – Grade 5

Group	Form 1			Form 2			Forms 3&6			Form 4			Form 5		
	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD
Total Group	1162	2.99	1.10	1145	2.92	1.08	2301	3.19	1.02	1156	2.90	1.16	1145	3.23	0.99
Male	580	2.86	1.10	580	2.72	1.07	1178	3.05	1.04	629	2.78	1.17	608	3.15	0.99
Female	578	3.13	1.07	564	3.14	1.05	1115	3.33	0.98	524	3.05	1.12	535	3.34	0.97
Unknown	4	2.00	1.63	1	0.00	.	8	2.63	1.41	3	1.33	2.31	2	1.50	2.12
American Indian	47	2.70	1.10	38	2.71	1.14	96	3.11	1.02	47	2.68	1.25	46	3.17	0.90
Asian	5	3.60	0.55	8	3.25	0.71	17	3.53	0.80	12	2.83	1.27	8	2.88	1.46
Black	12	2.83	0.94	17	3.06	0.97	24	2.63	1.44	12	3.17	1.03	16	3.19	0.98
Haw. Pac. Islander	4	2.75	1.26	1	2.00	.	5	3.20	1.79	3	2.00	1.00	4	3.50	0.58
Hispanic	161	2.90	1.07	156	2.87	1.00	304	3.05	1.10	144	2.87	1.25	143	3.11	1.06
White	915	3.03	1.10	905	2.94	1.10	1802	3.22	0.99	919	2.93	1.13	910	3.27	0.98
Multiracial	9	2.89	1.27	14	3.07	0.73	41	3.07	1.08	14	2.71	1.54	14	3.14	0.86
Unknown	9	2.67	1.32	6	1.67	1.21	12	3.00	1.13	5	1.60	1.52	4	2.25	1.71
Not Free Lunch	761	3.05	1.07	715	3.01	1.06	1468	3.23	1.00	733	2.99	1.14	729	3.31	0.98
Free Lunch	401	2.88	1.14	430	2.78	1.11	833	3.11	1.06	423	2.75	1.17	416	3.10	1.00
Not Special Education	1002	3.08	1.06	1012	3.00	1.07	1980	3.26	0.99	980	2.99	1.14	990	3.31	0.96
Special Education	160	2.41	1.12	133	2.38	1.06	321	2.73	1.11	176	2.41	1.17	155	2.78	1.05
Not English Lang. Learner	1140	3.01	1.09	1111	2.94	1.08	2249	3.20	1.01	1133	2.91	1.16	1115	3.25	0.99
English Lang. Learner	22	2.14	0.94	34	2.56	1.05	52	2.52	1.23	23	2.70	1.06	30	2.77	0.90

Table M7. Gender and Ethnicity Performance by 8-point Prompt – Grade 5

Group	Form 1			Form 2			Form 3 & 6			Form 4			Form 5		
	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD
Total Group	1162	5.00	1.85	1145	4.59	1.96	2301	4.91	1.85	1156	4.52	1.92	1145	4.66	1.92
Male	580	4.67	1.80	580	4.27	1.90	1178	4.63	1.82	629	4.29	1.89	608	4.31	1.92
Female	578	5.34	1.84	564	4.92	1.97	1115	5.20	1.83	524	4.82	1.89	535	5.07	1.84
Unknown	4	3.75	2.63	1	0.00	.	8	4.63	2.33	3	1.33	2.31	2	2.00	2.83
American Indian	47	4.70	2.03	38	3.63	2.07	96	4.19	1.69	47	3.68	1.93	46	3.61	1.74
Asian	5	6.80	1.30	8	5.25	1.67	17	5.82	1.67	12	4.42	2.15	8	5.00	2.27
Black	12	4.67	2.15	17	4.18	1.55	24	3.71	2.39	12	4.08	2.02	16	4.44	1.90
Haw. Pac. Islander	4	5.50	2.38	1	8.00	.	5	4.40	2.70	3	2.67	1.53	4	3.50	2.38
Hispanic	161	4.83	1.64	156	4.37	1.87	304	4.62	1.84	144	4.10	2.00	143	4.50	1.74
White	915	5.04	1.87	905	4.68	1.96	1802	5.01	1.83	919	4.66	1.87	910	4.76	1.94
Multiracial	9	4.89	1.45	14	4.14	2.03	41	4.49	1.82	14	4.21	2.01	14	4.57	1.50
Unknown	9	4.44	1.94	6	3.33	2.66	12	5.25	2.30	5	3.20	3.27	4	2.50	1.91
Not Free Lunch	761	5.14	1.83	715	4.83	1.98	1468	5.06	1.85	733	4.64	1.88	729	4.88	1.93
Free Lunch	401	4.74	1.87	430	4.18	1.86	833	4.63	1.82	423	4.32	1.96	416	4.28	1.85
Not Special Education	1002	5.20	1.80	1012	4.74	1.92	1980	5.08	1.81	980	4.72	1.85	990	4.81	1.90
Special Education	160	3.76	1.65	133	3.40	1.91	321	3.85	1.74	176	3.40	1.87	155	3.70	1.77
Not English Lang. Learner	1140	5.02	1.85	1111	4.63	1.95	2249	4.93	1.84	1133	4.54	1.91	1115	4.70	1.92
English Lang. Learner	22	3.95	1.33	34	3.21	1.87	52	3.79	1.79	23	3.57	1.83	30	3.43	1.74

Table M8. Gender and Ethnicity Performance by 12-point Prompt – Grade 7

Group	Form 1			Form 2			Form 3			Form 4			Form 5			Form 6		
	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD
Total Group	1168	7.68	2.86	1159	8.02	2.89	1149	7.73	2.77	1167	7.74	2.94	1175	7.61	2.88	1153	7.46	2.83
Male	634	7.13	2.85	597	7.38	2.71	629	7.13	2.59	601	7.29	2.82	587	6.92	2.82	616	6.74	2.78
Female	532	8.35	2.72	562	8.70	2.93	520	8.45	2.80	564	8.23	2.99	586	8.29	2.78	537	8.29	2.65
Unknown	2	4.00	5.66							2	7.00	0.00	2	9.50	3.54			
American Indian	42	5.90	3.04	40	6.18	2.33	35	6.51	2.37	39	5.79	2.79	37	5.86	2.96	41	6.41	2.91
Asian	6	10.00	1.90	10	8.60	2.01	12	9.42	1.68	6	7.67	2.25	14	8.50	2.47	4	8.50	1.73
Black	12	6.83	3.49	18	8.83	1.69	15	6.93	3.58	24	6.79	2.55	20	6.00	2.75	21	6.19	3.68
Haw. Pac. Islander	2	8.50	3.54	2	2.00	2.83	2	8.00	4.24	2	8.50	4.95	2	5.00	7.07	4	7.50	1.73
Hispanic	155	7.33	2.71	160	7.39	2.82	147	7.21	2.88	149	7.27	2.96	141	7.01	2.77	141	7.27	2.85
White	928	7.83	2.84	912	8.20	2.90	915	7.86	2.75	925	7.93	2.93	931	7.83	2.85	920	7.55	2.79
Multiracial	20	7.50	2.42	16	8.19	3.39	19	7.63	2.01	18	7.83	1.92	25	6.12	2.55	19	8.11	2.66
Unknown	3	3.00	4.36	1	8.00	.	4	6.25	1.71	4	6.50	1.73	5	7.40	2.97	3	8.00	4.00
Not Free Lunch	764	7.93	2.82	778	8.36	2.80	777	8.08	2.78	772	8.02	2.92	778	7.97	2.82	758	7.79	2.81
Free Lunch	404	7.20	2.89	381	7.32	2.96	372	6.99	2.60	395	7.22	2.91	397	6.90	2.88	395	6.84	2.75
Not Special Education	1003	8.03	2.77	1014	8.36	2.76	1005	7.99	2.71	1011	8.06	2.85	1034	7.91	2.80	1010	7.74	2.76
Special Education	165	5.55	2.51	145	5.66	2.68	144	5.94	2.49	156	5.67	2.64	141	5.41	2.48	143	5.48	2.48
Not English Lang. Learner	1136	7.74	2.85	1129	8.07	2.89	1129	7.76	2.76	1137	7.80	2.93	1154	7.65	2.86	1125	7.52	2.80
English Lang. Learner	32	5.63	2.51	30	6.03	2.37	20	5.90	2.40	30	5.70	2.41	21	5.10	2.77	28	5.11	3.06

Table M9. Gender and Ethnicity Performance by 4-point Prompt – Grade 7

Group	Form 1			Form 2			Form 3			Form 4			Form 5			Form 6		
	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD
Total Group	1168	2.99	1.11	1159	2.88	1.18	1149	2.95	1.16	1167	2.69	1.21	1175	2.84	1.14	1153	2.90	1.05
Male	634	2.87	1.15	597	2.78	1.15	629	2.75	1.17	601	2.55	1.18	587	2.66	1.17	616	2.75	1.14
Female	532	3.13	1.04	562	2.99	1.20	520	3.19	1.09	564	2.85	1.22	586	3.03	1.08	537	3.08	0.91
Unknown	2	2.00	2.83							2	1.50	0.71	2	3.50	0.71			
American Indian	42	2.52	1.25	40	2.43	1.13	35	2.63	1.24	39	2.13	1.32	37	2.19	1.39	41	2.61	1.02
Asian	6	3.83	0.41	10	3.00	1.15	12	3.33	0.89	6	2.33	1.37	14	3.14	0.95	4	3.25	0.96
Black	12	2.50	1.31	18	2.72	1.18	15	2.93	1.39	24	2.17	1.09	20	3.00	1.26	21	2.57	1.25
Haw. Pac. Islander	2	3.50	0.71	2	0.50	0.71	2	4.00	0.00	2	3.00	1.41	2	1.50	2.12	4	2.75	0.96
Hispanic	155	2.66	1.10	160	2.54	1.24	147	2.77	1.19	149	2.60	1.22	141	2.50	1.16	141	2.87	1.03
White	928	3.06	1.09	912	2.96	1.16	915	2.98	1.15	925	2.75	1.20	931	2.93	1.10	920	2.92	1.05
Multiracial	20	3.05	1.00	16	3.13	1.09	19	3.05	0.97	18	2.72	1.13	25	2.36	1.29	19	3.05	0.85
Unknown	3	2.00	2.00	1	4.00	.	4	2.00	0.82	4	2.00	0.82	5	2.80	0.84	3	3.00	1.00
Not Free Lunch	764	3.09	1.07	778	3.02	1.13	777	3.05	1.14	772	2.77	1.20	778	2.97	1.10	758	2.97	1.06
Free Lunch	404	2.79	1.17	381	2.59	1.23	372	2.73	1.17	395	2.53	1.20	397	2.60	1.18	395	2.77	1.02
Not Special Education	1003	3.11	1.07	1014	2.98	1.15	1005	3.04	1.13	1011	2.81	1.19	1034	2.92	1.11	1010	2.98	1.02
Special Education	165	2.22	1.08	145	2.14	1.16	144	2.33	1.14	156	1.90	1.01	141	2.26	1.18	143	2.39	1.10
Not English Lang. Learner	1136	3.01	1.11	1129	2.90	1.18	1129	2.96	1.16	1137	2.71	1.21	1154	2.86	1.13	1125	2.92	1.05
English Lang. Learner	32	2.28	1.11	30	1.97	1.07	20	2.20	0.89	30	2.10	1.09	21	2.05	1.24	28	2.36	1.03

Table M10. Gender and Ethnicity Performance by 8-point Prompt – Grade 7

Group	Form 1			Form 2			Form 3			Form 4			Form 5			Form 6		
	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD	N	Mean	STD
Total Group	1168	5.26	1.90	1159	5.39	1.90	1149	5.05	1.88	1167	5.08	2.06	1175	5.04	2.04	1153	4.83	1.97
Male	634	5.00	1.91	597	5.02	1.83	629	4.74	1.86	601	4.74	2.00	587	4.68	2.04	616	4.42	1.99
Female	532	5.58	1.84	562	5.80	1.89	520	5.42	1.84	564	5.43	2.07	586	5.39	1.98	537	5.30	1.83
Unknown	2	3.50	4.95							2	5.50	0.71	2	6.50	0.71			
American Indian	42	4.02	2.18	40	4.93	1.72	35	4.54	1.93	39	3.87	2.17	37	3.92	2.25	41	4.44	1.92
Asian	6	6.50	1.22	10	6.00	1.76	12	5.92	1.83	6	5.50	0.55	14	5.86	1.75	4	5.00	1.15
Black	12	3.67	2.35	18	5.61	2.06	15	5.13	2.13	24	4.54	1.67	20	4.45	2.21	21	3.81	2.11
Haw. Pac. Islander	2	7.00	1.41	2	2.50	3.54	2	5.00	4.24	2	7.00	1.41	2	2.50	3.54	4	3.50	1.73
Hispanic	155	4.86	1.98	160	4.91	1.94	147	4.73	1.97	149	4.70	2.04	141	4.56	2.12	141	4.33	1.86
White	928	5.39	1.83	912	5.48	1.89	915	5.11	1.86	925	5.20	2.05	931	5.19	1.98	920	4.95	1.97
Multiracial	20	5.75	1.77	16	6.00	1.59	19	5.05	1.27	18	4.78	2.26	25	4.04	1.93	19	4.84	1.64
Unknown	3	3.33	3.51	1	6.00	.	4	4.50	0.58	4	5.25	0.50	5	5.00	3.16	3	6.00	1.00
Not Free Lunch	764	5.44	1.84	778	5.59	1.86	777	5.29	1.86	772	5.26	2.03	778	5.29	1.97	758	5.08	1.98
Free Lunch	404	4.93	1.98	381	5.00	1.94	372	4.53	1.83	395	4.71	2.07	397	4.56	2.09	395	4.36	1.86
Not Special Education	1003	5.47	1.83	1014	5.61	1.80	1005	5.20	1.84	1011	5.32	1.99	1034	5.25	1.98	1010	4.99	1.90
Special Education	165	4.01	1.86	145	3.89	1.93	144	3.95	1.82	156	3.50	1.84	141	3.53	1.88	143	3.73	2.05
Not English Lang. Learner	1136	5.30	1.89	1129	5.42	1.90	1129	5.07	1.87	1137	5.11	2.05	1154	5.07	2.03	1125	4.86	1.96
English Lang. Learner	32	4.09	1.94	30	4.37	1.43	20	3.50	1.70	30	3.90	2.06	21	3.33	1.93	28	3.54	1.97

Appendix N: SAWS Field Test Rater Reliability

Table N1. Grade 3 Field Test Trait Rater Agreement and Weighted Kappa – 12-point Prompt

Form	Trait	N	Rater 1		Rater 2		CORR	Percentages of Agreement		Weighted Kappa
			Mean	SD	Mean	SD		Exact	Exact + adjacent agreement only	
1	Total	289	6.86	2.61	6.82	2.66	0.72	26.64	59.17	
	Idea Development	289	1.85	0.72	1.84	0.72	0.60	58.82	99.65	0.60
	Organization	289	1.68	0.76	1.70	0.78	0.59	55.36	98.62	0.59
	Voice	289	1.70	0.73	1.68	0.75	0.64	61.25	99.65	0.64
	Conventions	289	1.63	0.73	1.61	0.77	0.66	67.13	98.27	0.66
2	Total	274	7.16	2.59	6.91	2.55	0.68	24.09	51.82	
	Idea Development	274	1.92	0.76	1.88	0.73	0.51	55.47	96.72	0.51
	Organization	274	1.71	0.76	1.57	0.78	0.59	54.38	98.54	0.58
	Voice	274	1.80	0.72	1.78	0.73	0.55	56.93	98.54	0.55
	Conventions	274	1.73	0.72	1.68	0.74	0.60	60.22	98.91	0.60
3	Total	301	6.63	2.57	6.50	2.74	0.75	32.56	64.78	
	Idea Development	301	1.71	0.74	1.72	0.75	0.61	63.79	97.67	0.61
	Organization	301	1.65	0.72	1.58	0.76	0.61	62.46	98.01	0.60
	Voice	301	1.65	0.72	1.63	0.75	0.64	63.79	99.00	0.63
	Conventions	301	1.62	0.71	1.58	0.78	0.65	64.45	98.67	0.65
4	Total	287	7.35	2.53	7.12	2.59	0.73	28.22	58.89	
	Idea Development	287	1.97	0.70	1.89	0.74	0.54	56.45	98.26	0.53
	Organization	287	1.78	0.71	1.74	0.76	0.60	62.37	97.91	0.60
	Voice	287	1.81	0.74	1.77	0.72	0.62	62.37	98.95	0.62
	Conventions	287	1.78	0.74	1.72	0.76	0.62	61.32	98.61	0.62
5	Total	278	6.91	2.60	6.89	2.54	0.72	28.42	58.63	
	Idea Development	278	1.79	0.74	1.81	0.72	0.60	60.43	98.92	0.60
	Organization	278	1.67	0.76	1.65	0.74	0.57	57.55	98.20	0.57
	Voice	278	1.73	0.71	1.72	0.72	0.59	60.43	99.28	0.59
	Conventions	278	1.72	0.75	1.71	0.73	0.67	67.27	98.92	0.67
6	Total	294	6.64	2.52	6.42	2.49	0.78	32.65	62.59	
	Idea Development	294	1.81	0.73	1.71	0.73	0.65	63.95	99.32	0.64
	Organization	294	1.59	0.74	1.48	0.72	0.66	66.67	98.64	0.65
	Voice	294	1.68	0.70	1.66	0.71	0.66	67.01	99.66	0.66
	Conventions	294	1.55	0.70	1.57	0.75	0.65	64.29	99.66	0.65

Table N2. Grade 5 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	280	7.72	2.60	7.68	2.85	0.73	32.86	58.21	
	Idea Development	280	2.03	0.74	2.04	0.76	0.61	59.64	98.93	0.61
	Organization	280	1.92	0.76	1.88	0.82	0.62	63.57	96.07	0.62
	Voice	280	1.90	0.74	1.92	0.81	0.60	55.00	98.93	0.60
	Conventions	280	1.86	0.75	1.84	0.82	0.64	60.36	98.21	0.64
2	Total	275	6.73	2.65	6.65	2.58	0.66	22.91	55.27	
	Idea Development	275	1.74	0.76	1.76	0.75	0.51	53.82	97.09	0.51
	Organization	275	1.68	0.76	1.64	0.75	0.56	56.73	97.82	0.56
	Voice	275	1.66	0.75	1.63	0.71	0.53	54.55	98.18	0.52
	Conventions	275	1.65	0.70	1.62	0.71	0.58	60.00	99.27	0.58
3	Total	279	7.66	2.88	7.62	2.83	0.82	29.75	66.67	
	Idea Development	279	1.97	0.78	1.97	0.77	0.73	69.18	99.28	0.73
	Organization	279	1.90	0.81	1.85	0.83	0.73	65.23	99.28	0.73
	Voice	279	1.91	0.77	1.94	0.78	0.70	65.59	99.64	0.70
	Conventions	279	1.87	0.80	1.85	0.78	0.67	63.08	98.57	0.67
4	Total	286	7.34	2.73	7.41	2.72	0.74	33.22	59.79	
	Idea Development	286	1.89	0.75	1.92	0.76	0.64	63.99	98.25	0.64
	Organization	286	1.82	0.77	1.82	0.79	0.65	59.09	99.30	0.65
	Voice	286	1.83	0.73	1.88	0.76	0.60	58.74	98.95	0.60
	Conventions	286	1.80	0.74	1.79	0.80	0.65	61.89	98.95	0.65
5	Total	276	8.16	2.58	7.97	2.60	0.74	30.43	64.13	
	Idea Development	276	2.12	0.75	2.05	0.72	0.66	63.41	99.64	0.65
	Organization	276	2.01	0.77	1.96	0.76	0.67	65.22	98.55	0.67
	Voice	276	2.08	0.70	1.99	0.72	0.60	60.87	99.28	0.59
	Conventions	276	1.95	0.73	1.97	0.72	0.53	56.16	98.19	0.53
6	Total	275	6.77	2.70	6.65	2.78	0.81	34.18	65.45	
	Idea Development	275	1.72	0.74	1.71	0.79	0.69	65.45	99.64	0.69
	Organization	275	1.71	0.76	1.69	0.79	0.69	64.73	99.27	0.69
	Voice	275	1.63	0.74	1.64	0.76	0.69	66.91	99.27	0.69
	Conventions	275	1.71	0.77	1.61	0.76	0.71	66.55	99.64	0.71

Table N3. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 4-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	292	2.89	1.14	2.95	1.14	0.75	57.88	93.49	
	Response to Text	292	1.33	0.71	1.36	0.71	0.67	69.86	98.97	0.67
	Holistic	292	1.56	0.59	1.59	0.58	0.62	75.34	99.66	0.62
2	Total	283	3.02	1.08	2.93	1.10	0.74	60.42	92.58	
	Response to Text	283	1.47	0.66	1.43	0.67	0.62	68.55	99.29	0.62
	Holistic	283	1.54	0.57	1.50	0.57	0.65	77.03	100.00	0.65
3&6	Total	575	3.12	1.05	3.10	1.06	0.78	63.30	96.52	
	Response to Text	575	1.54	0.65	1.54	0.64	0.75	81.04	99.48	0.75
	Holistic	575	1.58	0.57	1.56	0.58	0.59	73.91	99.83	0.59
4	Total	284	2.87	1.10	2.88	1.15	0.71	56.69	91.20	
	Response to Text	284	1.36	0.72	1.37	0.75	0.65	69.37	97.89	0.65
	Holistic	284	1.51	0.55	1.51	0.55	0.61	76.41	100.00	0.61
5	Total	285	3.17	1.05	3.19	1.07	0.77	68.07	94.39	
	Response to Text	285	1.58	0.64	1.57	0.63	0.71	78.95	99.30	0.71
	Holistic	285	1.59	0.55	1.62	0.55	0.65	78.60	100.00	0.65

Table N4. Grade 5 Field Test Trait Rater Agreement and Weighted Kappa – 8-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	295	5.05	1.81	5.08	1.76	0.76	35.59	81.02	
	Response to Text	295	1.40	0.68	1.39	0.66	0.81	83.39	100.00	0.81
	Holistic	295	3.65	1.39	3.69	1.41	0.67	38.31	84.41	0.67
2	Total	288	4.59	2.06	4.44	2.06	0.88	40.97	88.54	
	Response to Text	288	1.35	0.84	1.32	0.83	0.87	82.64	99.65	0.87
	Holistic	288	3.24	1.46	3.12	1.47	0.82	45.49	94.10	0.82
3	Total	572	5.07	1.79	5.08	1.75	0.84	46.50	88.99	
	Response to Text	572	1.69	0.67	1.70	0.66	0.88	91.26	99.30	0.88
	Holistic	572	3.38	1.39	3.39	1.35	0.76	48.95	90.04	0.76
4	Total	285	4.73	1.88	4.57	1.84	0.69	30.88	74.04	
	Response to Text	285	1.16	0.80	1.08	0.83	0.48	57.54	90.88	0.48
	Holistic	285	3.58	1.40	3.49	1.33	0.71	41.05	86.67	0.71
5	Total	281	4.57	1.89	4.57	1.83	0.72	30.96	70.82	
	Response to Text	281	1.27	0.79	1.27	0.76	0.74	70.46	99.29	0.73
	Holistic	281	3.30	1.42	3.30	1.37	0.65	33.45	80.78	0.65

Table N5. 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	279	7.72	2.88	7.64	2.89	0.76	29.75	60.57	
	Idea Development	279	1.92	0.79	1.88	0.81	0.67	64.16	97.85	0.67
	Organization	279	1.96	0.83	1.96	0.82	0.70	63.80	98.57	0.70
	Voice	279	1.91	0.77	1.89	0.79	0.64	57.71	99.28	0.64
	Conventions	279	1.94	0.79	1.91	0.78	0.64	59.86	98.57	0.64
2	Total	289	7.88	3.07	7.92	3.03	0.77	32.53	60.21	
	Idea Development	289	2.00	0.84	2.00	0.84	0.70	64.01	97.92	0.70
	Organization	289	1.97	0.84	1.97	0.85	0.69	62.63	97.92	0.69
	Voice	289	1.98	0.82	2.01	0.80	0.70	64.01	98.62	0.69
	Conventions	289	1.93	0.85	1.94	0.84	0.65	59.52	96.89	0.65
3	Total	280	7.96	2.59	7.89	2.73	0.73	32.14	67.14	
	Idea Development	280	2.01	0.75	1.97	0.76	0.64	64.64	98.21	0.64
	Organization	280	1.99	0.72	1.99	0.77	0.67	71.07	97.50	0.67
	Voice	280	2.03	0.72	1.99	0.77	0.62	62.86	98.21	0.62
	Conventions	280	1.94	0.70	1.94	0.73	0.59	63.21	98.21	0.59
4	Total	287	7.84	2.89	7.85	2.89	0.77	31.71	58.89	
	Idea Development	287	2.05	0.77	2.02	0.79	0.67	59.93	100.00	0.67
	Organization	287	1.95	0.81	1.96	0.80	0.70	65.85	98.61	0.70
	Voice	287	1.96	0.80	1.97	0.78	0.68	62.02	99.30	0.68
	Conventions	287	1.88	0.77	1.90	0.79	0.65	59.93	99.30	0.65
5	Total	287	7.93	2.77	7.81	2.78	0.74	30.66	56.45	
	Idea Development	287	2.08	0.77	2.06	0.76	0.61	58.19	98.96	0.61
	Organization	287	1.94	0.77	1.93	0.76	0.69	65.51	99.65	0.69
	Voice	287	2.02	0.75	1.99	0.77	0.61	58.54	98.95	0.61
	Conventions	287	1.89	0.77	1.84	0.76	0.63	60.63	98.61	0.63
6	Total	283	7.32	2.61	7.45	2.69	0.72	32.16	61.48	
	Idea Development	283	1.86	0.74	1.90	0.74	0.64	63.96	98.94	0.64
	Organization	283	1.84	0.71	1.86	0.75	0.66	65.37	99.65	0.66
	Voice	283	1.86	0.71	1.88	0.73	0.57	60.78	98.23	0.57
	Conventions	283	1.76	0.73	1.81	0.74	0.63	63.25	98.94	0.63

Table N6. 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	
1	Total	289	3.00	1.08	2.97	1.07	0.78	61.25	96.19	0.74
	Response to Text	289	1.38	0.72	1.34	0.71	0.74	75.78	99.31	
	Holistic	289	1.63	0.54	1.64	0.54	0.63	78.89	100.00	
2	Total	285	2.95	1.16	2.91	1.16	0.80	63.16	95.09	0.78
	Response to Text	285	1.41	0.73	1.39	0.73	0.78	78.25	99.30	
	Holistic	285	1.54	0.57	1.52	0.58	0.62	74.74	100.00	
3	Total	287	2.94	1.14	2.99	1.13	0.76	55.40	93.73	0.62
	Response to Text	287	1.38	0.74	1.39	0.74	0.62	65.85	97.56	
	Holistic	287	1.55	0.55	1.60	0.54	0.71	82.23	100.00	
4	Total	290	2.78	1.15	2.80	1.14	0.77	59.66	93.10	0.70
	Response to Text	290	1.27	0.73	1.30	0.71	0.70	70.34	99.66	
	Holistic	290	1.51	0.55	1.50	0.57	0.67	79.31	100.00	
5	Total	295	2.90	1.14	2.96	1.16	0.75	60.00	91.53	0.60
	Response to Text	295	1.33	0.72	1.35	0.71	0.60	65.42	97.97	
	Holistic	295	1.58	0.59	1.61	0.58	0.74	82.37	100.00	
6	Total	287	2.80	1.08	2.91	1.13	0.68	52.61	91.29	0.51
	Response to Text	287	1.25	0.68	1.29	0.68	0.51	64.81	96.52	
	Holistic	287	1.54	0.59	1.62	0.58	0.67	77.70	99.65	

Table N7. Grade 7 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	
1	Total	290	5.41	1.84	5.39	1.87	0.84	42.76	84.14	
	Response to Text	290	1.61	0.62	1.54	0.64	0.75	81.03	99.66	0.75
	Holistic	290	3.81	1.37	3.85	1.40	0.80	48.62	91.72	0.80
2	Total	291	5.51	1.83	5.45	1.79	0.79	36.77	82.47	
	Response to Text	291	1.48	0.68	1.47	0.70	0.69	73.20	99.31	0.69
	Holistic	291	4.03	1.39	3.98	1.35	0.75	45.70	90.03	0.75
3	Total	293	5.04	1.92	5.01	1.83	0.79	35.84	80.89	
	Response to Text	293	1.32	0.73	1.26	0.72	0.66	66.55	99.32	0.66
	Holistic	293	3.72	1.41	3.75	1.32	0.77	43.69	91.81	0.77
4	Total	289	5.29	2.06	5.11	2.06	0.83	39.10	83.74	
	Response to Text	289	1.37	0.79	1.32	0.79	0.87	85.81	99.31	0.87
	Holistic	289	3.92	1.50	3.79	1.51	0.75	42.56	87.54	0.75
5	Total	292	5.10	2.08	4.87	2.01	0.85	39.38	80.82	
	Response to Text	292	1.47	0.68	1.37	0.70	0.68	75.00	97.95	0.67
	Holistic	292	3.63	1.56	3.50	1.47	0.83	47.60	91.78	0.83
6	Total	294	4.90	1.94	4.85	2.01	0.81	42.52	81.63	
	Response to Text	294	1.21	0.79	1.17	0.79	0.79	75.17	99.66	0.79
	Holistic	294	3.70	1.44	3.67	1.48	0.74	44.56	85.71	0.74

Appendix O: SAWS Operational Rater Reliability

Table O1. SAWS 2013 Overall Inter-Rater Reliability for 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 Total	1629	7.19	2.38	7.06	2.52	0.67	25.97	57.09	
	Idea Development	1629	1.89	0.69	1.87	0.70	0.56	61.57	98.77	0.45
	Organization	1629	1.83	0.69	1.81	0.71	0.55	60.04	98.59	0.44
	Voice	1629	1.76	0.69	1.72	0.75	0.54	57.58	98.04	0.41
	Conventions	1629	1.72	0.71	1.66	0.76	0.58	58.62	98.53	0.45
4	Prompt Total	1698	6.87	2.52	6.80	2.51	0.73	30.68	62.78	
	Idea Development	1698	1.77	0.73	1.77	0.72	0.59	62.01	98.29	0.48
	Organization	1698	1.61	0.79	1.57	0.78	0.69	70.32	97.11	0.60
	Voice	1698	1.74	0.73	1.73	0.73	0.59	60.78	98.53	0.47
	Conventions	1698	1.75	0.74	1.74	0.74	0.60	60.13	98.76	0.47
5	Prompt Total	1699	7.37	2.61	7.48	2.63	0.76	32.49	64.39	
	Idea Development	1699	1.95	0.72	2.01	0.72	0.66	67.57	98.82	0.54
	Organization	1699	1.84	0.75	1.88	0.77	0.67	66.39	98.29	0.56
	Voice	1699	1.83	0.76	1.86	0.76	0.66	63.74	99.18	0.54
	Conventions	1699	1.74	0.74	1.74	0.77	0.63	60.92	98.94	0.50
6	Prompt Total	1605	7.09	2.40	7.16	2.41	0.60	28.16	55.89	
	Idea Development	1605	1.80	0.69	1.84	0.67	0.50	58.19	98.69	0.39
	Organization	1605	1.79	0.68	1.79	0.69	0.52	60.00	98.44	0.42
	Voice	1605	1.79	0.69	1.82	0.69	0.50	58.57	98.01	0.40
	Conventions	1605	1.71	0.68	1.71	0.71	0.50	56.82	98.38	0.38
7	Prompt Total	1711	7.66	2.79	7.66	2.76	0.74	30.27	57.16	
	Idea Development	1711	1.96	0.77	1.97	0.75	0.60	59.44	98.07	0.46
	Organization	1711	1.91	0.78	1.90	0.79	0.64	62.36	98.01	0.52
	Voice	1711	1.95	0.78	1.97	0.77	0.65	61.13	98.89	0.51
	Conventions	1711	1.83	0.77	1.82	0.78	0.63	60.20	98.60	0.49
8	Prompt Total	1698	7.88	2.73	7.88	2.72	0.70	31.68	57.18	
	Idea Development	1698	2.05	0.75	2.02	0.74	0.61	60.54	98.59	0.47
	Organization	1698	1.98	0.76	2.00	0.75	0.63	62.60	98.35	0.51
	Voice	1698	1.96	0.74	1.97	0.75	0.60	60.78	98.23	0.47
	Conventions	1698	1.88	0.76	1.89	0.76	0.59	57.77	98.23	0.45

Appendix P: SAWS Field Test Subscale Correlations

Table P1. Grade 3 SAWS 2013 Field Test Subscale Correlations – 12-point prompt

Form	Score	Total	Idea Development	Organization	Voice	Conventions
1	Total	1.00	0.89	0.91	0.92	0.87
	Idea Development		1.00	0.78	0.78	0.66
	Organization			1.00	0.78	0.69
	Voice				1.00	0.74
	Conventions					1.00
2	Total	1.00	0.89	0.91	0.91	0.86
	Idea Development		1.00	0.76	0.77	0.66
	Organization			1.00	0.76	0.71
	Voice				1.00	0.71
	Conventions					1.00
3	Total	1.00	0.88	0.90	0.90	0.86
	Idea Development		1.00	0.75	0.76	0.64
	Organization			1.00	0.74	0.68
	Voice				1.00	0.70
	Conventions					1.00
4	Total	1.00	0.88	0.89	0.90	0.86
	Idea Development		1.00	0.75	0.74	0.64
	Organization			1.00	0.73	0.68
	Voice				1.00	0.71
	Conventions					1.00
5	Total	1.00	0.89	0.90	0.91	0.85
	Idea Development		1.00	0.75	0.76	0.63
	Organization			1.00	0.76	0.67
	Voice				1.00	0.71
	Conventions					1.00
6	Total	1.00	0.90	0.89	0.91	0.85
	Idea Development		1.00	0.75	0.78	0.65
	Organization			1.00	0.74	0.65
	Voice				1.00	0.69
	Conventions					1.00

Table P2. Grade 5 SAWS 2013 Field Test Subscale Correlations – 12-point prompt

Form	Score	Total	Idea Development	Organization	Voice	Conventions
1	Total	1.00	0.90	0.89	0.91	0.85
	Idea Development		1.00	0.76	0.79	0.67
	Organization			1.00	0.73	0.65
	Voice				1.00	0.71
	Conventions					1.00
2	Total	1.00	0.90	0.90	0.90	0.84
	Idea Development		1.00	0.77	0.78	0.64
	Organization			1.00	0.74	0.68
	Voice				1.00	0.69
	Conventions					1.00
3	Total	1.00	0.92	0.91	0.92	0.88
	Idea Development		1.00	0.79	0.82	0.72
	Organization			1.00	0.78	0.72
	Voice				1.00	0.75
	Conventions					1.00
4	Total	1.00	0.91	0.91	0.91	0.88
	Idea Development		1.00	0.78	0.80	0.72
	Organization			1.00	0.77	0.73
	Voice				1.00	0.74
	Conventions					1.00
5	Total	1.00	0.89	0.90	0.91	0.85
	Idea Development		1.00	0.75	0.77	0.64
	Organization			1.00	0.76	0.65
	Voice				1.00	0.70
	Conventions					1.00
6	Total	1.00	0.88	0.89	0.90	0.84
	Idea Development		1.00	0.73	0.76	0.63
	Organization			1.00	0.74	0.66
	Voice				1.00	0.69
	Conventions					1.00

Table P3. Grade 5 SAWS 2013 Field Test Subscale Correlations – 4-point prompt

Form	Score	Total	Response to Text	Holistic
1	Total	1.00	0.91	0.85
	Response to Text		1.00	0.54
	Holistic			1.00
2	Total	1.00	0.90	0.85
	Response to Text		1.00	0.53
	Holistic			1.00
3&6	Total	1.00	0.88	0.85
	Response to Text		1.00	0.49
	Holistic			1.00
4	Total	1.00	0.91	0.84
	Response to Text		1.00	0.54
	Holistic			1.00
5	Total	1.00	0.89	0.85
	Response to Text		1.00	0.51
	Holistic			1.00

Table P4. Grade 5 SAWS 2013 Field Test Subscale Correlations – 8-point prompt

Form	Score	Total	Response to Text	Holistic
1	Total	1.00	0.71	0.94
	Response to Text		1.00	0.44
	Holistic			1.00
2	Total	1.00	0.79	0.93
	Response to Text		1.00	0.51
	Holistic			1.00
3&6	Total	1.00	0.73	0.94
	Response to Text		1.00	0.45
	Holistic			1.00
4	Total	1.00	0.76	0.92
	Response to Text		1.00	0.45
	Holistic			1.00
5	Total	1.00	0.75	0.93
	Response to Text		1.00	0.45
	Holistic			1.00

Table P5. Grade 5 SAWS 2013 Field Test Subscale Correlations – 7-point prompt

Form	Score	Total	Idea Development	Organization	Voice	Conventions
1	Total	1.00	0.92	0.91	0.92	0.87
	Idea Development		1.00	0.80	0.81	0.70
	Organization			1.00	0.77	0.71
	Voice				1.00	0.73
	Conventions					1.00
2	Total	1.00	0.91	0.92	0.93	0.87
	Idea Development		1.00	0.80	0.82	0.70
	Organization			1.00	0.80	0.73
	Voice				1.00	0.74
	Conventions					1.00
3	Total	1.00	0.91	0.91	0.92	0.85
	Idea Development		1.00	0.79	0.80	0.66
	Organization			1.00	0.78	0.68
	Voice				1.00	0.72
	Conventions					1.00
4	Total	1.00	0.92	0.92	0.93	0.89
	Idea Development		1.00	0.81	0.83	0.73
	Organization			1.00	0.80	0.75
	Voice				1.00	0.77
	Conventions				0.77	1.00
5	Total	1.00	0.93	0.92	0.93	0.89
	Idea Development		1.00	0.82	0.85	0.74
	Organization			1.00	0.81	0.75
	Voice				1.00	0.78
	Conventions					1.00
6	Total	1.00	0.93	0.92	0.93	0.89
	Idea Development		1.00	0.82	0.84	0.73
	Organization			1.00	0.80	0.74
	Voice				1.00	0.78
	Conventions					1.00

Table P6. Grade 7 SAWS 2013 Field Test Subscale Correlations – 4-point prompt

Form	Score	Total	Response to Text	Holistic
1	Total	1.00	0.90	0.84
	Response to Text		1.00	0.52
	Holistic			1.00
2	Total	1.00	0.92	0.87
	Response to Text		1.00	0.61
	Holistic			1.00
3	Total	1.00	0.92	0.85
	Response to Text		1.00	0.58
	Holistic			1.00
4	Total	1.00	0.92	0.87
	Response to Text		1.00	0.61
	Holistic			1.00
5	Total	1.00	0.90	0.85
	Response to Text		1.00	0.54
	Holistic			1.00
6	Total	1.00	0.89	0.84
	Response to Text		1.00	0.50
	Holistic			1.00

Table P7. Grade 7 SAWS 2013 Field Test Subscale Correlations – 8-point prompt

Form	Score	Total	Response to Text	Holistic
1	Total	1.00	0.82	0.96
	Response to Text		1.00	0.64
	Holistic			1.00
2	Total	1.00	0.76	0.95
	Response to Text		1.00	0.51
	Holistic			1.00
3	Total	1.00	0.79	0.95
	Response to Text		1.00	0.55
	Holistic			1.00
4	Total	1.00	0.80	0.94
	Response to Text		1.00	0.56
	Holistic			1.00
5	Total	1.00	0.83	0.97
	Response to Text		1.00	0.67
	Holistic			1.00
6	Total	1.00	0.78	0.94
	Response to Text		1.00	0.50
	Holistic			1.00

Appendix Q: SAWS Operational Subscale Correlations

Table Q1. SAWS 2013 Subscale Correlations

Grade	Score	Total	Idea Development	Organization	Voice	Conventions
3	Total	1.00	0.90	0.90	0.90	0.85
	Idea Development		1.00	0.78	0.77	0.64
	Organization			1.00	0.75	0.67
	Voice				1.00	0.68
	Conventions					1.00
4	Total	1.00	0.87	0.78	0.89	0.82
	Idea Development		1.00	0.54	0.78	0.61
	Organization			1.00	0.55	0.49
	Voice				1.00	0.68
	Conventions					1.00
5	Total	1.00	0.89	0.89	0.90	0.84
	Idea Development		1.00	0.76	0.78	0.63
	Organization			1.00	0.73	0.66
	Voice				1.00	0.67
	Conventions					1.00
6	Total	1.00	0.91	0.90	0.91	0.85
	Idea Development		1.00	0.78	0.80	0.67
	Organization			1.00	0.76	0.69
	Voice				1.00	0.70
	Conventions					1.00
7	Total	1.00	0.91	0.91	0.92	0.87
	Idea Development		1.00	0.78	0.81	0.68
	Organization			1.00	0.76	0.71
	Voice				1.00	0.72
	Conventions					1.00
8	Total	1.00	0.92	0.91	0.92	0.88
	Idea Development		1.00	0.79	0.82	0.72
	Organization			1.00	0.78	0.74
	Voice				1.00	0.76
	Conventions					1.00