Technical Report

Proficiency Assessments for Wyoming Students (PAWS)

Reading and Mathematics: Grades 3–8

Science: Grades 4 and 8

2015-2016 Administration

Submitted to Wyoming Department of Education

Prepared by

Educational Testing Service

August 1, 2016

Final

Copyright © 2016 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/educators/state-assessment/paws/

Wyoming Department of Education

Jillian Balow

Superintendent of Wyoming Department of Education

Wyoming Department of Education Assessment Team

Deb Lindsey

Assessment Administrator

Jessica Steinbrenner

Assessment Supervisor

Emily Brantz and Michelle Carroll

Assessment Consultants

Educational Testing Service (ETS)

Lauren Hebert

Wyoming Program Manager

Rebecca Rabin-Reed

Wyoming Program Lead

Valerie Link

Senior Psychometrician

Henry Yoo

Psychometrician

Hyeonjoo Oh

Psychometric Manager

Mark Batleman

Principal Statistical Assoc. - Level 1

Natalie Hatrak

Principal Statistical Assoc. - Level 1

James Carroll

Principal Statistical Assoc. - Level 2

Su Liu

Data Analyst Manager

Patricia Baron

Lead Research Project Manager

Executive Summary

This report presents the technical results for the 2016 Proficiency Assessments for Wyoming Students (PAWS). The document covers the grades 3 to 8 Reading and Mathematics administration and the grades 4 and 8 Science administration. There were approximately 6,789 to 7,683 students in the total Wyoming student population at grades 3 to 8.

Structure of This Report

The initial chapter provides relevant policy decisions regarding the PAWS program, followed by brief descriptions of the PAWS as it was administered in 2016. The Reliability and Validity chapters present the evidence gathered to support the intended uses and interpretations of scores for the PAWS program. In short, the validity process began with test design and continued through the entire assessment process, including item development and field testing, analyses of item and test data, test scaling, scoring, and score reporting. Each of these processes is described in detail starting with the Test Design and Development chapter, and concluded in the Historical Comparisons chapter. Operational aspects of the program are discussed in the remaining chapters.

Elimination of SAWS Assessments

Student Assessment of Writing Skills (SAWS) at grades 3, 5, and 7 was eliminated from administration due to Enrolled Act 50 (EA50) of the 2015 general session of the Wyoming Legislature.

> ***Section 2***. *W.S. 21-2-204(c)(ii)(A)(III) is repealed.*
>
> ***Section 3.*** *Notwithstanding any other provision of law, the state board, the state superintendent and the department of education shall cease any development or administration of a separate writing assessment as a part of the statewide assessment system required under W.S. 21-2-304.*

Conclusion

The technical efforts conducted in 2015-2016 administration, described later in this report, demonstrated the fidelity of the program to its long-standing levels of validity and reliability. This quality was retained through diligent compliance to procedures and high caliber judgment and evaluation of numerous national professionals from the field of test measurement.

Table i provides raw score reliabilities of PAWS tests ranged from 0.89 to 0.93. Table ii provides a comparison of percentages of the students classified as Proficient and Advanced from 2014–2016 for Reading and Mathematics and 2013-2016 for Science. Table iii shows that cut point accuracy

and consistency ratings range from 0.86 to 0.92, with most results above 0.90 for accuracy and above .88 for consistency.

Table i. Observed Reliability Coefficients for the PAWS Tests

Grade	Reliability Coefficients		
	Reading	Mathematics	Science
3	0.90	0.92	
4	0.89	0.92	0.90
5	0.91	0.93	
6	0.92	0.92	
7	0.91	0.92	
8	0.90	0.92	0.89

Table i. Percentage of Students Classified as Proficient and Advanced for the PAWS Tests

Grade	Reading			Mathematics			Science			
	2014	2015	2016	2014	2015	2016	2013	2014	2015	2016
3	61.8	60.5	57.7	50.4	49.3	52.6				
4	63.7	60.1	65.2	46.6	50.1	55.0	57.5	52.5	51.3	54.4
5	58.1	58.3	60.7	54.0	52.2	56.0				
6	56.9	56.4	57.6	48.5	49.3	49.6				
7	59.7	56.3	59.7	43.6	43.1	49.0				
8	58.9	51.2	53.4	50.2	47.1	47.5	43.7	47.5	41.2	41.2

Table i. Cut Point Decision Accuracy and Consistency Indices for PAWS Tests

Grade	Reading		Mathematics		Science	
	Cut Point Accuracy	Cut point Consistency	Cut Point Accuracy	Cut point Consistency	Cut Point Accuracy	Cut point Consistency
3	0.90	0.87	0.92	0.89		
4	0.91	0.87	0.92	0.88	0.91	0.87
5	0.90	0.86	0.92	0.89		
6	0.91	0.88	0.92	0.88		
7	0.91	0.87	0.92	0.89		
8	0.90	0.86	0.92	0.89	0.91	0.88

Table of Contents

Executive Summary v
Structure of This Report v
Elimination of SAWS Assessments v
Conclusion v
1. OVERVIEW OF THE 2016 PAWS 1
1.1 Introduction 1
1.2 Background of PAWS 1
1.3 Overview of PAWS Test Components 3
1.4 Overview of the PAWS Design 4
1.5 State Policy on Student Participation 4
1.5.1 Students with Disabilities, 504 Plans, and English Language Learners 5
2. VALIDITY 9
2.1 Overview 9
2.1.1. Purpose of the PAWS 9
2.1.2. The Constructs to Be Measured 10
2.1.3. The Interpretations and Uses of the PAWS Scores 10
2.1.4. Intended Test Population 11
2.2 Evidence Based on Content-related Validity 11
2.2.1. Description of the State Standards 11
2.2.2. Specifications and Blueprints 12
2.2.3. Item Development Process 12
2.2.4. Item Review Process 12
2.2.5. Form Construction Process 12
2.2.6. Alignment Study 12
2.3 Evidence Based on Relations to Other Variables 13
2.3.1. Correlations between Content Areas 13
2.3.2. Differential Item Functioning Analyses 13
2.4 Evidence Based on Response Processes 14

2.5 Evidence Based on Internal Structure 15
2.5.1. Classical Statistics 15
2.5.2. Reliability 15
2.5.3. Dimensionality 16
2.6 Evidence Based on Consequences of Testing 16
3. PAWS TEST DESIGN AND DEVELOPMENT 18
3.1 Overview 18
3.2 Test Design and Blueprints 18
3.2.1. Purpose 18
3.2.2. Plan 19
3.3 Types of Items Used in PAWS 20
3.3.1. PAWS Reading Tests 20
3.3.2. PAWS Mathematics Tests 28
3.3.3. PAWS Science Tests 33
3.4 PAWS Test Development Process 36
3.5 Item and Test Form Development 37
3.5.1. Item Specifications 37
3.5.2. Item Difficulty Requirements 38
3.5.3. Item Graphics Requirements 38
3.6 Item Review 38
3.7 Field Testing 40
3.7.1. Classical Item Statistics 41
3.7.1.1. Item Difficulty 41
3.7.1.2. Item Discrimination 41
3.7.2. Differential Item Functioning 41
3.7.3. Item Response Theory (IRT) Analysis 43
3.7.4. Data Review Procedures 44
3.8 Test Form Construction 45
3.8.1. Construction of the Reading, Mathematics, and Science Forms 46
3.8.2. Final Review of Assembled Operational Tests 47

4. TEST ADMINISTRATION 48
4.1 Test Materials 48
4.2 Materials Return 48
4.3 Directions for Administration 48
4.3.1. Allowed Student Manipulatives 49
4.3.2. Test Security 49
4.4 Student Participation 50
4.5 PAWS Standard Accommodations 50
4.5.1. Students Eligible for Test Accommodations 52
4.5.2. Requirements for Use of Test Accommodations 52
4.5.3. Description of Standard Accommodations for Students with Disabilities 53
4.5.3.1. Presentation Accommodations 53
4.5.3.2. Response Accommodations 54
4.5.3.3. Setting Accommodations 55
4.5.3.4. Timing and Scheduling Accommodations 55
4.5.4. Description of Standard Accommodations for ELL 55
4.5.4.1. Presentation Accommodations 56
4.5.4.2. Setting Accommodations 56
4.5.4.3. Timing and Scheduling Accommodations 56
4.5.5. PAWS Monitoring of Appropriate Accommodations 56
4.5.5.1. Empirical Analysis of Accommodations 57
4.5.6. Selection and Administration of Accommodations 57
5. PROCESSING AND SCORING OF PAWS ITEMS 59
5.1 Overview 59
5.2 Receipt Control, Processing, Scanning, Editing 59
5.2.1 Receipt Control 59
5.2.2 Processing 60
5.2.3 Scanning 60
5.2.4 Resolutions 61
6. EQUATING AND SCALING 62

6.1 Overview 62
6.2 Item and Forms Development 62
6.3 IRT Model and Calibration 62
6.4 Fit Statistics for the Rasch Model 64
6.5 Reading and Mathematics Vertical Scales 65
6.6 Reading, Mathematics, and Science Equating Analyses 66
6.6.1. Calibration and Equating Process for the 2016 Administration 66
6.7 Translating Raw Scores to Scale Scores and Performance Levels 69
7. SCORE REPORTING 71
7.1 Overview 71
7.2 Scale Scores of each Subject 71
7.3 Performance Levels of each Subject 72
7.4 Domain Standard-Level Scale Scores and Performance Levels 73
7.5 Production of Printed Score Reports for PAWS 75
7.6 Assessment Score Reports: Supplement Guide for Districts and Schools for PAWS 75
8. RELIABILITY 76
8.1 Overview 76
8.2 Internal Consistency Reliability 76
8.3 Classical and Conditional Standard Errors of Measurement 77
8.4 Accuracy and Consistency of Classifications 81
9. QUALITY CONTROL PROCEDURES 86
9.1 Quality Control of Item Development 86
9.1.1. Item and Prompt Specifications 86
9.1.2. Item Writers 86
9.1.3. Internal Contractor Reviews 86
9.1.4. Content Review 87
9.1.5. Editorial Review 87
9.1.6. Fairness Review 87
9.1.7. Assessment Director Review 87
9.1.8. Data Review of Field Tested Items 87

9.1.9. Quality Control of the Item Bank 87
9.2 Quality Control of Test Materials 88
9.2.1. Collecting Test Materials 88
9.2.2. Processing Test Materials 89
9.3 Quality Control of Scanning 89
9.4 Quality Control of Psychometric Analyses 89
9.5 Quality Control of Reporting 90
9.5.1. Excluding Student Scores from Summary Reports 90
10. HISTORICAL COMPARISONS 91
11. REFERENCES 102
12. GLOSSARY OF TERMS 105
Appendix A: PAWS 2016 Reading, Mathematics, and Science Blueprints 110
Reading 110
Mathematics 122
Science 128
Appendix B: Sample PAWS Student Reports: Grade 4 Reading, Mathematics, and Science (exemplar for Grade 8) 132
Appendix C: Sample PAWS Student Reports: Grade 6 Reading, Mathematics (exemplar for Grade 3, 5, and 7) 134
Appendix D: PAWS Operational Subscale Correlations 136
Appendix E: DIF Results for 2016 Field Test Items 142
Reading 142
Mathematics 145
Science 148
Appendix F: Classical Item Statistics for 2016 Field Test Items 149
Reading 149
Mathematics 173
Science 191
Appendix G: Classical Item Statistics for 2016 Operational Items 197
Reading 197

Mathematics 209
Science 221
Appendix H: Rasch Difficulty, Standard Error, Fit Statistics, and *N*-counts for 2016 Field Test Items 225
Reading 225
Mathematics 249
Science 273
Appendix I: Rasch Difficulty, Standard Error, and Fit Statistics for 2016 Operational Items ... 281
Reading 281
Mathematics 293
Science 305
Appendix J: Frequency of Individual Accommodations for 2016 PAWS Tests 309
Reading 309
Mathematics 312
Science 315
Appendix K: Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status 318
Reading 318
Mathematics 318
Science 318
Appendix L: Scale Score Descriptive Statistics by Demographic Subgroup 319
Reading 320
Mathematics 323
Science 326
Appendix M: Raw Score to Scale Score Tables 327
Reading 327
Mathematics 339
Science 356
Appendix N: Performance Level Percentages by Demographic Subgroup 360
Reading 361

Mathematics 364
Science 367

LIST OF TABLES

Table 1. Statewide Participation in Reading 6
Table 2. Statewide Participation in Mathematics 7
Table 3. Statewide Participation in Science 8
Table 4. Reporting Strands Design for Reading Grade 3 22
Table 5. Reporting Strands Design for Reading Grade 4 23
Table 6. Reporting Strands Design for Reading Grade 5 24
Table 7. Reporting Strands Design for Reading Grade 6 25
Table 8. Reporting Strands Design for Reading Grade 7 26
Table 9. Reporting Strands Design for Reading Grade 8 27
Table 10. Reporting Strands Design for Mathematics Grade 3 29
Table 11. Reporting Strands Design for Mathematics Grade 4 30
Table 12. Reporting Strands Design for Mathematics Grade 5 30
Table 13. Reporting Strands Design for Mathematics Grade 6 31
Table 14. Reporting Strands Design for Mathematics Grade 7 31
Table 15. Reporting Strands Design for Mathematics Grade 8 32
Table 16. Reporting Strands Design for Science Grade 4 35
Table 17. Reporting Strands Design for Science Grade 8 35
Table 18. DIF Categories for Multiple-Choice Items 43
Table 19. Criteria to Evaluate Mean-Square Fit Statistics 64
Table 20. PAWS Scaling Constants and Lowest and Highest Obtainable Scale Scores 69
Table 21. Proficiency Level Ranges for Reading Grades 3–8 72
Table 22. Proficiency Level Ranges for Mathematics Grades 3–8 72
Table 23. Proficiency Level Ranges for Science Grades 4 and 8 72
Table 24. State Average Scores by Domain Standard-Level 74
Table 25. Reliability and Standard Errors of Measurement by Grade 78
Table 26. Reliability and Standard Errors of Measurement by Grade and Reading Domain 79
Table 27. Reliability and Standard Errors of Measurement by Grade and Mathematics Domain 80
Table 28. Reliability and Standard Errors of Measurement by Grade and Science Domain 80
Table 29. Decision Accuracy and Consistency Indices – Grade 3 84
Table 30. Decision Accuracy and Consistency Indices – Grade 4 84
Table 31. Decision Accuracy and Consistency Indices – Grade 5 84
Table 32. Decision Accuracy and Consistency Indices – Grade 6 84
Table 33. Decision Accuracy and Consistency Indices – Grade 7 85
Table 34. Decision Accuracy and Consistency Indices – Grade 8 85
Table 35. Year-to-Year Comparison of Scale Scores for the PAWS Reading 92
Table 36. Year-to-Year Comparison of Percentage of Proficient and Advanced Students for the PAWS Reading 92

Table 37. Year-to-Year Comparison of Scale Scores for the PAWS Mathematics..... 93
Table 38. Year-to-Year Comparison of Percentage of Proficient and Advanced Students for the PAWS Mathematics..... 93
Table 39. Year-to-Year Comparison of Scale Scores for the PAWS Science..... 94
Table 40. Year-to-Year Comparison of Percentage of Proficient and Advanced Students for the PAWS Science..... 94
Table A1. PAWS 2016 Reading Grade 3 Blueprint 110
Table A2. PAWS 2016 Reading Grade 4 Blueprint 112
Table A3. PAWS 2016 Reading Grade 5 Blueprint 114
Table A4. PAWS 2016 Reading Grade 6 Blueprint 116
Table A5. PAWS 2016 Reading Grade 7 Blueprint 118
Table A6. PAWS 2016 Reading Grade 8 Blueprint 120
Table A7. PAWS 2016 Mathematics Grade 3 Blueprint..... 122
Table A8. PAWS 2016 Mathematics Grade 4 Blueprint..... 123
Table A9. PAWS 2016 Mathematics Grade 5 Blueprint..... 124
Table A10. PAWS 2016 Mathematics Grade 6 Blueprint..... 125
Table A11. PAWS 2016 Mathematics Grade 7 Blueprint..... 126
Table A12. PAWS 2016 Mathematics Grade 8 Blueprint..... 127
Table A13. PAWS 2016 Science Grade 4 Blueprint 128
Table A14. PAWS 2016 Science Grade 8 Blueprint 130
Table D1. PAWS Grade 3 Total Test and Subscale Correlations..... 136
Table D2. PAWS Grade 4 Total Test and Subscale Correlations..... 137
Table D3. PAWS Grade 5 Total Test and Subscale Correlations..... 138
Table D4. PAWS Grade 6 Total Test and Subscale Correlations..... 139
Table D5. PAWS Grade 7 Total Test and Subscale Correlations..... 140
Table D6. PAWS Grade 8 Total Test and Subscale Correlations..... 141
Table E1. Reading Grade 3 DIF Summary Statistics for Embedded Field Test Items..... 142
Table E2. Reading Grade 4 DIF Summary Statistics for Embedded Field Test Items..... 142
Table E3. Reading Grade 5 DIF Summary Statistics for Embedded Field Test Items..... 143
Table E4. Reading Grade 6 DIF Summary Statistics for Embedded Field Test Items..... 143
Table E5. Reading Grade 7 DIF Summary Statistics for Embedded Field Test Items..... 144
Table E6. Reading Grade 8 DIF Summary Statistics for Embedded Field Test Items..... 144
Table E7. Mathematics Grade 3 DIF Summary Statistics for Embedded Field Test Items 145
Table E8. Mathematics Grade 4 DIF Summary Statistics for Embedded Field Test Items 145
Table E9. Mathematics Grade 5 DIF Summary Statistics for Embedded Field Test Items 146
Table E10. Mathematics Grade 6 DIF Summary Statistics for Embedded Field Test Items 146
Table E11. Mathematics Grade 7 DIF Summary Statistics for Embedded Field Test Items 147
Table E12. Mathematics Grade 8 DIF Summary Statistics for Embedded Field Test Items 147
Table E13. Science Grade 4 DIF Summary Statistics for Embedded Field Test Items..... 148

Table E14. Science Grade 8 DIF Summary Statistics for Embedded Field Test Items..... 148
Table F1. Reading Grade 3 Classical Statistics for Field Test Items..... 149
Table F2. Reading Grade 4 Classical Statistics for Field Test Items..... 153
Table F3. Reading Grade 5 Classical Statistics for Field Test Items..... 157
Table F4. Reading Grade 6 Classical Statistics for Field Test Items..... 161
Table F5. Reading Grade 7 Classical Statistics for Field Test Items..... 165
Table F6. Reading Grade 8 Classical Statistics for Field Test Items..... 169
Table F7. Mathematics Grade 3 Classical Statistics for Field Test Items 173
Table F8. Mathematics Grade 4 Classical Statistics for Field Test Items 176
Table F9. Mathematics Grade 5 Classical Statistics for Field Test Items 179
Table F10. Mathematics Grade 6 Classical Statistics for Field Test Items 182
Table F11. Mathematics Grade 7 Classical Statistics for Field Test Items 185
Table F12. Mathematics Grade 8 Classical Statistics for Field Test Items 188
Table F13. Science Grade 4 Classical Statistics for Field Test Items 191
Table F14. Science Grade 8 Classical Statistics for Field Test Items 194
Table G1. Reading Grade 3 Classical Statistics for Operational Items 197
Table G2. Reading Grade 4 Classical Statistics for Operational Items 199
Table G3. Reading Grade 5 Classical Statistics for Operational Items 201
Table G4. Reading Grade 6 Classical Statistics for Operational Items 203
Table G5. Reading Grade 7 Classical Statistics for Operational Items 205
Table G6. Reading Grade 8 Classical Statistics for Operational Items 207
Table G7. Mathematics Grade 3 Classical Statistics for Operational Items..... 209
Table G8. Mathematics Grade 4 Classical Statistics for Operational Items..... 211
Table G9. Mathematics Grade 5 Classical Statistics for Operational Items..... 213
Table G10. Mathematics Grade 6 Classical Statistics for Operational Items..... 215
Table G11. Mathematics Grade 7 Classical Statistics for Operational Items..... 217
Table G12. Mathematics Grade 8 Classical Statistics for Operational Items..... 219
Table G13. Science Grade 4 Classical Statistics for Operational Items 221
Table G14. Science Grade 8 Classical Statistics for Operational Items 223
Table H1. Reading Grade 3 IRT Statistics for Field Test Items..... 225
Table H2. Reading Grade 4 IRT Statistics for Field Test Items..... 229
Table H3. Reading Grade 5 IRT Statistics for Field Test Items..... 233
Table H4. Reading Grade 6 IRT Statistics for Field Test Items..... 237
Table H5. Reading Grade 7 IRT Statistics for Field Test Items..... 241
Table H6. Reading Grade 8 IRT Statistics for Field Test Items..... 245
Table H7. Mathematics Grade 3 IRT Statistics for Field Test Items..... 249
Table H8. Mathematics Grade 4 IRT Statistics for Field Test Items..... 253
Table H9. Mathematics Grade 5 IRT Statistics for Field Test Items..... 257
Table H10. Mathematics Grade 6 IRT Statistics for Field Test Items..... 261

Table H11. Mathematics Grade 7 IRT Statistics for Field Test Items..... 265
Table H12. Mathematics Grade 8 IRT Statistics for Field Test Items..... 269
Table H13. Science Grade 4 IRT Statistics for Field Test Items..... 273
Table H14. Science Grade 8 IRT Statistics for Field Test Items..... 277
Table I1. Reading Grade 3 IRT Statistics for Operational Items..... 281
Table I2. Reading Grade 4 IRT Statistics for Operational Items..... 283
Table I3. Reading Grade 5 IRT Statistics for Operational Items..... 285
Table I4. Reading Grade 6 IRT Statistics for Operational Items..... 287
Table I5. Reading Grade 7 IRT Statistics for Operational Items..... 289
Table I6. Reading Grade 8 IRT Statistics for Operational Items..... 291
Table I7. Mathematics Grade 3 IRT Statistics for Operational Items 293
Table I8. Mathematics Grade 4 IRT Statistics for Operational Items 295
Table I9. Mathematics Grade 5 IRT Statistics for Operational Items 297
Table I10. Mathematics Grade 6 IRT Statistics for Operational Items 299
Table I11. Mathematics Grade 7 IRT Statistics for Operational Items 301
Table I12. Mathematics Grade 8 IRT Statistics for Operational Items 303
Table I13. Science Grade 4 IRT Statistics for Operational Items..... 305
Table I14. Science Grade 8 IRT Statistics for Operational Items..... 307
Table J1. Frequency of IEP Student's Standard Accommodations for Reading: Presentation .. 309
Table J2. Frequency of IEP Student's Standard Accommodations for Reading: Response..... 310
Table J3. Frequency of IEP Student's Standard Accommodations for Reading: Setting..... 310
Table J4. Frequency of IEP Student's Standard Accommodations for Reading: Timing and Scheduling..... 311
Table J5. Frequency of ELLs Standard Accommodations for Reading: Presentation 311
Table J6. Frequency of ELLs Standard Accommodations for Reading: Setting..... 311
Table J7. Frequency of ELLs Standard Accommodations for Reading: Timing and Scheduling 311
Table J8. Frequency of IEP Student's Standard Accommodations for Mathematics: Presentation 312
Table J9. Frequency of IEP Student's Standard Accommodations for Mathematics: Response 313
Table J10. Frequency of IEP Student's Standard Accommodations for Mathematics: Setting . 314
Table J11. Frequency of IEP Student's Standard Accommodations for Mathematics: Timing and Scheduling..... 314
Table J12. Frequency of ELLs Standard Accommodations for Mathematics: Presentation 314
Table J13. Frequency of ELLs Standard Accommodations for Mathematics: Setting..... 314
Table J14. Frequency of ELLs Standard Accommodations for Mathematics: Timing and Scheduling..... 315
Table J15. Frequency of IEP Student's Standard Accommodations for Science: Presentation . 315
Table J16. Frequency of IEP Student's Standard Accommodations for Science: Response..... 316

Table J17. Frequency of IEP Student's Standard Accommodations for Science: Setting..... 316
Table J18. Frequency of IEP Student's Standard Accommodations for Science: Timing and Scheduling..... 317
Table J19. Frequency of ELLs Standard Accommodations for Science: Presentation 317
Table J20. Frequency of ELLs Standard Accommodations for Science: Setting..... 317
Table J21. Frequency of ELLs Standard Accommodations for Science: Timing and Scheduling 317
Table K1. Reading Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status 318
Table K2. Mathematics Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status 318
Table K3. Science Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status 318
Table L1. Summary Statistics of Reading, Mathematics, and Science Scale Score by Grade... 319
Table L2. Summary Statistics of Reading Grade 3 Scale Score..... 320
Table L3. Summary Statistics of Reading Grade 4 Scale Score..... 320
Table L4. Summary Statistics of Reading Grade 5 Scale Score..... 321
Table L5. Summary Statistics of Reading Grade 6 Scale Score..... 321
Table L6. Summary Statistics of Reading Grade 7 Scale Score..... 322
Table L7. Summary Statistics of Reading Grade 8 Scale Score..... 322
Table L8. Summary Statistics of Mathematics Grade 3 Scale Score 323
Table L9. Summary Statistics of Mathematics Grade 4 Scale Score 323
Table L10. Summary Statistics of Mathematics Grade 5 Scale Score 324
Table L11. Summary Statistics of Mathematics Grade 6 Scale Score 324
Table L12. Summary Statistics of Mathematics Grade 7 Scale Score 325
Table L13. Summary Statistics of Mathematics Grade 8 Scale Score 325
Table L14. Summary Statistics of Science Grade 4 Scale Score..... 326
Table L15. Summary Statistics of Science Grade 8 Scale Score..... 326
Table M1. Reading Grade 3 Raw Score to Scale Score 327
Table M2. Reading Grade 4 Raw Score to Scale Score 329
Table M3. Reading Grade 5 Raw Score to Scale Score 331
Table M4. Reading Grade 6 Raw Score to Scale Score 333
Table M5. Reading Grade 7 Raw Score to Scale Score 335
Table M6. Reading Grade 8 Raw Score to Scale Score 337
Table M7. Mathematics Grade 3 Raw Score to Scale Score 339
Table M8. Mathematics Grade 4 Raw Score to Scale Score 341
Table M9. Mathematics Grade 5 Raw Score to Scale Score 344
Table M10. Mathematics Grade 6 Raw Score to Scale Score 347
Table M11. Mathematics Grade 7 Raw Score to Scale Score 350

Table M12. Mathematics Grade 8 Raw Score to Scale Score 353
Table M13. Science Grade 4 Raw Score to Scale Score 356
Table M14. Science Grade 8 Raw Score to Scale Score 358
Table N1.Performance Levels of Reading, Mathematics, and Science by Grade Revised 360
Table N2. Performance Levels of Reading by Grade 3 Demographic Subgroup..... 361
Table N3. Performance Levels of Reading by Grade 4 Demographic Subgroup..... 361
Table N4. Performance Levels of Reading by Grade 5 Demographic Subgroup..... 362
Table N5. Performance Levels of Reading by Grade 6 Demographic Subgroup..... 362
Table N6. Performance Levels of Reading by Grade 7 Demographic Subgroup..... 363
Table N7. Performance Levels of Reading by Grade 8 Demographic Subgroup..... 363
Table N8. Performance Levels of Mathematics by Grade 3 Demographic Subgroup 364
Table N9. Performance Levels of Mathematics by Grade 4 Demographic Subgroup 364
Table N10. Performance Levels of Mathematics by Grade 5 Demographic Subgroup 365
Table N11. Performance Levels of Mathematics by Grade 6 Demographic Subgroup 365
Table N12. Performance Levels of Mathematics by Grade 7 Demographic Subgroup 366
Table N13. Performance Levels of Mathematics by Grade 8 Demographic Subgroup 366
Table N14. Performance Levels of Science by Grade 4 Demographic Subgroup..... 367
Table N15. Performance Levels of Science by Grade 8 Demographic Subgroup..... 367

LIST OF FIGURES

Figure 1. Graphical Description of Item Characteristic Curve (ICC)..... 63
Figure 2. Accuracy or Consistency around Critical Cut Point..... 82
Figure 3. Percentage of Proficient and Advanced Students for Reading Grade 3..... 95
Figure 4. Percentage of Proficient and Advanced Students for Reading Grade 4..... 95
Figure 5. Percentage of Proficient and Advanced Students for Reading Grade 5..... 96
Figure 6. Percentage of Proficient and Advanced Students for Reading Grade 6..... 96
Figure 7. Percentage of Proficient and Advanced Students for Reading Grade 7..... 97
Figure 8. Percentage of Proficient and Advanced Students for Reading Grade 8..... 97
Figure 9. Percentage of Proficient and Advanced Students for Mathematics Grade 3..... 98
Figure 10. Percentage of Proficient and Advanced Students for Mathematics Grade 4..... 98
Figure 11. Percentage of Proficient and Advanced Students for Mathematics Grade 5..... 99
Figure 12. Percentage of Proficient and Advanced Students for Mathematics Grade 6..... 99
Figure 13. Percentage of Proficient and Advanced Students for Mathematics Grade 7..... 100
Figure 14. Percentage of Proficient and Advanced Students for Mathematics Grade 8..... 100
Figure 15. Percentage of Proficient and Advanced Students for Science Grade 4..... 101
Figure 16. Percentage of Proficient and Advanced Students for Science Grade 8..... 101

1. OVERVIEW OF THE 2016 PAWS

1.1 Introduction

This report describes the technical characteristics of the 2016 Proficiency Assessments for Wyoming Students (PAWS). Primary purposes of the PAWS include improving teaching and learning, fostering school and program improvement, and measuring student performance indicators under the Wyoming Accountability in Education Act.

Beginning with the spring 2006 administration, PAWS became the official statewide assessment used to measure individual student achievement against the Wyoming Content and Performance Standards in English Language Arts and Mathematics at grades 3–8 and 11. The PAWS Reading and Mathematics tests meet all requirements of the No Child Left Behind Act of 2001 (NCLB). In 2008, a Science assessment was implemented at grades 4, 8, and 11 utilizing the 2008 Wyoming Science Content and Performance Standards. Further legislative action in 2013, Wyoming State Enrolled Act 65, removed grade 11 tests from the future administrations.

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. Recent legislation in 2014 (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 2014–2015 and each school year thereafter" (Section 3(a)). Further legislative action in 2015, Wyoming State Enrolled Act 50, removed SAWS from the future administrations.

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

1.2 Background of PAWS

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

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

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

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

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

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

- Timely and meaningful feedback to educators, parents, and students regarding student, school, district, and state performance, which could improve teaching and learning over the course of the school year.

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

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

> "W.S. 21-2-304(a)(iii) - By rule and regulation and in consultation with local school districts, prescribe uniform student content and performance standards for the common core of knowledge and the common core of skills specified under W.S. 21-9-101(b), and promulgate uniform standards . . .
>
> W.S.21-2-304(c) - The state board shall perform an ongoing review of state board duties prescribed by law and may make recommendations to the legislature on board duties. In addition and not less than once every five (5) years, the board shall evaluate and review the uniformity and quality of the content and performance standards imposed under W.S. 21-9-101 and 21-9-102 and the student content and performance standards promulgated under paragraph (a)(iii) of this section . . ."

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

1.3 Overview of PAWS Test Components

The test design for the spring 2016 administration of the PAWS included content area assessments in reading, mathematics, and science. For reading, mathematics and science, each test had two to three sessions. Multiple choice items were administered via pencil and paper in a consumable test booklet for students in grades 3-5 and via a separate answer document for students in grades 6-8.

1.4 Overview of the PAWS Design

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

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

The PAWS Reading and Mathematics assessments were used to measure individual student achievement against the newly adopted 2012 WyCPS in English language arts and mathematics. The WyCPS outlines knowledge and skills students are expected to acquire at each grade in order to succeed in school and at work.

The PAWS Science assessment provides additional skill-level reporting categories aligned to the 2008 Wyoming Science Content and Performance Standards 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 results are particularly intended to help educators make informed decisions about curriculum and instruction. Since PAWS is aligned to academic content and student performance standards, its results can reveal weaknesses and strengths in curricula or instructional methodology. Thus, they can also help educators target specific areas necessary for school and district improvement.

1.5 State Policy on Student Participation

With two exceptions, all public school students in grades 3 through 8 must participate in the regular PAWS 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 English Language Learner (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, but may take the Wyoming English proficiency assessment as a substitution for the ELA/Reading portions of PAWS. Additionally, students who are expelled during the test window and students who are educated in residential institutions outside of Wyoming do not participate in statewide assessments.

Students with significant cognitive disabilities were required to take the Wyoming Alternate Assessment for Students with Significant Cognitive Disabilities (Wy-ALT). All students will participate in the state accountability assessment program in one of three ways:

- Participation in PAWS regular assessment without accommodation
- Participation in PAWS regular assessment with accommodation
- Participation in Wy-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) team's recommendation. Students with 504 Plans and ELL also take the PAWS.

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

School districts may not exempt ELL students from the assessment, except for students who are in their first year of school in the United States. Students who are in their first year in U.S. schools may take the Wyoming ELL assessment (ACCESS for ELLs) instead of the reading component of PAWS, but they are not exempt from the mathematics and science tests. The Wyoming ELL assessment measures English language academic proficiency. After their first year, all students who are English learners are required to participate in PAWS reading in addition to mathematics and science tests.

Tables 1 through 3 provide data on the numbers of students tested in 2016. Additional information can be found on the WDE website: http://edu.wyoming.gov. Note that total number of students for both Reading grade 7, Mathematics grades 3, 4 and 7, and Science Grade 4 from Tables 1 and 2 will be slightly different than other Tables because one additional student record was included after final item analysis. Mathematics grade 4 had two additional students.

Table 1. Statewide Participation in Reading

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
Total	7683	100	7482	100	7323	100	7013	100	7052	100	6793	100
Male	3984	51.9	3876	51.8	3686	50.3	3671	52.3	3638	51.6	3529	52.0
Female	3696	48.1	3603	48.2	3633	49.6	3341	47.6	3413	48.4	3264	48.0
Unknown	3	0.0	3	0.0	4	0.1	1	0.0	1	0.0	.	.
Amer. Indian	295	3.8	287	3.8	285	3.9	263	3.8	262	3.7	254	3.7
Asian	55	0.7	60	0.8	55	0.8	66	0.9	78	1.1	62	0.9
African American	89	1.2	82	1.1	77	1.1	95	1.4	73	1.0	66	1.0
Haw./Pac. Islander	11	0.1	8	0.1	15	0.2	9	0.1	10	0.1	10	0.1
Hispanic/Latino	997	13.0	1089	14.6	1126	15.4	1000	14.3	963	13.7	931	13.7
White	6025	78.4	5811	77.7	5613	76.6	5437	77.5	5515	78.2	5329	78.4
Multiracial	192	2.5	136	1.8	144	2.0	141	2.0	145	2.1	137	2.0
Unknown	19	0.2	9	0.1	8	0.1	2	0.0	6	0.1	4	0.1
Free/Reduced Lunch	2893	37.7	2717	36.3	2624	35.8	2472	35.2	2451	34.8	2236	32.9
Not Free/Reduced Lunch	4790	62.3	4765	63.7	4699	64.2	4541	64.8	4601	65.2	4557	67.1
Special Education	1066	13.9	979	13.1	990	13.5	952	13.6	870	12.3	778	11.5
Not Special Education	6617	86.1	6503	86.9	6333	86.5	6061	86.4	6182	87.7	6015	88.5
English Language Learner	338	4.4	187	2.5	150	2.0	106	1.5	109	1.5	137	2.0
Not English Language Learner	7345	95.6	7295	97.5	7173	98.0	6907	98.5	6943	98.5	6656	98.0

Table 2. Statewide Participation in Mathematics

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
Total	7683	100	7485	100	7325	100	7011	100	7055	100	6804	100
Male	3982	51.8	3879	51.8	3688	50.3	3668	52.3	3638	51.6	3535	52.0
Female	3700	48.2	3603	48.1	3633	49.6	3342	47.7	3416	48.4	3269	48.0
Unknown	1	0.0	3	0.0	4	0.1	1	0.0	1	0.0	.	.
Amer. Indian	293	3.8	287	3.8	284	3.9	261	3.7	262	3.7	254	3.7
Asian	60	0.8	62	0.8	58	0.8	66	0.9	80	1.1	65	1.0
African American	90	1.2	83	1.1	79	1.1	95	1.4	72	1.0	66	1.0
Haw./Pac. Islander	11	0.1	8	0.1	15	0.2	9	0.1	11	0.2	10	0.1
Hispanic/Latino	1004	13.1	1095	14.6	1126	15.4	1004	14.3	973	13.8	939	13.8
White	6016	78.3	5801	77.5	5611	76.6	5433	77.5	5506	78.0	5329	78.3
Multiracial	191	2.5	137	1.8	144	2.0	141	2.0	145	2.1	137	2.0
Unknown	18	0.2	12	0.2	8	0.1	2	0.0	6	0.1	4	0.1
Free/Reduced Lunch	2898	37.7	2726	36.4	2624	35.8	2473	35.3	2449	34.7	2238	32.9
Not Free/Reduced Lunch	4785	62.3	4759	63.6	4701	64.2	4538	64.7	4606	65.3	4566	67.1
Special Education	1059	13.8	981	13.1	985	13.4	949	13.5	867	12.3	781	11.5
Not Special Education	6624	86.2	6504	86.9	6340	86.6	6062	86.5	6188	87.7	6023	88.5
English Language Learner	346	4.5	188	2.5	157	2.1	109	1.6	114	1.6	140	2.1
Not English Language Learner	7337	95.5	7297	97.5	7168	97.9	6902	98.4	6941	98.4	6664	97.9

Table 3. Statewide Participation in Science

	Grade 4		Grade 8	
	N	Percentage	*N*	Percentage
Total	7477	100	6789	100
Male	3877	51.9	3524	51.9
Female	3597	48.1	3265	48.1
Unknown	3	0.0	.	.
Amer. Indian	283	3.8	253	3.7
Asian	62	0.8	65	1.0
African American	83	1.1	66	1.0
Haw./Pac. Islander	8	0.1	10	0.1
Hispanic/Latino	1093	14.6	936	13.8
White	5802	77.6	5320	78.4
Multiracial	136	1.8	135	2.0
Unknown	10	0.1	4	0.1
Free/Reduced Lunch	2713	36.3	2231	32.9
Not Free/Reduced Lunch	4764	63.7	4558	67.1
Special Education	982	13.1	779	11.5
Not Special Education	6495	86.9	6010	88.5
English Language Learner	185	2.5	140	2.1
Not English Language Learner	7292	97.5	6649	97.9

2. VALIDITY

2.1 Overview

Validity refers to the degree to which each interpretation or use of a test score is supported by evidence that is gathered (American Educational Research Association [AERA], American Psychological Association [APA], and National Council on Measurement in Education [NCME], 2014; ETS, 2015). 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, 2014; ETS, 2015; Kane, 2006). The process begins with test design and continues through the entire assessment process including item development and field testing, analyses of item and test data, test scaling, scoring, and score reporting.

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

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

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

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

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

2.1.1. Purpose of the PAWS

The purposes of the PAWS are multifold, as outlined in Chapters 1 and 3. The assessment is intended to comply with federal and state mandates, to inform ongoing instruction, and to help

teachers plan instruction for the following year. Additionally, the PAWS in grades 3 through 8 is used in determining Adequate Yearly Progress (AYP) that applies toward meeting the requirement of the federal NCLB Act of 2001.

2.1.2. The Constructs to Be Measured

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

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

2.1.3. The Interpretations and Uses of the PAWS Scores

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

Subscore results compare an individual student's scale score to the average scale score for the state as a whole. Subscores should be cautiously used to draw inferences about a student's achievement in each of several specific knowledge or skill areas covered by each test. There are limitations to the inferences drawn from the subscores by domain, given that there are relatively few items addressing each domain. A detailed description of the uses and applications of PAWS scores is presented in Chapter 7. Examples of individual student reports are provided in Appendix B demonstrating the report for reading, mathematics, and science for grades 4 and 8 and Appendix C showing the report for reading and mathematics for grades 3, 5, 6, and 7.

The tests that make up the PAWS assessments 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, 2014, Standard 1.4).

2.1.4. Intended Test Population

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

2.2 Evidence Based on Content-related Validity

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

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

The procedures used for test administration and test scoring are also to be considered in evaluating test content. As Kane (2006, p. 29) has noted, although evidence that appropriate administration and scoring procedures have been used does not provide compelling evidence to support a particular score interpretation or use, such evidence may prove useful in refuting rival explanations of test results. Evidence based on content includes the following:

2.2.1. Description of the State Standards

As noted in Chapter 1, Wyoming adopted rigorous content standards in 2008 for science. In 2012, the Wyoming State Legislature adopted the 2012 WyCPS for the PAWS assessment

program for reading and mathematics. These standards, which are the Common Core State Standards, were designed to guide instruction and learning for all students in the state and to bring Wyoming students to world–class levels of achievement.

2.2.2. Specifications and Blueprints

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

2.2.3. Item Development Process

A detailed description of the content and psychometric criteria applicable to the construction of the 2016 PAWS is presented in Chapter 3.

2.2.4. Item Review Process

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

2.2.5. Form Construction Process

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

2.2.6. Alignment Study

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

2.3 Evidence Based on Relations to Other Variables

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

2.3.1. Correlations between Content Areas

To the degree that students' content area scores correlate as expected, evidence of the validity in regarding those scores as measures of the intended constructs is provided. PAWS Reading, Mathematics, and Science tests and subscale inter-correlations are presented in Appendix D. There are strong relationships between the PAWS Reading, Mathematics, and Science scores. In the grades where science was tested, it tended to be more strongly related to both reading and mathematics than reading was to mathematics, though the average difference was small. For reading, this was probably because the science items were tied to common passages rather than being discrete and independent items, thus requiring more reading ability. For mathematics, science items often involve mathematical functions or terms, thus giving students with higher levels of mathematical ability an advantage in answering them. The strong relationships between the scale 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 co-vary together as would be expected. All can be seen to have strong relationships with the other subscales within each of the subjects, indicating that the subscales are measuring different yet related areas of knowledge.

2.3.2. Differential Item Functioning Analyses

Analyses of Differential Item Functioning (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 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 E. 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 demographic characteristics. Due to small case counts, DIF analyses for ethnicities were performed only where sample size was sufficient.

2.4 Evidence Based on Response Processes

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

Prior to the transition to the 2012 WyCPS, the WDE and ETS determined the need for new items and new item types in order to align the test blueprint with the new curriculum standards. Aligning the new PAWS grades 3-8 Reading and Mathematics assessment blueprints to the 2012 WyCPS is part of the chain of validity evidence supporting assessment of standards to measure student progress. To assess the appropriateness of the new items and item types for Wyoming students, ETS conducted a series of cognitive labs to learn how students in grades 3, 5 and 8 responded to a sample of the items under development. Results from the cognitive labs study provide diagnostic feedback and support for subsequent item development. The cognitive labs were conducted in two phases, the first in December 2012 and the second in March 2013. A full report was submitted to the WDE in 2015 (Baron, 2015).

The main findings from the cognitive lab were as follows.

- The items seemed to differentiate student mathematics performance well. Items elicited a range of student responses with some students having difficulty and others answering with little difficulty.
- Student mistakes in mathematics were most often due to being unfamiliar with the mathematics concepts.
- For reading, on the whole, the students had no problems reading the passages and understanding what the questions were asking.
- Students appeared to understand the need to refer to the text to support their answers, but did not always do so successfully (i.e., they misremembered or misinterpreted the text).
- Students in grade 3 were more successful reading and answering questions on the narrative passage compared to informational passage.

The results of the study, including all completed summary templates, were shared with the assessment development team for review. The assessment development experts found the results mostly confirmed their expectations of how students would respond to the new items. Furthermore, the cognitive labs did not identify specific problems with the items that needed to be corrected.

2.5 Evidence Based on Internal Structure

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

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

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

2.5.1. Classical Statistics

Point biserial correlations calculated for the items in a test show the degree to which the items discriminate between students with low and high scores on a test. To the degree that the correlations are high, evidence that the items assess the same construct is provided. The point biserials for the items in the PAWS are presented in Appendices F (field test) and G (operational).

Also germane to the validity of a score interpretation are the ranges of item difficulty for the items on which a test score will be based. The finding that items have difficulties spanning the range of examinee ability provides evidence that the items adequately measures examinees at all levels of ability. Information on average item score (i.e., *p*-values) is given in Appendices F (field test) and G (operational); the distributions of item *b*-values are given in Appendices H (field test) and I (operational). A detailed description of *p*-values and *b*-values can be found in Section 3.7.

2.5.2. Reliability

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

Overall reliability — The reliability analyses on each of the PAWS assessments are presented in Section 8.2. The results indicate that the reliabilities for all PAWS were medium-high to high, ranging from 0.89 to 0.93.

Reliability of performance classifications — The methodology used for estimating the reliability of classification decisions is described in Section 8.4, Accuracy and Consistency of Classifications. These levels of cut point accuracy and consistency are high as ranged from 0.86 to 0.92, and they are consistent with levels seen in previous years.

2.5.3. Dimensionality

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

2.6 Evidence Based on Consequences of Testing

As observed in the *Standards* (AERA, APA, & NCME, 2014), tests are commonly administered "in the expectation that some benefit will be realized from the interpretation and use of the scores intended by the test developers" (p. 19). When this is the case, expected benefits evidence will provide support for intended use of the scores. The WDE and ETS are in the process of determining what kinds of information can be gathered to assess the consequences of administration of the PAWS. One source of information for this purpose is the recently completed *Instructional Impact Study* (Baron, 2015).

During the transition to the 2012 WyCPS; hereafter referred to as the *new standards* and the 2006 standards-aligned PAWS, the WDE commissioned the *Instructional Impact Study*. The purpose of the study was to investigate the experience of Wyoming teachers, principals and curriculum directors (hereafter referred to as *educators*) during transition and implementation of the new standards. Development of the survey was informed by teacher focus groups, the Wyoming Technical Advisory Committee and staff from the Assessment and Standards teams of the WDE. The survey included selected-response and open-ended questions and focused on educators' perspectives with respect to the timing of the implementation, professional support and resource availability, materials alignment, and student struggles during the transition. Results include comments collected from survey respondents and focus group participants.

Educators indicated that they understand that the PAWS is aligned to the new standards, and teachers described some frustration during the transition period due to a need for aligned curriculum, materials and classroom assessments. Principals and curriculum directors are experiencing a process of transition versus teachers who are experiencing a more immediate impact when standards are adopted. The evidence indicates that leadership has been working on re-writing curriculum, trying to better understand the standards, align curriculum within and across grades. Meanwhile the teachers in the classrooms are aware that the transition to the new standards will get easier over time; they are experiencing some struggles in the process. Students at the higher grades have not had the scope and sequence that is expected at the grade level; teachers' experiences across districts differ from each other, and this adds to the frustration during the transition period. Some teachers remarked that finding materials during the multi-year transition is getting easier, but many noted that they need more time and better materials to successfully implement the new standards. The results of the impact study may provide some indication of district-level differences in progress toward full implementation, and alignment of PAWS to the 2012 WyCPS will allow some measurement of student progress during the transition period.

3. PAWS TEST DESIGN AND DEVELOPMENT

3.1 Overview

The Wyoming PAWS statewide assessments adhere to the principles of sound and ethical test construction set forth in the *Standards* (AERA, APA, & NCME, 2014). These assessments comply with the requirements of NCLB (P.L. 107–110) and were designed to provide teachers with information to improve instruction based on the 2012 WyCPS for reading and mathematics and 2008 Wyoming Science Content and Performance Standards.

3.2 Test Design and Blueprints

3.2.1. Purpose

Standards-based educational reform began in Wyoming in 1997–1998, with adoption of rigorous academic content standards in language arts,[1] mathematics, science, and social studies.[2] Wyoming educators have continued the other earlier efforts to implement standards-based curriculum and assessment to meet the goals of improving teaching and the academic achievement of all of our students.

In 2004, the Wyoming Legislature passed a law describing the purpose and implementation of a statewide assessment system (§21-2-304) in order to meet the requirements of NCLB.[3] 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 grade, school, district, and state levels. As previously noted, the primary purpose of the PAWS is to foster program improvement at the school, district, and state levels that supports the teaching and learning that takes place in Wyoming public classrooms. The construction of PAWS also ensures that it meets NCLB requirements. Improvement of teaching and learning in schools and fostering school program improvement are the primary purposes of statewide assessment of student performance in Wyoming.

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

---

[1] As previously noted, Wyoming test focused on only the Reading Language Arts Standards.
[2] Social studies is not presently tested in the PAWS assessments.
[3] The decision was made based on the recommendations of the Wyoming Statewide Task Force on Student Assessment and Education Accountability.

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

The documents were open to public comment during the fall of 2004 and again in 2012 for the 2012 WyCPS. From these documents arose the guiding principle of the design of PAWS 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 2012 WyCPS and 2008 Wyoming Science Content and Performance Standards led to the development of the PAWS blueprints and specifications.

3.2.2. Plan

The first step in test development is to create item and test specifications. WDE's test specifications reflect skill expectations that are outlined in 2012 WyCPS and 2008 Wyoming Science 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 development and 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 WDE and staff at ETS, based on the 2012 WyCPS and 2008 Wyoming Science 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, and science standards to be tested
- Number of items and points per test form
- Percentage and/or number of items and points per content standard
- Distribution of multiple item types (multiple choice and constructed response)
- Proposed distribution of items by cognitive complexity, i.e., percentage of items with low, moderate, or high levels of cognitive complexity
- Approximate time requirements for each assessment

3.3 Types of Items Used in PAWS

Consistent with Wyoming state law, the PAWS assessments have been composed solely of multiple choice items since the spring 2013 administration. 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 PAWS assessment is used to measure individual student achievement against the 2012 WyCPS in Reading and Mathematics. For Science, the Wyoming Science Content and Performance Standards adopted in 2008 remain in place. From 2012 to 2014, PAWS Reading and Mathematics blueprints were revised and test items developed to better align these assessments with the 2012 WyCPS. The newly developed blueprints are detailed in Appendix A.

The 2012 WyCPS and 2008 Wyoming Science 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 (Webb, Alt, Ely, & Vesperman, 2005). The categories *low complexity*, *moderate complexity*, and *high complexity* form an ordered description of the cognitive load involved in responding to the item.

3.3.1. PAWS Reading Tests

The Wyoming Language Arts Content and Performance Standards include an expectation that all students will become effective 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 2014, the reconfigured writing test was removed from the PAWS assessment. In 2015, Enrolled Act 50 removed the writing test from the statewide assessment system altogether). The 2012 WyCPS requires 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 school quality as measured by federal or state accountability.

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 2016 PAWS Reading test was based on the 2015 PAWS Reading Blueprint (see Appendix A). The content of the test is aligned to the reading content standards of the Wyoming Language Arts Content and Performance Standards. The PAWS assessment is designed to assess overall literacy skills in the following skill-reporting categories:

- Determine information's relevance and importance, and select and apply information for a task within a functional text;
- Understand main points and supporting details, recognize expositional 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.

Four content standards are assessed for each grade for grades 3–8:

- Reading Literature
- Reading Informational Text
- Integration of Knowledge and Ideas
- Language

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

The 2015-2016 PAWS Reading blueprints and reporting categories for each of the grade levels are provided in Appendix A. As noted in the blueprints, the percentage of assessment coverage of text type reflects the emphasis of instruction in Wyoming classrooms across grades. Tables 4–9 provide the number of items for each reading assessment by reporting strand, for the assessment overall and for the vertical scale set.

Table 4. Reporting Strands Design for Reading Grade 3

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	12	8	10	5
Lit: Craft and Structure (LTCR)	6	2	6	2
Inf.: Key Ideas and Details (INKY)	9	4	9	4
Inf.: Craft and Structure (INCR)	7	3	7	3
Integration of Knowledge and Ideas (INTG)	8	-	9	3
Language (LANG)	8	3	9	3
Totals	50	20	50	20

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	24	40	20	25
Lit: Craft and Structure (LTCR)	12	10	12	10
Inf.: Key Ideas and Details (INKY)	18	20	18	20
Inf.: Craft and Structure (INCR)	14	15	14	15
Integration of Knowledge and Ideas (INTG)	16	-	18	15
Language (LANG)	16	15	18	15

Table 5. Reporting Strands Design for Reading Grade 4

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	10	5	12	4
Lit: Craft and Structure (LTCR)	7	3	6	3
Inf.: Key Ideas and Details (INKY)	11	6	10	4
Inf.: Craft and Structure (INCR)	6	3	7	2
Integration of Knowledge and Ideas (INTG)	7	-	7	3
Language (LANG)	9	3	8	4
Totals	50	20	50	20

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	20	25	24	20
Lit: Craft and Structure (LTCR)	14	15	12	15
Inf.: Key Ideas and Details (INKY)	22	30	20	20
Inf.: Craft and Structure (INCR)	12	15	14	10
Integration of Knowledge and Ideas (INTG)	14	-	14	15
Language (LANG)	18	15	16	20

Table 6. Reporting Strands Design for Reading Grade 5

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	12	6	12	5
Lit: Craft and Structure (LTCR)	6	3	6	2
Inf.: Key Ideas and Details (INKY)	12	7	14	5
Inf.: Craft and Structure (INCR)	8	3	6	3
Integration of Knowledge and Ideas (INTG)	7	-	6	3
Language (LANG)	9	3	10	4
Totals	54	22	54	22

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	22	27	22	23
Lit: Craft and Structure (LTCR)	11	14	11	9
Inf.: Key Ideas and Details (INKY)	22	32	26	23
Inf.: Craft and Structure (INCR)	15	14	11	14
Integration of Knowledge and Ideas (INTG)	13	-	11	14
Language (LANG)	17	14	19	18

Table 7. Reporting Strands Design for Reading Grade 6

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	12	6	13	5
Lit: Craft and Structure (LTCR)	7	4	7	3
Inf.: Key Ideas and Details (INKY)	13	6	13	5
Inf.: Craft and Structure (INCR)	7	4	7	3
Integration of Knowledge and Ideas (INTG)	8	-	8	3
Language (LANG)	9	3	8	4
Totals	56	23	56	23

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	21	26	23	22
Lit: Craft and Structure (LTCR)	13	17	13	13
Inf.: Key Ideas and Details (INKY)	23	26	23	22
Inf.: Craft and Structure (INCR)	13	17	13	13
Integration of Knowledge and Ideas (INTG)	14	-	14	13
Language (LANG)	16	13	14	17

Table 8. Reporting Strands Design for Reading Grade 7

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	10	5	11	4
Lit: Craft and Structure (LTCR)	7	3	7	3
Inf.: Key Ideas and Details (INKY)	16	10	16	7
Inf.: Craft and Structure (INCR)	8	3	7	3
Integration of Knowledge and Ideas (INTG)	6	-	6	2
Language (LANG)	9	4	9	4
Totals	56	25	56	23

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	18	20	20	17
Lit: Craft and Structure (LTCR)	13	12	13	13
Inf.: Key Ideas and Details (INKY)	29	40	29	30
Inf.: Craft and Structure (INCR)	14	12	13	13
Integration of Knowledge and Ideas (INTG)	11	-	11	9
Language (LANG)	16	16	16	17

Table 9. Reporting Strands Design for Reading Grade 8

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	11	6	10	5
Lit: Craft and Structure (LTCR)	7	2	7	3
Inf.: Key Ideas and Details (INKY)	12	8	12	5
Inf.: Craft and Structure (INCR)	10	3	10	4
Integration of Knowledge and Ideas (INTG)	6	-	7	2
Language (LANG)	10	4	10	4
Totals	56	23	56	23

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Lit: Key Ideas and Details (LTKY)	20	26	18	22
Lit: Craft and Structure (LTCR)	13	9	13	13
Inf.: Key Ideas and Details (INKY)	21	35	21	22
Inf.: Craft and Structure (INCR)	18	13	18	17
Integration of Knowledge and Ideas (INTG)	11	-	13	9
Language (LANG)	18	17	18	17

3.3.2. PAWS Mathematics Tests

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

The structure of the operational 2016 PAWS Mathematics test is detailed in the 2014 PAWS Mathematics Blueprints (see Appendix A). The content of the test is aligned to the five content standards within the 2012 WyCPS in Mathematics for grades 3–5:

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

Five content standards within the 2012 WyCPS in Mathematics for grades 6 and 7:

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

Five content standards within the 2012 WyCPS in Mathematics for grade 8:

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

The 2014 PAWS Mathematics blueprints and reporting categories for each of the grade levels are provided in Appendix A. As noted in the tables below, the percentage of assessment coverage of each content standard reflects the emphasis of instruction in Wyoming classrooms across grades.

For example, at grade 8 the emphasis is placed upon Expressions and Equations. Calculator use is not permitted for the grades 3–5 assessments. Tables 10–15 provide the number of items for each mathematics assessment by reporting strand and for the assessment overall

Table 10. Reporting Strands Design for Mathematics Grade 3

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Operations and Algebraic Thinking (ALGE)	20	8	20	8
Number Operations–Base Ten (BTEN)	6	3	6	3
Number Operations–Fractions (FRCT)	6	2	6	2
Measurement and Data (MEAS)	12	5	12	5
Geometry (GEOM)	6	2	6	2
Total	50	20	50	20
	Percentages of Items			
Strand	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Operations and Algebraic Thinking (ALGE)	40	40	40	40
Number Operations–Base Ten (BTEN)	12	15	12	15
Number Operations–Fractions (FRCT)	12	10	12	10
Measurement and Data (MEAS)	24	25	24	25
Geometry (GEOM)	12	10	12	10

Table 11. Reporting Strands Design for Mathematics Grade 4

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Operations and Algebraic Thinking (ALGE)	13	5	13	5
Number Operations–Base Ten (BTEN)	10	4	10	4
Number Operations–Fractions (FRCT)	20	8	20	8
Measurement and Data (MEAS)	10	4	10	4
Geometry (GEOM)	6	3	6	3
Total	59	24	59	24
	Percentages of Items			
Strand	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Operations and Algebraic Thinking (ALGE)	22	21	22	21
Number Operations–Base Ten (BTEN)	17	17	17	17
Number Operations–Fractions (FRCT)	34	33	34	33
Measurement and Data (MEAS)	17	17	17	17
Geometry (GEOM)	10	13	10	13

Table 12. Reporting Strands Design for Mathematics Grade 5

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Operations and Algebraic Thinking (ALGE)	6	3	6	2
Number Operations–Base Ten (BTEN)	16	6	16	6
Number Operations–Fractions (FRCT)	19	8	19	8
Measurement and Data (MEAS)	12	5	12	5
Geometry (GEOM)	6	2	6	3
Total	59	24	59	24
	Percentages of Items			
Strand	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Operations and Algebraic Thinking (ALGE)	10	13	10	8
Number Operations–Base Ten (BTEN)	27	25	27	25
Number Operations–Fractions (FRCT)	32	33	32	33
Measurement and Data (MEAS)	20	21	20	21
Geometry (GEOM)	10	8	10	13

Table 13. Reporting Strands Design for Mathematics Grade 6

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Geometry (GEOM)	6	3	6	3
Ratios and Proportional Relationships (RELT)	10	4	10	4
The Number System (NMBR)	15	6	15	6
Expressions and Equations (EQTN)	20	8	20	8
Statistics and Probability (STPR)	8	3	8	3
Total	59	24	59	24

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Geometry (GEOM)	10	13	10	13
Ratios and Proportional Relationships (RELT)	17	17	17	17
The Number System (NMBR)	25	25	25	25
Expressions and Equations (EQTN)	34	33	34	33
Statistics and Probability (STPR)	14	13	14	13

Table 14. Reporting Strands Design for Mathematics Grade 7

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Geometry (GEOM)	9	4	9	4
Ratios and Proportional Relationships (RELT)	13	5	13	5
The Number System (NMBR)	10	4	10	4
Expressions and Equations (EQTN)	18	7	18	7
Statistics and Probability (STPR)	9	4	9	4
Total	59	24	59	24

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Geometry (GEOM)	15	17	15	17
Ratios and Proportional Relationships (RELT)	22	21	22	21
The Number System (NMBR)	17	17	17	17
Expressions and Equations (EQTN)	31	29	31	29
Statistics and Probability (STPR)	15	17	15	17

Table 15. Reporting Strands Design for Mathematics Grade 8

Strand	Number of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Geometry (GEOM)	16	6	16	6
The Number System (NMBR)	6	3	6	2
Expressions and Equations (EQTN)	23	9	23	9
Statistics and Probability (STPR)	6	2	6	3
Functions (FNCT)	14	6	14	6
Total	65	26	65	26

Strand	Percentages of Items			
	2015		2016	
	Total Test Items	Anchor Items	Total Test Items	Anchor Items
Geometry (GEOM)	25	23	25	23
The Number System (NMBR)	9	12	9	8
Expressions and Equations (EQTN)	35	35	35	35
Statistics and Probability (STPR)	9	8	9	12
Functions (FNCT)	22	23	22	23

3.3.3. PAWS Science Tests

The 2008 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 2008 Wyoming Science Content and Performance Standards requires schools and districts to include instruction and monitoring of student achievement in the areas of the history and nature of science at the local level, but these measures are not assessed by the PAWS at present.

In order to accurately reflect the expectations of the 2008 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 2008 Wyoming Science Content and Performance Standards establish the expectations that teachers will judge where students are performing in relation to the benchmarks, and ultimately, the standards. To evaluate students' mastery against the 2008 Wyoming Science Content and Performance Standards, teachers are expected 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 2008 Wyoming Science Content and Performance Standards, notably, the science performance inferences to be drawn from the results.

Based on this design, the PAWS Science assessment items are written to measure students' mastery of science inquiry skills within the context of the benchmarks from Standard I: Concepts and Processes. The items are distributed equally among the physical science, life science, and earth/space science benchmarks. Over the course of a two-year cycle, each of the inquiry skills is assessed within the context of each benchmark in Standard I Concepts and Processes. All too often, students' understanding of core concepts and scientific theories is measured without careful attention to how students internalize core assumptions, apply important ideas, or make connections to relevant everyday experiences. Without measurement of such epistemological standards, teachers will not know whether students have a firm foundation on which to base scientific arguments.

The design of both the 2008 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 2016 PAWS Science test was based on the 2015 PAWS Science Blueprint. The content of the test is aligned to the Science as Inquiry content standard of the 2008 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 2008 Wyoming Science Content and Performance Standard I: Concepts and Processes, and a score analysis is reported in each of the following areas:

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

The number of items assessing each skill-reporting category and content standard is constant across all grade levels; they are provided in Appendix A. Tables 16 and 17 provide the number of items (and points) for each Science assessment, by reporting strand for the assessment overall, and for the anchor item set. The anchor item set is utilized for year-to-year equating. These tables include similar information for 2015 and 2016 for comparison purposes. There is some fluctuation between the percentage of the total raw score represented by the reporting strand and the percentage of anchor item points for the strand. Across all grades, the representation of the anchor sets remained stable although the number of anchor items was reduced to ensure equal representation.

Table 16. Reporting Strands Design for Science Grade 4

Strand	Number of Items 2015 Total Test Items	Number of Items 2015 Anchor Items	Number of Items 2016 Total Test Items	Number of Items 2016 Anchor Items
Life Science (LIFE)	16	7	16	6
Physical Science (PHYS)	18	7	18	7
Earth and Space Science (ESCI)	16	6	16	7
Totals	50	20	50	20

Strand	Percentages of Items 2015 Total Test Items	Percentages of Items 2015 Anchor Items	Percentages of Items 2016 Total Test Items	Percentages of Items 2016 Anchor Items
Life Science (LIFE)	32	35	32	30
Physical Science (PHYS)	36	35	36	35
Earth and Space Science (ESCI)	32	30	32	35

Table 17. Reporting Strands Design for Science Grade 8

Strand	Number of Items 2015 Total Test Items	Number of Items 2015 Anchor Items	Number of Items 2016 Total Test Items	Number of Items 2016 Anchor Items
Life Science (LIFE)	16	7	17	7
Physical Science (PHYS)	18	7	18	7
Earth and Space Science (ESCI)	16	6	15	6
Totals	50	20	50	20

Strand	Percentages of Items 2015 Total Test Items	Percentages of Items 2015 Anchor Items	Percentages of Items 2016 Total Test Items	Percentages of Items 2016 Anchor Items
Life Science (LIFE)	32	35	34	35
Physical Science (PHYS)	36	35	36	35
Earth and Space Science (ESCI)	32	30	30	30

3.4 PAWS Test Development Process

A state committee consisting of regional representatives utilized national and regional documents to establish that the rigor of the Wyoming language arts standards are consistent with these documents, and adjustments were made as deemed appropriate by the state committees.[4] The 2012 WyCPS in English Language Arts address three content standards: (1) Reading, (2) Writing, and (3) Speaking and Listening. Content standard 2 (Writing) and 3, (Speaking and Listening), are not currently assessed by PAWS.

As noted elsewhere in this report, multiple choice items (only) were used on the PAWS Reading, Mathematics, and Science portions.

The 2012 WyCPS in Mathematics 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.

The 2008 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.

Initial creation of blueprints, item and passage specifications, and assessment descriptions took place in the fall of 2012. Development of these documents has been an ongoing process, and they guided the development, review, and field testing of items for use on the PAWS assessments.

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

[4] These documents included the following publications:

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

their statistics from field testing (see Section 3.7.4, Data Review Procedures). 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 2016 forms considering both the content and psychometric properties of each item selected.

3.5 Item and Test Form Development

In this section, the general process for item development is described. Using the 2012WyCPS and 2008 Wyoming Science Content and Performance Standards as a foundation, test blueprints were developed by the WDE setting forth the number of items for each Reading, Mathematics, or Science content standard. These blueprints were initially developed in the fall of 2012 for Reading and Mathematics and 2006 for Science. They have been refined during the course of the program, balancing the need to provide a high level of information about student ability to inform instruction against the desire to impinge upon instructional time as little as possible.

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

3.5.1. Item Specifications

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

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

3.5.2. Item Difficulty Requirements

The Rasch measurement model was used to develop the scale for each of the PAWS Reading, Mathematics, and Science 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., reading, mathematics, or science). See Section 6.3 for further detail on the Rasch model.

3.5.3. Item Graphics Requirements

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

3.6 Item Review

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

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

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

- Knowledge of the 2012 WyCPS and 2008 Wyoming Science 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 receive 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 are evaluated by WDE and ETS. All of the feedback generated by the reviewers is utilized to make final decisions on which items to accept and what revisions to include in the version of the item that is 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 provides input on potential bias and/or sensitivity in the test content. With regard to fairness and content, panelists suggest revision or deletion of items as they deem necessary. Any items surviving this rigorous examination becomes part of the pool of items eligible for field testing.

3.7 Field Testing

During the 2016 PAWS administration, field-test items were embedded within each operational exam for Reading, Mathematics, and Science. For each subject and grade, there were ten different forms, each containing a different set of field-test items. Some field-test items were used in more than one form. Each form within a subject/grade contained the same number of field-test items, administered in the same positions across forms. Forms were spiraled within classroom and school in order that randomly equivalent samples of students would receive each of the forms; each form was responded to by approximately 700-800 students, more if the item appeared on more than one form. 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.

Students' 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 future use.

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

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

3.7.1. Classical Item Statistics

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

- *N*-counts for each statistic;
- Item difficulty (or average item score);
- Item discrimination (or point biserial correlation);
- Multiple choice item distractor discrimination;
- Multiple choice item response; and
- DIF statistics (Mantel & Haenszel, 1959) by gender and ethnicity.

3.7.1.1. Item Difficulty

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

3.7.1.2. Item Discrimination

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

3.7.2. Differential Item Functioning

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

If the item is differentially more difficult for an identifiable subgroup when conditioned on ability, it may be measuring something different from the intended construct. However, it is important to recognize that DIF-flagged items might be related to actual differences in relevant knowledge or skills (item impact) or statistical Type I error. As a result, DIF statistics are used to identify items that are potentially functioning differentially. Subsequent review by content experts and bias/sensitivity committees are required to determine the source and meaning of

performance differences. For the spring 2016 PAWS Reading, Mathematics, and Science tests, DIF analyses were conducted for gender groups (Male/Female) and race/ethnicity groups (White/American Indian (Native American, White/Asian, White/Black (African American), and White/Hispanic (Latino)) where sample size was sufficient.

The Mantel-Haenszel procedure (Mantel & Haenszel, 1959) was conducted to detect DIF for multiple-choice items. As part of the Mantel-Haenszel procedure, the statistic described by Holland and Thayer (1988), known as MH D-DIF, was used.

The formula for the estimate of constant odds ratio is:

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

where

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

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

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

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

N_m = total group at ability level m.

This statistic is expressed as the differences between members of the “focal group” (female, Asian, African American, Hispanic/Latino, & 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. Items are

classified into one of three categories and assigned values of A, B, or C based on these DIF statistics. Category A items contain negligible DIF. Category B items exhibit slight or moderate DIF. Category C items have moderate to large values of DIF. Negative values imply that, conditional on the matching variable, the focal group has a lower mean item score than the reference group. In contrast, a positive value implies that, conditional on total test score, the reference group has lower mean item score than the focal group. The flagging criteria for multiple-choice items are provided in Table 18.

Table 18. 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 E summarizes the number and percentage of items by DIF category from the 2016 field test items for each grade and content area. The 2016 operational tests are composed of items that were piloted in years prior to 2016, which were reviewed and approved by Content Review, Bias and Fairness Review, and Data Review Committees.

3.7.3. Item Response Theory (IRT) Analysis

Rasch model was used to scale the PAWS. IRT model 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 (*b*-value). The probability of a correct response to an item increases as the student’s ability increases. See Chapter 6 for further details on the Rasch model. The results of the Rasch IRT analyses of the field test items can be found in Appendix H.

3.7.4. Data Review Procedures

Following the spring 2016 PAWS administration the statistics discussed above were computed for each item field tested. These statistics will be compiled into books along with images of the items for use in data review meetings. Each item will appear on one page of the data review book with its statistics on the opposite page. An item with any statistics outside pre-established limits will have an appropriate annotation.

Field test items are evaluated by panels of Wyoming state educators selected by the WDE. Each data review panel consists of 8–12 educators with experience in the target grade and subject. Items field tested during the 2016 administration were reviewed in July 2016 by a panel in Sheridan, Wyoming.

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

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

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

The data review meetings begin with a training session led by an ETS assessment lead and psychometrician. This session covers the statistics that the panelists will be using as they evaluate each item, the meaning of each in the context of evaluating item quality and suitability for use on future operational exam forms, and the role of the panelists' expertise in the data review process.

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

and/or with distractors with positive discriminations were culled out, along with items flagged for possible DIF.

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

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

The results of the Rasch IRT analyses of the field test items can be found in Appendix H, the PAWS classical analysis results appear in Appendix F, and PAWS DIF in Appendix E. Items accepted at data review from the 2016 administration are eligible for use as operational items beginning with the 2017 administration.

3.8 Test Form Construction

After each administration, analyses were conducted by the ETS psychometrician to determine the statistical properties of all items that were present on any of the forms (both operational items and field test items). This includes estimation of Rasch difficulty parameters on the current scale for all items. Thus, all items that have been field tested or used operationally were equated to the original scales and have known Rasch difficulty. Therefore, when forms were constructed for the 2016 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 base PAWS tests, as test characteristic curves (TCCs), information, and standard error curves could be evaluated to help ensure statistical comparability.

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

ETS utilized proprietary test construction software for the construction of the 2016 forms to the base scale for reading, mathematics, and science. The ETS psychometrician utilized the test content blueprint and the preliminary statistical targets in a configuration file for each grade and subject test being constructed. The 2016 blueprints can be found in Appendix A. The blueprints remain unchanged from 2014 for reading and mathematics and 2013 for science. The TCCs, information, and standard error curves from the base year administration, 2014 for reading and mathematics and 2013 for science, constituted the statistical targets for the 2016 forms.

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

The assessment development leads assembled a draft form conforming to the blueprint and statistical targets; then it was reviewed and edited by the psychometrician. The test construction software provided real-time feedback on the psychometric properties of the form, allowing the psychometrician and content staff to immediately see the results of a proposed change in the items on the form. Finally, the software noted the items' positions, used to minimize the difference between the items' position on the 2016 form and its position on the form from its most recent use.

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

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

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

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

3.8.2. Final Review of Assembled Operational Tests

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

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

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

4. TEST ADMINISTRATION

4.1 Test Materials

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

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

4.2 Materials Return

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

4.3 Directions for Administration

The *PAWS Directions for Administration* provided the guidelines for planning and managing the PAWS administration for district and school administrators. The *PAWS Directions for Administration* provided specific directions for test administrators, from scheduling and timing for sessions and preparing students to testing students from special populations. Two half-day comprehensive Building Coordinator training sessions conducted jointly by the WDE and ETS were held in January 2016 prior to the 2016 testing window. All test administrators around the state were expected to view the Building Coordinator Training Video before the test window opened. Building principals required test administrators as well as anyone handling test materials to sign off assuring they had been trained on test security and how to administer the test. These certification documents were retained in the school and were available to the WDE upon request.

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

6–8 of Mathematics were administered in three untimed sessions, one non-calculator and two calculator sessions.

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

4.3.1. Allowed Student Manipulatives

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

4.3.2. Test Security

PAWS test security guidelines strictly prohibit the photocopying of all or any part of a test booklet, and require that all violations of the WDE's guidelines be reported to the WDE immediately. The reporting of violations to the WDE ensured that test scores could be invalidated if necessary. All test booklets were considered secure materials. The PAWS Building Coordinators were required to document the receipt of secure materials, check the lists of students, and return all test materials to ETS for scoring.

The specific procedures that were to be followed during any test administration and used in the handling of documentation were outlined in the 2016 *PAWS Directions for Administration*. 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 and after testing.
- Not provide answers to any test items to any students.
- Not change or otherwise alter a student's answer.
- Follow the suggested time periods as closely as possible in order to maintain uniformity in the test administration. (Note: PAWS is an untimed test.)
- Follow the *Directions for Administration* manual explicitly.
- Follow all Ethics and Security Requirements as outlined in the 2016 *PAWS Directions for Administration*. If there is a violation, the students' materials *will not be scored* and the school will not be able to count the student(s) for participation.

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

4.4 Student Participation

As noted previously, all Wyoming students in grades 3 through 8 were required to participate in the regular PAWS tests, the PAWS with appropriate accommodations, or the Wy-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 (except English learners in their first year of education in the U.S., students who were expelled, and students educated in out-of-state residential institutions). Students with disabilities, who were on a 504 Plan, or who were ELL were allowed to be provided with standard accommodations during the administration of PAWS consistent with guidance provided by the WDE. Students with significant cognitive disabilities were required to take the Wy-ALT as determined by their IEP teams.

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

- Participation in PAWS regular assessment without accommodation
- Participation in PAWS regular assessment with standard accommodation
- Participation in Wy-ALT

4.5 PAWS 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 for students with disabilities, for students on a 504 Plan, and ELLs. The WDE recognizes that the proper administration of standard accommodations allows these students access to the test, resulting in the students' ability to demonstrate their knowledge and skills consistent with the measured test constructs in each content area. Often the conditions under which the test was standardized differ from those present when accommodations were used. These differences, in some cases like reading the

reading passages, reached a level sufficient to jeopardize the validity of interpretations. However, based on available evidence, the standard accommodations allowed for PAWS were considered incidental to the construct intended to be measured by the test (AERA, APA, & NCME, 2014, p.67) by the WDE. Thus, students using accommodations received scores on PAWS 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* (AERA, APA, & NCME, 2014).

The administration of standard accommodations during PAWS 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 WDE in conjunction with the Council of Chief State School Officers (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 standard accommodations on the WDE website (See http://edu.wyoming.gov/educators/state-assessment/paws/).

Two addendums related to the administration of standard accommodations were distributed through postings on the WDE website including the *Wyoming Statewide Assessment System 2016 Standard Accommodations*
https://edu.wyoming.gov/downloads/assessments/2016StandardAccommodations.pdf

and the 2016 *PAWS Frequently Asked Questions (FAQ)*
https://edu.wyoming.gov/downloads/assessments/paws-frequently-asked-questions-2015-16.pdf. The *Wyoming Statewide Assessment System 2016 Standard Accommodations* document provides information about the administration of standard accommodations and also identifies the allowable standard accommodations, divided into four categories (presentation, response, setting, and timing and scheduling). The FAQ document provides information about the administration and documentation of standard accommodations as well as detailed information regarding specific accommodations including the administration of standard accommodations for ELL students, best practices associated with the selection and administration of accommodations, and a specific list of standard accommodations for ELL students.

4.5.1. Students Eligible for Test Accommodations

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

4.5.2. Requirements for Use of Test Accommodations

For students with disabilities, the selection of accommodations for the general assessment was the responsibility of a student's IEP team or 504 Plan committee. Guidance was provided in the *Wyoming Accommodations Manual for Instruction and Assessment: How to Select, Administer, and Evaluate Use of Accommodations for Instruction and Assessment of Students with Disabilities* (January 2006). Currently permitted are standard accommodations for students with disabilities, 504 Plans, or who were ELLs and were listed in the 2016 *Directions for Administration*. 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 2016 PAWS standard accommodations.
4. Standard accommodations were administered as described in the *Wyoming Statewide Assessment System 2016 Standard Accommodations* https://edu.wyoming.gov/downloads/assessments/2016StandardAccommodations.pdf and the *Wyoming Accommodations Manual for Instruction and Assessment* https://edu.wyoming.gov/downloads/assessments/PAWS_Accommodations_Manual.pdf .
5. The accommodations provided were effective in providing access to the test and had been regularly used by the student during instruction and classroom assessment.
6. The accommodations were administered by a trained Test Administrator or access assistant who was familiar to the student.

Accommodations could not:

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

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

1. Understood the procedures for administering standard accommodations; and
2. Has effectively administered the accommodation(s) to the student during instruction and/or assessment; and
3. Has attended a 2016 PAWS Training or has viewed the 2016 PAWS Training online and submitted record of the training to the building principal; and

PAWS administrations were untimed for all students. Large print, English/Spanish audio (Mathematics and Science only), and braille versions of PAWS were available for all grade levels and content areas.

4.5.3. Description of Standard Accommodations for Students with Disabilities

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

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

Appropriate documentation and monitoring of the standardized use of accommodations was required of test administrators, test coordinators, and/or principals.The Special Programs Unit reviewed documentation of accommodations during on-site monitoring visits. The following assessment accommodations were allowable for students with an IEP or 504 Plan.

4.5.3.1. Presentation Accommodations

1. Student uses a Braille Special Test Form.
2. Student uses a Large Print Special Test Form.
3. Student uses an Audio Special Test Form.
4. Student uses magnification devices.
5. Student uses color overlays to reduce glare or enhance text.
6. Student uses templates to reduce the amount of visible print.

7. Student uses tactile graphics.

8. Sign language interpreter signs directions in all content areas and/or signs test questions as written in all content areas EXCEPT reading. The interpreter may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Interpreters need to be familiar with the terminology and symbols specific to the content. It is recommended that one interpreter be provided for each individual student.

9. A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.

10. A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.

11. Student asks for clarification of directions (not test questions or answer choices).

12. Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.

13. Student uses text-to-speech software in all content areas EXCEPT reading.

4.5.3.2. Response Accommodations

14. A certified staff member or access assistant scribes what a student dictates through alternate augmentative communications (AAC), pointing, sign language, or speech. The scribe may not edit or alter the student's work in any way and must record, word for word, exactly what the student has dictated. A scribe must allow the student to review and edit what that student has written. The student's final response must be transcribed by a certified staff member or access assistant into the Student Test and Answer Book on the pages in which the student's response is to be written.

15. A student types responses using a word processor. Dictionary and synonym/thesaurus devices MUST be disabled. The margins for word-processed documents should match the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages in which the student's response is to be written.

16. Student uses speech-to-text conversion or voice recognition in all content areas. The margins for this document should match as closely as possible the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages in which the student's response is to be written.

17. Student uses a Brailler. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book or answer document on the pages in which the student's response is to be written.
18. Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book or answer document on the pages in which the student's response is to be written.
19. A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book or answer document.
20. Student uses visual organizers including graph paper, place markers, and templates. Student uses a pencil to underline text. Highlighters CANNOT be used in the Student Test and Answer Book or answer document.

4.5.3.3. Setting Accommodations

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

4.5.3.4. Timing and Scheduling Accommodations

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

4.5.4. Description of Standard Accommodations for ELL

Schools could not exempt ELL students from the PAWS 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, 2016, could be waived from taking the PAWS Reading content assessments with an exemption approved by the Wyoming Department of Education. Students who received this exemption took the ACCESS for ELLs assessment instead of the Reading portion of PAWS, but they were not exempted from the mathematics and science portions of PAWS.

ELL students could be provided with accommodations during PAWS 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 2016 administration.

4.5.4.1. Presentation Accommodations

25. A certified staff member or access assistant translates written directions to the student.
26. A certified staff member or access assistant re-reads, simplifies, or clarifies directions in English or in the student's primary language (NOT test questions or answer choices) without clueing correct responses.
27. A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT reading. 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.

4.5.4.2. Setting Accommodations

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

4.5.4.3. Timing and Scheduling Accommodations

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

4.5.5. PAWS 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 IEP files looking for evidence that IEP teams have made sound accommodations decisions to enable students with disabilities to gain access to instructional content and assessment measures. In addition, general and special education teachers, administrators, and service providers are interviewed to provide further information about school and district practices regarding accommodations. Failure to provide

accommodations listed in a student's IEP or failure to thoughtfully consider accommodations for a student or students may contribute to a finding of noncompliance, thus requiring the district to address the issue through the creation and implementations of a Corrective Action Plan (CAP). Monitoring of standard accommodations for ELL's was provided by the Local Education Agency.

4.5.5.1. Empirical Analysis of Accommodations

IEP and 504 Plan students comprised approximately 11%–14% of students at each grade level, with between 60%–80% of those IEP and 504 students receiving testing accommodations (depending on grade and subject). While Wyoming allows 31 specific accommodations on PAWS as described herein, the overwhelming majority across all content areas were provided as auditory presentations (e.g., reading directions, reading questions, clarifying directions, or the audio form), setting accommodations (i.e., testing in a separate location), or an accommodation in timing/scheduling (e.g., extended time, multiple breaks, test over multiple days). This breakdown by specific accommodation also provides a baseline for monitoring accommodations in future years. Frequency tables for accommodations provided during the 2016 PAWS for Reading, Mathematics, and Science for all grades are presented in Appendix J. In general, students with disabilities who did not receive accommodations had higher mean scale scores. Mean scale scores for students with disabilities broken down by accommodation status are presented in Appendix K.

4.5.6. Selection and Administration of Accommodations

An important question regarding the use of accommodations in large-scale assessment is whether the resultant student scores mean the same thing as scores resulting from non-accommodated assessment (Kim, Wang, Zhao, & Li, 2006). In other words, do the accommodations yield meaningful, valid scores of the level of a student's subject mastery? It is also imperative to know the effect of including scores of accommodated students in test calibration, 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 ELLs, and for the continued evaluation of the standard accommodations with regard to current research.

Standard accommodations were implemented for students with disabilities, students with 504 Plans, and the 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. WDE and ETS 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.

5. PROCESSING AND SCORING OF PAWS ITEMS

5.1 Overview

This chapter describes the receipt control, scanning, and scoring procedures used at ETS for the 2016 PAWS.

At the close of testing, the PAWS Student Test and Answer Books (grades 3–5) and Answer Documents (grades 6–8) were returned to ETS. Upon receipt, they were scanned into ETS's electronic imaging system. Subsequent processing of student responses necessary to score those responses and to produce reports used these images rather than the paper documents. After scanning, the physical documents were put into archival storage. Student responses to PAWS multiple-choice test items were machine-scored. Correct answers were assigned a score of one point and incorrect answers were assigned a score of zero points.

5.2 Receipt Control, Processing, Scanning, Editing

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

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

5.2.1 Receipt Control

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

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

5.2.2 Processing

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

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

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

5.2.3 Scanning

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

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

The images produced by the scanner included document identification and all information gridded by the test-taker and were stored as 8-bit (256 level) grayscale images. The scanning program checked the validity of the document identification using optical mark recognition (OMR), skunk codes, and optical character recognition (OCR) module codes to ensure that the booklet that was being scanned was the correct booklet. The scanning program also compared the actual number of pages scanned to the number of pages expected for the document according

to its identification. These two checks ensured that the correct document was being imaged and that the entire document was imaged. Finally, the skunk and module codes acted as reference points indicating the orientation of the document as it moved through the scanner.

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

5.2.4 Resolutions

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

Another step in the paper resolution 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 resolution 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.

6. EQUATING AND SCALING

6.1 Overview

This chapter covers equating procedure of the 2016 PAWS Reading, Mathematics, and Science assessments. Translation of raw scores to scale scores along with descriptive statistics for all of the 2016 PAWS scales is also described.

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

6.2 Item and Forms Development

Kolen and Brennan (2014) state that equating adjusts for differences in test form difficulty, not for differences in content. Reading and Mathematics have a new 2012 style guide based on the 2012 WyCPS. Science items have been developed to the same style guide since 2005 (with minor updates throughout), and tests have used comparable blueprints since the first operational administration in 2008. Science test blueprints did change between the 2012 and 2013 administrations with the removal of constructed response items from the assessments. There were no changes in 2016.

6.3 IRT Model and Calibration

Rasch model (Rasch, 1980; also known as one parameter logistic model) was used to calibrate the 2016 PAWS assessments for dichotomous items. This measurement model is used regularly to construct test forms, for scaling and equating, and to develop and maintain large item banks. All test analyses, including item model fit analysis, equating, and performance prediction were accomplished within this framework. The statistical software used to calibrate the PAWS operational and field test items that were used in the spring 2016 administration was *WINSTEPS Version 3.68.1* (Linacre, 2007).

The most basic expression of the Rasch model is in the Item Characteristic Curve (ICC). It conceptualizes the probability of a correct response to an item as a function of the student's ability level and the difficulty of the item. The probability of a correct response is bounded by 1

(certainty of a correct response) and 0 (certainty of an incorrect response). The ability scale is theoretically unbounded. In practice, the ability scale ranges from approximately –4 to +4 logits for heterogeneous ability groups. The probability of an examinee with ability $\hat{\theta}$ answering item *i* with difficulty D_i correctly is shown in the equation below:

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

As an example, consider Figure 5.1, in which the response probability curve for a dichotomous item is depicted with a Rasch difficulty (D_i) of 0.85. When a person answers a dichotomous item with a difficulty that is at the same level as their ability (ability is represented by 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).

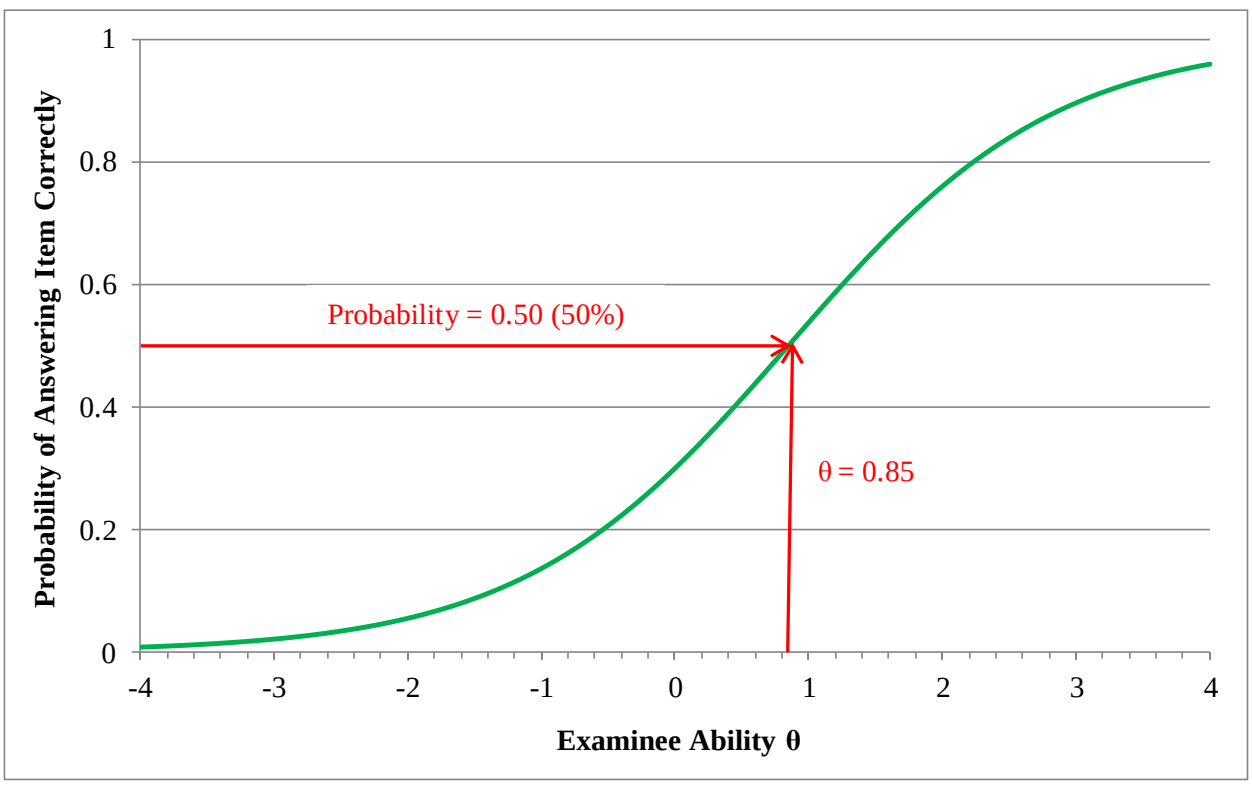

Figure 1. Graphical Description of Item Characteristic Curve (ICC)

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

6.4 Fit Statistics for the Rasch Model

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

Outfit mean-squares are influenced by outliers and are usually easy to diagnose. Infit mean-squares, on the other hand, are influenced by response patterns and are harder to diagnose and remedy. Table 19 presents guidelines for evaluating mean-square fit statistics (Linacre, 2007).

Table 19. Criteria to Evaluate Mean-Square Fit Statistics

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

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

Appendix I provides Rasch difficulty estimates, standard errors, and infit and outfit statistics for 2016 PAWS operational items. The majority of fit statistics 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 equating the 2016 PAWS Reading, Mathematics and Science operational tests. Operational classical item statistics are presented in Appendix G.

Appendix H provides IRT statistics and *N*-counts for items field-tested in 2016. Item fit is a factor that is considered during test construction, and items with less than optimal fit statistics that survive data review are not likely to be used on future PAWS forms.

6.5 Reading and Mathematics Vertical Scales

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

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

It is important to note that vertical scaling produces scales that are *linked* across adjacent grades as opposed to scales that are *equated*. Linked scales are comparable, but have a weaker relationship than equated scales. This relationship is strongest across adjacent grades and weakens as the gap between the grades being compared widens. This is due to the fact that the tests from adjacent grades cover different subject matter that is specific to their targeted grades. For an equating relationship to exist, the test forms that are being equated should cover the same subject matter. Thus, test forms from the same grade and subject are equated from year to year while test forms from adjacent grades (within grades 3 to 8) and the same subject are linked via the vertical scale.

New scales measuring students' academic performance on 2012 WyCPS were established in 2014 for PAWS Reading and Mathematics assessments. The vertical scales for reading and mathematics were developed in 2014 using an operational and embedded vertical anchor test design. The vertical linking items were embedded within the PAWS 2014 assessment test booklets in the field test positions. The vertical linking items did not count toward a student's scale score.

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

A Rasch model was used for calibration and vertical scaling of 2014 PAWS Reading and Mathematics tests. A common-item nonequivalent groups design in which students in adjacent grade levels respond to the same items was used to collect data to build a vertical scale for PAWS Reading and Mathematics assessments. The linkages between adjacent grades were established by fixing the item difficulty parameters of the vertical anchor sets in the upper grade to the values obtained from the calibration of these items in the lower grade. Following the decision from Technical Advisory Committee (TAC) meeting, the grade 3 reading scale was

chosen to be the base scale for Reading assessments. The grade 4 scale was linked to the grade 3, the grade 5 to the grade 4, etc. As a result, reading and mathematics test scores in grades 3–8 are directly comparable across adjacent grades. More information regarding the reading and mathematics vertical scales can be found in the PAWS 2014 Calibration and Scaling Reports for Reading and Mathematics.

6.6 Reading, Mathematics, and Science Equating Analyses

As was previously mentioned, the PAWS Reading, Mathematics, and Science assessments for 2016 were post-equated, meaning that the item and test statistics used to generate the scoring tables (i.e., tables displaying the relationship between specific raw scores and scale scores for a particular grade level and subject) came from the present (spring 2016) administration. All tests were equated to the existing scale, and so scale scores on the 2016 administration used the same metric as scale scores for the same grade level and subject from previous administrations of the PAWS. The Science tests were only given at grades 4 and 8 and therefore were not vertically scaled.

6.6.1. Calibration and Equating Process for the 2016 Administration

The procedures for equating the 2016 forms of the Reading, Mathematics, and Science forms to the preexisting scales were similar to those used in previous years. To establish a strong relationship between the 2015 and 2016 forms, each 2016 form had approximately 30% of its items drawn from the set of 2015 operational items. Other items were drawn from the item bank, which was composed of all items used operationally (with the exception of those items released publicly as sample PAWS items), and items field tested and accepted at data review.

The tests were equated via common item equating to a calibrated item pool (Kolen & Brennan, 2014). All items were drawn from previous years' administrations and can potentially function as anchor items with their parameters being drawn from their most recent operational use[5]. Only core items that had been operationally used in a previous form and deemed to have reliable bank parameter values were selected for the anchor set. Other items that were only field tested in previous administrations were excluded from the anchor set.

There were some items that were identified as possibly having unreliable statistics from their most recent use and such items were removed from being anchor items. Some items had been

[5] 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.

modified since their most recent use (primarily older items modified to bring them in line with current PAWS item style guidelines). Their previous statistics in the item bank might not be comparable to the statistics of the new modified version of the item. These items were removed from the anchor set. All items that were not used as linking items had their parameters freely estimated while holding the parameters of the remaining anchor items fixed.

Though item parameters from IRT models (including Rasch) 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 was 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 2016 administration data differed significantly from their known values used for equating).

Anchor stability analysis was conducted to identify items that were not suitable for use as anchor items. The Robust-Z statistic (Huynh & Meyer, 2010) was utilized to identify items that exhibited instability in their difficulty parameter estimates (multiple choice items) for the 2016 calibration as compared to their difficulty estimates from their most recent previous use.

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

Robust-Z is defined as

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

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

Items with a Robust-Z that exceeded 1.645 were deemed to have drifted in difficulty and were considered for being eliminated from the anchor set in the previous protocol. However, ETS retained all flagged item(s) in the anchor set if the items were not identified to be flawed by content experts. This departure from previous protocol was accepted by the WDE based on

Yen's (2007) white paper. In the second round of anchor stability checks, anchored items' displacement values were also examined. Linacre (2007, p. 362) describes displacement statistic as:

> *. . . the size of the change in the parameter estimate that would be observed in the next estimation iteration if this parameter was free (unanchored) and all other parameter estimates were anchored at their current values. For a parameter (item or person) that is anchored in the main estimation, (the displacement value) indicates the size of disagreement between an estimate based on the current data and the anchor value.*

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

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

$$TrueScore = \sum_{i=1}^{n} P_i(\theta)$$

where $P_i(\theta)$ implies probability of a correct response for each of the items. That is, the true score of a student with ability θ is the sum of the item characteristic curves. And true score is set to each achievable raw score point to find its theta equivalent.

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

6.7 Translating Raw Scores to Scale Scores and Performance Levels

Scale scores on the PAWS Reading, Mathematics, and Science tests ranged generally from 300 to 975 for grades 3–8; the specific minimum and maximum possible scale scores varied by grade and subject. Appendix L provides scale score descriptive statistics for the 2016 PAWS Reading, Mathematics, and Science tests. The following formulae were used to convert the underlying PAWS IRT Reading, Mathematics, and Science scales to the PAWS reporting scale:

$$PAWS\ Scale\ Score = \hat{\theta} \times Slope + Intercept$$

$$PAWS\ Scale\ SEM = SEM(\hat{\theta}) \times Slope$$

where $\hat{\theta}$ was the *IRT* ability estimate, and *SEM*($\hat{\theta}$) was the estimated conditional standard error of measurement (SEM) of the ability estimate $\hat{\theta}$. Table 20 also contains the slope, intercept, and lowest obtainable scale score (LOSS) and highest obtainable scale score (HOSS) for the PAWS Reading, Mathematics, and Science scales.

The raw score to scale score conversion tables for the 2016 PAWS Reading, Mathematics, and Science tests can be found in Appendix M. Conditional standard error estimates and performance levels for the scale scores are also included in these tables.

Table 20. PAWS Scaling Constants and Lowest and Highest Obtainable Scale Scores

Grade	Scaling constant	LOSS	HOSS
	Reading		
Grade 3	Scale Score = $\hat{\theta}$ * 43.89281 + 553.1639	375	800
Grade 4	Scale Score = $\hat{\theta}$ * 43.89281 + 553.1639	400	825
Grade 5	Scale Score = $\hat{\theta}$ * 43.89281 + 553.1639	425	850
Grade 6	Scale Score = $\hat{\theta}$ * 43.89281 + 553.1639	450	875
Grade 7	Scale Score = $\hat{\theta}$ * 43.89281 + 553.1639	475	900
Grade 8	Scale Score = $\hat{\theta}$ * 43.89281 + 553.1639	500	925
	Mathematics		
Grade 3	Scale Score = $\hat{\theta}$ * 43.4074 + 570.41	375	850
Grade 4	Scale Score = $\hat{\theta}$ * 43.4074 + 570.41	400	875
Grade 5	Scale Score = $\hat{\theta}$ * 43.4074 + 570.41	425	900

Grade 6	Scale Score = $\hat{\theta}$ * 43.4074 + 570.41	450	925
Grade 7	Scale Score = $\hat{\theta}$ * 43.4074 + 570.41	475	950
Grade 8	Scale Score = $\hat{\theta}$ * 43.4074 + 570.41	500	975
	Science		
Grades 4 and 8	Scale Score = $\hat{\theta}$ * 48.21 + 637.5	300	975

7. SCORE REPORTING

7.1 Overview

A thorough understanding of the results of the PAWS assessments is essential for all members of the school community (parents, teachers, administrators, and students) in order to report on student achievement and to help all students to meet grade level expectations. This level of assessment literacy is only possible if professional educators are well versed in assessment practice and assessment results are presented clearly. Sample student reports are located in Appendices B and C for PAWS. Appendix B shows the PAWS Reading, Mathematics, and Science reports for grades 4 and 8. Appendix C demonstrates the report for PAWS Reading and Mathematics for grades 3, 5, 6, and 7.

The following reporting information is provided:

- Scale Scores of each Subject and Domain Standard
- Performance Levels of each Subject and Domain Standard
- State Average Scores of each Domain Standard
- Production of PAWS Individual Student Score Reports

7.2 Scale Scores of each Subject

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

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

7.3 Performance Levels of each Subject

Performance classifications are determined by applying the appropriate scale score cuts established from the PAWS standard setting activities described in the 2014 Standard Setting Summary Proficiency Assessments for Wyoming Students (PAWS for Reading and Mathematics) (Baron, 2014) and 2008 (Science) PAWS standard setting reports (Pearson, 2008). Tables 21–23 provide the scale score ranges for the PAWS Reading, Mathematics, and Science tests.

Table 21. Proficiency Level Ranges for Reading Grades 3–8

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

Table 22. Proficiency Level Ranges for Mathematics Grades 3–8

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

Table 23. Proficiency Level Ranges for Science Grades 4 and 8

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 reading, mathematics, and science are included on the Student Score Report.

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

7.4 Domain Standard-Level Scale Scores and Performance Levels

Domain standard-level scores, provided on the individual student report, (i.e., subscale scores, reporting categories, or strands) are provided in the form of scale scores. The domain standard-level scores were produced in the same way as the overall test scale scores—a raw score to scale score table for each domain standard within a particular form/grade/content area combination was derived using the Rasch item parameters of the items that mapped to that standard.

Detailed information regarding domain standard-level categories for reading, mathematics, and science can be found in the blueprints in Appendix A. For Reading, scale scores are provided for Literature: Key Ideas and Details, Literature: Craft and Structure, Informational: Key Ideas and Details, Informational: Craft and Structure, Integration of Knowledge and Ideas, and Language. In Mathematics, scale scores are provided for Grade 3-5: Operations & Algebraic Thinking, Number Operations – Base Ten, Number Operations – Fractions, Measurement & Data, and Geometry, Grades 6-7: Ratios & Proportional Relationships, The Number System, Expressions & Equation, Geometry, and Statistics & Probability, and Grade 8: Functions, The Number System, Expressions & Equation, Geometry, and Statistics & Probability. For Science, scale scores are given for Life Science, Physical Science, and Earth and Space Science.

Domain standard-level scale scores and associated error ranges (student scale score +/– one CSEM) are provided electronically to the WDE for further aggregated reporting. Assuming a normal distribution of scale scores, the probability that a student's true score will be in the range indicated by the error bar is approximately 68%. Since measurement error is related to the number of reliable items making up the measure, the error ranges for the domain standard-level scores will generally be larger than those for the overall content-level scale score. On the individual student report, domain standard-level scale score is reported for the individual student and the state average for comparison purposes. The state averages are provided in Table 24.

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

Table 24. State Average Scores by Domain Standard-Level

Reading						
			Grade			
Strand	3	4	5	6	7	8
LTKY	599	621	630	636	651	656
LTCR	594	621	621	630	650	656
INKY	599	624	632	640	653	658
INCR	600	620	623	637	649	653
INTG	599	622	623	635	653	655
LANG	600	617	633	639	652	659
Mathematics						
			Grade			
Strand	3	4	5	6	7	8
GEOM	600	640	657	680	700	708
MEAS	605	647	665	-	-	-
ALGE	607	648	659	-	-	-
BTEN	598	643	666	-	-	-
FRCT	602	650	668			-
RELT	-	-	-	680	700	-
FNCT	-	-	-	-	-	710
NMBR	-	-	-	682	699	709
EQTN	-	-	-	681	701	710
STPR	-	-	-	681	697	708
Science						
			Grade			
Strand	3	4	5	6	7	8
LIFE	-	676	-	-	-	649
PHYS	-	673	-	-	-	649
ESCI	-	675	-	-	-	647

Performance classifications are also reported for domain standard-level scores (Strand Performance Indicators). Strand Performance Indicators are generated by using the domain standard-level standard error (*SE*) to determine a threshold for domain standard-level. For example, if the difference between the total score proficiency theta cut score and the domain standard-level theta score is greater than a specified value, which is a threshold multiplied by the domain standard-level error (i.e., 1.5 * domain standard-level theta *SE*), the resulting domain standard-level classification would be 'Above Average' or "★★★".

The rules used to classify strands are as follows:

- $\theta_{rc} - \theta_p < -1.5 * SE_{rc}$ indicates "★" or "Below Average" performance
- -1.5 * $SE_{rc} \leq \theta_{rc} - \theta_p \leq 1.5 * SE_{rc}$, designates a "★★" or "Average" performance,
- $\theta_{rc} - \theta_p > 1.5 * SE_{rc}$ indicates "★★★" or Above Average" performance, and

where θ_{rc} is the test taker's theta score on a domain standard-level, θ_p is the proficiency theta cut, and SE_{rc} is the standard error for the test taker's theta score on the domain standard-level.

7.5 Production of Printed Score Reports for PAWS

In final preparation for the production and printing of the PAWS 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 QC reports.

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

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

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

7.6 Assessment Score Reports: Supplement Guide for Districts and Schools for PAWS

The 2016 Wyoming State Assessment Program Score Reports: Interpretation Guide for Teachers 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 reports. It was available on the WDE websites, and was intended for use by all users of the data from the PAWS assessment http://www.wyedu.ets.org/rsc/pdf/paws_interpretation_guide_teachers.pdf .

8. RELIABILITY

8.1 Overview

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

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

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

This chapter describes the reliability analyses of the 2016 PAWS operational assessments. Internal consistency, interrater reliabilities, classical and conditional standard errors of measurement, and accuracy and consistency of classification results are included.

8.2 Internal Consistency Reliability

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

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

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

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

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

N = the number of items on the test.

Based on the total test, overall coefficient alpha statistics suggest reasonable internal consistency reliability for PAWS assessments at all grades and subjects. Coefficient alphas were mostly above 0.90 and never lower than 0.89 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 25, including coefficient alpha and the standard error of measurement for each grade and content area. Tables 26–28 provide coefficient alpha and the SEM for each domain within a grade and content area.

8.3 Classical and Conditional Standard Errors of Measurement

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

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

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

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

In the 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 scale score units towards the center of the scale where there are more items and more test information and larger at the extremes where there are fewer items and less test information.

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

Overall Coefficient Alpha and SEM results for PAWS assessments are presented in Table 25. Conditional SEMs for all achievable scores on the assessment are included with the raw score to scale score tables in Appendix M.

Table 25. Reliability and Standard Errors of Measurement by Grade

Grade	*N* Counts	Possible Points	Cronbach's Alpha	SEM
		Reading		
3	7683	50	0.90	2.91
4	7482	50	0.89	2.92
5	7323	54	0.91	2.88
6	7013	56	0.92	3.11
7	7051	56	0.91	3.18
8	6793	56	0.90	3.04
		Mathematics		
3	7683	50	0.92	2.94
4	7485	59	0.92	3.08
5	7325	59	0.93	3.31
6	7011	59	0.92	3.31
7	7054	59	0.92	3.37
8	6804	65	0.92	3.54
		Science		
4	7477	50	0.90	3.00
8	6789	50	0.89	3.25

Table 26. Reliability and Standard Errors of Measurement by Grade and Reading Domain

Grade	Domain	N Counts	Possible Points	Cronbach's Alpha	SEM
3	LTKY	7683	10	0.64	1.26
	LTCR	7683	6	0.48	1.01
	INKY	7683	9	0.65	1.30
	INCR	7683	7	0.49	1.21
	INTG	7683	9	0.66	1.28
	LANG	7683	9	0.67	1.15
4	LTKY	7482	12	0.69	1.39
	LTCR	7482	6	0.47	1.02
	INKY	7482	10	0.66	1.28
	INCR	7482	7	0.48	1.08
	INTG	7482	7	0.56	1.13
	LANG	7482	8	0.62	1.05
5	LTKY	7323	12	0.65	1.36
	LTCR	7323	6	0.55	0.94
	INKY	7323	14	0.77	1.49
	INCR	7323	6	0.59	0.90
	INTG	7323	6	0.47	0.95
	LANG	7323	10	0.66	1.28
6	LTKY	7013	13	0.75	1.45
	LTCR	7013	7	0.61	1.00
	INKY	7013	13	0.76	1.47
	INCR	7013	7	0.57	1.18
	INTG	7013	8	0.68	1.13
	LANG	7013	8	0.56	1.19
7	LTKY	7051	11	0.65	1.24
	LTCR	7051	7	0.57	1.05
	INKY	7051	16	0.78	1.69
	INCR	7051	7	0.56	1.13
	INTG	7051	6	0.46	1.10
	LANG	7051	9	0.68	1.30
8	LTKY	6793	10	0.70	1.20
	LTCR	6793	7	0.47	1.16
	INKY	6793	12	0.65	1.42
	INCR	6793	10	0.60	1.33
	INTG	6793	7	0.47	1.16
	LANG	6793	10	0.69	1.17

Table 27. Reliability and Standard Errors of Measurement by Grade and Mathematics Domain

Grade	Domain	N Counts	Possible Points	Cronbach's Alpha	SEM
3	GEOM	7683	6	0.50	0.99
	MEAS	7683	12	0.72	1.48
	ALGE	7683	20	0.86	1.76
	BTEN	7683	6	0.66	0.93
	FRCT	7683	6	0.62	1.05
4	GEOM	7485	6	0.49	1.00
	MEAS	7485	10	0.64	1.38
	ALGE	7485	13	0.75	1.35
	BTEN	7485	10	0.68	1.13
	FRCT	7485	20	0.83	1.90
5	GEOM	7325	6	0.53	1.03
	MEAS	7325	12	0.77	1.39
	ALGE	7325	6	0.53	1.03
	BTEN	7325	16	0.80	1.65
	FRCT	7325	19	0.85	1.86
6	GEOM	7011	6	0.63	1.03
	RELT	7011	10	0.72	1.27
	NMBR	7011	15	0.76	1.67
	EQTN	7011	20	0.80	1.92
	STPR	7011	8	0.51	1.26
7	GEOM	7054	9	0.57	1.31
	RELT	7054	13	0.75	1.60
	NMBR	7054	10	0.71	1.35
	EQTN	7054	18	0.79	1.83
	STPR	7054	9	0.64	1.32
8	GEOM	6804	16	0.74	1.78
	FNCT	6804	14	0.71	1.72
	NMBR	6804	6	0.52	1.11
	EQTN	6804	23	0.81	2.14
	STPR	6804	6	0.51	1.05

Table 28. Reliability and Standard Errors of Measurement by Grade and Science Domain

Grade	Domain	*N* Counts	Possible Points	Cronbach's Alpha	SEM
4	LIFE	7477	16	0.76	1.62
	PHYS	7477	18	0.75	1.85
	ESCI	7477	16	0.74	1.73
8	LIFE	6789	17	0.76	1.86
	PHYS	6789	18	0.76	1.91
	ESCI	6789	15	0.69	1.78

8.4 Accuracy and Consistency of Classifications

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

Every discrete test administration will result in some error in the classification of examinees. When an assessment uses performance classifications as the primary method to report test results, accuracy and consistency of decisions become important indicators about the quality of the assessment. This section includes the estimates of decision consistency and accuracy for the PAWS assessments administered in March 2016.

The *accuracy* of decisions is represented by the agreement between the classifications based on students' observed scores on the actual test form and the classifications that would have been made based on students' true scores. True scores are assumed to be errorless but are unknown. They can, however, be estimated based on the expected values of test scores over all possible forms of the test. A false positive decision results when a true score corresponds to a classification below a critical cut score (e.g., "does not meet standard"), but the observed score corresponds to a "meets standard" classification. A false negative decision results when a true score "meets standard," but the observed score corresponds to a "does not meet standard" classification. Decision *consistency* is the agreement between two non-overlapping and equally difficult forms of the test. This index is estimated using response data from the actual test form and a hypothetical alternate form, based on the actual test form's estimated reliability.

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

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

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

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

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

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

Figure 2. Accuracy or Consistency around Critical Cut Point

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 29–34 present the results of the decision accuracy and consistency of the PAWS cut scores for Reading, Mathematics, and Science.

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

Negative value is the proportion of student scores misclassified to the category *"Does Not Achieve Proficiency"* when student scores actually do meet proficiency.

Overall, cut point accuracy and consistency ratings range from 0.86 to 0.92, with most results above 0.90 for accuracy and above .88 for consistency. All false negative and false positive results are at or below 0.17. These results suggest acceptable levels of reliability at the cut points for all PAWS assessments.

Table 29. 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	7683	0.76	0.12	0.12	0.67	0.16	0.17	0.90	0.87
Mathematics	7683	0.80	0.10	0.10	0.72	0.14	0.14	0.92	0.89

Table 30. 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	7482	0.75	0.12	0.12	0.66	0.17	0.17	0.91	0.87
Mathematics	7485	0.82	0.09	0.09	0.74	0.13	0.13	0.92	0.88
Science	7477	0.78	0.11	0.11	0.70	0.15	0.15	0.91	0.87

Table 31. 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	7323	0.77	0.12	0.11	0.69	0.16	0.15	0.90	0.86
Mathematics	7325	0.82	0.09	0.09	0.75	0.12	0.13	0.92	0.89

Table 32. 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	7013	0.78	0.11	0.10	0.70	0.15	0.15	0.91	0.88
Mathematics	7011	0.82	0.09	0.09	0.75	0.12	0.13	0.92	0.88

Table 33. 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	7052	0.78	0.11	0.11	0.69	0.15	0.15	0.91	0.87
Mathematics	7055	0.80	0.10	0.10	0.72	0.14	0.14	0.92	0.89

Table 34. 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	6793	0.77	0.13	0.09	0.69	0.16	0.15	0.90	0.86
Mathematics	6804	0.80	0.10	0.10	0.72	0.14	0.14	0.92	0.89
Science	6789	0.77	0.12	0.11	0.68	0.16	0.16	0.91	0.88

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 its 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 was 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 has developed an item utilization plan to guide the development of the items for each content area. Item writing emphasis was determined in consultation with the WDE. Adherence to the specifications ensured the maintenance of quality and consistency of the item development process.

9.1.2. Item Writers

The items for the PAWS were written by item writers having a thorough understanding of the Wyoming Content and Performance Standards. The item writers were 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 were invited to participate in an extensive training program for item writers.

9.1.3. Internal Contractor Reviews

Once items were written, ETS assessment specialists ensured each item underwent an intensive internal review process. Every step of this process was designed to produce items exceeding industry standards for quality. It included 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 ensured the items and related materials complied with ETS's written guidelines for clarity, style, accuracy, and appropriateness and with approved item specifications. The artwork and graphics for the items were created during the internal content review period so assessment specialists could evaluate the correctness and appropriateness of the art early in the item development process. ETS selected visuals relevant to the item content and that were 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 involved 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 included 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 was 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, conducted this bias and sensitivity review. These staff members were 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 tested newly developed items to obtain statistical information about item performance. This information was used to evaluate items that are candidates for use in operational test forms. The item statistics were examined carefully at data review meetings, where content experts discussed items that have poor statistics and do not meet the psychometric criteria for item quality. The WDE defined the criteria for acceptable or unacceptable item statistics. This ensured that the item had an appropriate level of difficulty for the target population. The content experts made recommendations about whether to accept or reject each item for inclusion in the PAWS 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 delivered the prompts to the WDE through an electronic item bank. The item bank database was 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 was crucial. The ETS internal item bank database resided on a server within the ETS firewall. Access to the SQL, the server database, was strictly controlled by means of system administration. The electronic item banking application included 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 were authorized by a designated administrator at the WDE and ETS.

9.2 Quality Control of Test Materials

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

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

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

9.2.1. Collecting Test Materials

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

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

9.2.2. Processing Test Materials

Upon receipt of the test materials, ETS used precise inventory and test processing systems, in addition to quality assurance procedures, to maintain an up-to-date accounting of all the testing materials within their facilities. The materials were removed carefully from the shipping cartons and examined for a number of conditions, including physical damage, shipping errors, and omissions. A visual inspection to compare the number of students recorded on the Header sheets with the number of test and answer books or answer documents in the stack was also conducted.

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

9.3 Quality Control of Scanning

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

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

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

9.4 Quality Control of Psychometric Analyses

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

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

9.5 Quality Control of Reporting

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

- Comparing report formats to input sources from the WDE-approved samples
- Validating and verifying the report data by querying the appropriate student data
- Proofreading 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 were provided to the WDE for review and approval. ETS posted a PDF of the sample reports via a secure site. The WDE and ETS reviewed and signed off on the reports after a thorough review. Upon the WDE's approval of the reports, ETS proceeded with production.

9.5.1. Excluding Student Scores from Summary Reports

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

10. HISTORICAL COMPARISONS

Historical comparisons of the PAWS test results are routinely performed to identify trends in examinee performance, in terms of percentage of students meeting standards. As this is the third administration of the PAWS Reading and Mathematics under the new standards 2012 WyCPS, only the performance results from last three years are included in Tables 35–38.

The percentages of students in the equating population classified as Proficient and Advanced from Reading test increased for all grades except grade 3 compare to 2015 administration. For grade 3, the percentages of students Proficient and Advanced decreased from 60.5% in 2015 to 57.7% in 2016, a decrease of 2.8%. Grade 4 had a largest increase of 5.1% from 60.1% to 65.2%. Grade 7 had a second largest increase of 3.4% from 56.3% to 59.7%.

The percentages of students in the equating population classified as Proficient and Advanced from Mathematics test increased for all grades compare to 2015 administration. For grades 3, 4 and 5, the percentage of students classified as Proficient and Advanced moderately increased from 2015 to 2016, 3.3%, 4.9%, and 3.8% respectively. Grade 7 had the largest increase of 5.9% from 43.1% in 2015 to 49.0% in 2016. For all grades, the percentage of students classified as Proficient and Advanced students are at the maximum of the observed values for each grade.

Tables 39–40 provide a comparison of percentages of the students classified as “Proficient and Advanced” from 2008 to 2016 for PAWS Science. The percentages of students in the equating population classified as Proficient and Advanced from 2016 administration increased for Grade 4 compare to 2015 administration. Grade 4 had an increase of 3.1% from 51.3% in 2015 to 54.4% in 2016. Grade 8 had no change in the percentage of students classified as Proficient and Advanced. For all grades, the percentage of Proficient and Advanced students is within previously observed values for the specific grade.

Figures 3 through 14 display the PAWS percentages of students in the equating populations classified as “Proficient and Advanced” from 2014 through 2016 for each Reading and Mathematics grade level. Figures 15 through 16 display the PAWS percentages of students in the equating populations classified as “Proficient and Advanced” from 2008 through 2016 for each Science grade level. The results for 2010 were not provided due to federal exemption for reporting scores.

Table 35. Year-to-Year Comparison of Scale Scores for the PAWS Reading

	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8		
Year	*N*	Mean	*SD*	*N*	Mean	*SD*	*N*	Mean	*SD*	*N*	Mean	*SD*	*N*	Mean	*SD*	*N*	Mean	*SD*
2014	7365	600.0	50.0	7022	616.0	46.8	7075	626.5	47.4	6758	636.3	48.8	6796	649.3	44.6	6781	661.1	47.7
2015	7541	599.3	48.0	7316	616.3	50.5	6966	628.8	49.9	7102	637.4	47.5	6766	648.8	46.0	6788	654.7	47.9
2016	7683	598.9	46.5	7482	620.8	47.1	7323	630.7	50.1	7013	637.3	51.3	7052	650.9	47.4	6793	655.0	44.8

Table 36. Year-to-Year Comparison of Percentage of Proficient and Advanced Students for the PAWS Reading

Grade	2014	2015	2016	Min	Max	Median	2016 Difference from Median	2016 Difference from 2015
3	61.8	60.5	57.7	57.7	61.8	60.5	-2.8	-2.8
4	63.7	60.1	65.2	60.1	65.2	63.7	1.5	5.1
5	58.1	58.3	60.7	58.1	60.7	58.3	2.4	2.4
6	56.9	56.5	57.6	56.4	57.6	56.9	0.7	1.2
7	58.8	56.3	59.7	56.3	59.7	59.7	0.0	3.4
8	57.7	51.2	53.4	51.2	58.9	53.4	0.0	2.2

Table 37. Year-to-Year Comparison of Scale Scores for the PAWS Mathematics

	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*
2014	7369	600.0	50.0	7026	634.7	45.1	7077	659.7	50.6	6760	678.1	48.2	6799	691.8	45.5	6784	707.5	45.0
2015	7547	602.4	51.2	7319	639.6	50.6	6975	662.3	54.1	7107	679.4	48.9	6767	693.9	45.4	6802	706.7	44.2
2016	7683	604.1	54.2	7485	647.2	52.1	7325	665.8	55.2	7011	680.4	48.3	7055	699.9	47.3	6804	709.4	43.4

Table 38. Year-to-Year Comparison of Percentage of Proficient and Advanced Students for the PAWS Mathematics

Grade	2014	2015	2016	Min	Max	Median	2016 Difference from Median	2016 Difference from 2015
3	50.4	49.3	52.6	49.3	52.6	50.4	2.2	3.3
4	46.6	50.2	55.0	46.6	55.0	50.1	4.9	4.9
5	54.0	52.4	56.0	52.2	56.0	54.0	2.0	3.8
6	48.5	49.3	49.6	48.5	49.6	49.3	0.3	0.3
7	42.8	43.1	49.0	43.1	49.0	43.6	5.4	5.9
8	49.2	47.0	47.5	47.1	50.2	47.5	0.0	0.4

Table 39. Year-to-Year Comparison of Scale Scores for the PAWS Science

	Grade 4			Grade 8		
Year	*N*	Mean	*SD*	*N*	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
2014	7022	669.7	46.4	6770	650.9	45.5
2015	7307	668.9	47.0	6790	648.0	46.7
2016	7477	673.4	51.7	6789	647.6	47.1

Table 40. Year-to-Year Comparison of Percentage of Proficient and Advanced Students for the PAWS Science

Grade	2008	2009	2010	2011	2012	2013	2014	2015	2016	Min	Max	Median	2016 Difference from Median	2015 Difference from 2016
4	50.9	50.5	N/A	54.5	63.3	57.5	52.5	51.3	54.4	50.5	63.3	53.5	0.9	3.1
8	46.4	42.9	N/A	50.7	51.2	43.7	46.8	41.2	41.2	41.2	51.2	45.1	-3.9	0.0

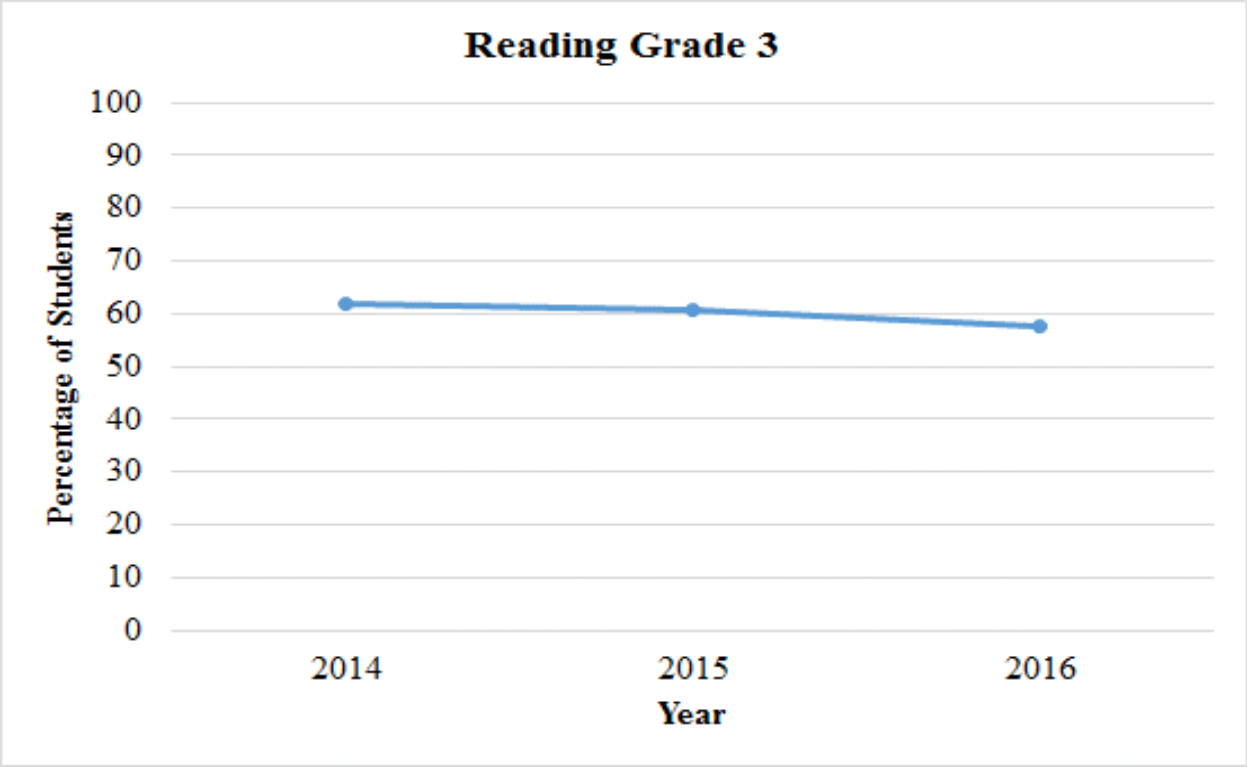

Figure 3. Percentage of Proficient and Advanced Students for Reading Grade 3

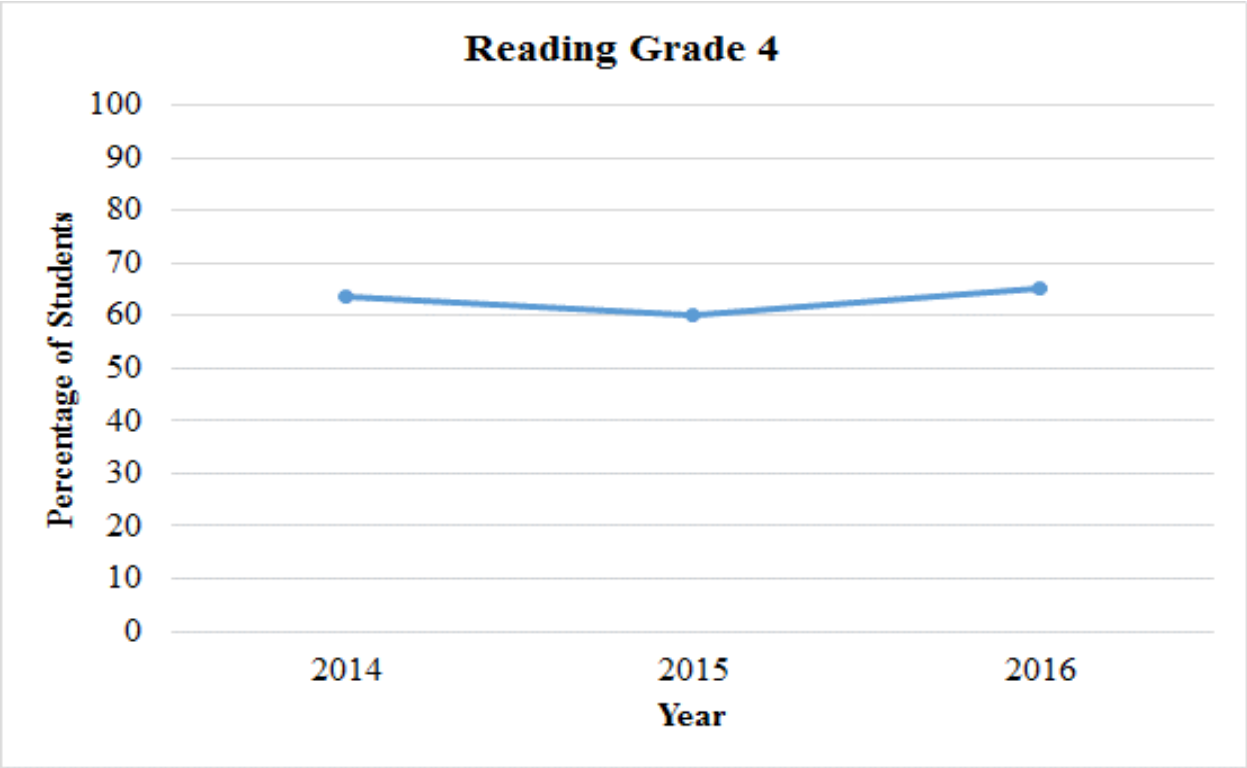

Figure 4. Percentage of Proficient and Advanced Students for Reading Grade 4

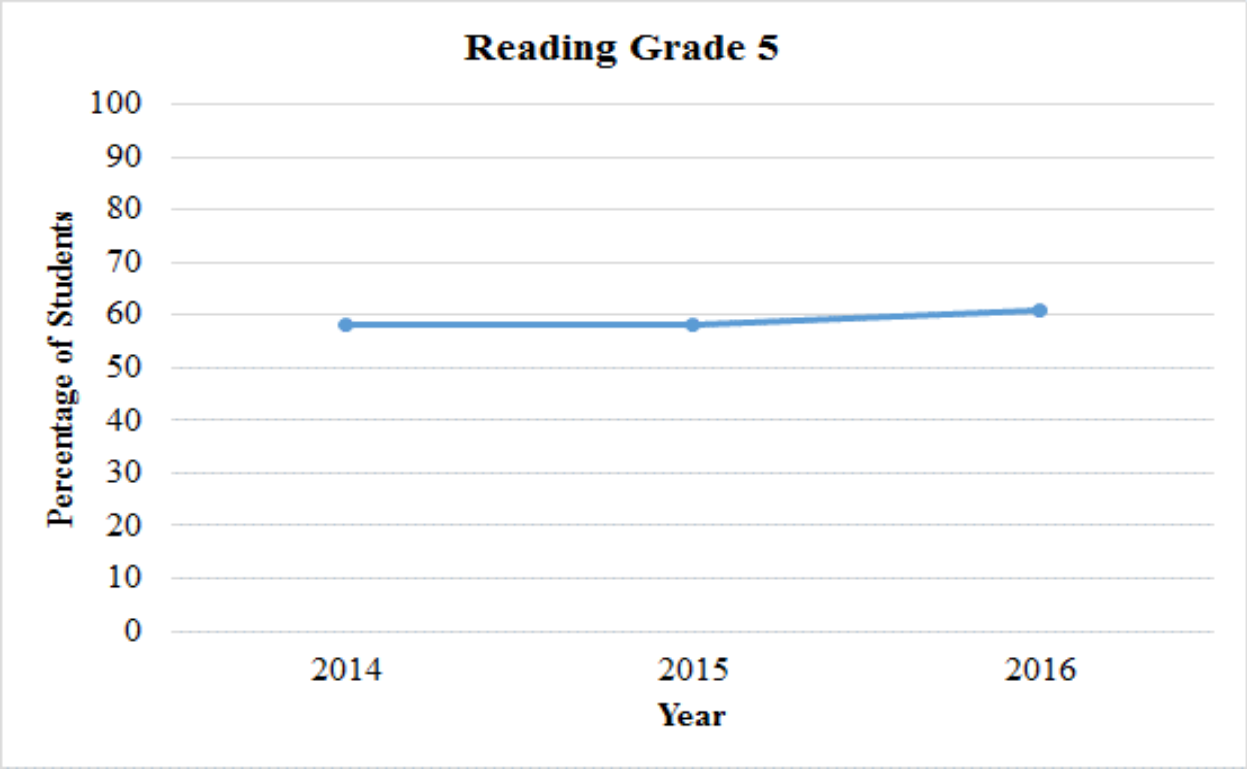

Figure 5. Percentage of Proficient and Advanced Students for Reading Grade 5

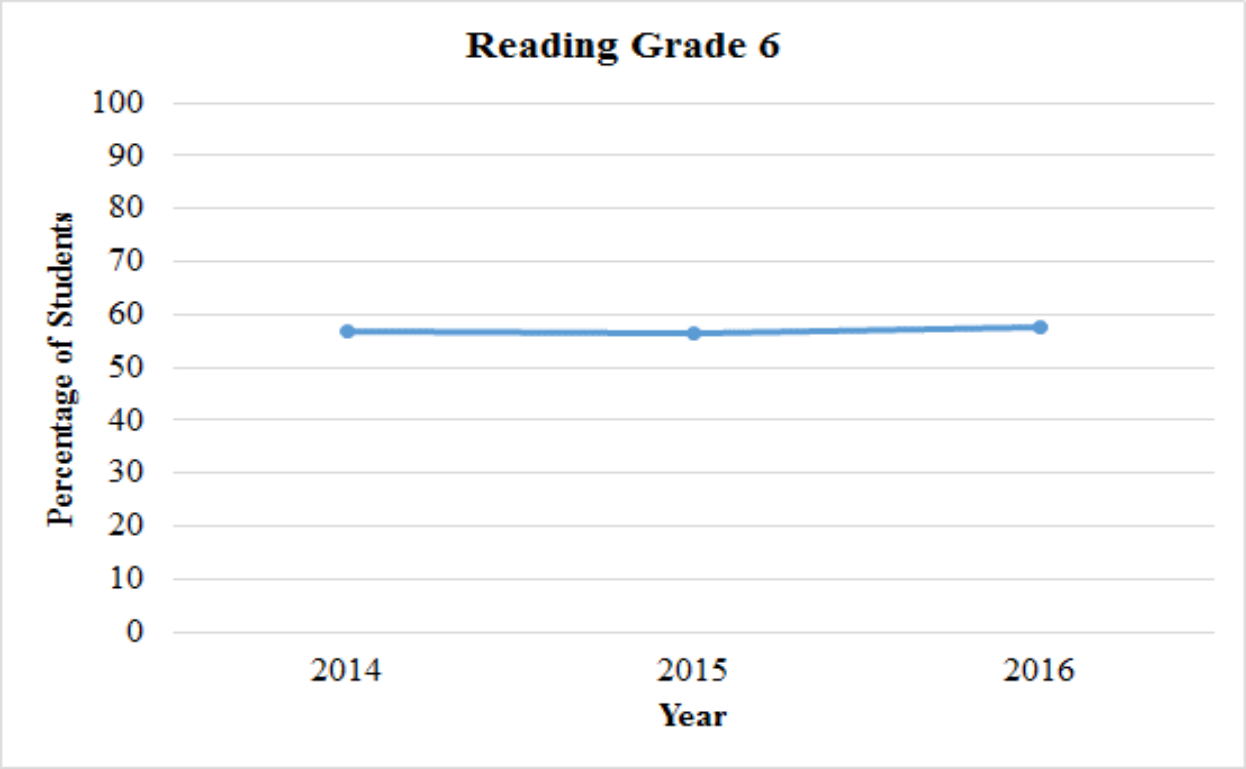

Figure 6. Percentage of Proficient and Advanced Students for Reading Grade 6

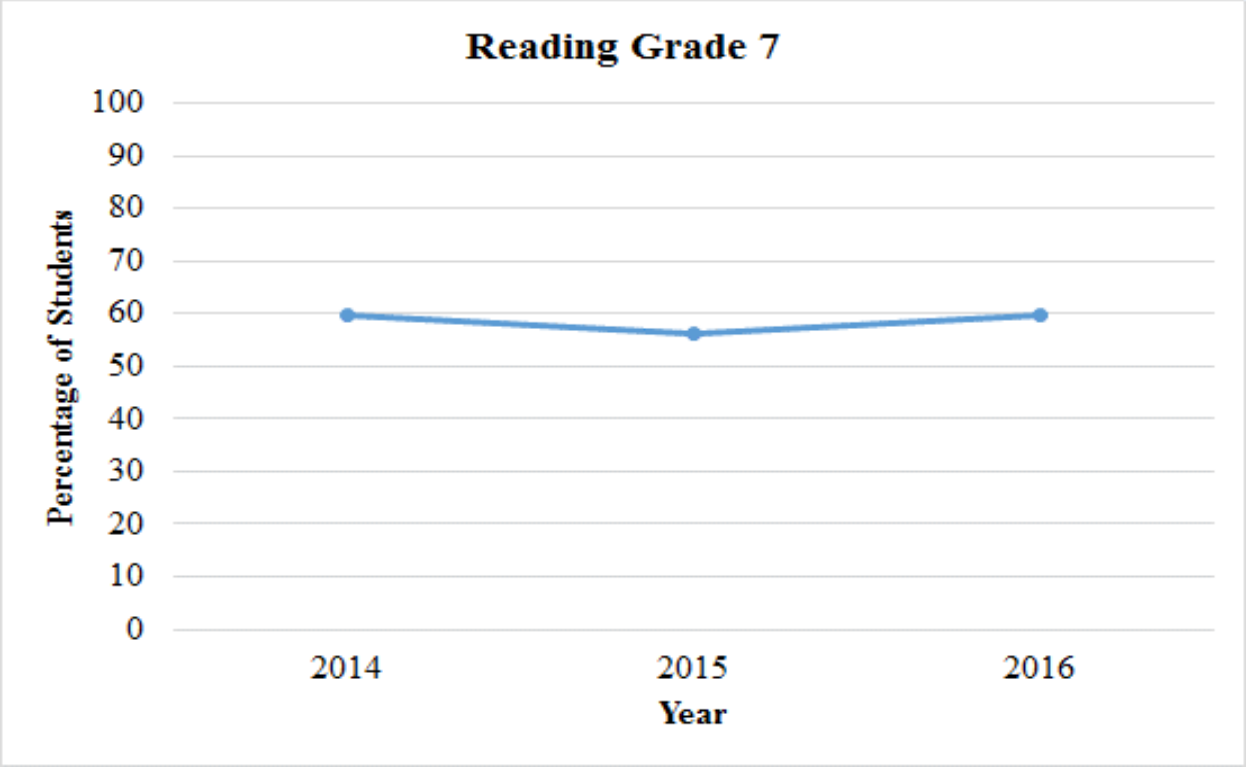

Figure 7. Percentage of Proficient and Advanced Students for Reading Grade 7

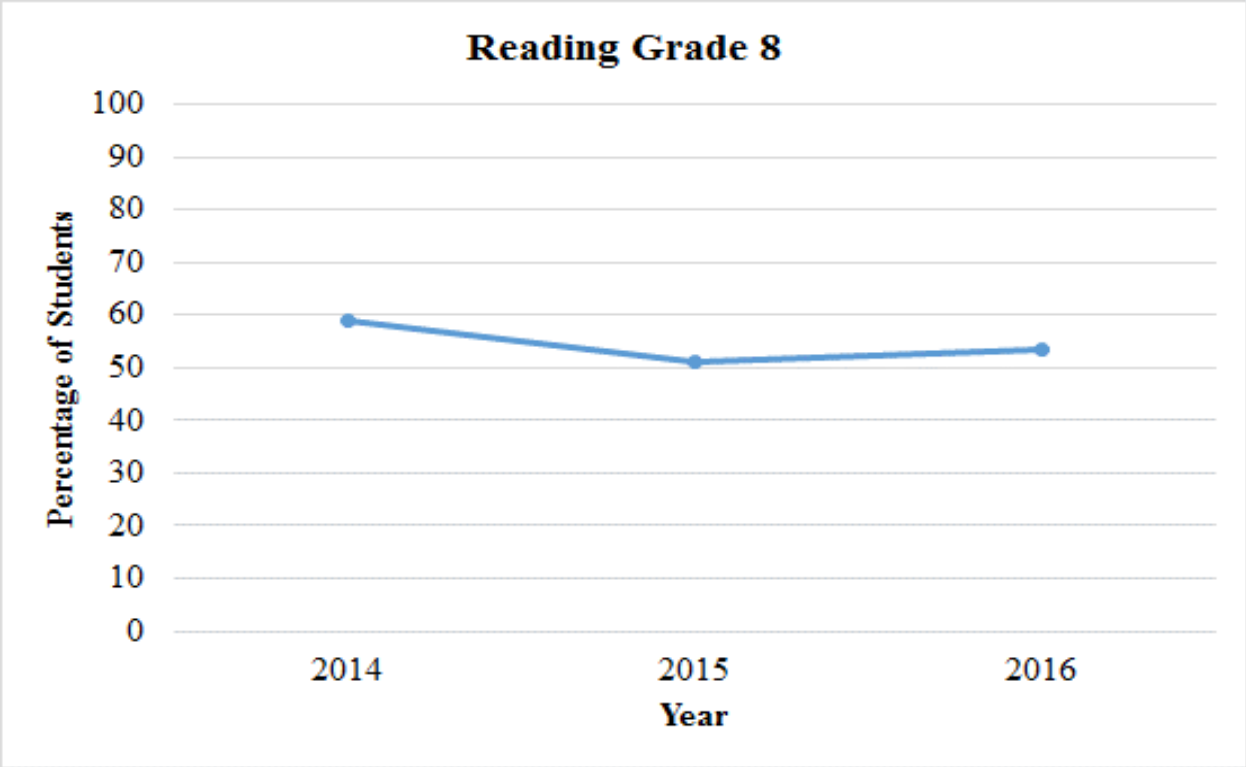

Figure 8. Percentage of Proficient and Advanced Students for Reading Grade 8

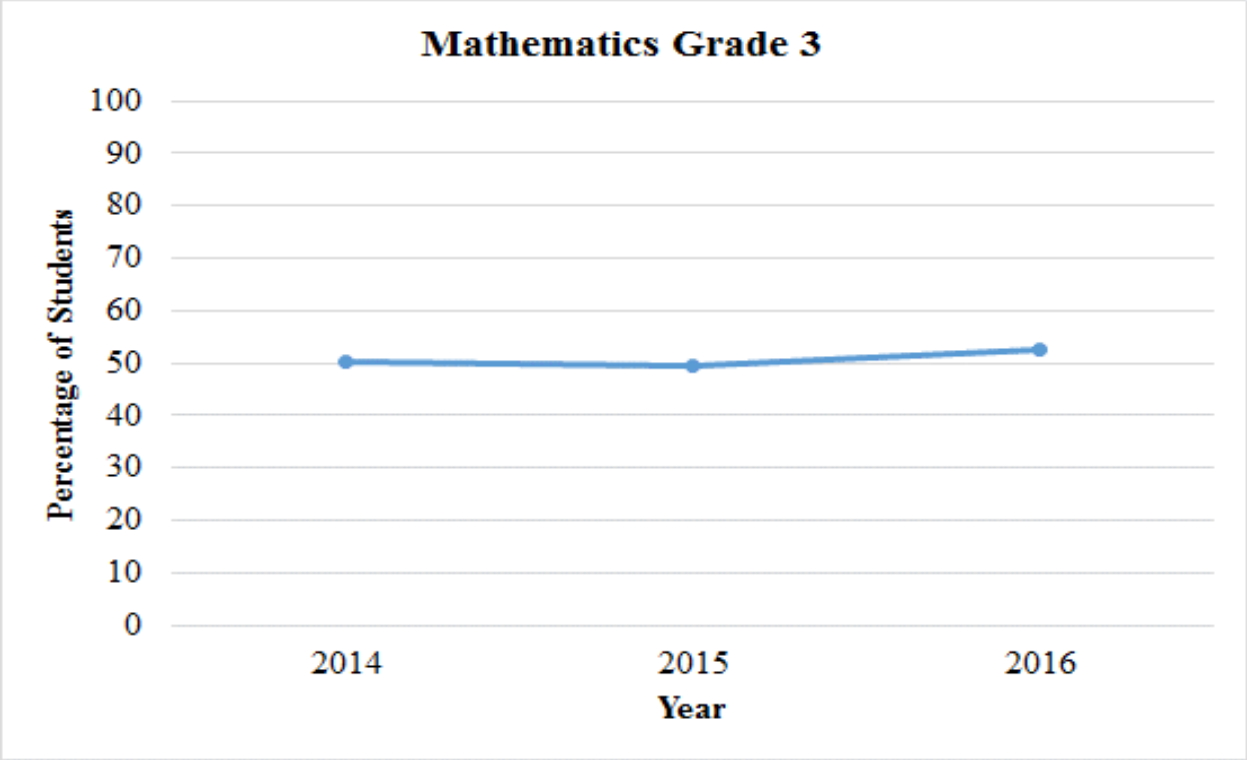

Figure 9. Percentage of Proficient and Advanced Students for Mathematics Grade 3

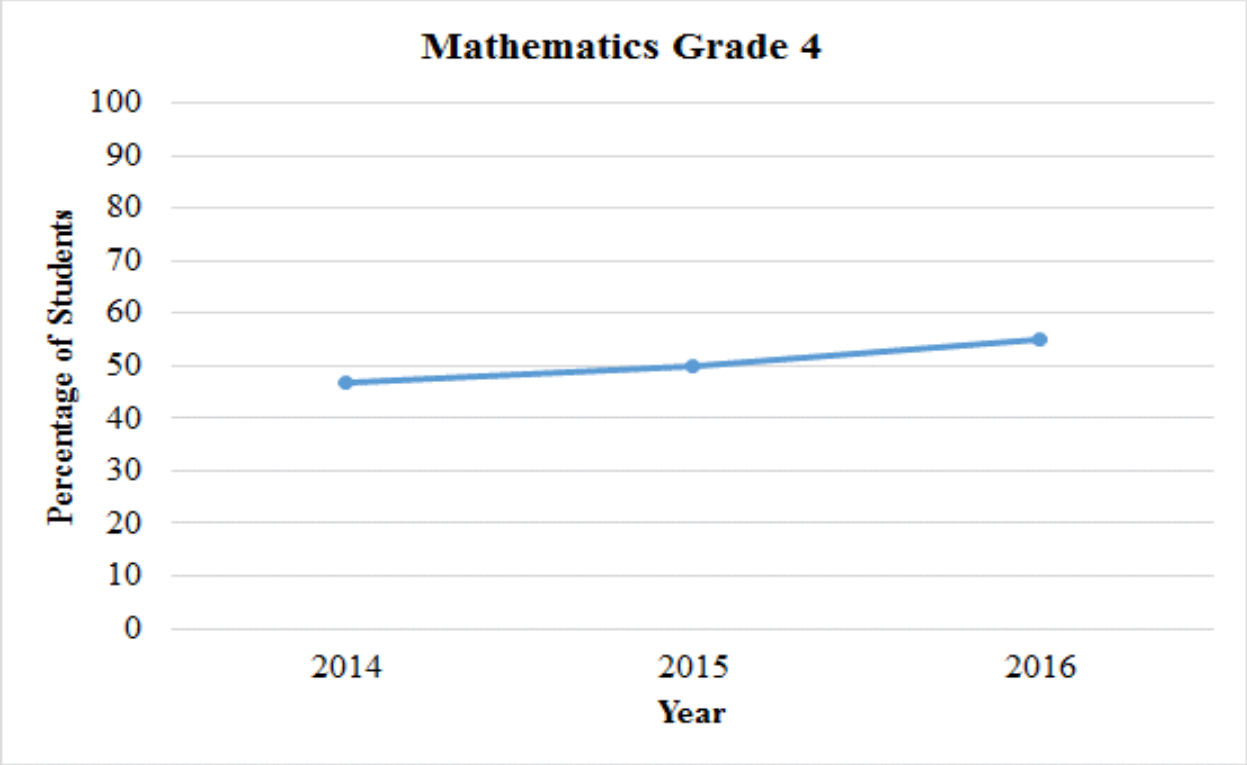

Figure 10. Percentage of Proficient and Advanced Students for Mathematics Grade 4

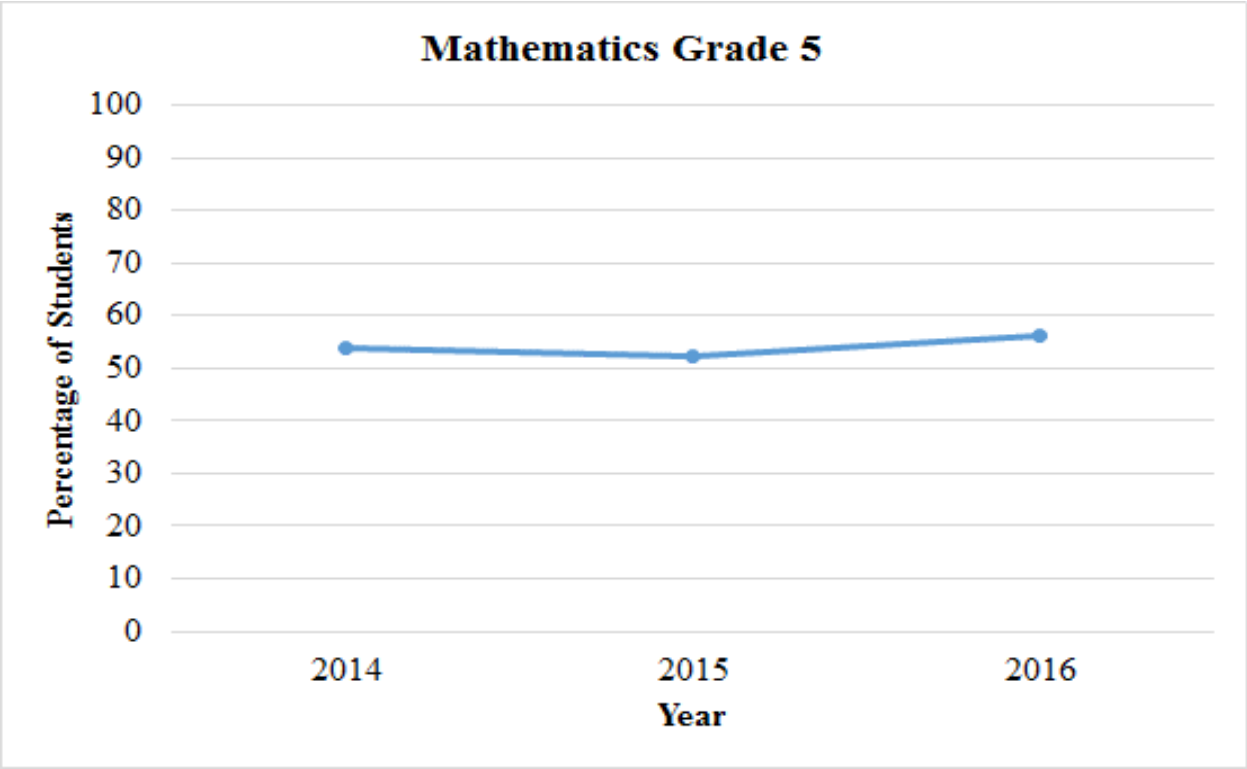

Figure 11. Percentage of Proficient and Advanced Students for Mathematics Grade 5

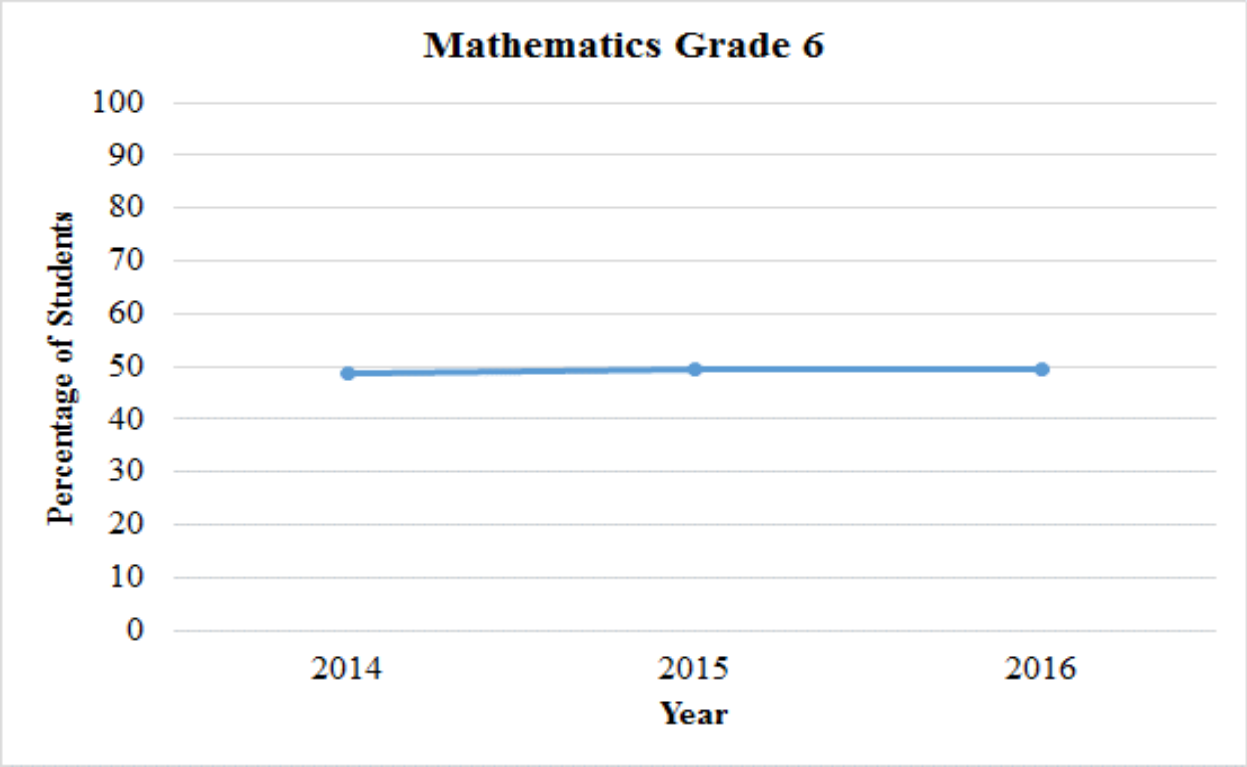

Figure 12. Percentage of Proficient and Advanced Students for Mathematics Grade 6

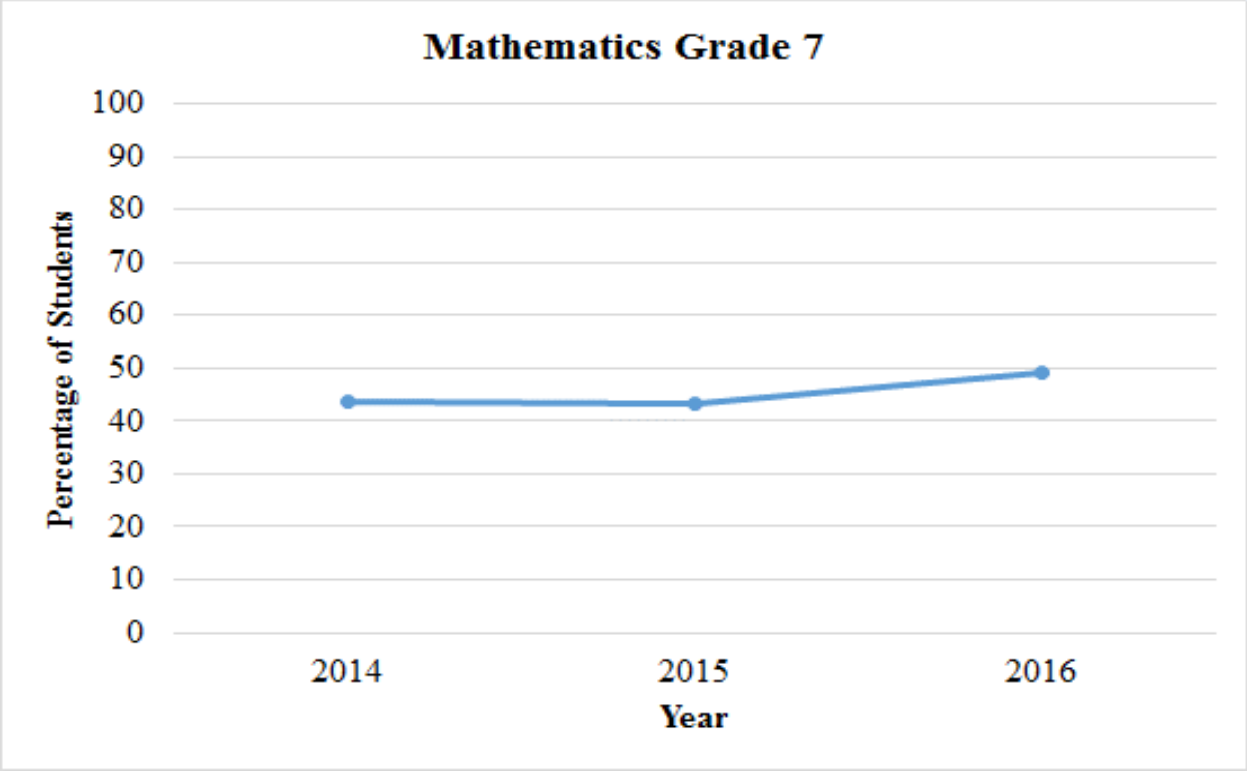

Figure 13. Percentage of Proficient and Advanced Students for Mathematics Grade 7

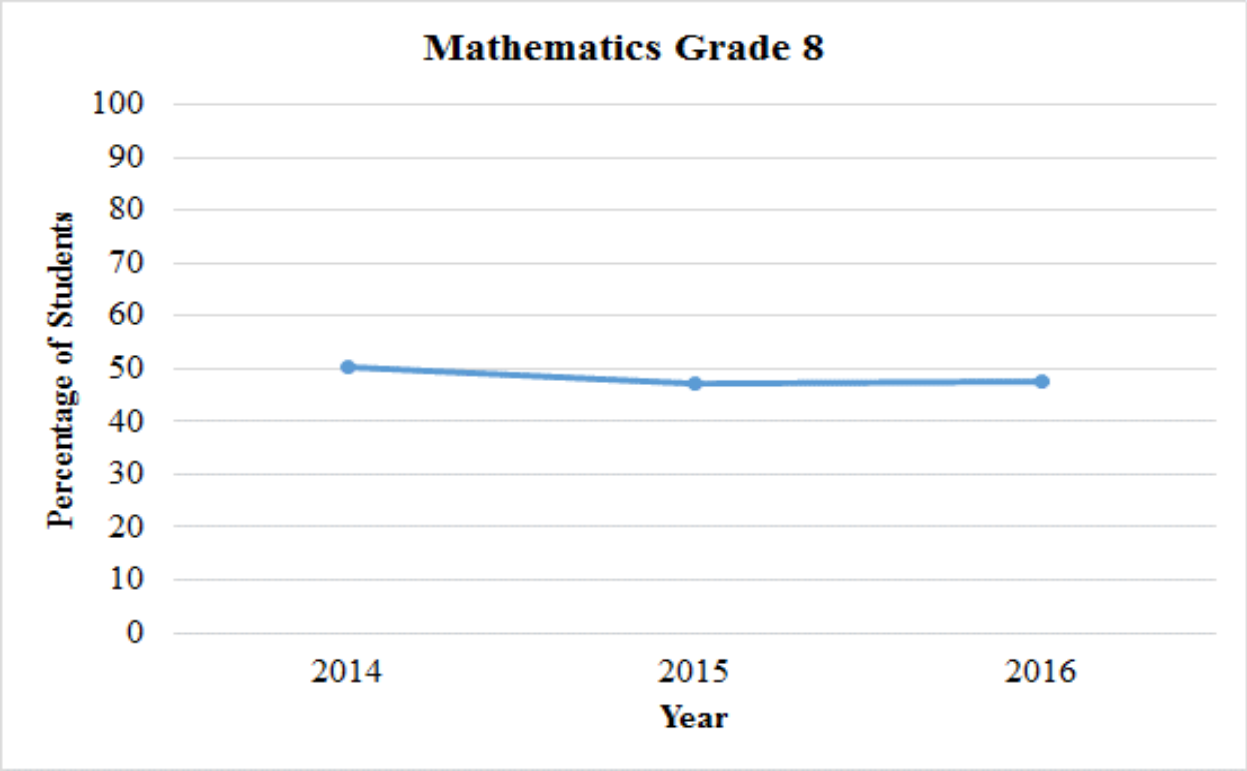

Figure 14. Percentage of Proficient and Advanced Students for Mathematics Grade 8

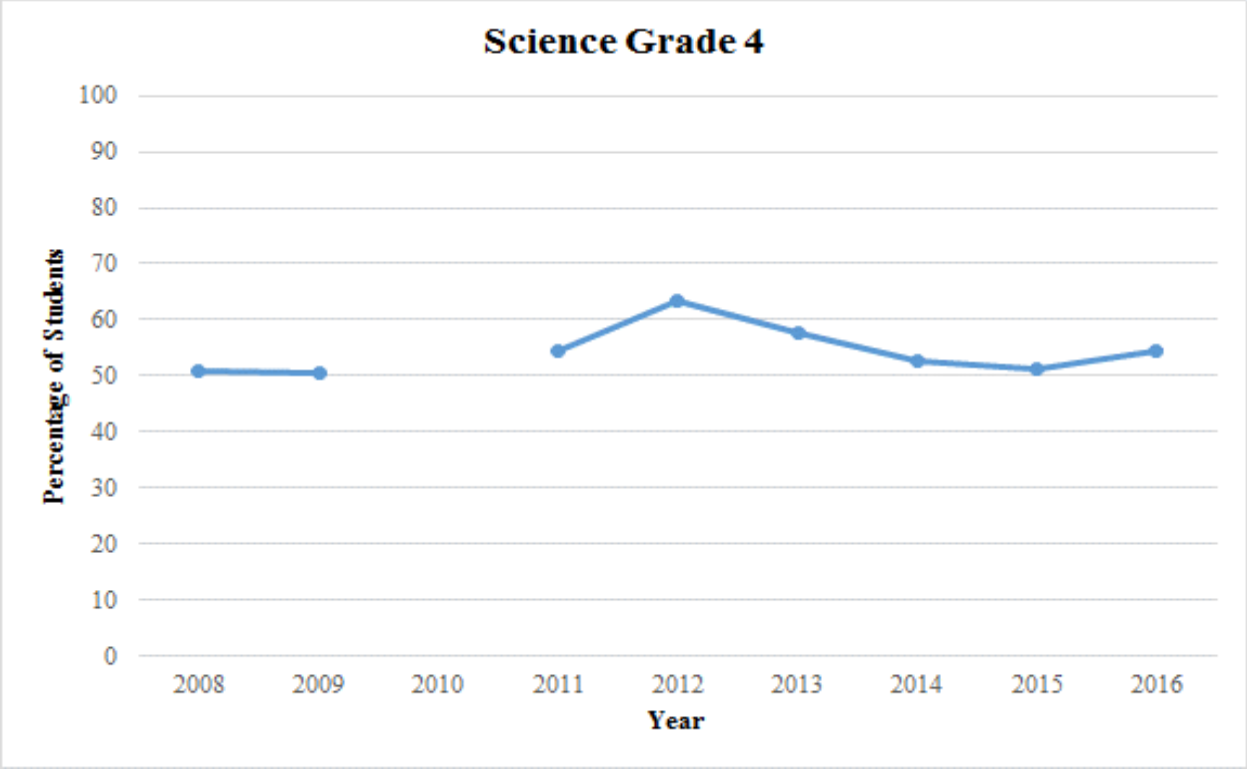

Figure 15. Percentage of Proficient and Advanced Students for Science Grade 4

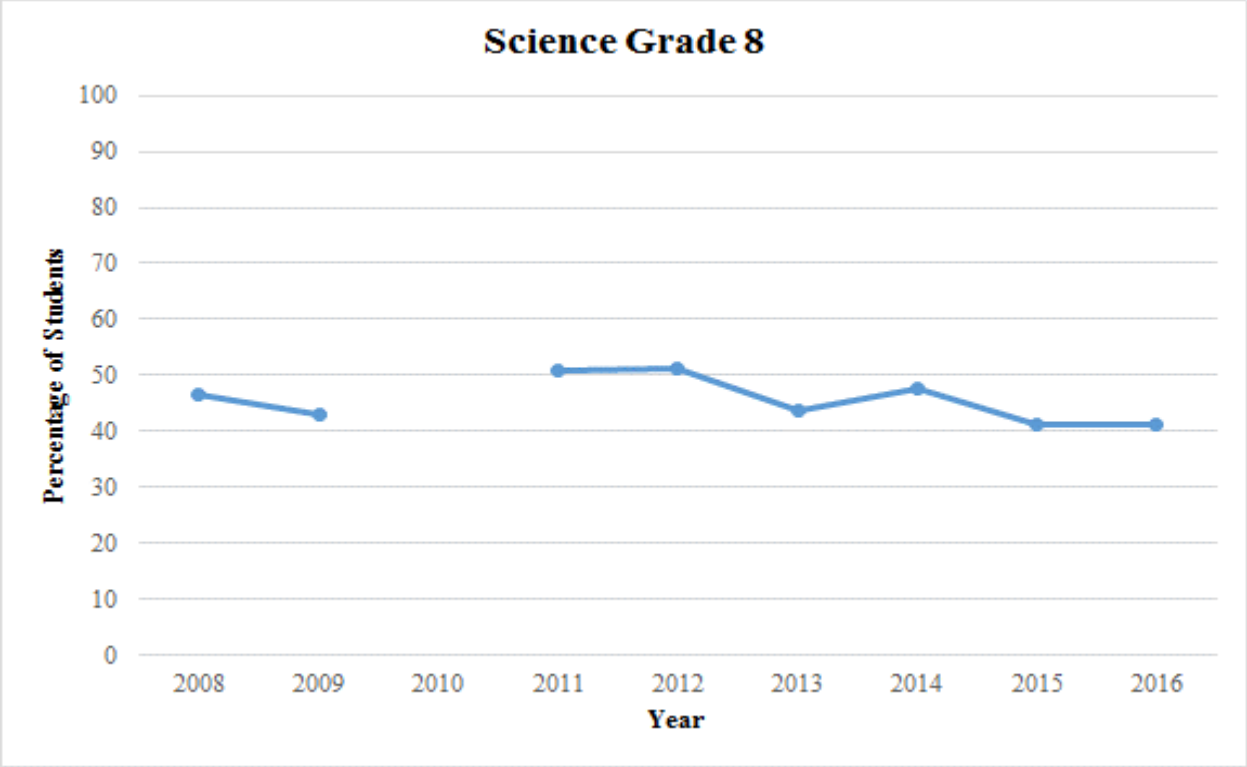

Figure 16. Percentage of Proficient and Advanced Students for Science Grade 8

11. REFERENCES

American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Andrich, A. (1988). *Rasch models for measurement*. Newbury Park, CA: SAGE.

Andrich, A. (1989). Distinctions between assumptions and requirements in measurement in the social sciences, in J. A. Keats, R. Taft, R. A. Heath, and H. H. Lovibond (Eds.) *Mathematical and theoretical systems* (pp. 7–16). North-Holland: Elsevier Science.

Baron, P. (2014). 2014 Standard Setting Summary Proficiency Assessments for Wyoming Students (PAWS) and Student Assessment of Writing Skills (SAWS). (Technical Report). Princeton, NJ: Educational Testing Service.

Baron, P. (2015). Wyoming Instructional Impact Study Final Report. (Technical Report). Princeton, NJ: Educational Testing Service.

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

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Embretson, S., & Reise, S. (2000). *Item response theory for psychologists.* City, NJ: Erlbaum.

Educational Testing Service (2015). *ETS standards for quality and fairness.* Princeton, NJ: Educational Testing Service.

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

Holland, P. W., & Thayer, D. T. (1988). Differential item performances and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Eds.), *Test Validity*, (pp. 129–145). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Huynh, H., & Meyer, P. (2010). Use of Robust z in detecting unstable items in item response theory models. *Practical Assessment, Research and Evaluation, 15*(2). Retrieved from: http://pareonline.net/getvn.asp?v=15and*N*=2

IEP definition (2014). Retrieved from http://www.ncld.org/students-disabilities/iep-504-plan/what-is-iep).

Kane, M. (2006). Validation. In R. Brennan (Ed.), *Educational Measurement* (4th ed.). Washington, DC: American Council on Education and National Council on Measurement in Education.

Karkee, T., Lewis, D. M., & Barton, K. (2005, April). *The effect of including or excluding students with testing accommodations on IRT calibrati*ons. Paper presented at the meeting of the American Educational Research Association, Montreal, Canada.

Kim, D., Wang, S., Zhao, Y., & Li, T. (2006, April). *Validation and invariance of factor structure of a statewide reading comprehension test under accommodation and non-accommodation condition*s. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). New York, NY: Springer Science + Business Media.

Linacre, J. M. (2007) *WINSTEPS Rasch measurement computer program and manual (PDF file) v 3.64.2*. Chicago, IL: Winsteps.com

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

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18,* 5–11.

Nitko, A. J. (2004). *Educational assessments of students*. Englewood Cliffs, NJ: Prentice Hall.

No Child Left Behind Act of 2001, Pub. L. No. 107–110, 115 Stat. 1425 (2002).

Pearson. (2008) *Wyoming Technical Report: Science Grades 4, 8, and 11 Standard Setting*. Iowa City, IA: Pearson.

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, NY; Oxford University Press.

Ryan, J. P. (1983). Introduction to latent trait analysis and item response theory, in W. E. Hathaway (Ed.), *Testing in the Schools. New Directions for Testing and Measurement* (pp. 48–64). San Francisco, CA: Jossey-Bass.

Webb, N. L., Alt, M., Ely, R., and Vesperman, B. (2005). *Web Alignment Tool (WAT) Training Manual.* Wisconsin Center of Educational Research, University of Wisconsin-Madison. Retrieved from http://wat.wceruw.org/Training%20Manual%202.1%20Draft%20091205.doc

Wyoming Department of Education (2006). *Wyoming accommodations manual for instruction and assessment: How to select, administer, and evaluate use of accommodations for instruction and assessment of students with disabilities*. Cheyenne, WY: Author.

Yen, W. (2007). Removing outlier anchor items from IRT equating for educational achievement tests. Unpublished paper.

Young, M. J. (2006). Vertical scales. In S.M. Downing and T.M. Haladyna (Eds.), *Handbook of Test Development* (pp.469–485). Mahwah, NJ: Erlbaum.

Young, M. J., & Yoon, B. (1998, April). *Estimating the consistency and accuracy of classifications in a standards-referenced assessment.* (Technical Report 475). Center for the Study of Evaluation, Standards, and Student Testing. Los Angeles, CA: University of California, Los Angeles.

U.S. Department of Education, National Center for Education Statistics (NCES), (2001). The Condition of Education 2001. NCES 2001–072, Washington, DC: U.S. Government Printing Office.

12. GLOSSARY OF TERMS

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

2012 WyCPS - 2012 Wyoming Content Performance Standards for reading and mathematics

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.

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.

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 the 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.

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

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.

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

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

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

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

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

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

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

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

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

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

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

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

ETS – Educational Testing Service, current vendor for the PAWS.

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

Form - Operational items and embedded field test items that uniquely define a (test) booklet.

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

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

Item - A test question. Examples of formats are multiple-choice, open-ended (constructed response), and extended response. For PAWS, only multiple-choice items are used.

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

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

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

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

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

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

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

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

MC - Multiple-choice item (worth 1-point)

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

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

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

PAWS- Proficiency Assessments for Wyoming Students

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

Performance 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.

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

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

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

Reliability - The extent to which test scores are reproducible. If a class of students theoretically took the same test twice in one day and each student's score was the same on the second

administration of the test as on the first, the test would be perfectly reliable (1.00). Of course, perfection is not possible and reliabilities in the 0.90s are considered good.

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

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 spread of test scores around the mean. Assuming a normal distribution, if you know the mean and standard deviation of a distribution, you can determine what proportion of scores falls within one standard deviation of the mean.

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

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

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

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

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

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

Appendix A: PAWS 2016 Reading, Mathematics, and Science Blueprints

Reading

Table A1. PAWS 2016 Reading Grade 3 Blueprint

Strand / Standard	PAWS 3rd Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis	PAWS Administration 2015 to Present Reading Wyoming Department of Education
Reading Literature				
	Key Ideas and Details	10-12	Approx. 36%	
RL3.1	Ask and answer questions to demonstrate understanding of a text, referring explicitly to the text as the basis for answers.			
RL3.2	Recount stories, including fables, folktales, and myths from diverse cultures; determine the central message, lesson, or moral and explain how it is conveyed through key details in the text.			
RL3.3	Describe characters in a story (e.g., their traits, motivations, or feelings) and explain how their actions contribute to the sequence of events.			
	Craft and Structure	6-8		
RL3.4	Determine the meaning of words and phrases as they are used in a text, distinguishing literal from non-literal language.			
RL3.5	Refer to parts of stories, dramas, and poems when writing or speaking about a text, using terms such as chapter, scene, and stanza; describe how each successive part builds on earlier sections.			
RL3.6	Distinguish their own point of view from that of the narrator or those of the characters.			
Reading Informational Text				
	Key Ideas and Details	7-9	Approx. 32%	
RI3.1	Ask and answer questions to demonstrate understanding of a text, referring explicitly to the text as the basis for the answers.			
RI3.2	Determine the main idea of a text; recount the key details and explain how they support the main idea.			
RI3.3	Describe the relationship between a series of historical events, scientific ideas or concepts, or steps in technical procedures in a text, using language that pertains to time, sequence, and cause/effect.			
	Craft and Structure	7-9		
RI3.4	Determine the meaning of general academic and domain-specific words and phrases in a text relevant to a grade 3 topic or subject area.			
RI3.5	Use text features and search tools (e.g., key words, sidebars, hyperlinks) to locate information relevant to a given topic efficiently.			
RI3.6	Distinguish their own point of view from that of the author of a text.			

Strand / Standard	PAWS 3rd Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis
	Integration of Knowledge and Ideas*	7-9	
RL3.7	Explain how specific aspects of a text's illustrations contribute to what is conveyed by the words in a story (e.g., create mood, emphasize aspects of a character or setting).		Approx. 16%
RL3.9	Compare and contrast the themes, settings, and plots of stories written by the same author about the same or similar characters (e.g., in books from a series).		
RI3.7	Use information gained from illustrations (e.g., maps, photographs) and the words in a text to demonstrate understanding of the text (e.g., where, when, why, and how key events occur).		
RI3.8	Describe the logical connection between particular sentences and paragraphs in a text (e.g., comparison, cause/effect, first/second/third in a sequence).		
RI3.9	Compare and contrast the most important points and key details presented in two texts on the same topic.		
	Language*	7-9	
L3.4	Determine or clarify the meaning of unknown and multiple-meaning word and phrases based on grade 3 reading and content, choosing flexibly from a range of strategies.		Approx. 16%
L3.4.a	Use sentence-level context as a clue to the meaning of a word or phrase.		
L3.4.b	Determine the meaning of the new word formed when a known affix is added to a known word (e.g., agreeable/disagreeable, comfortable/uncomfortable, care/careless, heat/preheat).		
L3.4.c	Use a known root word as a clue to the meaning of an unknown word with the same root (e.g., company, companion).		
L3.5	Demonstrate understanding of word relationships and nuances in word meanings.		
L3.5.a	Distinguish the literal and nonliteral meanings of words and phrases in context (e.g., take steps).		
L3.5.c	Distinguish shades of meaning among related words that describe states of mind or degrees of certainty (e.g., knew, believed, suspected, heard, wondered).		
L3.6	Acquire and use accurately grade-appropriate conversational, general academic, and domain-specific words and phrases, including those that signal spatial and temporal relationships (e.g., After dinner that night we went looking for them).		
*Integration and Language reporting categories have items from both RL and RI Strands (both literary and informational passages). The goal will be to strike an overall balance of approx. 50% per genre on the test form.		50	100%

PAWS Administration 2015 to Present Reading Wyoming Department of Education

Table A2. PAWS 2016 Reading Grade 4 Blueprint

Strand / Standard	PAWS 4th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis	PAWS Administration 2015 to Present Reading Wyoming Department of Education
Reading Literature				
	Key Ideas and Details	10-12	Approx. 35%	
RL4.1	Refer to details and examples in a text when explaining what the text says explicitly and when drawing inferences from the text.			
RL4.2	Determine a theme of a story, drama, or poem from details in the text; summarize the text.			
RL4.3	Describe in depth a character, setting, or event in a story or drama, drawing on specific details in the text (e.g., a character's thoughts, words, or actions).			
	Craft and Structure	6-8		
RL4.4	Determine the meaning of words and phrases as they are used in a text, including those that allude to significant characters found in mythology (e.g., Herculean).			
RL4.5	Explain major differences between poems, drama, and prose, and refer to the structural elements of poems (e.g., verse, rhythm, meter) and drama (e.g., casts of characters, settings, descriptions, dialogue, stage directions) when writing or speaking about a text.			
RL4.6	Compare and contrast the point of view from which different stories are narrated, including the difference between first- and third-person narrations.			
Reading Informational Text				
	Key Ideas and Details	10-12	Approx. 35%	
RI4.1	Refer to details and examples in a text when explaining what the text says explicitly and when drawing inferences from the text.			
RI4.2	Determine the main idea of a text and explain how it is supported by key details; summarize the text.			
RI4.3	Explain events, procedures, ideas, or concepts in a historical, scientific, or technical text, including what happened and why, based on specific information in the text.			
	Craft and Structure	6-8		
RI4.4	Determine the meaning of general academic and domain-specific words or phrases in a text relevant to a grade 4 topic or subject area.			
RI4.5	Describe the overall structure (e.g., chronology, comparison, cause/effect, problem/solution) of events, ideas, concepts, or information in a text or part of a text.			
RI4.6	Compare and contrast a firsthand and secondhand account of the same event or topic; describe the differences in focus and the information provided.			

Strand / Standard	PAWS 4th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis
	Integration of Knowledge and Ideas*	6-8	
RL4.9	Compare and contrast the treatment of similar themes and topics (e.g., opposition of good and evil) and patterns of events (e.g., the quest) in stories, myths, and traditional literature from different cultures.		Approx. 14%
RI4.7	Interpret information presented visually, orally, or quantitatively (e.g., in charts, graphs, diagrams, time lines, animations, or interactive elements on Web pages) and explain how the information contributes to an understanding of the text in which it appears.		
RI4.8	Explain how an author uses reasons and evidence to support particular points in a text.		
RI4.9	Integrate information from two texts on the same topic in order to write or speak about the subject knowledgeably.		
	Language*	7-9	
L4.4	Determine or clarify the meaning of unknown and multiple-meaning words and phrases based on grade 4 reading and content, choosing flexibly from a range of strategies.		Approx. 16%
L4.4.a	Use context (e.g., definitions, examples, or restatements in text) as a clue to the meaning of a word or phrase.		
L4.4.b	Use common, grade-appropriate Greek and Latin affixes and roots as clues to the meaning of a word (e.g., telegraph, photograph, autograph).		
L4.4.c	Consult reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation and determine or clarify the precise meaning of key words and phrases.		
L4.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.		
L4.5.a	Explain the meaning of simple similes and metaphors (e.g., as pretty as a picture) in context.		
L4.5.b	Recognize and explain the meaning of common idioms, adages, and proverbs.		
L4.5.c	Demonstrate understanding of words by relating them to their opposites (antonyms) and to words with similar but not identical meanings (synonyms).		
L4.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases, including those that signal precise actions, emotions, or states of being (e.g., quizzed, whined, stammered) and that are basic to a particular topic (e.g., wildlife, conservation, and endangered when discussing animal preservation).		
*Integration and Language reporting categories have items from both RL and RI Strands (both literary and informational passages). The goal will be to strike an overall balance of approx. 50% per genre on the test form.		50	100%

PAWS Administration 2015 to Present Reading Wyoming Department of Education

Table A3. PAWS 2016 Reading Grade 5 Blueprint

Strand / Standard	PAWS 5th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis	PAWS Administration 2015 to Present Reading Wyoming Department of Education
Reading Literature				
	Key Ideas and Details	11-13	Approx. 35%	
RL5.1	Quote accurately from a text when explaining what the text says explicitly and when drawing inferences from the text.			
RL5.2	Determine a theme of a story, drama, or poem from details in the text, including how characters in a story or drama respond to challenges or how the speaker in a poem reflects upon a topic; summarize the text.			
RL5.3	Compare and contrast two or more characters, settings, or events in a story or drama, drawing on specific details in the text (e.g., how characters interact).			
	Craft and Structure	6-8		
RL 5.4	Determine the meaning of words and phrases as they are used in a text, including figurative language such as metaphors and similes.			
RL5.5	Explain how a series of chapters, scenes, or stanzas fits together to provide the overall structure of a particular story, drama, or poem.			
RL 5.6	Describe how a narrator's or speaker's point of view influences how events are described.			
Reading Informational Text				
	Key Ideas and Details	12-14	Approx. 36%	
RI5.1	Quote accurately from a text when explaining what the text says explicitly and when drawing inferences from the text.			
RI5.2	Determine two or more main ideas of a text and explain how they are supported by key details; summarize the text.			
RI5.3	Explain the relationships or interactions between two or more individuals, events, ideas, or concepts in a historical, scientific, or technical text based on specific information in the text.			
	Craft and Structure	6-8		
RI5.4	Determine the meaning of general academic and domain-specific words and phrases in a text relevant to a grade 5 topic or subject area.			
RI5.5	Compare and contrast the overall structure (e.g., chronology, comparison, cause/effect, problem/solution) of events, ideas, concepts, or information in two or more texts.			
RI5.6	Analyze multiple accounts of the same event or topic, noting important similarities and differences in the point of view they represent.			

Strand / Standard	PAWS 5th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis
	Integration of Knowledge and Ideas*	6-8	
RL5.9	Compare and contrast stories in the same genre (e.g., mysteries and adventure stories) on their approaches to similar themes and topics.		Approx. 13%
RI5.7	Draw on information from multiple print or digital sources, demonstrating the ability to locate an answer to a question quickly or to solve a problem efficiently.		
RI5.8	Explain how an author uses reasons and evidence to support particular points in a text, identifying which reasons and evidence support which point(s).		
RI5.9	Integrate information from several texts on the same topic in order to write or speak about the subject knowledgeably.		
	Language*	8-10	
L5.4	Determine or clarify the meaning of unknown and multiple-meaning words and phrases based on grade 5 reading and content, choosing flexibly from a range of strategies.		Approx. 16%
L5.4.a	Use context (e.g., cause/effect relationships and comparisons in text) as a clue to the meaning of a word or phrase.		
L5.4.b	Use common, grade-appropriate Greek and Latin affixes and roots as clues to the meaning of a word (e.g., photograph, photosynthesis).		
L5.4.c	Consult reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation and determine or clarify the precise meaning of key words and phrases.		
L5.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.		
L5.5.a	Interpret figurative language, including similes and metaphors, in context.		
L5.5.b	Recognize and explain the meaning of common idioms, adages, and proverbs.		
L5.5.c	Use the relationship between particular words (e.g., synonyms, antonyms, homographs) to better understand each of the words.		
L5.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases, including those that signal contrast, addition, and other logical relationships (e.g., however, although, nevertheless, similarly, moreover, in addition).		
*Integration and Language reporting categories have items from both RL and RI Strands (both literary and informational passages). The goal will be to strike an overall balance of approx. 50% per genre on the test form.		54	100%

PAWS Administration 2015 to Present Reading Wyoming Department of Education

Table A4. PAWS 2016 Reading Grade 6 Blueprint

Strand / Standard	PAWS 6th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis	PAWS Administration 2015 to Present Reading Wyoming Department of Education
Reading Literature				
	Key Ideas and Details	12-14	Approx. 36%	
RL6.1	Cite textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.			
RL6.2	Determine a theme or central idea of a text and how it is conveyed through particular details; provide a summary of the text distinct from personal opinions or judgments.			
RL6.3	Describe how a particular story's or drama's plot unfolds in a series of episodes as well as how the characters respond or change as the plot moves toward a resolution.			
	Craft and Structure	6-8		
RL6.4	Determine the meaning of words and phrases as they are used in a text, including figurative and connotative meanings; analyze the impact of a specific word choice on meaning and tone.			
RL6.5	Analyze how a particular sentence, chapter, scene, or stanza fits into the overall structure of a text and contributes to the development of the theme, setting, or plot.			
RL6.6	Explain how an author develops the point of view of the narrator or speaker in a text.			
Reading Informational Text				
	Key Ideas and Details	12-14	Approx. 36%	
RI6.1	Cite textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.			
RI6.2	Determine a central idea of a text and how it is conveyed through particular details; provide a summary of the text distinct from personal opinions or judgments.			
RI6.3	Analyze in detail how a key individual, event, or idea is introduced, illustrated, and elaborated in a text (e.g., through examples or anecdotes).			
	Craft and Structure	6-8		
RI6.4	Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings.			
RI6.5	Analyze how a particular sentence, paragraph, chapter, or section fits into the overall structure of a text and contributes to the development of the ideas.			
RI6.6	Determine an author's point of view or purpose in a text and explain how it is conveyed in the text.			

Strand / Standard	PAWS 6th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis
	Integration of Knowledge and Ideas*	7-9	
RL6.9	Compare and contrast texts in different forms or genres (e.g., stories and poems; historical novels and fantasy stories) in terms of their approaches to similar themes and topics.		Approx. 14%
RI6.7	Integrate information presented in different media or formats (e.g., visually, quantitatively) as well as in words to develop a coherent understanding of a topic or issue.		
RI6.8	Trace and evaluate the argument and specific claims in a text, distinguishing claims that are supported by reasons and evidence from claims that are not.		
RI6.9	Compare and contrast one author's presentation of events with that of another (e.g., a memoir written by and a biography on the same person).		
	Language*	7-9	
L6.4	Determine or clarify the meaning of unknown and multiple-meaning words and phrases based on grade 6 reading and content, choosing flexibly from a range of strategies.		Approx. 14%
L6.4.a	Use context (e.g., the overall meaning of a sentence or paragraph; a word's position or function in a sentence) as a clue to the meaning of a word or phrase.		
L6.4.b	Use common, grade-appropriate Greek or Latin affixes and roots as clues to the meaning of a word (e.g., audience, auditory, audible).		
L6.4.c	Consult reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation of a word or determine or clarify its precise meaning or its part of speech.		
L6.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.		
L6.5.a	Interpret figures of speech (e.g., personification) in context.		
L6.5.b	Use the relationship between particular words (e.g., cause/effect, part/whole, item/category) to better understand each of the words.		
L6.5.c	Distinguish among the connotations (associations) of words with similar denotations (definitions) (e.g., stingy, scrimping, economical, unwasteful, thrifty).		
L6.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases; gather vocabulary knowledge when considering a word or phrase important to comprehension or expression.		
*Integration and Language reporting categories have items from both RL and RI Strands (both literary and informational passages). The goal will be to strike an overall balance of approx. 50% per genre on the test form.		56	100%

PAWS Administration 2015 to Present Reading Wyoming Department of Education

Table A5. PAWS 2016 Reading Grade 7 Blueprint

Strand / Standard	PAWS 7th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis	PAWS Administration 2015 to Present Reading Wyoming Department of Education
Reading Literature				
	Key Ideas and Details	9-11	Approx. 32%	
RL7.1	Cite several pieces of textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.			
RL7.2	Determine a theme or central idea of a text and analyze its development over the course of the text; provide an objective summary of the text.			
RL7.3	Analyze how particular elements of a story or drama interact (e.g., how setting shapes the characters or plot).			
	Craft and Structure	7-9		
RL7.4	Determine the meaning of words and phrases as they are used in a text, including figurative and connotative meanings; analyze the impact of rhymes and other repetitions of sounds (e.g., alliteration) on a specific verse or stanza of a poem or section of a story or drama.			
RL7.5	Analyze how a drama's or poem's form or structure (e.g., soliloquy, sonnet) contributes to its meaning.			
RL7.6	Analyze how an author develops and contrasts the points of view of different characters or narrators in a text.			
Reading Informational Text				
	Key Ideas and Details	14-16	Approx. 39%	
RI7.1	Cite several pieces of textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.			
RI7.2	Determine two or more central ideas in a text and analyze their development over the course of the text; provide an objective summary of the text.			
RI7.3	Analyze the interactions between individuals, events, and ideas in a text (e.g., how ideas influence individuals or events, or how individuals influence ideas or events).			
	Craft and Structure	6-8		
RI7.4	Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings; analyze the impact of a specific word choice on meaning and tone.			
RI7.5	Analyze the structure an author uses to organize a text, including how the major sections contribute to the whole and to the development of the ideas.			
RI7.6	Determine an author's point of view or purpose in a text and analyze how the author distinguishes his or her position from that of others.			

Strand / Standard	PAWS 7th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis
	Integration of Knowledge and Ideas*	6-8	
RL7.9	Compare and contrast a fictional portrayal of a time, place, or character and a historical account of the same period as a means of understanding how authors of fiction use or alter history.		Approx. 13%
RI7.8	Trace and evaluate the argument and specific claims in a text, assessing whether the reasoning is sound and the evidence is relevant and sufficient to support the claims.		
RI7.9	Analyze how two or more authors writing about the same topic shape their presentations of key information by emphasizing different evidence or advancing different interpretations of facts.		
	Language*	8-10	
L7.4.a	Use context (e.g., the overall meaning of a sentence or paragraph; a word's position or function in a sentence) as a clue to the meaning of a word or phrase.		Approx. 16%
L7.4.b	Use common, grade-appropriate Greek or Latin affixes and roots as clues to the meaning of a word (e.g., belligerent, bellicose, rebel).		
L7.4.c	Consult general and specialized reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation of a word or determine or clarify its precise meaning or its part of speech.		
L7.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.		
L7.5.a	Interpret figures of speech (e.g., literary, biblical, and mythological allusions) in context.		
L7.5.b	Use the relationship between particular words (e.g., synonym/antonym, analogy) to better understand each of the words.		
L7.5.c	Distinguish among the connotations (associations) of words with similar denotations (definitions) (e.g., refined, respectful, polite, diplomatic, condescending).		
L7.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases; gather vocabulary knowledge when considering a word or phrase important to comprehension or expression.		
*Integration and Language reporting categories have items from both RL and RI Strands (both literary and informational passages). The goal will be to strike an overall balance of approx. 50% per genre on the test form.		56	100%

PAWS Administration 2015 to Present Reading Wyoming Department of Education

Table A6. PAWS 2016 Reading Grade 8 Blueprint

Strand / Standard	PAWS 8th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis	PAWS Administration 2015 to Present Reading Wyoming Department of Education
Reading Literature				
	Key Ideas and Details	9-11	Approx. 32%	
RL8.1	Cite the textual evidence that most strongly supports an analysis of what the text says explicitly as well as inferences drawn from the text.			
RL8.2	Determine a theme or central idea of a text and analyze its development over the course of the text, including its relationship to the characters, setting, and plot; provide an objective summary of the text.			
RL8.3	Analyze how particular lines of dialogue or incidents in a story or drama propel the action, reveal aspects of a character, or provoke a decision.			
	Craft and Structure	7-9		
RL8.4	Determine the meaning of words and phrases as they are used in a text, including figurative and connotative meanings; analyze the impact of specific word choices on meaning and tone, including analogies or allusions to other texts.			
RL8.5	Compare and contrast the structure of two or more texts and analyze how the differing structure of each text contributes to its meaning and style.			
RL8.6	Analyze how differences in the points of view of the characters and the audience or reader (e.g., created through the use of dramatic irony) create such effects as suspense or humor.			
Reading Informational Text				
	Key Ideas and Details	12-14	Approx. 39%	
RI8.1	Cite the textual evidence that most strongly supports an analysis of what the text says explicitly as well as inferences drawn from the text.			
RI8.2	Determine a central idea of a text and analyze its development over the course of the text, including its relationship to supporting ideas; provide an objective summary of the text.			
RI8.3	Analyze how a text makes connections among and distinctions between individuals, ideas, or events (e.g., through comparisons, analogies, or categories).			
	Craft and Structure	8-10		
RI8.4	Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings; analyze the impact of specific word choices on meaning and tone, including analogies or allusions to other texts.			
RI8.5	Analyze in detail the structure of a specific paragraph in a text, including the role of particular sentences in developing and refining a key concept.			
RI8.6	Determine an author's point of view or purpose in a text and analyze how the author acknowledges and responds to conflicting evidence or viewpoints.			

Strand / Standard	PAWS 8th Grade Reading Blueprint Cluster Heading	Items Per Strand	PAWS Emphasis
	Integration of Knowledge and Ideas*	6-8	
RL8.9	Analyze how a modern work of fiction draws on themes, patterns of events, or character types from myths, traditional stories, or religious works such as the Bible, including describing how the material is rendered new.		Approx. 13%
RI8.8	Delineate and evaluate the argument and specific claims in a text, assessing whether the reasoning is sound and the evidence is relevant and sufficient; recognize when irrelevant evidence is introduced.		
RI8.9	Analyze a case in which two or more texts provide conflicting information on the same topic and identify where the texts disagree on matters of fact or interpretation.		
	Language*	8-10	
L8.4.a	Use context (e.g., the overall meaning of a sentence or paragraph; a word's position or function in a sentence) as a clue to the meaning of a word or phrase.		Approx. 16%
L8.4.b	Use common, grade-appropriate Greek or Latin affixes and roots as clues to the meaning of a word (e.g., precede, recede, secede).		
L8.4.c	Consult general and specialized reference materials (e.g., dictionaries, glossaries, thesauruses), both print and digital, to find the pronunciation of a word or determine or clarify its precise meaning or its part of speech.		
L8.5	Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.		
L8.5.a	Interpret figures of speech (e.g. verbal irony, puns) in context.		
L8.5.b	Use the relationship between particular words to better understand each of the words.		
L8.5.c	Distinguish among the connotations (associations) of words with similar denotations (definitions) (e.g., bullheaded, willful, firm, persistent, resolute).		
L8.6	Acquire and use accurately grade-appropriate general academic and domain-specific words and phrases; gather vocabulary knowledge when considering a word or phrase important to comprehension or expression.		
*Integration and Language reporting categories have items from both RL and RI Strands (both literary and informational passages). The goal will be to strike an overall balance of approx. 50% per genre on the test form.		56	100%

PAWS Administration 2015 to Present Reading Wyoming Department of Education

Mathematics

Table A7. PAWS 2016 Mathematics Grade 3 Blueprint

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

Table A8. PAWS 2016 Mathematics Grade 4 Blueprint

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

PAWS Administration 2014 to Present Mathematics Wyoming Department of Education

Table A9. PAWS 2016 Mathematics Grade 5 Blueprint

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

PAWS Administration 2014 to Present Mathematics Wyoming Department of Education

Table A10. PAWS 2016 Mathematics Grade 6 Blueprint

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

PAWS Administration 2014 to Present Mathematics Wyoming Department of Education

Table A11. PAWS 2016 Mathematics Grade 7 Blueprint

Domain / Standard	PAWS 7th Grade Mathematics Blueprint Cluster Heading	Focus m = major s = supporting a = additional	Items Per Domain	# of Items / Cluster Heading	PAWS Emphasis
Ratios and Proportional Relationships			13		
7.RP.1 7.RP.2 7.RP.3	Analyze proportional relationships and use them to solve real-world and mathematical problems.	m		13	22%
The Number System			10		
7.NS.1 7.NS.2 7.NS.3	Apply and extend previous understandings of operations with fractions to add, subtract, multiply, and divide rational numbers.	m		10	17%
Expressions and Equations			18		
7.EE.1 7.EE.2	Use properties of operations to generate equivalent expressions.	m		5-7	31%
7.EE.3 7.EE.4	Solve real-life and mathematical problems using numerical and algebraic expressions and equations.	m		11-13	
Geometry			9		
7.G.1 7.G.2 7.G.3	Draw, construct, and describe geometrical figures and describe the relationships between them.	a		1-3	15%
7.G.4 7.G.5 7.G.6	Solve real-life and mathematical problems involving angle measure, area, surface area, and volume.	a		6-8	
Statistics and Probability			9		
7.SP.1 7.SP.2	Use random sampling to draw inferences about a population.	s		2-4	15%
7.SP.3 7.SP.4	Draw informal comparative inferences about two populations.	a		1-3	
7.SP.5 7.SP.6 7.SP.7 7.SP.8	Investigate chance processes and develop, use, and evaluate probability models.	s		3-5	
			59	59	100%

PAWS Administration 2014 to Present Mathematics Wyoming Department of Education

Table A12. PAWS 2016 Mathematics Grade 8 Blueprint

Domain / Standard	PAWS 8th Grade Mathematics Blueprint Cluster Heading	Focus m = major s = supporting a = additional	Items Per Domain	# of Items / Cluster Heading	PAWS Emphasis	
The Number System			6			PAWS Administration 2014 to Present Mathematics Wyoming Department of Education
8.NS.1 8.NS.2	Know that there are numbers that are not rational, and approximate them by rational numbers.	s		6	9%	
Expressions and Equations			23			
8.EE.1 8.EE.2 8.EE.3 8.EE.4	Work with radicals and integer exponents.	m		6-8	35%	
8.EE.5 8.EE.6	Understand the connections between proportional relationships, lines, and linear equations.	m		5-7		
8.EE.7 8.EE.8	Analyze and solve linear equations and pairs of simultaneous linear equations.	m		9-11		
Functions			14			
8.F.1 8.F.2 8.F.3	Define, evaluate, and compare functions.	m		7-9	22%	
8.F.4 8.F.5	Use functions to model relationships between quantities.	s		5-7		
Geometry			16			
8.G.1 8.G.2 8.G.3 8.G.4 8.G.5	Understand congruence and similarity using physical models, transparencies, or geometry software.	m		6-8	25%	
8.G.6 8.G.7 8.G.8	Understand and apply the Pythagorean Theorem.	m		5-7		
8.G.9	Solve real-world and mathematical problems involving volume of cylinders, cones, and spheres.	a		2-4		
Statistics and Probability			6			
8.SP.1 8.SP.2 8.SP.3 8.SP.4	Investigate patterns of association in bivariate data.	s		6	9%	
			65	65	100%	

Science

Table A13. PAWS 2016 Science Grade 4 Blueprint

Domain / Standard	PAWS 4th Grade Science Blueprint Benchmarks	# of Items Aligning Skill & Benchmark	Items Per Branch	# of Items / Benchmark	PAWS Emphasis	PAWS Administration 2014 to Present Science Wyoming Department of Education
SKILLS: I.1 Observe and Question I.2 Design and Conduct a Scientific Investigation I.3 Organize and Represent Data I.4 Draw Conclusions and Make Connections						
Concepts and Processes						
Life Science			16			
4.1.1	Characteristics of Organisms: Students describe observable characteristics of living things, including structures that serve specific functions and everyday behaviors.	I.1: 0-2 I.2: 1-2 I.3: 0-1 I.4: 0-1		5-6	32%	
4.1.2	Life Cycles of Organisms: Students sequence life cycles of living things, and recognize that plants and animals resemble their parents.	I.1: 0-1 I.2: 1-2 I.3: 0-1 I.4: 1-2		5-6		
4.1.3	Organisms and Their Environments: Students show connections between living things, their basic needs, and the environments.	I.1: 0-1 I.2: 1-2 I.3: 0-1 I.4: 0-1		4-5		
Earth and Space Science			16			
4.1.4	Properties of Earth Materials: Students investigate water, air, rocks, and soils to compare basic properties of earth materials.	I.1: 0-1 I.2: 0-1 I.3: 0-1 I.4: 1-2		4-5	32%	
4.1.5	Objects in the Sky: Students describe observable objects in the sky and their patterns of movement.	I.1: 0-2 I.2: 0-1 I.3: 0-1 I.4: 1-2		5-6		
4.1.6	Changes in Earth and Sky: Students describe observable changes in earth and sky, including rapid and gradual changes to the earth's surface, and daily and seasonal changes in the weather.	I.1: 0-1 I.2: 1-2 I.3: 1-2 I.4: 1-2		5-6		

Domain / Standard	PAWS 4th Grade Science Blueprint Benchmarks	# of Items Aligning Skill & Benchmark	Items Per Branch	# of Items / Benchmark	PAWS Emphasis
SKILLS: I.1 Observe and Question I.2 Design and Conduct a Scientific Investigation I.3 Organize and Represent Data I.4 Draw Conclusions and Make Connections					
Concepts and Processes					
Physical Science			18		
4.1.7	Properties of Objects: Students classify objects by properties that can be observed, measured, and recorded, including color, shape, size, weight, volume, texture, and temperature.	I.1: 0-1 I.2: 1-2 I.3: 0-1 I.4: 0-1		4-5	36%
4.1.8	Changes in States of Matter: Students demonstrate that the processes of heating and cooling can change matter from one state to another.	I.1: 1-2 I.2: 0-1 I.3: 0-1 I.4: 1-2		4-5	
4.1.9	Physical Phenomena: Students investigate physical phenomena commonly encountered in daily life, including light, heat, electricity, sound, and magnetism.	I.1: 0-1 I.2: 1-2 I.3: 0-1 I.4: 1-2		4-5	
4.1.10	Position and Motion of Objects: Students demonstrate that pushing and pulling can change the position and motion of objects.	I.1: 0-1 I.2: 0-1 I.3: 0-1 I.4: 0-1		4-5	
Science as Inquiry					
4.2.1	Students research answers to science questions and present findings through appropriate means.	Not Assessed			
4.2.2	Students use the inquiry process to conduct simple scientific investigations: 1) Collect and organize data; 2) Use data to construct simple graphs, charts, diagrams, and/or model; 3) Draw conclusions and accurately communicate results, making connections to daily life; 4) Pose or identify questions and make predictions; and 5) Conduct investigations to answer questions and check predictions.	Assessed with Concepts & Processes			
4.2.3	Students identify and use appropriate scientific equipment.				
4.2.4	Students properly use safety equipment and recognize hazards and safety symbols while practicing standard safety procedures.				
			50	50	100%

PAWS Administration 2014 to Present Science Wyoming Department of Education

Table A14. PAWS 2016 Science Grade 8 Blueprint

Domain / Standard	PAWS 8th Grade Science Blueprint Benchmarks	# of Items Aligning Skill & Benchmark	Items Per Branch	# of Items / Benchmark	PAWS Emphasis	PAWS Administration 2014 to Present Science Wyoming Department of Education
SKILLS: I.1 Observe and Question I.2 Design and Conduct a Scientific Investigation I.3 Organize and Represent Data I.4 Draw Conclusions and Make Connections						
Concepts and Processes						
Life Science			16			
8.1.1	Levels of Organization in Living Systems: Students model the cell as the basic unit of a living system. They realize that all functions that sustain life act within a single cell and cells differentiate into specialized cells, tissues, organs, and organ systems.	I.1: 0-2 I.2: 1-2 I.3: 0-1 I.4: 0-1		2-3	32%	
8.1.2	Reproduction and Heredity: Students describe reproduction as a characteristic of all living systems, which is essential to the continuation of species, and identify and interpret traits, patterns of inheritance, and the interaction between genetics and environment.	I.1: 0-1 I.2: 0-1 I.3: 0-1 I.4: 0-1		2-3		
8.1.3	Evolution as a Theory: Students explain evolution as a theory and apply the theory to the diversity of species, which results from natural selection and the acquisition of unique characteristics through biological adaptation.	I.1: 0-1 I.2: 0-1 I.3: 0-1 I.4: 0-1		2-3		
8.1.4	Diversity of Organisms: Students investigate the interconnectedness of organisms, identifying similarity and diversity of organisms through a classification system of hierarchical relationships and structural homologies.	I.1: 0-1 I.2: 0-1 I.3: 0-1 I.4: 0-1		2-3		
8.1.5	Behavior and Adaptation: Students recognize behavior as a response of an organism to an internal or environmental stimulus and connect the characteristics and behaviors of an organism to biological adaptation.	I.1: 0-1 I.2: 0-1 I.3: 0-1 I.4: 0-1		2-3		
8.1.6	Interrelationships of Populations and Ecosystems: Students illustrate populations of organisms and their interconnection within an ecosystem, identifying relationships among producers, consumers, and decomposers.	I.1: 0-2 I.2: 0-1 I.3: 0-1 I.4: 1-2		2-3		

Domain / Standard	PAWS 8th Grade Science Blueprint Benchmarks	# of Items Aligning Skill & Benchmark	Items Per Branch	# of Items / Benchmark	PAWS Emphasis
SKILLS: I.1 Observe and Question I.2 Design and Conduct a Scientific Investigation I.3 Organize and Represent Data I.4 Draw Conclusions and Make Connections					
Concepts and Processes					
Earth and Space Science			16		
8.1.7	The Earth in the Solar System: Students describe Earth as the third planet in the Solar System and understand the effects of the sun as a major source of energy, gravitational forces, and motions of objects in the Solar System.	I.1: 0-1 I.2: 1-2 I.3: 0-1 I.4: 1-2		5-6	32%
8.1.8	The Structure of the Earth System: Students examine the structure of the Earth, identifying layers of the Earth, considering plate movement and its effect, and recognizing landforms resulting from constructive and destructive forces.	I.1: 1-2 I.2: 1-2 I.3: 0-1 I.4: 1-2		5-6	
8.1.9	The Earth's History: Students systematize the Earth's history in terms of geologic evidence, comparing past and present Earth processes and identifying catastrophic events and fossil evidence.	I.1: 0-1 I.2: 0-1 I.3: 1-2 I.4: 1-2		5-6	
Physical Science			18		
8.1.10	The Structure and Properties of Matter: Students identify characteristic properties of matter such as density, solubility, and boiling point and understand that elements are the basic components of matter.	I.1: 0-1 I.2: 1-2 I.3: 0-1 I.4: 0-1		3-4	36%
8.1.11	Physical and Chemical Changes in Matter: Students evaluate chemical and physical changes, recognizing that chemical change forms compounds with different properties and that physical change alters the appearance but not the composition of a substance.	I.1: 1-2 I.2: 0-1 I.3: 0-1 I.4: 1-2		3-4	
8.1.12	Forms and Uses of Energy: Students investigate energy as a property of substances in a variety of forms with a range of uses.	I.1: 0-1 I.2: 1-2 I.3: 0-1 I.4: 1-2		3-4	
8.1.13	The Conservation of Matter and Energy: Students identify supporting evidence to explain conservation of matter and energy, indicating that matter or energy cannot be created or destroyed but is transferred from one object to another.	I.1: 0-1 I.2: 0-1 I.3: 0-1 I.4: 0-1		3-4	
8.1.14	Effects of Motions and Forces: Students describe motion of an object by position, direction, and speed, and identify the effects of force and inertia on an object.	I.1: 0-1 I.2: 1-2 I.3: 0-1 I.4: 1-2		3-4	

PAWS Administration 2014 to Present Science Wyoming Department of Education

Appendix B: Sample PAWS Student Reports: Grade 4 Reading, Mathematics, and Science (exemplar for Grade 8)

WYOMING
DEPARTMENT OF EDUCATION

PAWS Student Report

First Name: Abigail xxxxxxx
Middle Initial: L
Last Name: Buchbaumxxxxxxxxxxxx xxxx
Grade: 4
Birthdate: 7/14/2006
Student ID: 123456

Test Window: 3/2/16 - 3/27/16
School: Buffalo Ridge Elementary xxxxxxxxxxxxxxxx
District: Laramie xxxxxxxxxxxx xxxx

YOUR CHILD'S PAWS RESULTS

READING

460*

Your child's score: 460
Performance Level: Below Basic
State Percentile Rank: 56

Below Basic	Basic	Proficient	Advanced
(375-552)	(553-589)	(590-640)	(641-800)

MATHEMATICS

850*

Your child's score: 850
Performance Level: Advanced
State Percentile Rank: 75

Below Basic	Basic	Proficient	Advanced
(375-549)	(550-598)	(599-659)	(660-850)

SCIENCE

563*

Your child's score: 563
Performance Level: Below Basic
State Percentile Rank: 57

Below Basic	Basic	Proficient	Advanced
(300-611)	(612-665)	(666-725)	(726-975)

* The dotted lines indicate the range of likely scores your child would receive if your child took many versions of the test.

Lexile® Measure 123L **Quantile® Measure** 321Q

For more information, and to search for books by Lexile measure, visit www.Lexile.com.
For more information about Quantile measures, visit the Math @ Home section at www.Quantiles.com.

Glossary of Terms

Your child's score: is represented by a scale score. A scale score is the raw score (total number of points earned) transformed into a score on a scale.

State Percentile Rank: Indicates your child's performance in relation to other Wyoming students in the same grade. The number shows the percentage of Wyoming students in the same grade who obtained scores equal to or less than your child's score.

Lexile Measure: Helps readers select materials at their reading level. This can serve as a guide in selecting books for your child.

Quantile Measure: Helps you identify math activities to do at home. These activities will help your child practice mathematical skills leading to increased mathematical understanding.

Topics Tested: Describes your child's performance in sub-categories (domains) of each content area. The student performance column indicates relative performance within each content area. The scale score column indicates the performance in relation to the overall scale. The state average score shows how the Wyoming student population performed in each of the measured domains.

Not Scored (NS): Your child did not complete enough of the test to calculate a score.

Not Tested (NT): Your child did not take this part of the assessment.

Not Applicable (NA): This content area is not tested at this grade level.

TOPICS TESTED

	Domain Content Area	Student Performance	Student Scale Score	State Average Score
Reading	Literature: Key Ideas and Details	★★★	315	315
	Literature: Craft and Structure	★★★	306	306
	Informational Text: Key Ideas and Details	★★	415	415
	Informational Text: Craft and Structure	★★★	808	808
	Integration of Knowledge and Ideas	★★★	406	406
	Language	★	460	460
Math	Geometry	★★	813	813
	Measurement and Data	★	610	610
	Operations and Algebraic Thinking	★★	320	320
	Number Operations – Base 10	★★★	510	510
	Number Operations – Fractions	★★★	306	306
Science	Life Science	★★	813	813
	Physical Science	★	610	610
	Earth and Space Science	★★	320	320

The stars indicate your child's relative performance on each domain within reading, mathematics, or science. One star is below average and three stars is above average.

For Additional Resources and Information -
Visit the Wyoming Department of Education online

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

A Guide to the PAWS Score Report

This report provides information about your child's achievement on the Proficiency Assessments for Wyoming Students (PAWS). This report will help you understand your child's performance in reading and math (grades 3-8) and science (grades 4 and 8) for the 2015-2016 school year. As indicated, several types of information are provided on this report. One is a description of your child's performance level (advanced, proficient, basic, or below basic) in each content area. Another is a scale score your child earned for those content areas. You can also see your child's performance in each of the content area domains (subcategories). Overall, these results are designed to provide your child's school with information about how well your child is learning the Wyoming Content and Performance Standards.

The charts at the top right of the first page indicate the range of the scale scores depicting your child's scale score in each of the content areas. Different scales are used for each content area, so the numerical scores from each content area are not comparable. The placement of the score in each proficiency band is important information for you and your child's teacher. The "Topics Tested" box on the lower right side of the front page contains more detailed information about your child's performance. This information helps you and your child's teacher identify specific areas of strength and areas for improvement within the PAWS topics. Your child's performance is also expressed as a number of stars showing your child's strengths and weaknesses in each topic . Lexile and Quantile scores are also provided for your child. The websites indicated will provide valuable tools and resources which can support your child's academic growth.

Performance Level Descriptors

	Reading	Mathematics	Science
Advanced	Students demonstrate thorough ability to comprehend implied main ideas, answer questions, and explain structural comparisons between one or more texts. They understand complex words and phrases.	Students demonstrate an in-depth understanding of whole numbers and measurements to solve problems using the four operations; whole number place value; equivalent fractions or decimals; fraction operations; classifying and analyzing geometric shapes.	Students demonstrate in-depth science knowledge in complex situations to analyze, construct, and use information and ideas. They evaluate conclusions based on evidence to explain phenomena and generalize cause and effect relationships.
Proficient	Students demonstrate adequate ability to comprehend main ideas, answer questions about text, and explain structural comparisons between one or more texts. They understand grade-appropriate words and phrases.	Students demonstrate a solid understanding of whole numbers and measurements to solve problems using the four operations; whole number place value; equivalent fractions or decimals; fraction operations; identifying or classifying geometric shapes.	Students demonstrate solid understanding of concepts and skills to analyze information and use ideas. They can describe and recall facts, perform investigations, record results, predict outcomes and provide solutions based on evidence.
Basic	Students demonstrate partial ability to comprehend main ideas, answer questions, and explain structural comparisons between one or more texts. They understand simple words and phrases	Students demonstrate a partial understanding of whole numbers and measurements to solve problems using the four operations; whole number place value; equivalent fractions or decimals; fraction operations; identifying or classifying geometric shapes.	Students demonstrate partial understanding or require assistance to use concepts and skills to analyze facts and ideas. They demonstrate or explain basic scientific principles and observations with support.
Below Basic	Students require extensive support or provide little or no evidence in meeting the standard.	Students require extensive support or provide little or no evidence in meeting the standard.	Students require extensive support or provide little or no evidence in meeting the standard.

What Can You Do at Home?

Reading

- Read for at least thirty minutes per night.
- Read aloud to your child.
- Model being a reader.
- Ask who, what, where, when, why and how, questions about the book your child is reading.
- Ask your child to jot notes about his or her reading.

Mathematics

- Be familiar with what your child is learning at school and work on those concepts in different contexts.
- Have a specific place for homework at home where you can check your child's progress and concepts being taught at any time.
- Practice fluency facts.

Science

- Promote investigative activities that happen outside.
- Turn everyday household activities into experiments (baking, cleaning, etc.).
- Form hypotheses when trying to work out everyday problems.

Appendix C: Sample PAWS Student Reports: Grade 6 Reading, Mathematics (exemplar for Grade 3, 5, and 7)

WYOMING
DEPARTMENT OF EDUCATION

PAWS Student Report

First Name: Abigail xxxxxxx
Middle Initial: L
Last Name: Buchbaumxxxxxxxxxxxx xxxx

Grade: 6
Birthdate: 7/14/2003
Student ID: 123456

Test Window: 3/2/16 - 5/1/16
School: Buffalo Ridge Elementary xxxxxxxxxxxxxxx
District: Laramie xxxxxxxxxxxx xxxx

YOUR CHILD'S PAWS RESULTS

READING

Your child's score: 460
Performance Level: Below Basic
State Percentile Rank: 56

460*

Below Basic	Basic	Proficient	Advanced
(375-552)	(553-589)	(590-640)	(641-800)

MATHEMATICS

Your child's score: 850
Performance Level: Advanced
State Percentile Rank: 76

850*

Below Basic	Basic	Proficient	Advanced
(375-549)	(550-598)	(599-659)	(660-850)

SCIENCE

Your child's score: NA

* The dotted lines indicate the range of likely scores your child would receive if your child took many versions of the test.

Lexile® Measure 123L **Quantile® Measure** 321Q

For more information, and to search for books by Lexile measure, visit www.Lexile.com.
For more information about Quantile measures, visit the Math @ Home section at www.Quantiles.com.

Glossary of Terms

Your child's score: is represented by a scale score. A scale score is the raw score (total number of points earned) transformed into a score on a scale.

State Percentile Rank: Indicates your child's performance in relation to other Wyoming students in the same grade. The number shows the percentage of Wyoming students in the same grade who obtained scores equal to or less than your child's score.

Lexile Measure: Helps readers select materials at their reading level. This can serve as a guide in selecting books for your child.

Quantile Measure: Helps you identify math activities to do at home. These activities will help your child practice mathematical skills leading to increased mathematical understanding.

Topics Tested: Describes your child's performance in sub-categories (domains) of each content area. The student performance column indicates relative performance within each content area. The scale score column indicates the performance in relation to the overall scale. The state average score shows how the Wyoming student population performed in each of the measured domains.

Not Scored (NS): Your child did not complete enough of the test to calculate a score.

Not Tested (NT): Your child did not take this part of the assessment.

Not Applicable (NA): This content area is not tested at this grade level.

TOPICS TESTED

	Domain Content Area	Student Performance	Student Scale Score	State Average Score
Reading	Literature: Key Ideas and Details	★★★	315	315
	Literature: Craft and Structure	★★★	306	306
	Informational Text: Key Ideas and Details	★★	415	415
	Informational Text: Craft and Structure	★★★	808	808
	Integration of Knowledge and Ideas	★★★	406	406
	Language	★	460	460
Math	Geometry	★★	813	813
	Ratio and Proportional Relationships	★	610	610
	The Number System	★★	320	320
	Expressions and Equations	★★★	510	510
	Statistics and Probability	★★★	306	306
Science	NA			

The stars indicate your child's relative performance on each domain within reading, mathematics, or science. One star is below average and three stars is above average.

For Additional Resources and Information -
Visit the Wyoming Department of Education online

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

A Guide to the PAWS Score Report

This report provides information about your child's achievement on the Proficiency Assessments for Wyoming Students (PAWS). This report will help you understand your child's performance in reading and math (grades 3-8) and science (grades 4 and 8) for the 2015-2016 school year. As indicated, several types of information are provided on this report. One is a description of your child's performance level (advanced, proficient, basic, or below basic) in each content area. Another is a scale score your child earned for those content areas. You can also see your child's performance in each of the content area domains (subcategories). Overall, these results are designed to provide your child's school with information about how well your child is learning the Wyoming Content and Performance Standards.

The charts at the top right of the first page indicate the range of the scale scores depicting your child's scale score in each of the content areas. Different scales are used for each content area, so the numerical scores from each content area are not comparable. The placement of the score in each proficiency band is important information for you and your child's teacher. The "Topics Tested" box on the lower right side of the front page contains more detailed information about your child's performance. This information helps you and your child's teacher identify specific areas of strength and areas for improvement within the PAWS topics. Your child's performance is also expressed as a number of stars showing your child's strengths and weaknesses in each topic . Lexile and Quantile scores are also provided for your child. The websites indicated will provide valuable tools and resources which can support your child's academic growth.

Performance Level Descriptors

	Reading	Mathematics	Science
Advanced	Students demonstrate thorough comprehension by supporting inferences, analyzing points of view, and making comparisons between themes, topics, and text structures within/across texts. They understand complex words and phrases.	Students demonstrate an in-depth understanding of ratios, positive rational numbers, and geometric properties to solve problems; negative numbers; writing, interpreting, and using expressions, equations, and inequalities; summarizing data.	Not Applicable.
Proficient	Students demonstrate adequate comprehension by supporting inferences, explaining points of view, and making comparisons between themes, topics, and text structures within/across texts. They understand grade-appropriate words and phrases.	Students demonstrate a solid understanding of ratios, positive rational numbers, and geometric properties to solve problems; negative numbers; writing, interpreting, and using expressions, equations, and inequalities; summarizing data.	Not Applicable.
Basic	Students demonstrate partial comprehension by supporting inferences, identifying points of view, and making comparisons between themes, topics, and text structures within/across texts. They understand simple words and phrases.	Students demonstrate a partial understanding of ratios, positive rational numbers, and geometric properties to solve problems; negative numbers; writing, interpreting, and using expressions, equations, and inequalities; summarizing data.	Not Applicable.
Below Basic	Students require extensive support or provide little or no evidence in meeting the standard.	Students require extensive support or provide little or no evidence in meeting the standard.	Not Applicable.

What Can You Do at Home?

Reading

- Read for at least thirty minutes per night.
- Read aloud to your child.
- Model being a reader.
- Ask who, what, where, when, why and how, questions about the book your child is reading.
- Ask your child to jot notes about his or her reading.

Mathematics

- Be familiar with what your child is learning at school and work on those concepts in different contexts.
- Have a specific place for homework at home where you can check your child's progress and concepts being taught at any time.
- Practice fluency facts.

Science

- Promote investigative activities that happen outside.
- Turn everyday household activities into experiments (baking, cleaning, etc.).
- Form hypotheses when trying to work out everyday problems.

Appendix D: PAWS Operational Subscale Correlations

Table D1. PAWS Grade 3 Total Test and Subscale Correlations

	Reading Total	LTKY	LTCR	INKY	INCR	INTG	LANG	Mathematics Total	GEOM	MEAS	ALGE	BTEN	FRCT
Reading Total	1.00												
LTKY	0.81	1.00											
LTCR	0.72	0.56	1.00										
INKY	0.83	0.58	0.52	1.00									
INCR	0.75	0.51	0.46	0.58	1.00								
INTG	0.84	0.61	0.53	0.63	0.55	1.00							
LANG	0.82	0.59	0.54	0.60	0.54	0.61	1.00						
Mathematics Total	0.73	0.57	0.51	0.62	0.54	0.64	0.60	1.00					
GEOM	0.51	0.39	0.36	0.43	0.37	0.45	0.42	0.68	1.00				
MEAS	0.66	0.50	0.44	0.56	0.49	0.58	0.53	0.89	0.54	1.00			
ALGE	0.70	0.55	0.49	0.59	0.50	0.61	0.58	0.94	0.56	0.76	1.00		
BTEN	0.59	0.46	0.41	0.49	0.43	0.51	0.49	0.80	0.48	0.66	0.71	1.00	
FRCT	0.48	0.35	0.33	0.42	0.36	0.42	0.38	0.72	0.44	0.59	0.57	0.49	1.00

Table D2. PAWS Grade 4 Total Test and Subscale Correlations

	Reading Total	LTKY	LTCR	INKY	INCR	INTG	LANG	Mathematics Total	GEOM	MEAS	ALGE	BTEN	FRCT	Science Total	LIFE	PHYS	ESCI
Reading Total	1.00																
LTKY	0.86	1.00															
LTCR	0.71	0.54	1.00														
INKY	0.85	0.64	0.53	1.00													
INCR	0.75	0.56	0.48	0.58	1.00												
INTG	0.77	0.58	0.48	0.59	0.50	1.00											
LANG	0.81	0.64	0.52	0.61	0.54	0.55	1.00										
Mathematics Total	0.70	0.58	0.52	0.61	0.53	0.55	0.56	1.00									
GEOM	0.38	0.30	0.29	0.34	0.30	0.29	0.30	0.59	1.00								
MEAS	0.53	0.43	0.40	0.48	0.41	0.41	0.41	0.81	0.42	1.00							
ALGE	0.66	0.56	0.48	0.57	0.49	0.52	0.53	0.87	0.43	0.62	1.00						
BTEN	0.60	0.50	0.43	0.51	0.43	0.46	0.50	0.79	0.38	0.55	0.70	1.00					
FRCT	0.64	0.52	0.48	0.55	0.48	0.50	0.51	0.93	0.47	0.70	0.73	0.65	1.00				
Science Total	0.79	0.67	0.56	0.69	0.61	0.60	0.64	0.75	0.42	0.60	0.69	0.61	0.69	1.00			
LIFE	0.75	0.64	0.52	0.65	0.58	0.57	0.61	0.67	0.38	0.52	0.62	0.55	0.61	0.90	1.00		
PHYS	0.71	0.59	0.51	0.62	0.54	0.54	0.57	0.70	0.40	0.57	0.63	0.55	0.64	0.92	0.75	1.00	
ESCI	0.71	0.60	0.50	0.62	0.54	0.54	0.56	0.68	0.37	0.55	0.62	0.56	0.62	0.90	0.73	0.73	1.00

Table D3. PAWS Grade 5 Total Test and Subscale Correlations

	Reading Total	LTKY	LTCR	INKY	INCR	INTG	LANG	Mathematics Total	GEOM	MEAS	ALGE	BTEN	FRCT
Reading Total	1.00												
LTKY	0.84	1.00											
LTCR	0.77	0.61	1.00										
INKY	0.90	0.67	0.62	1.00									
INCR	0.78	0.59	0.55	0.65	1.00								
INTG	0.71	0.54	0.49	0.58	0.51	1.00							
LANG	0.83	0.63	0.59	0.67	0.59	0.52	1.00						
Mathematics Total	0.70	0.55	0.52	0.65	0.55	0.52	0.59	1.00					
GEOM	0.51	0.40	0.39	0.48	0.41	0.38	0.43	0.71	1.00				
MEAS	0.59	0.47	0.43	0.55	0.48	0.44	0.50	0.86	0.56	1.00			
ALGE	0.57	0.45	0.43	0.54	0.44	0.43	0.47	0.75	0.48	0.59	1.00		
BTEN	0.63	0.50	0.47	0.59	0.50	0.48	0.53	0.90	0.57	0.70	0.65	1.00	
FRCT	0.64	0.50	0.48	0.60	0.50	0.48	0.54	0.93	0.60	0.74	0.64	0.78	1.00

Table D4. PAWS Grade 6 Total Test and Subscale Correlations

	Reading Total	LTKY	LTCR	INKY	INCR	INTG	LANG	Mathematics Total	GEOM	RELT	NMBR	EQTN	STPR
Reading Total	1.00												
LTKY	0.88	1.00											
LTCR	0.81	0.68	1.00										
INKY	0.90	0.72	0.67	1.00									
INCR	0.78	0.60	0.56	0.63	1.00								
INTG	0.84	0.69	0.64	0.71	0.60	1.00							
LANG	0.77	0.60	0.57	0.62	0.55	0.57	1.00						
Mathematics Total	0.74	0.63	0.60	0.69	0.57	0.62	0.58	1.00					
GEOM	0.54	0.45	0.42	0.50	0.43	0.44	0.42	0.77	1.00				
RELT	0.67	0.57	0.54	0.63	0.53	0.57	0.52	0.83	0.58	1.00			
NMBR	0.61	0.52	0.49	0.57	0.45	0.51	0.47	0.87	0.62	0.63	1.00		
EQTN	0.69	0.59	0.56	0.64	0.53	0.58	0.55	0.92	0.66	0.72	0.72	1.00	
STPR	0.57	0.48	0.45	0.53	0.45	0.47	0.45	0.74	0.52	0.57	0.55	0.61	1.00

Table D5. PAWS Grade 7 Total Test and Subscale Correlations

	Reading Total	LTKY	LTCR	INKY	INCR	INTG	LANG	Mathematics Total	GEOM	RELT	NMBR	EQTN	STPR
Reading Total	1.00												
LTKY	0.81	1.00											
LTCR	0.78	0.60	1.00										
INKY	0.91	0.67	0.65	1.00									
INCR	0.79	0.56	0.57	0.66	1.00								
INTG	0.72	0.50	0.49	0.59	0.52	1.00							
LANG	0.83	0.59	0.58	0.69	0.61	0.54	1.00						
Mathematics Total	0.74	0.56	0.57	0.67	0.57	0.57	0.65	1.00					
GEOM	0.53	0.39	0.41	0.48	0.41	0.41	0.49	0.77	1.00				
RELT	0.66	0.50	0.50	0.60	0.51	0.51	0.57	0.89	0.61	1.00			
NMBR	0.65	0.49	0.50	0.60	0.50	0.49	0.57	0.86	0.59	0.71	1.00		
EQTN	0.64	0.48	0.49	0.58	0.49	0.48	0.56	0.91	0.62	0.73	0.72	1.00	
STPR	0.67	0.50	0.51	0.61	0.52	0.51	0.58	0.79	0.54	0.65	0.63	0.63	1.00

Table D6. PAWS Grade 8 Total Test and Subscale Correlations

	Reading Total	LTKY	LTCR	INKY	INCR	INTG	LANG	Mathematics Total	GEOM	FNCT	NMBR	EQTN	STPR	Science Total	LIFE	PHYS	ESCI
Reading Total	1.00																
LTKY	0.84	1.00															
LTCR	0.72	0.57	1.00														
INKY	0.84	0.63	0.51	1.00													
INCR	0.82	0.61	0.51	0.62	1.00												
INTG	0.74	0.54	0.46	0.57	0.53	1.00											
LANG	0.81	0.61	0.52	0.60	0.59	0.53	1.00										
Mathematics Total	0.69	0.54	0.48	0.60	0.58	0.53	0.56	1.00									
GEOM	0.59	0.46	0.40	0.51	0.49	0.46	0.47	0.86	1.00								
FNCT	0.59	0.45	0.41	0.51	0.50	0.45	0.48	0.86	0.64	1.00							
NMBR	0.46	0.36	0.32	0.40	0.39	0.36	0.36	0.71	0.54	0.53	1.00						
EQTN	0.63	0.49	0.43	0.55	0.53	0.49	0.52	0.93	0.72	0.75	0.61	1.00					
STPR	0.58	0.47	0.40	0.49	0.49	0.44	0.49	0.71	0.55	0.57	0.44	0.60	1.00				
Science Total	0.75	0.59	0.52	0.65	0.62	0.57	0.60	0.79	0.67	0.68	0.53	0.73	0.62	1.00			
LIFE	0.68	0.54	0.47	0.59	0.57	0.52	0.55	0.71	0.61	0.62	0.47	0.66	0.55	0.91	1.00		
PHYS	0.68	0.54	0.48	0.59	0.58	0.52	0.54	0.74	0.64	0.63	0.50	0.68	0.56	0.91	0.74	1.00	
ESCI	0.66	0.52	0.46	0.57	0.54	0.50	0.53	0.68	0.57	0.59	0.45	0.63	0.55	0.88	0.71	0.71	1.00

Appendix E: DIF Results for 2016 Field Test Items[6]

Reading

Table E1. Reading Grade 3 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	0	0	0	0
B-	2	1.4	0	0	0	0	1	0.7	0	0
A	131	93.6	0	0	0	0	55	39.3	0	0
B+	7	5.0	0	0	0	0	0	0	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL *N*	0	0	140	100	140	100	84	60	140	100
TOTAL	140	100	140	100	140	100	140	100	140	100

Table E2. Reading Grade 4 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	1	0.7	0	0
B-	3	2.1	0	0	0	0	2	1.4	0	0
A	131	93.6	0	0	0	0	122	87.1	0	0
B+	6	4.3	0	0	0	0	1	0.7	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL *N*	0	0	140	100	140	100	14	10	140	100
TOTAL	140	100	140	100	140	100	140	100	140	100

[6] Not all percentages will sum to a total of 100 due to rounding.

Table E3. Reading Grade 5 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	0	0	0	0
B-	4	2.9	0	0	0	0	5	3.6	0	0
A	117	85.4	0	0	0	0	127	92.7	0	0
B+	15	10.9	0	0	0	0	4	2.9	0	0
C+	1	0.7	0	0	0	0	1	0.7	0	0
SMALL *N*	0	0	137	100	137	100	0	0	137	100
TOTAL	137	100	137	100	137	100	137	100	137	100

Table E4. Reading Grade 6 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	2	1.5	0	0	0	0	0	0	0	0
B-	3	2.2	0	0	0	0	2	1.5	0	0
A	122	89.7	0	0	0	0	57	41.9	0	0
B+	8	5.9	0	0	0	0	1	0.7	0	0
C+	1	0.7	0	0	0	0	0	0	0	0
SMALL *N*	0	0	136	100	136	100	76	55.9	136	100
TOTAL	136	100	136	100	136	100	136	100	136	100

Table E5. Reading Grade 7 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	1	0.7	0	0
B-	5	3.6	0	0	0	0	0	0	0	0
A	106	76.3	0	0	0	0	27	19.4	0	0
B+	25	18	0	0	0	0	1	0.7	0	0
C+	3	2.2	0	0	0	0	0	0	0	0
SMALL *N*	0	0	139	100	139	100	110	79.1	139	100
TOTAL	139	100	139	100	139	100	139	100	139	100

Table E6. Reading Grade 8 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	1	0.7	0	0	0	0	0	0	0	0
B-	6	4.3	0	0	0	0	3	2.1	0	0
A	122	87.1	0	0	0	0	38	27.1	0	0
B+	11	7.9	0	0	0	0	1	0.7	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL *N*	0	0	140	100	140	100	98	70	140	100
TOTAL	140	100	140	100	140	100	140	100	140	100

Mathematics

Table E7. Mathematics Grade 3 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	0	0	0	0
B-	2	1.7	0	0	0	0	4	3.3	0	0
A	114	95.0	0	0	0	0	67	55.8	0	0
B+	3	2.5	0	0	0	0	1	0.8	0	0
C+	1	0.8	0	0	0	0	0	0	0	0
SMALL *N*	0	0	120	100	120	100	48	40.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table E8. Mathematics Grade 4 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	0	0	0	0
B-	3	2.5	0	0	0	0	3	2.5	0	0
A	111	92.5	0	0	0	0	81	67.5	0	0
B+	6	5.0	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	36	30.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table E9. Mathematics Grade 5 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	1	0.8	0	0	0	0	0	0	0	0
B-	6	5.0	0	0	0	0	1	0.8	0	0
A	109	90.8	0	0	0	0	93	77.5	0	0
B+	4	3.3	0	0	0	0	2	1.7	0	0
C+	0	0	0	0	0	0	0	0	0	0
SMALL *N*	0	0	120	100	120	100	24	20.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table E10. Mathematics Grade 6 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	0	0	0	0
B-	6	5.0	0	0	0	0	1	0.8	0	0
A	110	91.7	0	0	0	0	47	39.2	0	0
B+	4	3.3	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	72	60.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table E11. Mathematics Grade 7 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	0	0	0	0
B-	5	4.2	0	0	0	0	0	0	0	0
A	111	92.5	0	0	0	0	24	20.0	0	0
B+	3	2.5	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	96	80.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table E12. Mathematics Grade 8 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
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	114	95	0	0	0	0	23	19.2	0	0
B+	2	1.7	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	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Science

Table E13. Science Grade 4 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	0	0	0	0
B-	9	7.5	0	0	0	0	2	1.7	0	0
A	110	91.7	0	0	0	0	106	88.3	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	12	10.0	120	100
TOTAL	120	100	120	100	120	100	120	100	120	100

Table E14. Science Grade 8 DIF Summary Statistics for Embedded Field Test Items

DIF Category	Male vs. Female		White vs. Asian		White vs. African American		White vs. Hispanic/Latino		White vs. Native American	
	N	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage	*N*	Percentage
C-	0	0	0	0	0	0	0	0	0	0
B-	3	2.5	0	0	0	0	1	0.8	0	0
A	112	93.3	0	0	0	0	11	9.2	0	0
B+	5	4.2	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

Appendix F: Classical Item Statistics for 2016 Field Test Items

Reading

Table F1. Reading Grade 3 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH345371	802	0.63	0.42
	VH345370	802	0.45	0.35
	VH345359	802	0.69	0.35
	VH345362	802	0.57	0.18
	VH345376	802	0.47	0.34
	VH345375	802	0.32	0.12
	VH343052	802	0.75	0.31
	VH343057	802	0.16	0.18
	VH343076	802	0.65	0.43
	VH343072	802	0.85	0.49
	VH343085	802	0.63	0.29
	VH343081	802	0.58	0.27
	VH345489	802	0.57	0.45
	VH345515	802	0.65	0.35
2	VH345372	774	0.64	0.37
	VH345373	774	0.56	0.29
	VH345365	774	0.56	0.44
	VH345367	774	0.65	0.31
	VH345378	774	0.47	0.38
	VH345377	774	0.24	0.12
	VH343064	774	0.69	0.26
	VH343063	774	0.73	0.50
	VH343078	774	0.56	0.26
	VH343080	774	0.71	0.52
	VH343087	774	0.29	0.18
	VH343088	774	0.45	0.29
	VH345525	774	0.84	0.38
	VH345520	774	0.47	0.15
3	VH331034	771	0.62	0.36
	VH331037	771	0.75	0.31
	VH331077	771	0.59	0.41
	VH331079	771	0.51	0.28
	VH331092	771	0.70	0.34
	VH331100	771	0.50	0.38
	VH324712	771	0.37	0.21
	VH324717	771	0.23	0.06
	VH324741	771	0.21	0.02
	VH324735	771	0.82	0.45
	VH324755	771	0.36	0.32
	VH324757	771	0.47	0.17
	VH345533	771	0.81	0.56
	VH345539	771	0.79	0.54

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
4	VH331036	768	0.36	0.14
	VH331040	768	0.74	0.39
	VH331076	768	0.47	0.21
	VH331073	768	0.60	0.36
	VH331086	768	0.50	0.33
	VH331088	768	0.51	0.54
	VH324713	768	0.88	0.43
	VH324715	768	0.36	0.34
	VH324726	768	0.39	0.15
	VH324745	768	0.36	0.27
	VH324760	768	0.74	0.46
	VH324754	768	0.38	0.20
	VH345612	768	0.80	0.44
	VH345617	768	0.59	0.45
5	VH346222	775	0.58	0.31
	VH346212	775	0.91	0.47
	VH346299	775	0.87	0.52
	VH346290	775	0.29	0.11
	VH346342	775	0.59	0.28
	VH346337	775	0.42	0.26
	VH345196	775	0.24	0.03
	VH345197	775	0.33	0.15
	VH345201	775	0.66	0.42
	VH345200	775	0.25	0.16
	VH345206	775	0.53	0.43
	VH345203	775	0.59	0.31
	VH345608	775	0.33	0.20
	VH345614	775	0.78	0.51
6	VH346227	757	0.40	0.24
	VH346215	757	0.36	0.34
	VH346286	757	0.21	-0.08
	VH346233	757	0.67	0.33
	VH346333	757	0.55	0.30
	VH346331	757	0.40	0.17
	VH345198	757	0.73	0.42
	VH345199	757	0.61	0.48
	VH345195	757	0.71	0.48
	VH345207	757	0.20	0.03
	VH345192	757	0.72	0.32
	VH345204	757	0.51	0.16
	VH345605	757	0.87	0.51
	VH345579	757	0.89	0.48

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
7	VH330263	759	0.59	0.19
	VH330269	759	0.80	0.43
	VH330268	759	0.54	0.16
	VH330266	759	0.54	0.35
	VH330258	759	0.52	0.35
	VH330271	759	0.32	0.22
	VH345380	759	0.72	0.59
	VH345379	759	0.56	0.36
	VH345385	759	0.55	0.24
	VH345384	759	0.73	0.52
	VH345389	759	0.59	0.33
	VH345392	759	0.64	0.38
	VH345545	759	0.80	0.55
	VH345535	759	0.84	0.54
8	VH330275	749	0.43	0.26
	VH330270	749	0.70	0.42
	VH330274	749	0.55	0.32
	VH330278	749	0.49	0.33
	VH330262	749	0.69	0.23
	VH330273	749	0.81	0.51
	VH345382	749	0.63	0.54
	VH345383	749	0.40	0.36
	VH345387	749	0.46	0.37
	VH345386	749	0.44	0.29
	VH345390	749	0.78	0.45
	VH345391	749	0.72	0.37
	VH345567	749	0.75	0.51
	VH345587	749	0.64	0.46
9	VH318252	760	0.76	0.45
	VH318247	760	0.91	0.50
	VH318255	760	0.62	0.30
	VH318271	760	0.76	0.42
	VH318230	760	0.56	0.28
	VH318222	760	0.85	0.48
	VH313589	760	0.81	0.40
	VH313601	760	0.60	0.38
	VH313596	760	0.75	0.38
	VH313587	760	0.68	0.37
	VH313593	760	0.43	0.25
	VH313614	760	0.93	0.45
	VH345600	760	0.92	0.47
	VH345595	760	0.74	0.49

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
10	VH318283	766	0.65	0.36
	VH318277	766	0.46	0.31
	VH318261	766	0.87	0.52
	VH318268	766	0.75	0.42
	VH318285	766	0.17	0.02
	VH318281	766	0.70	0.41
	VH313599	766	0.46	0.16
	VH313585	766	0.70	0.35
	VH313618	766	0.48	0.34
	VH313592	766	0.68	0.47
	VH313616	766	0.77	0.36
	VH313603	766	0.57	0.31
	VH345574	766	0.72	0.56
	VH345560	766	0.89	0.37

Table F2. Reading Grade 4 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH330664	780	0.62	0.39
	VH330647	780	0.68	0.43
	VH330691	780	0.31	0.20
	VH330801	780	0.70	0.50
	VH330845	780	0.47	0.40
	VH330848	780	0.81	0.50
	VH331827	780	0.77	0.43
	VH331824	780	0.75	0.42
	VH331854	780	0.61	0.38
	VH331846	780	0.58	0.27
	VH331861	780	0.59	0.33
	VH331863	780	0.43	0.28
	VH325008	780	0.66	0.37
	VH324983	780	0.56	0.46
2	VH330634	738	0.57	0.31
	VH330597	738	0.79	0.39
	VH330811	738	0.76	0.39
	VH330696	738	0.54	0.25
	VH330841	738	0.73	0.40
	VH330839	738	0.73	0.30
	VH331842	738	0.74	0.42
	VH331836	738	0.36	0.12
	VH331849	738	0.44	0.15
	VH331851	738	0.83	0.50
	VH331865	738	0.56	0.34
	VH331866	738	0.54	0.22
	VH324993	738	0.53	0.31
	VH325010	738	0.74	0.45
3	VH345656	743	0.53	0.29
	VH345659	743	0.61	0.31
	VH345665	743	0.75	0.36
	VH345666	743	0.78	0.40
	VH345671	743	0.77	0.29
	VH345674	743	0.67	0.48
	VH346003	743	0.66	0.43
	VH346001	743	0.65	0.41
	VH346029	743	0.66	0.51
	VH345999	743	0.38	0.06
	VH346002	743	0.36	0.16
	VH345996	743	0.53	0.26
	VH325048	743	0.91	0.44
	VH325062	743	0.60	0.35

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
4	VH345651	740	0.58	0.33
	VH345650	740	0.78	0.49
	VH345668	740	0.84	0.45
	VH345664	740	0.35	0.25
	VH345672	740	0.26	0.13
	VH345676	740	0.53	0.24
	VH346034	740	0.68	0.42
	VH346024	740	0.52	0.40
	VH346007	740	0.24	0.12
	VH346033	740	0.55	0.40
	VH346028	740	0.73	0.49
	VH346020	740	0.60	0.25
	VH325053	740	0.52	0.33
	VH325066	740	0.78	0.53
5	VH311932	741	0.44	0.23
	VH311948	741	0.95	0.25
	VH311937	741	0.57	0.35
	VH311969	741	0.55	0.30
	VH311966	741	0.40	0.15
	VH311901	741	0.87	0.42
	VH332787	741	0.35	0.11
	VH332764	741	0.72	0.38
	VH332833	741	0.54	0.16
	VH332844	741	0.19	-0.02
	VH332901	741	0.21	0.13
	VH332876	741	0.65	0.47
	VH325055	741	0.72	0.31
	VH325027	741	0.58	0.34
6	VH311955	745	0.62	0.31
	VH311934	745	0.51	0.22
	VH311943	745	0.69	0.45
	VH311941	745	0.78	0.45
	VH312255	745	0.59	0.45
	VH311911	745	0.68	0.35
	VH332770	745	0.62	0.21
	VH332780	745	0.52	0.31
	VH332824	745	0.68	0.41
	VH332812	745	0.69	0.27
	VH332909	745	0.55	0.44
	VH332866	745	0.72	0.27
	VH325041	745	0.72	0.39
	VH325073	745	0.93	0.40

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
7	VH346125	750	0.80	0.41
	VH346137	750	0.65	0.38
	VH346153	750	0.86	0.45
	VH346150	750	0.34	0.31
	VH346183	750	0.59	0.43
	VH346188	750	0.51	0.38
	VH346064	750	0.44	0.24
	VH346055	750	0.40	0.19
	VH346056	750	0.51	0.23
	VH346065	750	0.29	0.23
	VH346067	750	0.34	0.04
	VH346052	750	0.77	0.38
	VH325022	750	0.75	0.50
	VH325032	750	0.76	0.43
8	VH346120	750	0.60	0.41
	VH346134	750	0.64	0.40
	VH346149	750	0.77	0.48
	VH346155	750	0.51	0.39
	VH346185	750	0.64	0.28
	VH346184	750	0.46	0.28
	VH346062	750	0.66	0.43
	VH346059	750	0.48	0.10
	VH346061	750	0.59	0.26
	VH346068	750	0.64	0.15
	VH346073	750	0.47	0.29
	VH346043	750	0.53	0.12
	VH325019	750	0.65	0.49
	VH325003	750	0.90	0.41
9	VH311841	752	0.82	0.45
	VH311778	752	0.51	0.16
	VH311804	752	0.85	0.42
	VH311800	752	0.69	0.20
	VH311846	752	0.64	0.30
	VH311772	752	0.75	0.38
	VH331783	752	0.68	0.47
	VH331761	752	0.44	0.36
	VH331767	752	0.80	0.47
	VH331775	752	0.77	0.31
	VH331771	752	0.81	0.46
	VH331733	752	0.50	0.20
	VH325059	752	0.79	0.46
	VH325017	752	0.38	0.13

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
10	VH311791	743	0.33	0.08
	VH311786	743	0.79	0.42
	VH311812	743	0.40	0.22
	VH311833	743	0.65	0.30
	VH311818	743	0.49	0.18
	VH311822	743	0.70	0.43
	VH331745	743	0.86	0.45
	VH331785	743	0.69	0.38
	VH331778	743	0.48	0.29
	VH331790	743	0.43	0.19
	VH331772	743	0.55	0.33
	VH331792	743	0.58	0.29
	VH325023	743	0.81	0.50
	VH325039	743	0.40	0.23

Table F3. Reading Grade 5 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH345910	794	0.58	0.36
	VH345954	794	0.81	0.56
	VH346000	794	0.71	0.31
	VH345994	794	0.82	0.50
	VH346066	1520	0.78	0.44
	VH346070	794	0.75	0.51
	VH332285	794	0.82	0.20
	VH332273	794	0.57	0.15
	VH332267	794	0.74	0.37
	VH332306	794	0.46	0.45
	VH332280	794	0.70	0.47
	VH332248	794	0.88	0.50
	VH321425	794	0.43	0.41
	VH321429	794	0.57	0.26
2	VH346066	1520	0.78	0.44
	VH345940	726	0.59	0.27
	VH345974	726	0.55	0.48
	VH346031	726	0.77	0.38
	VH346006	726	0.68	0.32
	VH346075	726	0.55	0.42
	VH332303	726	0.79	0.29
	VH332288	726	0.85	0.50
	VH332292	726	0.82	0.30
	VH332298	726	0.33	0.14
	VH332312	726	0.85	0.49
	VH332296	726	0.67	0.36
	VH321434	726	0.91	0.42
	VH321777	726	0.94	0.48
3	VH345951	1437	0.38	0.24
	VH345929	720	0.80	0.42
	VH345933	720	0.68	0.44
	VH345943	720	0.22	0.06
	VH345966	720	0.43	0.18
	VH345914	720	0.62	0.26
	VH333338	720	0.60	0.08
	VH333335	720	0.51	0.35
	VH333563	720	0.85	0.42
	VH333527	720	0.97	0.32
	VH333569	720	0.85	0.43
	VH333574	720	0.85	0.34
	VH321628	720	0.61	0.28
	VH321690	720	0.92	0.29

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
4	VH345939	717	0.47	0.05
	VH345951	1437	0.38	0.24
	VH345973	717	0.60	0.44
	VH345947	717	0.62	0.45
	VH345963	717	0.75	0.41
	VH345958	717	0.45	0.24
	VH333347	717	0.52	0.20
	VH333344	717	0.36	-0.02
	VH333497	717	0.86	0.48
	VH333493	717	0.43	0.22
	VH333572	717	0.74	0.40
	VH333578	717	0.77	0.38
	VH321670	717	0.81	0.56
	VH321766	717	0.85	0.48
5	VH314088	1450	0.43	0.26
	VH332343	724	0.73	0.45
	VH332339	724	0.26	0.30
	VH332359	724	0.69	0.38
	VH332353	724	0.86	0.43
	VH332366	724	0.74	0.38
	VH332385	724	0.81	0.45
	VH314102	724	0.48	0.33
	VH314080	724	0.90	0.55
	VH314083	724	0.75	0.39
	VH314106	724	0.43	0.28
	VH314090	724	0.78	0.33
	VH321688	724	0.86	0.54
	VH321544	724	0.61	0.29
6	VH332335	726	0.89	0.46
	VH332336	726	0.47	0.12
	VH332345	726	0.73	0.45
	VH332347	726	0.60	0.26
	VH332387	726	0.89	0.40
	VH332378	726	0.78	0.31
	VH314085	726	0.66	0.39
	VH314104	726	0.39	0.18
	VH314078	726	0.69	0.42
	VH314101	726	0.75	0.43
	VH314088	1450	0.43	0.26
	VH314094	726	0.66	0.43
	VH321443	726	0.86	0.38
	VH321456	726	0.91	0.47

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
7	VH321302	727	0.47	0.22
	VH321311	727	0.57	0.32
	VH321372	727	0.55	0.30
	VH321375	727	0.69	0.44
	VH321393	727	0.56	0.25
	VH321384	727	0.76	0.38
	VH313647	727	0.81	0.41
	VH313648	727	0.86	0.35
	VH313644	727	0.53	0.20
	VH313652	727	0.68	0.29
	VH313650	727	0.76	0.27
	VH313637	727	0.51	0.23
	VH321439	727	0.88	0.53
	VH321468	727	0.80	0.32
8	VH321297	729	0.83	0.51
	VH321318	729	0.78	0.47
	VH321367	729	0.88	0.39
	VH321346	729	0.91	0.41
	VH321380	729	0.70	0.33
	VH321386	729	0.59	0.32
	VH313641	729	0.57	0.09
	VH313645	729	0.87	0.34
	VH313643	729	0.50	0.26
	VH313649	729	0.42	0.23
	VH313640	729	0.67	0.31
	VH313638	729	0.81	0.32
	VH321609	729	0.85	0.54
	VH321656	729	0.72	0.34
9	VH321130	728	0.48	0.34
	VH321077	728	0.71	0.22
	VH321089	728	0.53	0.23
	VH321132	728	0.63	0.38
	VH321062	728	0.55	0.45
	VH321068	728	0.81	0.52
	VH333040	728	0.87	0.39
	VH332958	728	0.66	0.32
	VH333069	728	0.66	0.38
	VH333033	728	0.92	0.35
	VH332986	728	0.78	0.49
	VH332955	728	0.61	0.28
	VH321663	728	0.94	0.45
	VH321686	728	0.85	0.57

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
10	VH321096	732	0.48	0.21
	VH321082	732	0.64	0.44
	VH321123	732	0.56	0.44
	VH321133	732	0.51	0.23
	VH321128	732	0.67	0.34
	VH321072	732	0.32	0.27
	VH332960	732	0.64	0.41
	VH333072	732	0.55	0.31
	VH332994	732	0.77	0.31
	VH333038	732	0.86	0.30
	VH333083	732	0.36	0.27
	VH333056	732	0.79	0.38
	VH321471	732	0.76	0.40
	VH321676	732	0.86	0.41

Table F4. Reading Grade 6 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH315111	781	0.76	0.29
	VH315264	781	0.60	0.45
	VH315235	781	0.80	0.37
	VH315267	781	0.46	0.24
	VH315118	781	0.78	0.41
	VH315255	781	0.82	0.47
	VH333639	781	0.60	0.31
	VH333641	781	0.64	0.42
	VH333670	781	0.41	0.31
	VH333667	781	0.55	0.37
	VH333689	781	0.50	0.41
	VH333685	781	0.37	0.17
	VH315975	781	0.76	0.40
	VH315969	781	0.80	0.44
2	VH315258	677	0.46	0.30
	VH315243	677	0.91	0.36
	VH315247	677	0.71	0.48
	VH315124	677	0.72	0.29
	VH315256	677	0.82	0.34
	VH315090	677	0.71	0.32
	VH333622	677	0.47	0.38
	VH333607	677	0.69	0.38
	VH333674	677	0.40	0.30
	VH333657	677	0.35	0.17
	VH333679	677	0.53	0.39
	VH333693	677	0.60	0.43
	VH315492	677	0.37	0.33
	VH315438	677	0.60	0.34
3	VH325342	678	0.74	0.36
	VH325365	678	0.37	0.15
	VH325329	678	0.61	0.27
	VH325384	678	0.45	0.21
	VH325321	678	0.62	0.43
	VH325304	678	0.91	0.31
	VH334475	678	0.52	0.34
	VH334630	678	0.64	0.36
	VH334617	678	0.89	0.29
	VH334624	678	0.74	0.41
	VH334389	678	0.69	0.54
	VH334393	678	0.90	0.38
	VH315446	678	0.58	0.42
	VH315409	678	0.68	0.38

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
4	VH325336	699	0.80	0.35
	VH325360	699	0.61	0.32
	VH325368	699	0.77	0.40
	VH325300	699	0.79	0.31
	VH325308	699	0.48	0.35
	VH325318	699	0.64	0.38
	VH334398	699	0.67	0.51
	VH334395	699	0.85	0.46
	VH334465	699	0.64	0.37
	VH334480	699	0.88	0.43
	VH334333	699	0.52	0.30
	VH334518	699	0.91	0.38
	VH315458	699	0.74	0.40
	VH315977	699	0.79	0.42
5	VH315342	1390	0.69	0.42
	VH315337	1390	0.51	0.43
	VH332436	1390	0.78	0.40
	VH315340	705	0.73	0.45
	VH315344	705	0.55	0.27
	VH315338	705	0.66	0.50
	VH315332	705	0.80	0.53
	VH332428	705	0.79	0.43
	VH332505	705	0.54	0.43
	VH332506	705	0.49	0.21
	VH332419	705	0.76	0.45
	VH332442	705	0.87	0.35
	VH315467	705	0.68	0.34
	VH315429	705	0.49	0.01
6	VH315342	1390	0.69	0.42
	VH315339	685	0.80	0.23
	VH315337	1390	0.51	0.43
	VH315341	685	0.86	0.33
	VH315333	685	0.48	-0.01
	VH315331	685	0.46	0.26
	VH332436	1390	0.78	0.40
	VH332430	685	0.50	0.33
	VH332500	685	0.57	0.40
	VH332518	685	0.58	0.48
	VH332413	685	0.61	0.53
	VH332423	685	0.56	0.33
	VH315964	685	0.48	0.44
	VH315400	685	0.82	0.56

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
7	VH318937	689	0.42	0.39
	VH318930	689	0.60	0.41
	VH319047	689	0.50	0.31
	VH318974	689	0.69	0.41
	VH319070	689	0.85	0.49
	VH319062	689	0.61	0.44
	VH315315	689	0.74	0.51
	VH315307	689	0.60	0.36
	VH315304	689	0.54	0.43
	VH315313	689	0.68	0.47
	VH315323	689	0.56	0.46
	VH315325	689	0.64	0.35
	VH315967	689	0.65	0.47
	VH315960	689	0.45	0.41
8	VH318912	699	0.88	0.48
	VH318924	699	0.79	0.50
	VH319038	699	0.43	0.15
	VH318954	699	0.66	0.35
	VH319079	699	0.75	0.37
	VH319083	699	0.90	0.39
	VH315310	699	0.65	0.40
	VH315308	699	0.42	0.25
	VH315327	699	0.83	0.51
	VH315318	699	0.58	0.35
	VH315299	699	0.89	0.28
	VH315297	699	0.61	0.38
	VH315389	699	0.53	0.35
	VH315518	699	0.75	0.45
9	VH332497	1400	0.81	0.46
	VH332485	700	0.78	0.45
	VH332558	700	0.22	0.04
	VH332542	700	0.55	0.31
	VH332502	700	0.62	0.40
	VH332530	700	0.80	0.56
	VH333943	700	0.60	0.41
	VH334052	700	0.28	0.15
	VH334048	700	0.75	0.34
	VH333809	700	0.44	0.20
	VH333802	700	0.79	0.39
	VH334043	700	0.83	0.45
	VH315499	700	0.92	0.41
	VH315534	700	0.76	0.28

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
10	VH332536	700	0.39	0.20
	VH332513	700	0.74	0.45
	VH332550	700	0.80	0.53
	VH332497	1400	0.81	0.46
	VH332517	700	0.35	0.09
	VH332507	700	0.42	0.24
	VH333816	700	0.55	0.40
	VH333821	700	0.76	0.46
	VH333930	700	0.44	0.36
	VH334058	700	0.65	0.27
	VH333791	700	0.61	0.38
	VH333807	700	0.60	0.34
	VH315510	700	0.60	0.35
	VH315415	700	0.76	0.49

Table F5. Reading Grade 7 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH345749	782	0.52	0.39
	VH345746	782	0.60	0.55
	VH345769	782	0.52	0.33
	VH345766	782	0.49	0.26
	VH345774	782	0.71	0.43
	VH345780	782	0.46	0.24
	VH345485	782	0.61	0.10
	VH345516	782	0.79	0.42
	VH345440	782	0.44	0.22
	VH345538	782	0.77	0.49
	VH345519	1466	0.71	0.31
	VH345564	782	0.63	0.29
	VH316374	782	0.77	0.37
	VH316736	782	0.32	0.24
2	VH345519	1466	0.71	0.31
	VH345750	684	0.48	0.25
	VH345753	684	0.27	0.14
	VH345764	684	0.44	0.33
	VH345771	684	0.70	0.42
	VH345778	684	0.57	0.31
	VH345775	684	0.30	0.02
	VH345493	684	0.85	0.27
	VH345528	684	0.54	0.23
	VH345518	684	0.64	0.38
	VH345544	684	0.43	0.26
	VH345546	684	0.45	0.16
	VH316391	684	0.82	0.36
	VH316770	684	0.41	0.41
3	VH332947	703	0.56	0.39
	VH332935	703	0.74	0.40
	VH332922	703	0.61	0.47
	VH332809	703	0.88	0.36
	VH332915	703	0.86	0.36
	VH332891	703	0.85	0.19
	VH332641	703	0.73	0.26
	VH332639	703	0.75	0.30
	VH332676	703	0.63	0.38
	VH332617	703	0.78	0.31
	VH332647	703	0.84	0.41
	VH332663	703	0.86	0.40
	VH316266	703	0.76	0.52
	VH316281	703	0.71	0.51

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
4	VH332910	699	0.92	0.33
	VH332940	699	0.58	0.23
	VH332943	699	0.53	0.36
	VH332929	699	0.66	0.46
	VH332846	699	0.75	0.48
	VH332799	699	0.72	0.46
	VH332602	699	0.83	0.44
	VH332599	699	0.86	0.25
	VH332637	699	0.58	0.32
	VH332672	699	0.77	0.36
	VH332593	699	0.74	0.18
	VH332671	699	0.60	0.15
	VH316274	699	0.74	0.37
	VH316403	699	0.32	0.24
5	VH333944	705	0.80	0.42
	VH333950	705	0.78	0.39
	VH333917	705	0.49	0.18
	VH333931	705	0.85	0.51
	VH333961	705	0.46	0.31
	VH333908	705	0.56	0.38
	VH345770	705	0.70	0.40
	VH345765	705	0.53	0.15
	VH345895	705	0.29	0.11
	VH345777	705	0.71	0.47
	VH345913	705	0.36	0.08
	VH345904	705	0.79	0.39
	VH316398	705	0.73	0.43
	VH316457	705	0.82	0.48
6	VH333955	681	0.65	0.43
	VH333946	681	0.69	0.29
	VH333941	681	0.69	0.43
	VH333956	681	0.53	0.42
	VH333921	681	0.56	0.38
	VH333878	681	0.69	0.29
	VH345758	681	0.38	0.10
	VH345756	681	0.76	0.40
	VH345883	681	0.60	0.19
	VH345779	681	0.50	0.46
	VH345945	681	0.54	0.21
	VH348029	681	0.50	0.31
	VH316428	681	0.79	0.58
	VH316781	681	0.46	0.24

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
7	VH333919	709	0.64	0.11
	VH333959	709	0.60	0.16
	VH333954	709	0.57	0.34
	VH333965	709	0.60	0.37
	VH333880	709	0.33	0.04
	VH333859	709	0.47	0.44
	VH344777	709	0.69	0.17
	VH344780	709	0.76	0.43
	VH344754	709	0.65	0.40
	VH344763	709	0.61	0.41
	VH344765	709	0.82	0.35
	VH344749	709	0.66	0.43
	VH316423	709	0.74	0.45
	VH316807	709	0.66	0.45
8	VH333925	692	0.75	0.38
	VH333947	692	0.69	0.20
	VH333957	692	0.61	0.40
	VH333891	692	0.92	0.37
	VH333964	692	0.85	0.41
	VH333847	692	0.58	0.37
	VH344770	692	0.72	0.23
	VH344752	692	0.62	0.36
	VH344764	692	0.68	0.36
	VH344775	692	0.45	0.27
	VH344760	692	0.48	0.22
	VH344746	692	0.34	0.27
	VH316751	692	0.65	0.39
	VH316466	692	0.65	0.42
9	VH334238	697	0.48	0.34
	VH334213	697	0.45	0.19
	VH334227	697	0.63	0.34
	VH334200	697	0.76	0.52
	VH334204	697	0.43	0.27
	VH334199	697	0.78	0.41
	VH332675	697	0.64	0.37
	VH332634	697	0.55	0.32
	VH332636	697	0.54	0.37
	VH332631	697	0.57	0.37
	VH332622	697	0.49	0.29
	VH332613	697	0.71	0.52
	VH316459	697	0.54	0.22
	VH316687	697	0.48	0.30

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
10	VH334205	699	0.34	0.30
	VH334219	699	0.58	0.39
	VH334209	699	0.32	0.01
	VH334223	699	0.58	0.30
	VH334242	699	0.34	0.37
	VH334216	699	0.48	0.16
	VH332626	699	0.44	0.27
	VH332673	699	0.66	0.51
	VH332677	699	0.41	0.37
	VH332618	699	0.71	0.42
	VH332628	699	0.46	0.30
	VH332678	699	0.60	0.48
	VH316709	699	0.38	0.42
	VH316467	699	0.54	0.39

Table F6. Reading Grade 8 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH345147	779	0.80	0.48
	VH345177	779	0.57	0.26
	VH345151	779	0.64	0.42
	VH345172	779	0.83	0.39
	VH345124	779	0.67	0.46
	VH345121	779	0.60	0.45
	VH342979	779	0.70	0.45
	VH342994	779	0.23	-0.05
	VH342980	779	0.48	0.28
	VH343009	779	0.52	0.25
	VH342990	779	0.66	0.41
	VH342964	779	0.69	0.42
	VH325149	779	0.73	0.50
	VH325228	779	0.33	0.00
2	VH345141	660	0.41	0.22
	VH345135	660	0.77	0.49
	VH345170	660	0.90	0.46
	VH345181	660	0.78	0.52
	VH345175	660	0.82	0.44
	VH345154	660	0.60	0.35
	VH342968	660	0.84	0.32
	VH342983	660	0.69	0.36
	VH342971	660	0.40	0.04
	VH343003	660	0.67	0.42
	VH342981	660	0.50	0.25
	VH343004	660	0.75	0.42
	VH325227	660	0.39	0.34
	VH325157	660	0.55	0.24
3	VH343387	666	0.47	0.13
	VH343380	666	0.76	0.45
	VH343529	666	0.50	0.34
	VH343514	666	0.63	0.45
	VH350931	666	0.30	0.04
	VH343550	666	0.50	0.24
	VH343222	666	0.86	0.53
	VH343231	666	0.50	0.32
	VH343179	666	0.53	0.32
	VH343239	666	0.51	0.40
	VH343219	666	0.73	0.45
	VH343082	666	0.50	0.14
	VH325233	666	0.69	0.39
	VH325123	666	0.50	0.30

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
4	VH343422	675	0.41	0.32
	VH343313	675	0.72	0.31
	VH343521	675	0.39	-0.04
	VH343503	675	0.72	0.51
	VH343545	675	0.43	0.33
	VH343538	675	0.45	0.25
	VH343053	675	0.36	0.33
	VH343235	675	0.57	0.20
	VH343073	675	0.49	0.32
	VH343148	675	0.68	0.48
	VH343047	675	0.65	0.35
	VH342960	675	0.89	0.31
	VH325091	675	0.66	0.29
	VH325218	675	0.85	0.34
5	VH326366	679	0.83	0.36
	VH326386	679	0.61	0.36
	VH326377	679	0.51	0.31
	VH326381	679	0.76	0.29
	VH326380	679	0.54	0.47
	VH326372	679	0.73	0.42
	VH343939	679	0.29	0.25
	VH343960	679	0.57	0.21
	VH344147	679	0.44	0.29
	VH344131	679	0.51	0.27
	VH344167	679	0.35	-0.09
	VH344176	679	0.81	0.44
	VH325164	679	0.51	0.27
	VH325144	679	0.72	0.24
6	VH326376	656	0.45	0.17
	VH326384	656	0.64	0.48
	VH326379	656	0.67	0.39
	VH326388	656	0.54	0.32
	VH326374	656	0.65	0.25
	VH326361	656	0.86	0.48
	VH343950	656	0.22	0.06
	VH343876	656	0.60	0.44
	VH344092	656	0.59	0.35
	VH344067	656	0.57	0.31
	VH344180	656	0.39	0.26
	VH344153	656	0.37	0.13
	VH325199	656	0.79	0.53
	VH325171	656	0.66	0.38

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
7	VH345018	653	0.34	0.16
	VH345022	653	0.79	0.38
	VH345004	653	0.85	0.50
	VH345012	653	0.83	0.52
	VH345026	653	0.64	0.11
	VH345028	653	0.89	0.46
	VH343135	653	0.68	0.27
	VH343132	653	0.45	0.34
	VH343191	653	0.74	0.36
	VH343196	653	0.68	0.40
	VH343126	653	0.85	0.47
	VH343129	653	0.72	0.03
	VH325186	653	0.35	0.31
	VH325211	653	0.57	0.26
8	VH344994	679	0.87	0.35
	VH344998	679	0.84	0.34
	VH345031	679	0.71	0.24
	VH345020	679	0.85	0.42
	VH345016	679	0.59	0.35
	VH344978	679	0.58	0.10
	VH343193	679	0.58	0.34
	VH343184	679	0.62	0.31
	VH343176	679	0.70	0.36
	VH343186	679	0.46	0.20
	VH343124	679	0.60	0.29
	VH343115	679	0.77	0.38
	VH325231	679	0.32	0.31
	VH325195	679	0.69	0.41
9	VH345688	667	0.66	0.49
	VH345726	667	0.93	0.36
	VH345730	667	0.62	0.38
	VH345706	667	0.50	0.29
	VH345694	667	0.80	0.30
	VH345734	667	0.59	0.46
	VH343247	667	0.61	0.39
	VH343240	667	0.50	0.36
	VH343260	667	0.65	0.43
	VH343230	667	0.77	0.44
	VH343234	667	0.63	0.35
	VH343263	667	0.62	0.18
	VH325216	667	0.20	-0.06
	VH325130	667	0.82	0.46

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
10	VH345686	679	0.35	0.23
	VH345723	679	0.63	0.24
	VH345735	679	0.48	0.26
	VH345718	679	0.22	0.07
	VH345703	679	0.33	0.17
	VH345733	679	0.42	0.28
	VH343254	679	0.73	0.33
	VH343257	679	0.90	0.43
	VH343232	679	0.67	0.29
	VH343250	679	0.75	0.47
	VH343237	679	0.67	0.35
	VH343224	679	0.63	0.07
	VH325220	679	0.58	0.34
	VH325191	679	0.73	0.42

Mathematics

Table F7. Mathematics Grade 3 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH327773	836	0.39	0.39
	VH331639	836	0.18	0.19
	VH327891	836	0.43	0.43
	VH326762	836	0.44	0.46
	VH326729	836	0.50	0.38
	VH337124	836	0.59	0.28
	VH327898	836	0.51	0.39
	VH318564	836	0.37	0.38
	VH336404	836	0.25	0.01
	VH318704	836	0.41	0.30
	VH303943	836	0.73	0.54
	VH303635	836	0.75	0.35
2	VH318837	775	0.39	0.36
	VH303570	775	0.47	0.11
	VH327878	775	0.38	0.23
	VH336409	775	0.87	0.30
	VH326721	775	0.45	0.26
	VH336574	775	0.27	0.16
	VH273044	775	0.31	0.21
	VH331638	775	0.72	0.37
	VH303611	775	0.58	0.44
	VH273115	775	0.50	0.33
	VH327000	775	0.17	0.26
	VH303869	775	0.46	0.30
3	VH318818	783	0.61	0.49
	VH310953	783	0.15	0.23
	VH303648	783	0.74	0.40
	VH326759	783	0.42	0.33
	VH310966	783	0.47	0.34
	VH326993	783	0.47	0.36
	VH327895	783	0.25	0.19
	VH318584	783	0.57	0.41
	VH303879	783	0.29	0.07
	VH303664	783	0.43	0.38
	VH310884	783	0.56	0.32
	VH326786	783	0.36	0.06
4	VH273074	776	0.61	0.28
	VH336532	776	0.26	-0.09
	VH327797	776	0.40	0.28
	VH303523	776	0.47	0.35
	VH331640	776	0.43	0.38
	VH327871	776	0.64	0.45
	VH326997	776	0.67	0.13
	VH336381	776	0.58	0.10
	VH273083	776	0.33	0.09
	VH318718	776	0.56	0.48
	VH304079	776	0.33	0.20
	VH310932	776	0.40	0.35

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
5	VH303507	769	0.76	0.51
	VH310964	769	0.23	0.34
	VH327801	769	0.63	0.32
	VH335375	769	0.45	0.07
	VH310863	769	0.45	0.36
	VH318631	769	0.59	0.51
	VH327913	769	0.40	0.29
	VH327886	769	0.46	0.31
	VH336373	769	0.26	0.12
	VH327903	769	0.56	0.54
	VH326797	769	0.48	0.28
	VH303698	769	0.40	0.14
6	VH327839	773	0.40	0.38
	VH310911	773	0.21	0.04
	VH327815	773	0.52	0.40
	VH336427	773	0.63	0.22
	VH327922	773	0.43	0.29
	VH331717	773	0.63	0.46
	VH326974	773	0.31	0.53
	VH318645	773	0.54	0.55
	VH303657	773	0.59	0.31
	VH273148	773	0.64	0.46
	VH326790	773	0.77	0.31
	VH310857	773	0.22	0.17
7	VH327819	757	0.74	0.51
	VH310958	757	0.18	0.07
	VH303596	757	0.22	0.02
	VH335365	757	0.82	0.32
	VH326990	757	0.53	0.52
	VH273027	757	0.36	0.29
	VH303935	757	0.22	0.38
	VH273095	757	0.24	0.07
	VH310949	757	0.90	0.22
	VH326794	757	0.39	0.18
	VH303950	757	0.69	0.57
	VH336376	757	0.37	0.30
8	VH273033	737	0.19	0.23
	VH303643	737	0.58	0.54
	VH327923	737	0.57	0.48
	VH303560	737	0.41	0.18
	VH318657	737	0.43	0.40
	VH272617	737	0.20	0.23
	VH310937	737	0.55	0.41
	VH272624	737	0.70	0.48
	VH326761	737	0.22	0.20
	VH273103	737	0.49	0.36
	VH303940	737	0.67	0.46
	VH304127	737	0.50	0.19

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
9	VH318768	741	0.40	0.34
	VH326465	741	0.32	0.31
	VH303624	741	0.66	0.45
	VH335356	741	0.30	0.15
	VH318803	741	0.59	0.32
	VH303949	741	0.77	0.45
	VH272510	741	0.16	0.35
	VH318734	741	0.64	0.38
	VH303681	741	0.55	0.22
	VH318609	741	0.62	0.44
	VH326964	741	0.30	0.25
	VH310934	741	0.83	0.42
10	VH318777	735	0.53	0.41
	VH303651	735	0.71	0.46
	VH303554	735	0.47	0.28
	VH335372	735	0.64	0.14
	VH326984	735	0.61	0.42
	VH318334	735	0.92	0.33
	VH272517	735	0.68	0.42
	VH310928	735	0.62	0.25
	VH273136	735	0.55	0.50
	VH273090	735	0.36	-0.13
	VH303946	735	0.69	0.44
	VH303629	735	0.64	0.50

Table F8. Mathematics Grade 4 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH327010	822	0.42	0.48
	VH331635	822	0.35	0.25
	VH313887	822	0.25	0.18
	VH305767	822	0.43	0.27
	VH305882	822	0.54	0.42
	VH313961	822	0.38	0.31
	VH326696	822	0.31	0.26
	VH327015	822	0.22	0.18
	VH335023	822	0.31	0.11
	VH326687	822	0.40	0.33
	VH329506	822	0.50	0.48
	VH317610	822	0.63	0.49
2	VH305896	754	0.45	0.54
	VH305749	754	0.08	0.00
	VH313944	754	0.42	0.31
	VH331636	754	0.70	0.42
	VH334941	754	0.41	0.35
	VH334950	754	0.39	0.29
	VH326661	754	0.63	0.09
	VH317460	754	0.69	0.46
	VH329325	754	0.70	0.49
	VH317521	754	0.58	0.21
	VH306121	754	0.24	0.05
	VH326676	754	0.41	0.34
3	VH335221	747	0.69	0.30
	VH331152	747	0.27	0.14
	VH313929	747	0.49	0.37
	VH327013	747	0.37	0.43
	VH306097	747	0.62	0.53
	VH313822	747	0.44	0.43
	VH326691	747	0.38	0.18
	VH317762	747	0.13	0.22
	VH313766	747	0.51	0.33
	VH319829	747	0.65	0.45
	VH335019	747	0.26	-0.02
	VH317589	747	0.64	0.48
4	VH331156	742	0.34	0.19
	VH326717	742	0.23	0.29
	VH326999	742	0.28	0.34
	VH313814	742	0.23	0.12
	VH305890	742	0.46	0.26
	VH326671	742	0.49	0.24
	VH327038	742	0.28	0.20
	VH317816	742	0.56	0.50
	VH305918	742	0.43	0.27
	VH313880	742	0.32	0.33
	VH319812	742	0.50	0.29
	VH317601	742	0.51	0.51

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
5	VH313730	735	0.30	0.17
	VH313920	735	0.46	0.39
	VH331634	735	0.35	0.41
	VH334637	735	0.31	0.13
	VH317504	735	0.73	0.51
	VH317471	735	0.35	0.30
	VH326668	735	0.19	0.23
	VH317809	735	0.43	0.21
	VH306130	735	0.56	0.30
	VH329260	735	0.53	0.58
	VH334982	735	0.54	0.29
	VH313889	735	0.54	0.28
6	VH334652	740	0.72	0.33
	VH326714	740	0.49	0.35
	VH305785	740	0.60	0.33
	VH327006	740	0.42	0.47
	VH305906	740	0.65	0.47
	VH317694	740	0.21	0.35
	VH327067	740	0.49	0.48
	VH317724	740	0.55	0.53
	VH306127	740	0.70	0.51
	VH317577	740	0.66	0.49
	VH329431	740	0.89	0.40
	VH313898	740	0.63	0.35
7	VH313955	746	0.35	0.29
	VH305760	746	0.51	0.40
	VH326711	746	0.34	0.23
	VH327002	746	0.79	0.40
	VH317685	746	0.51	0.45
	VH317539	746	0.78	0.41
	VH326680	746	0.34	0.18
	VH327022	746	0.50	0.23
	VH319890	746	0.91	0.29
	VH327046	746	0.52	0.44
	VH313751	746	0.62	0.47
	VH317629	746	0.69	0.40
8	VH331637	747	0.64	0.40
	VH326998	747	0.55	0.40
	VH326709	747	0.36	0.44
	VH335041	747	0.20	0.09
	VH305790	747	0.37	0.44
	VH317703	747	0.69	0.37
	VH317782	747	0.47	0.58
	VH327054	747	0.44	0.47
	VH313740	747	0.51	0.35
	VH306155	747	0.29	0.27
	VH317741	747	0.54	0.38
	VH317801	747	0.29	0.37

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
9	VH313837	737	0.34	0.12
	VH331160	737	0.41	0.17
	VH313967	737	0.30	0.25
	VH335038	737	0.44	0.24
	VH327062	737	0.40	0.46
	VH305909	737	0.44	0.41
	VH317710	737	0.54	0.46
	VH319860	737	0.80	0.52
	VH306115	737	0.66	0.53
	VH335031	737	0.32	0.20
	VH319897	737	0.87	0.36
	VH317772	737	0.51	0.31
10	VH335034	715	0.49	0.18
	VH313801	715	0.24	0.12
	VH305755	715	0.54	0.47
	VH331141	715	0.31	0.16
	VH327030	715	0.47	0.35
	VH317690	715	0.69	0.42
	VH317513	715	0.49	0.38
	VH305784	715	0.25	0.07
	VH306147	715	0.30	0.09
	VH329239	715	0.87	0.35
	VH317729	715	0.40	0.32
	VH314029	715	0.28	0.18

Table F9. Mathematics Grade 5 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH319261	815	0.29	0.22
	VH335026	815	0.35	0.17
	VH332183	815	0.51	0.18
	VH333962	815	0.33	0.36
	VH314306	815	0.23	0.22
	VH326916	815	0.54	0.48
	VH336215	815	0.22	0.14
	VH326780	815	0.57	0.51
	VH334660	815	0.33	0.26
	VH319641	815	0.47	0.28
	VH092639	815	0.65	0.48
	VH335003	815	0.15	0.05
2	VH326904	728	0.87	0.29
	VH336104	728	0.48	0.25
	VH332140	728	0.63	0.51
	VH317267	728	0.45	0.47
	VH334657	728	0.27	0.34
	VH334292	728	0.45	0.33
	VH314260	728	0.77	0.24
	VH319556	728	0.50	0.54
	VH326751	728	0.82	0.34
	VH334986	728	0.35	0.10
	VH332121	728	0.47	0.42
	VH314191	728	0.56	0.47
3	VH317107	713	0.40	0.33
	VH334073	713	0.24	0.14
	VH335036	713	0.43	0.27
	VH334937	713	0.52	0.48
	VH314297	713	0.74	0.31
	VH314361	713	0.34	0.16
	VH317292	713	0.47	0.27
	VH326804	713	0.20	-0.03
	VH095032	713	0.53	0.47
	VH319663	713	0.44	0.18
	VH336288	713	0.17	0.26
	VH334942	713	0.36	0.29
4	VH314353	712	0.41	0.27
	VH092839	712	0.52	0.27
	VH334324	712	0.21	0.10
	VH326935	712	0.18	0.13
	VH319572	712	0.73	0.28
	VH326775	712	0.80	0.47
	VH317339	712	0.50	0.24
	VH336229	712	0.50	0.42
	VH332193	712	0.47	0.33
	VH336353	712	0.40	0.14
	VH335051	712	0.56	0.39
	VH333915	712	0.32	0.28

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
5	VH332144	717	0.37	0.31
	VH319266	717	0.36	0.27
	VH317013	717	0.81	0.30
	VH334396	717	0.27	0.16
	VH326943	717	0.46	0.32
	VH314337	717	0.27	0.17
	VH319583	717	0.59	0.35
	VH335048	717	0.67	0.37
	VH333922	717	0.11	-0.07
	VH336339	717	0.30	0.33
	VH326796	717	0.53	0.43
	VH314178	717	0.71	0.36
6	VH332143	746	0.58	0.60
	VH319433	746	0.23	0.13
	VH334079	746	0.22	0.07
	VH314214	746	0.76	0.33
	VH314320	746	0.29	0.06
	VH326965	746	0.56	0.33
	VH317122	746	0.52	0.39
	VH326785	746	0.33	0.18
	VH336088	746	0.34	0.34
	VH319630	746	0.49	0.27
	VH336271	746	0.25	0.41
	VH335044	746	0.37	0.40
7	VH332139	730	0.51	0.54
	VH317020	730	0.58	0.41
	VH326987	730	0.49	0.19
	VH333949	730	0.30	0.26
	VH314314	730	0.25	0.03
	VH326748	730	0.60	0.31
	VH335040	730	0.57	0.30
	VH319654	730	0.23	0.36
	VH317145	730	0.53	0.50
	VH326811	730	0.26	0.22
	VH099656	730	0.49	0.34
	VH334282	730	0.25	0.32
8	VH334959	723	0.23	0.28
	VH319473	723	0.45	0.11
	VH317046	723	0.23	0.19
	VH334354	723	0.23	0.40
	VH314286	723	0.32	0.21
	VH336361	723	0.53	0.25
	VH317322	723	0.33	0.24
	VH332141	723	0.44	0.40
	VH334313	723	0.50	0.33
	VH319532	723	0.47	0.37
	VH333894	723	0.50	0.25
	VH326813	723	0.07	0.15

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
9	VH334973	722	0.61	0.33
	VH317141	722	0.69	0.42
	VH334626	722	0.19	0.39
	VH314274	722	0.59	0.42
	VH332207	722	0.74	0.47
	VH326774	722	0.67	0.50
	VH317303	722	0.33	0.35
	VH104496	722	0.31	0.07
	VH332126	722	0.55	0.57
	VH319545	722	0.20	0.31
	VH326899	722	0.67	0.41
	VH314204	722	0.33	0.40
10	VH334348	719	0.36	0.32
	VH335010	719	0.66	0.48
	VH317138	719	0.36	0.36
	VH314281	719	0.53	0.41
	VH334373	719	0.26	0.15
	VH335006	719	0.34	0.33
	VH336248	719	0.18	0.18
	VH326722	719	0.71	0.35
	VH333824	719	0.68	0.07
	VH326814	719	0.67	0.46
	VH332138	719	0.50	0.52
	VH319509	719	0.35	0.14

Table F10. Mathematics Grade 6 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH135854	825	0.43	0.19
	VF862767	825	0.64	0.37
	VH316009	825	0.31	0.31
	VF882940	825	0.53	0.43
	VH331472	825	0.38	0.25
	VH329926	825	0.76	0.39
	VH333341	825	0.46	0.08
	VH314512	825	0.49	0.48
	VH331938	825	0.34	0.23
	VH329962	825	0.31	0.37
	VH330025	825	0.24	0.01
	VH332000	825	0.41	0.25
2	VH316001	683	0.61	0.48
	VH319189	683	0.46	0.42
	VH320665	683	0.54	0.38
	VH326082	683	0.53	0.37
	VH333594	683	0.42	0.49
	VH314593	683	0.32	0.21
	VH329993	683	0.59	0.50
	VH330017	683	0.56	0.33
	VH136338	683	0.54	0.26
	VH331623	683	0.59	0.38
	VH314476	683	0.33	0.41
	VH329597	683	0.42	0.48
3	VH320741	684	0.81	0.38
	VH316017	684	0.54	0.46
	VH325587	684	0.43	0.47
	VH319122	684	0.53	0.58
	VH330042	684	0.40	0.11
	VH331487	684	0.53	0.37
	VH332053	684	0.21	0.17
	VH331630	684	0.51	0.06
	VH329607	684	0.65	0.36
	VH331620	684	0.50	0.40
	VH331633	684	0.37	0.25
	VH314495	684	0.44	0.41
4	VH326139	681	0.51	0.21
	VH319131	681	0.43	0.47
	VH316021	681	0.19	-0.17
	VF741705	681	0.70	0.35
	VH329539	681	0.67	0.48
	VH314392	681	0.17	0.24
	VH332012	681	0.48	0.33
	VH331627	681	0.24	0.03
	VH331936	681	0.20	-0.12
	VH329981	681	0.57	0.39
	VH314542	681	0.39	0.15
	VH332001	681	0.21	0.13

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
5	VH325593	688	0.19	0.19
	VH319203	688	0.60	0.21
	VF741703	688	0.46	0.27
	VH325591	688	0.40	0.27
	VH331624	688	0.07	-0.01
	VH314559	688	0.34	0.14
	VH331547	688	0.32	0.28
	VH331985	688	0.41	0.33
	VH331946	688	0.37	0.31
	VH331521	688	0.39	0.21
	VH329584	688	0.49	0.41
	VH314487	688	0.36	0.33
6	VH121079	702	0.73	0.16
	VH118695	702	0.39	0.30
	VH326102	702	0.35	0.19
	VH320813	702	0.40	0.37
	VH314584	702	0.52	0.42
	VH329936	702	0.33	0.22
	VH331988	702	0.30	0.27
	VH331980	702	0.37	0.33
	VH331626	702	0.32	0.20
	VH333348	702	0.16	0.18
	VH314517	702	0.44	0.23
	VH330051	702	0.43	0.39
7	VH326121	673	0.30	0.18
	VH319180	673	0.42	0.35
	VH325579	673	0.50	0.32
	VH121052	673	0.35	0.26
	VH331968	673	0.34	0.08
	VH329789	673	0.61	0.38
	VH333625	673	0.22	0.21
	VH331594	673	0.25	0.15
	VH331619	673	0.36	0.35
	VH332010	673	0.28	0.25
	VH314506	673	0.50	0.47
	VH330012	673	0.52	0.26
8	VH316165	686	0.44	0.39
	VH316030	686	0.43	0.26
	VH320829	686	0.42	0.42
	VH319153	686	0.64	0.43
	VH141195	686	0.28	0.00
	VH333343	686	0.25	0.35
	VH331992	686	0.11	0.02
	VH314590	686	0.42	0.29
	VH331625	686	0.27	0.05
	VH332080	686	0.29	0.08
	VH314597	686	0.31	0.25
	VH331585	686	0.42	0.25

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
9	VF862661	688	0.62	0.42
	VH316181	688	0.45	0.37
	VH326109	688	0.29	0.14
	VH316141	688	0.35	0.19
	VH314579	688	0.26	0.19
	VH331610	688	0.27	0.32
	VH333616	688	0.40	0.31
	VH331939	688	0.67	0.34
	VH331929	688	0.40	0.05
	VH333334	688	0.39	0.39
	VH331516	688	0.51	0.28
	VH332031	688	0.48	0.36
10	VH316125	701	0.39	0.38
	VH320721	701	0.61	0.36
	VH316016	701	0.73	0.47
	VH325575	701	0.83	0.30
	VH332100	701	0.42	0.30
	VH331448	701	0.35	0.12
	VH333293	701	0.28	0.34
	VH331951	701	0.52	0.24
	VH331621	701	0.29	0.26
	VH332098	701	0.36	0.36
	VH329531	701	0.41	0.24
	VH331495	701	0.33	0.21

Table F11. Mathematics Grade 7 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH306429	820	0.36	0.39
	VH303120	820	0.30	0.32
	VH306193	820	0.24	0.17
	VH325443	820	0.43	0.40
	VH312445	820	0.32	0.31
	VH330863	820	0.26	0.11
	VH328433	820	0.27	0.38
	VH333671	820	0.49	0.22
	VH335315	820	0.20	0.04
	VH333706	820	0.58	0.38
	VH328319	820	0.50	0.42
	VF492957	820	0.55	0.24
2	VH325549	681	0.29	0.25
	VH325400	681	0.40	0.20
	VH306275	681	0.53	0.39
	VH306418	681	0.24	0.08
	VH331221	681	0.66	0.39
	VH328916	681	0.41	0.38
	VH335579	681	0.49	0.43
	VH328327	681	0.48	0.36
	VH331013	681	0.26	0.21
	VH312469	681	0.34	0.03
	VH335465	681	0.22	0.12
	VH331383	681	0.45	0.28
3	VH325435	683	0.36	0.22
	VH303346	683	0.35	0.01
	VH306390	683	0.45	0.53
	VH325467	683	0.42	0.50
	VH312723	683	0.33	0.25
	VH328394	683	0.27	0.38
	VH333655	683	0.47	0.43
	VH335583	683	0.35	0.35
	VH331398	683	0.39	0.24
	VH330880	683	0.20	0.20
	VH335784	683	0.48	0.19
	VH312436	683	0.19	0.02
4	VH325544	700	0.20	0.20
	VF492587	700	0.22	0.22
	VH306332	700	0.75	0.48
	VH306208	700	0.55	0.55
	VH331271	700	0.69	0.47
	VH312742	700	0.51	0.23
	VH328920	700	0.28	0.34
	VH100013	700	0.27	0.39
	VH335476	700	0.42	0.36
	VH333687	700	0.39	0.32
	VH331403	700	0.48	0.40
	VH330884	700	0.46	0.38

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
5	VH325421	699	0.50	0.18
	VH325475	699	0.07	0.19
	VH303234	699	0.42	0.30
	VH306239	699	0.39	0.28
	VH330956	699	0.45	0.48
	VH312497	699	0.12	0.18
	VH331263	699	0.40	0.23
	VH335467	699	0.37	0.10
	VH328943	699	0.26	0.16
	VH328326	699	0.50	0.37
	VH331269	699	0.67	0.42
	VH333637	699	0.19	0.06
6	VH306229	672	0.39	0.46
	VH307249	672	0.41	0.39
	VH325528	672	0.24	0.11
	VH303324	672	0.29	0.09
	VH330883	672	0.60	0.47
	VH331280	672	0.62	0.38
	VH328421	672	0.29	0.31
	VH312744	672	0.29	0.24
	VH333644	672	0.22	0.03
	VH312448	672	0.65	0.21
	VH336013	672	0.76	0.41
	VH328884	672	0.39	0.33
7	VH303131	720	0.46	0.25
	VH325538	720	0.56	0.41
	VH325396	720	0.35	0.39
	VH306236	720	0.71	0.41
	VH328317	720	0.54	0.39
	VH328875	720	0.27	0.19
	VH331207	720	0.35	0.34
	VH335778	720	0.38	0.34
	VH312895	720	0.23	0.16
	VF866380	720	0.13	0.05
	VH335481	720	0.55	0.45
	VH330965	720	0.43	0.46
8	VH325533	683	0.15	-0.08
	VH306422	683	0.39	0.39
	VH303297	683	0.30	0.29
	VH303147	683	0.38	0.36
	VH335614	683	0.51	0.38
	VH331003	683	0.44	0.35
	VH328888	683	0.28	0.20
	VH335799	683	0.46	0.26
	VH328414	683	0.33	0.33
	VH331240	683	0.27	0.10
	VH312475	683	0.25	0.28
	VH312631	683	0.58	0.40

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
9	VH306233	694	0.54	0.41
	VH095539	694	0.17	0.12
	VH303186	694	0.33	0.12
	VH325449	694	0.62	0.47
	VH331267	694	0.52	0.35
	VF882725	694	0.39	0.27
	VH312486	694	0.47	0.30
	VH335314	694	0.62	0.42
	VH333677	694	0.40	0.31
	VH328932	694	0.48	0.47
	VH335317	694	0.33	0.41
	VH312734	694	0.36	0.40
10	VH306378	702	0.62	0.53
	VH306201	702	0.30	0.32
	VH303156	702	0.23	0.12
	VH325459	702	0.49	0.32
	VH336010	702	0.47	0.39
	VH330990	702	0.50	0.53
	VH328322	702	0.37	0.28
	VH312539	702	0.21	-0.05
	VH331234	702	0.20	0.20
	VH333640	702	0.28	0.07
	VH335824	702	0.28	0.19
	VH312439	702	0.34	0.31

Table F12. Mathematics Grade 8 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH273319	779	0.52	0.44
	VH273219	779	0.24	0.08
	VF812655	779	0.50	0.32
	VF823461	779	0.49	0.34
	VH327154	779	0.36	0.21
	VH335264	779	0.39	0.41
	VH326522	779	0.57	0.40
	VH326397	779	0.59	0.46
	VH327295	779	0.31	0.34
	VH279222	779	0.41	0.20
	VH328152	779	0.46	0.43
	VH275351	779	0.52	0.46
2	VH273345	662	0.35	0.39
	VH273404	662	0.30	0.28
	VH273241	662	0.41	0.33
	VH278285	662	0.32	0.25
	VH279127	662	0.39	0.11
	VH326444	662	0.34	0.29
	VH277021	662	0.56	0.37
	VH335216	662	0.59	0.34
	VH335113	662	0.51	0.29
	VH327161	662	0.27	0.30
	VH327423	662	0.61	0.49
	VH328223	662	0.47	0.37
3	VH273284	660	0.51	0.52
	VH277940	660	0.64	0.30
	VH273342	660	0.38	0.48
	VH273400	660	0.36	0.32
	VH275357	660	0.54	0.34
	VH326529	660	0.13	0.04
	VH335279	660	0.56	0.19
	VH327122	660	0.27	0.31
	VH279238	660	0.42	0.11
	VH328168	660	0.52	0.45
	VH327173	660	0.56	0.42
	VH326508	660	0.43	0.24
4	VH120095	674	0.34	0.23
	VH273378	674	0.61	0.33
	VH273231	674	0.24	0.12
	VH275661	674	0.54	0.34
	VH327094	674	0.50	0.35
	VH335194	674	0.29	0.34
	VH326525	674	0.32	0.34
	VH326402	674	0.20	0.14
	VH335226	674	0.46	0.37
	VH328252	674	0.65	0.34
	VH279210	674	0.47	0.36
	VH327547	674	0.56	0.27

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
5	VH278724	673	0.39	0.23
	VH279104	673	0.22	0.12
	VH273412	673	0.27	0.25
	VH273267	673	0.71	0.47
	VH328109	673	0.47	0.36
	VH327481	673	0.30	0.32
	VH326462	673	0.51	0.28
	VH275352	673	0.60	0.34
	VH327109	673	0.62	0.46
	VH335237	673	0.30	0.24
	VH326518	673	0.36	0.43
	VH335277	673	0.69	0.37
6	VH275564	650	0.52	0.40
	VH273383	650	0.34	0.42
	VH273236	650	0.19	0.19
	VH279008	650	0.36	0.33
	VH279185	650	0.45	0.25
	VH335272	650	0.38	0.42
	VH326399	650	0.27	0.26
	VH327545	650	0.65	0.28
	VH335121	650	0.48	0.30
	VH335242	650	0.20	0.12
	VH327164	650	0.52	0.52
	VH328117	650	0.65	0.49
7	VH277469	662	0.39	0.25
	VH275779	662	0.41	0.38
	VH273280	662	0.50	0.36
	VH273369	662	0.39	0.45
	VH328183	662	0.40	0.31
	VH327150	662	0.58	0.42
	VH335206	662	0.47	0.45
	VH326505	662	0.36	0.17
	VH327439	662	0.31	0.38
	VH327101	662	0.32	0.29
	VH276482	662	0.35	0.50
	VH326449	662	0.39	-0.02
8	VH275743	708	0.68	0.45
	VH275759	708	0.52	0.33
	VH273295	708	0.27	0.16
	VH273335	708	0.46	0.30
	VH335208	708	0.43	0.35
	VH335262	708	0.66	0.39
	VH326498	708	0.31	0.22
	VH327176	708	0.62	0.39
	VH328125	708	0.58	0.37
	VH326531	708	0.27	0.30
	VH327542	708	0.40	0.39
	VH275489	708	0.21	0.30

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
9	VH273262	670	0.53	0.27
	VH279045	670	0.41	0.28
	VH273388	670	0.23	0.28
	VH277658	670	0.33	0.34
	VH327429	670	0.34	0.18
	VH326511	670	0.28	0.23
	VH328260	670	0.58	0.48
	VH327115	670	0.54	0.52
	VH335243	670	0.63	0.36
	VH327335	670	0.47	0.31
	VH275483	670	0.25	0.36
	VH327497	670	0.47	0.35
10	VH273397	666	0.29	0.17
	VH273366	666	0.23	0.37
	VH273257	666	0.51	0.55
	VH278508	666	0.38	0.38
	VH327449	666	0.33	0.38
	VH328245	666	0.56	0.34
	VH327551	666	0.79	0.39
	VH275440	666	0.49	0.23
	VH279229	666	0.24	0.18
	VH335202	666	0.43	0.29
	VH335214	666	0.48	0.26
	VH326513	666	0.22	0.09

Science

Table F13. Science Grade 4 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH260643	826	0.30	0.06
	VH260644	826	0.47	0.45
	VH260636	826	0.18	-0.20
	VH260662	826	0.28	0.08
	VH260514	826	0.51	0.30
	VH097892	826	0.88	0.36
	VH097820	826	0.65	0.10
	VH256003	826	0.58	0.35
	VH256810	826	0.65	0.42
	VH255995	826	0.32	0.33
	VH256791	826	0.20	-0.01
	VH256803	826	0.25	0.11
2	VH260516	762	0.65	0.33
	VH260654	762	0.44	0.43
	VH260512	762	0.43	0.25
	VH264414	762	0.31	0.14
	VH260651	762	0.38	0.22
	VH097815	762	0.77	0.37
	VH097878	762	0.94	0.37
	VH256019	762	0.63	0.48
	VH256022	762	0.81	0.22
	VH256027	762	0.61	0.34
	VH256821	762	0.72	0.39
	VH256852	762	0.27	0.15
3	VH320328	741	0.34	0.34
	VH320263	741	0.25	0.12
	VH320261	741	0.45	0.26
	VH320349	741	0.61	0.36
	VH320281	741	0.22	-0.10
	VH311274	741	0.31	0.26
	VH311280	741	0.84	0.51
	VH264410	741	0.73	0.32
	VH264364	741	0.56	0.42
	VH264406	741	0.37	0.16
	VH264348	741	0.70	0.34
	VH264404	741	0.40	0.32
4	VH320338	733	0.55	0.29
	VH320242	733	0.46	0.21
	VH320273	733	0.39	0.20
	VH320256	733	0.37	0.29
	VH320333	733	0.71	0.31
	VH311283	733	0.29	0.15
	VH311275	733	0.65	0.43
	VH253167	733	0.47	0.30
	VH253123	733	0.57	0.41
	VH253274	733	0.64	0.27
	VH253209	733	0.50	0.25
	VH253228	733	0.62	0.36

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
5	VH264369	732	0.69	0.43
	VH264373	732	0.75	0.37
	VH264376	732	0.41	0.25
	VH264397	732	0.22	0.17
	VH264412	732	0.71	0.39
	VH320346	732	0.51	0.15
	VH320294	732	0.61	0.35
	VH253127	732	0.63	0.40
	VH253180	732	0.39	0.22
	VH253201	732	0.72	0.21
	VH253194	732	0.61	0.27
	VH253271	732	0.47	0.37
6	VH311558	753	0.93	0.25
	VH311552	753	0.62	0.36
	VH311529	753	0.54	0.19
	VH311553	753	0.40	0.22
	VH311510	753	0.61	0.27
	VH311264	753	0.46	0.36
	VH311255	753	0.51	0.38
	VH260801	753	0.58	0.44
	VH260767	753	0.34	0.24
	VH260807	753	0.60	0.41
	VH260785	753	0.48	0.46
	VH260824	753	0.34	0.16
7	VH311495	746	0.72	0.38
	VH311498	746	0.25	0.11
	VH311543	746	0.20	0.07
	VH311501	746	0.35	0.01
	VH311526	746	0.50	0.12
	VH256383	746	0.37	0.10
	VH256373	746	0.51	0.25
	VH267013	746	0.51	0.31
	VH267016	746	0.47	0.41
	VH267026	746	0.77	0.47
	VH266958	746	0.56	0.53
	VH266924	746	0.37	0.34
8	VH267032	739	0.75	0.48
	VH266994	739	0.75	0.29
	VH267007	739	0.50	0.29
	VH266971	739	0.48	0.29
	VH267052	739	0.24	0.05
	VH256405	739	0.83	0.47
	VH256416	739	0.70	0.39
	VH264469	739	0.29	0.10
	VH264455	739	0.57	0.30
	VH264464	739	0.17	-0.18
	VH264443	739	0.37	0.28
	VH264458	739	0.23	0.00

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
9	VH264488	732	0.55	0.34
	VH264484	732	0.21	-0.04
	VH264486	732	0.23	0.06
	VH264461	732	0.30	0.12
	VH264474	732	0.51	0.35
	VH256376	732	0.52	0.26
	VH256424	732	0.44	0.33
	VH320167	732	0.57	0.38
	VH320192	732	0.76	0.42
	VH320146	732	0.49	0.41
	VH320130	732	0.27	0.04
	VH320182	732	0.50	0.39
10	VH267198	713	0.46	0.28
	VH267368	713	0.42	0.17
	VH267359	713	0.41	0.04
	VH267376	713	0.88	0.38
	VH267389	713	0.45	0.24
	VH256391	713	0.69	0.30
	VH256369	713	0.59	0.36
	VH320124	713	0.61	0.42
	VH320119	713	0.45	0.39
	VH320173	713	0.74	0.50
	VH320190	713	0.26	-0.03
	VH325817	713	0.45	0.29

Table F14. Science Grade 8 Classical Statistics for Field Test Items

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
1	VH272770	774	0.85	0.38
	VH272762	774	0.87	0.37
	VH272778	774	0.66	0.36
	VH272766	774	0.65	0.30
	VH272782	774	0.79	0.41
	VH260222	774	0.29	0.20
	VH260216	774	0.40	0.24
	VH260290	774	0.40	0.21
	VH260307	774	0.53	0.34
	VH260339	774	0.31	0.24
	VH260355	774	0.48	0.21
	VH260370	774	0.80	0.33
2	VH260318	658	0.61	0.26
	VH260326	658	0.65	0.27
	VH260329	658	0.79	0.47
	VH260347	658	0.40	0.23
	VH260361	658	0.71	0.20
	VH260229	658	0.39	0.20
	VH260244	658	0.45	0.29
	VH272750	658	0.80	0.41
	VH272759	658	0.78	0.37
	VH272765	658	0.76	0.31
	VH272761	658	0.84	0.30
	VH272772	658	0.56	0.44
3	VH304307	664	0.63	0.32
	VH304378	664	0.54	0.44
	VH304364	664	0.78	0.20
	VH304374	664	0.77	0.41
	VH304354	664	0.46	0.27
	VH260236	664	0.35	0.04
	VH260257	664	0.69	0.28
	VH267619	664	0.33	0.18
	VH267562	664	0.72	0.31
	VH267471	664	0.59	0.12
	VH267634	664	0.29	0.18
	VH267596	664	0.23	0.23
4	VH267601	670	0.32	-0.02
	VH267491	670	0.59	0.28
	VH267352	670	0.36	0.17
	VH267437	670	0.31	0.20
	VH267625	670	0.70	0.44
	VH260247	670	0.35	0.32
	VH260266	670	0.46	0.18
	VH304390	670	0.43	0.17
	VH304370	670	0.59	0.36
	VH304400	670	0.40	0.28
	VH304392	670	0.54	0.39
	VH304395	670	0.28	0.12

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
5	VH268068	685	0.58	0.35
	VH268093	685	0.70	0.39
	VH268157	685	0.21	-0.04
	VH313550	685	0.37	0.12
	VH268128	685	0.74	0.28
	VH260259	685	0.34	0.07
	VH260262	685	0.74	0.30
	VH310799	685	0.78	0.43
	VH310801	685	0.63	0.43
	VH310773	685	0.68	0.31
	VH310788	685	0.51	0.40
	VH310812	685	0.34	0.24
6	VH310771	650	0.48	0.34
	VH310779	650	0.45	0.32
	VH310782	650	0.26	0.20
	VH310816	650	0.30	0.14
	VH310822	650	0.47	0.34
	VH260417	650	0.31	0.27
	VH260427	650	0.22	-0.03
	VH268075	650	0.57	0.45
	VH268148	650	0.56	0.32
	VH268153	650	0.60	0.42
	VH268119	650	0.48	0.15
	VH268155	650	0.62	0.47
7	VH305912	652	0.64	0.35
	VH305886	652	0.78	0.27
	VH305898	652	0.61	0.30
	VH305911	652	0.25	0.22
	VH305917	652	0.31	0.04
	VH260431	652	0.41	0.19
	VH260477	652	0.65	0.39
	VH271293	652	0.68	0.38
	VH271303	652	0.71	0.42
	VH271336	652	0.52	0.12
	VH271325	652	0.58	0.36
	VH271363	652	0.64	0.45
8	VH271321	706	0.63	0.29
	VH271329	706	0.20	0.09
	VH271311	706	0.52	0.24
	VH271285	706	0.58	0.42
	VH271366	706	0.57	0.42
	VH260423	706	0.53	0.39
	VH260469	706	0.71	0.29
	VH305902	706	0.56	0.23
	VH305880	706	0.61	0.40
	VH305894	706	0.31	0.11
	VH305908	706	0.47	0.34
	VH305913	706	0.51	0.47

Form	Accession Number	N	Average Item Score	Point Biserial Corr.
9	VH095960	656	0.81	0.31
	VH095948	656	0.49	0.21
	VH095976	656	0.24	0.07
	VH096248	656	0.69	0.38
	VH096235	656	0.73	0.41
	VH260453	656	0.62	0.47
	VH260440	656	0.39	0.18
	VH305813	656	0.75	0.38
	VH305598	656	0.45	0.36
	VH305787	656	0.38	0.32
	VH305819	656	0.31	0.06
	VH305822	656	0.60	0.40
10	VH305589	674	0.54	0.30
	VH305772	674	0.33	0.12
	VH305792	674	0.24	0.26
	VH305609	674	0.60	0.47
	VH305821	674	0.31	0.24
	VH260475	674	0.56	0.31
	VH260438	674	0.48	0.36
	VH096228	674	0.48	0.10
	VH096002	674	0.87	0.31
	VH096008	674	0.64	0.43
	VH095997	674	0.35	0.15
	VH095991	674	0.54	0.29

Appendix G: Classical Item Statistics for 2016 Operational Items

Reading

Table G1. Reading Grade 3 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF394057	7683	0.71	0.29
VF394053	7683	0.78	0.47
VF394041	7683	0.72	0.30
VF394045	7683	0.65	0.39
VF394046	7683	0.85	0.51
VF394050	7683	0.83	0.34
VF394049	7683	0.81	0.35
VF394051	7683	0.76	0.26
VF389477	7683	0.65	0.53
VF389620	7683	0.66	0.41
VF389446	7683	0.56	0.31
VF389470	7683	0.69	0.51
VF389473	7683	0.60	0.45
VF389457	7683	0.44	0.33
VF389165	7683	0.70	0.51
VF821120	7683	0.76	0.48
VF821218	7683	0.79	0.45
VF821123	7683	0.72	0.37
VF821312	7683	0.66	0.39
VF821272	7683	0.90	0.50
VF821338	7683	0.69	0.44
VF821362	7683	0.63	0.45
VF883330	7683	0.46	0.45
VF882884	7683	0.47	0.41
VF883326	7683	0.66	0.33
VF883549	7683	0.75	0.55
VF883561	7683	0.52	0.30
VF883364	7683	0.87	0.50
VF883614	7683	0.78	0.43
VF883619	7683	0.63	0.48
VF883622	7683	0.67	0.35
VH143172	7683	0.62	0.48
VH143162	7683	0.88	0.42
VH143197	7683	0.52	0.32
VH143131	7683	0.75	0.34

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH143274	7683	0.71	0.35
VH143252	7683	0.60	0.46
VF497759	7683	0.48	0.45
VF497725	7683	0.68	0.58
VF497758	7683	0.60	0.44
VF497767	7683	0.49	0.41
VF497766	7683	0.39	0.23
VF497718	7683	0.70	0.32
VH152948	7683	0.60	0.38
VH153021	7683	0.79	0.39
VH153042	7683	0.82	0.49
VF885214	7683	0.69	0.45
VF885379	7683	0.55	0.34
VF885192	7683	0.86	0.55
VF885434	7683	0.65	0.50

Table G2. Reading Grade 4 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF880215	7482	0.76	0.47
VF880204	7482	0.45	0.43
VF880210	7482	0.57	0.33
VF880326	7482	0.68	0.34
VF880311	7482	0.57	0.43
VF880321	7482	0.52	0.36
VF880314	7482	0.33	0.21
VF880343	7482	0.62	0.39
VF880345	7482	0.60	0.41
VF880350	7482	0.67	0.41
VH145672	7482	0.65	0.31
VH145677	7482	0.43	0.36
VH145638	7482	0.71	0.49
VH145684	7482	0.32	0.27
VH145660	7482	0.53	0.33
VH145702	7482	0.53	0.33
VH145595	7482	0.67	0.47
VH150628	7482	0.73	0.40
VH150639	7482	0.86	0.45
VH150646	7482	0.68	0.40
VH150612	7482	0.74	0.39
VH150620	7482	0.59	0.34
VH150605	7482	0.85	0.47
VH150637	7482	0.84	0.46
VF884843	7482	0.86	0.53
VF884830	7482	0.68	0.45
VF884836	7482	0.64	0.52
VF884910	7482	0.83	0.36
VF884900	7482	0.86	0.50
VF884925	7482	0.74	0.43
VF884913	7482	0.52	0.40
VF497338	7482	0.58	0.20
VF497297	7482	0.92	0.44
VF497314	7482	0.74	0.55
VF497322	7482	0.69	0.48
VF497303	7482	0.84	0.48
VF497330	7482	0.79	0.47
VF497326	7482	0.63	0.53
VF497327	7482	0.73	0.39
VF407243	7482	0.76	0.39

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF407287	7482	0.87	0.48
VF407232	7482	0.85	0.53
VF407295	7482	0.92	0.42
VF407297	7482	0.73	0.44
VF407235	7482	0.56	0.27
VF407298	7482	0.67	0.36
VF407282	7482	0.90	0.51
VF885226	7482	0.88	0.44
VF885195	7482	0.84	0.46
VH152031	7482	0.54	0.38

Table G3. Reading Grade 5 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH134512	7323	0.86	0.45
VH134422	7323	0.78	0.58
VH134412	7323	0.59	0.32
VH134521	7323	0.84	0.25
VH134580	7323	0.72	0.36
VH143307	7323	0.74	0.44
VH143304	7323	0.60	0.49
VH143328	7323	0.54	0.40
VH143318	7323	0.85	0.37
VH143331	7323	0.82	0.40
VF407319	7323	0.69	0.36
VF407388	7323	0.72	0.40
VF407329	7323	0.79	0.49
VF407332	7323	0.79	0.42
VF407360	7323	0.91	0.37
VF407355	7323	0.84	0.46
VH144472	7323	0.88	0.30
VH144483	7323	0.76	0.42
VH144474	7323	0.66	0.50
VH144509	7323	0.66	0.32
VH144512	7323	0.82	0.42
VH144593	7323	0.89	0.43
VF496865	7323	0.70	0.49
VF496879	7323	0.65	0.42
VF496213	7323	0.71	0.37
VF496206	7323	0.65	0.40
VF496209	7323	0.79	0.48
VF496212	7323	0.82	0.48
VF496221	7323	0.93	0.46
VF496200	7323	0.83	0.41
VF880864	7323	0.89	0.47
VF882769	7323	0.73	0.45
VF882762	7323	0.51	0.36
VF882790	7323	0.52	0.40
VF909893	7323	0.60	0.34
VF882786	7323	0.71	0.32
VH136637	7323	0.81	0.42
VH136631	7323	0.83	0.46
VH136633	7323	0.77	0.40
VH136634	7323	0.67	0.44

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH136607	7323	0.92	0.35
VF497025	7323	0.69	0.52
VF497030	7323	0.65	0.48
VF497037	7323	0.86	0.54
VF497012	7323	0.72	0.48
VF497016	7323	0.76	0.51
VF885197	7323	0.86	0.56
VF885217	7323	0.78	0.51
VF885221	7323	0.61	0.33
VH150300	7323	0.63	0.40
VH150408	7323	0.71	0.45
VH150398	7323	0.45	0.41
VH150312	7323	0.68	0.53
VH150267	7323	0.56	0.29

Table G4. Reading Grade 6 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF497047	7013	0.76	0.34
VF497049	7013	0.72	0.29
VF497035	7013	0.81	0.42
VF497034	7013	0.82	0.43
VF497041	7013	0.85	0.48
VF496873	7013	0.68	0.43
VF496204	7013	0.64	0.49
VF496208	7013	0.64	0.44
VF496863	7013	0.64	0.49
VF496191	7013	0.60	0.49
VF496867	7013	0.71	0.46
VF496415	7013	0.67	0.37
VF496172	7013	0.66	0.45
VF496055	7013	0.48	0.43
VF496083	7013	0.87	0.50
VF496097	7013	0.73	0.52
VF496065	7013	0.70	0.48
VF496071	7013	0.74	0.48
VF496100	7013	0.74	0.53
VH151373	7013	0.69	0.45
VH151417	7013	0.61	0.36
VH151420	7013	0.74	0.56
VH151386	7013	0.70	0.49
VH151408	7013	0.68	0.46
VH153064	7013	0.58	0.38
VH153081	7013	0.58	0.39
VH153209	7013	0.85	0.49
VH153197	7013	0.65	0.51
VH153196	7013	0.76	0.51
VF495908	7013	0.81	0.44
VF495961	7013	0.63	0.40
VF495968	7013	0.46	0.39
VF495918	7013	0.72	0.43
VF495945	7013	0.62	0.35
VF495925	7013	0.75	0.44
VF806599	7013	0.85	0.52
VF806591	7013	0.73	0.53
VF806593	7013	0.63	0.44
VF806588	7013	0.63	0.31
VF806587	7013	0.56	0.34

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH129551	7013	0.80	0.49
VH129569	7013	0.89	0.52
VH129567	7013	0.81	0.47
VH129555	7013	0.74	0.51
VH129558	7013	0.55	0.40
VH129579	7013	0.59	0.42
VH129572	7013	0.83	0.49
VF884733	7013	0.66	0.40
VF884772	7013	0.41	0.38
VF884844	7013	0.66	0.43
VF884880	7013	0.75	0.42
VH152232	7013	0.51	0.43
VF884630	7013	0.64	0.31
VF884988	7013	0.69	0.51
VF884628	7013	0.54	0.32
VH152242	7013	0.88	0.49

Table G5. Reading Grade 7 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF496949	7051	0.83	0.34
VF496944	7051	0.66	0.34
VF496901	7051	0.68	0.23
VF496913	7051	0.87	0.46
VF496908	7051	0.92	0.36
VF496903	7051	0.82	0.46
VF496895	7051	0.51	0.47
VF496900	7051	0.88	0.47
VF497956	7051	0.69	0.26
VF497955	7051	0.83	0.48
VF497958	7051	0.71	0.51
VF497951	7051	0.70	0.42
VF497961	7051	0.61	0.49
VF497978	7051	0.55	0.43
VF497974	7051	0.59	0.40
VF865426	7051	0.83	0.43
VF865413	7051	0.76	0.43
VF865473	7051	0.51	0.49
VF865494	7051	0.55	0.26
VF865477	7051	0.79	0.40
VF906623	7051	0.47	0.35
VF865627	7051	0.75	0.45
VF820260	7051	0.56	0.49
VF820251	7051	0.65	0.35
VF820333	7051	0.55	0.26
VF820315	7051	0.66	0.45
VF820301	7051	0.78	0.45
VF820351	7051	0.55	0.38
VF820210	7051	0.66	0.43
VF498023	7051	0.83	0.54
VF497995	7051	0.72	0.60
VF498018	7051	0.80	0.49
VF497980	7051	0.49	0.43
VF498062	7051	0.60	0.39
VF498011	7051	0.57	0.34
VF498051	7051	0.63	0.41
VF498054	7051	0.60	0.48
VF498057	7051	0.64	0.44
VF498063	7051	0.49	0.36
VF498034	7051	0.66	0.47

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF498052	7051	0.41	0.30
VF820464	7051	0.64	0.32
VF820412	7051	0.78	0.39
VF820449	7051	0.72	0.39
VF820435	7051	0.40	0.31
VF820391	7051	0.67	0.39
VF820404	7051	0.46	0.46
VF864902	7051	0.70	0.53
VF864898	7051	0.83	0.42
VF865057	7051	0.72	0.58
VF865072	7051	0.73	0.55
VF865094	7051	0.33	0.26
VF865104	7051	0.56	0.54
VF885398	7051	0.64	0.44
VF885820	7051	0.50	0.53
VH150842	7051	0.76	0.42

Table G6. Reading Grade 8 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH152367	6793	0.59	0.48
VH152400	6793	0.76	0.34
VH152415	6793	0.66	0.23
VH152422	6793	0.83	0.47
VH152362	6793	0.84	0.44
VH152391	6793	0.83	0.54
VH152434	6793	0.74	0.45
VF497207	6793	0.87	0.46
VF497196	6793	0.82	0.36
VF497178	6793	0.56	0.31
VF497213	6793	0.83	0.41
VF497203	6793	0.66	0.48
VF497209	6793	0.71	0.34
VF497173	6793	0.46	0.32
VF497229	6793	0.69	0.31
VF497244	6793	0.81	0.40
VF497242	6793	0.49	0.35
VF497255	6793	0.87	0.37
VF497271	6793	0.80	0.51
VF867197	6793	0.75	0.49
VF867239	6793	0.65	0.42
VF867326	6793	0.74	0.53
VF867274	6793	0.77	0.40
VF867355	6793	0.33	0.29
VF497096	6793	0.67	0.44
VF497103	6793	0.78	0.45
VF497098	6793	0.52	0.31
VF497114	6793	0.76	0.43
VF497115	6793	0.76	0.40
VF819971	6793	0.44	0.25
VF820011	6793	0.61	0.22
VF820174	6793	0.73	0.51
VF820025	6793	0.76	0.42
VF820170	6793	0.43	0.35
VF820236	6793	0.45	0.34
VF820159	6793	0.72	0.31
VF497329	6793	0.79	0.48
VF497328	6793	0.60	0.37
VF497325	6793	0.70	0.45
VF497349	6793	0.89	0.44

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF497370	6793	0.66	0.32
VF497363	6793	0.59	0.35
VF820734	6793	0.87	0.44
VF820781	6793	0.65	0.47
VF820727	6793	0.61	0.50
VF820720	6793	0.59	0.37
VF820786	6793	0.65	0.19
VF820792	6793	0.72	0.37
VF820796	6793	0.48	0.28
VF820801	6793	0.89	0.47
VH151102	6793	0.70	0.46
VH151057	6793	0.87	0.44
VH151086	6793	0.86	0.53
VF883716	6793	0.82	0.39
VF883823	6793	0.78	0.45
VF883653	6793	0.89	0.46

Mathematics

Table G7. Mathematics Grade 3 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH095532	7683	0.82	0.42
VH103654	7683	0.81	0.44
VF393788	7683	0.75	0.45
VF821723	7683	0.50	0.43
VF803172	7683	0.51	0.45
VF406331	7683	0.76	0.52
VF387500	7683	0.67	0.49
VF821698	7683	0.50	0.44
VF493136	7683	0.79	0.48
VF492342	7683	0.79	0.54
VF406297	7683	0.51	0.32
VF865405	7683	0.63	0.45
VH103554	7683	0.69	0.47
VF866235	7683	0.66	0.53
VH095414	7683	0.43	0.38
VF737752	7683	0.55	0.59
VF740960	7683	0.51	0.43
VF866360	7683	0.89	0.50
VF866898	7683	0.80	0.41
VF740890	7683	0.50	0.42
VH095347	7683	0.65	0.55
VH095406	7683	0.60	0.41
VH103449	7683	0.79	0.28
VF865570	7683	0.63	0.58
VF821652	7683	0.65	0.48
VF866354	7683	0.79	0.57
VH098018	7683	0.59	0.16
VF803266	7683	0.57	0.61
VF865389	7683	0.68	0.54
VF819676	7683	0.88	0.30
VF494693	7683	0.57	0.50
VF394378	7683	0.39	0.23
VH103497	7683	0.55	0.57
VF867001	7683	0.64	0.55
VF803121	7683	0.54	0.38
VF393751	7683	0.79	0.48
VF865316	7683	0.90	0.47
VF493127	7683	0.58	0.48

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH094951	7683	0.49	0.54
VF821767	7683	0.45	0.35
VF394232	7683	0.86	0.42
VF394239	7683	0.61	0.62
VF493461	7683	0.43	0.31
VF867224	7683	0.79	0.44
VF394358	7683	0.60	0.56
VF493364	7683	0.56	0.43
VF866264	7683	0.74	0.51
VF493387	7683	0.78	0.42
VH098050	7683	0.53	0.44
VH095303	7683	0.36	0.45

Table G8. Mathematics Grade 4 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH097534	7485	0.83	0.33
VH097139	7485	0.70	0.57
VF492371	7485	0.92	0.26
VH118288	7485	0.53	0.34
VF866677	7485	0.79	0.42
VF822854	7485	0.93	0.36
VF800697	7485	0.71	0.09
VF492358	7485	0.90	0.33
VF815962	7485	0.72	0.42
VH103556	7485	0.55	0.31
VH097502	7485	0.82	0.39
VF867084	7485	0.44	0.51
VF823145	7485	0.82	0.42
VF801214	7485	0.48	0.50
VH097287	7485	0.49	0.60
VF823138	7485	0.68	0.54
VF497391	7485	0.71	0.55
VF864145	7485	0.48	0.43
VF492392	7485	0.69	0.39
VF493144	7485	0.64	0.54
VF493149	7485	0.89	0.33
VF492339	7485	0.90	0.42
VF493228	7485	0.62	0.61
VF393714	7485	0.55	0.41
VF866857	7485	0.67	0.46
VF741948	7485	0.38	0.28
VF492376	7485	0.79	0.39
VF864149	7485	0.51	0.45
VF493219	7485	0.77	0.35
VF863975	7485	0.84	0.39
VF880294	7485	0.63	0.58
VF741944	7485	0.78	0.30
VF816041	7485	0.66	0.43
VF801835	7485	0.68	0.31
VH103619	7485	0.41	0.34
VH097472	7485	0.44	0.40
VH097255	7485	0.82	0.46
VF815957	7485	0.68	0.43
VF801227	7485	0.41	0.43
VF492338	7485	0.71	0.42
VF866381	7485	0.62	0.57

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF492354	7485	0.81	0.41
VF493301	7485	0.49	0.35
VH097368	7485	0.47	0.54
VH097423	7485	0.89	0.23
VF393648	7485	0.40	0.38
VF823371	7485	0.74	0.38
VF864100	7485	0.77	0.42
VF492364	7485	0.76	0.49
VF816057	7485	0.70	0.48
VF867088	7485	0.45	0.45
VF866870	7485	0.40	0.33
VH097521	7485	0.71	0.52
VF493262	7485	0.71	0.40
VF497402	7485	0.55	0.48
VF880252	7485	0.78	0.53
VF493361	7485	0.51	0.23
VF816151	7485	0.83	0.41
VF815909	7485	0.85	0.48

Table G9. Mathematics Grade 5 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF491888	7325	0.86	0.44
VF492269	7325	0.72	0.48
VF819900	7325	0.84	0.28
VF491924	7325	0.70	0.47
VF492083	7325	0.40	0.54
VF864604	7325	0.62	0.47
VF741081	7325	0.63	0.41
VF491942	7325	0.75	0.46
VF815846	7325	0.52	0.38
VF797963	7325	0.62	0.51
VF491626	7325	0.77	0.48
VH094283	7325	0.40	0.33
VF492313	7325	0.74	0.54
VF741371	7325	0.83	0.41
VF823759	7325	0.37	0.43
VH095125	7325	0.35	0.32
VF819994	7325	0.57	0.43
VF866061	7325	0.58	0.49
VH095132	7325	0.44	0.53
VF492296	7325	0.41	0.29
VF740936	7325	0.82	0.44
VF741539	7325	0.82	0.46
VH099648	7325	0.31	0.39
VF492255	7325	0.70	0.53
VH103584	7325	0.70	0.58
VF802894	7325	0.60	0.43
VF492113	7325	0.68	0.46
VF491914	7325	0.63	0.34
VF492077	7325	0.65	0.55
VF491635	7325	0.52	0.56
VH093957	7325	0.61	0.46
VF491992	7325	0.67	0.39
VF491755	7325	0.44	0.51
VF492099	7325	0.60	0.56
VH095017	7325	0.61	0.53
VF798083	7325	0.65	0.53
VH094425	7325	0.56	0.55
VH094296	7325	0.62	0.23
VF491804	7325	0.64	0.64
VF492374	7325	0.74	0.41
VF823764	7325	0.82	0.40

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF491932	7325	0.62	0.53
VF491630	7325	0.51	0.41
VF492397	7325	0.49	0.47
VF492304	7325	0.77	0.50
VF491905	7325	0.71	0.49
VF880813	7325	0.57	0.50
VF492435	7325	0.44	0.30
VH103721	7325	0.50	0.51
VF864609	7325	0.52	0.56
VF802778	7325	0.32	0.34
VF492001	7325	0.68	0.33
VF866103	7325	0.62	0.39
VF491753	7325	0.66	0.34
VF819989	7325	0.41	0.40
VF491761	7325	0.66	0.46
VF815982	7325	0.39	0.47
VF802014	7325	0.62	0.60
VF797938	7325	0.69	0.54

Table G10. Mathematics Grade 6 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH135098	7011	0.79	0.41
VF862804	7011	0.83	0.26
VF862699	7011	0.70	0.47
VF811515	7011	0.45	0.45
VF492647	7011	0.76	0.48
VF492066	7011	0.52	0.47
VF882963	7011	0.47	0.24
VF862786	7011	0.61	0.54
VF809839	7011	0.65	0.58
VH135857	7011	0.75	0.38
VF492192	7011	0.58	0.37
VF741723	7011	0.62	0.39
VF492759	7011	0.44	0.35
VF810696	7011	0.50	0.50
VF492030	7011	0.67	0.42
VF812185	7011	0.60	0.41
VF492093	7011	0.49	0.51
VF492160	7011	0.45	0.38
VH118676	7011	0.58	0.49
VH140963	7011	0.50	0.48
VF492388	7011	0.67	0.40
VF741557	7011	0.42	0.44
VH121098	7011	0.62	0.27
VF809034	7011	0.75	0.39
VF741572	7011	0.56	0.42
VF741928	7011	0.85	0.40
VF492722	7011	0.42	0.41
VF865682	7011	0.30	0.34
VF866265	7011	0.75	0.32
VF491935	7011	0.72	0.52
VF866278	7011	0.49	0.36
VF865635	7011	0.59	0.45
VF797964	7011	0.38	0.24
VF822007	7011	0.39	0.42
VF883067	7011	0.66	0.41
VF491931	7011	0.88	0.47
VH121115	7011	0.75	0.52
VF741533	7011	0.44	0.44
VH121032	7011	0.60	0.57
VH136168	7011	0.69	0.49
VF866221	7011	0.53	0.42

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF821920	7011	0.62	0.35
VF797977	7011	0.77	0.44
VF821998	7011	0.80	0.42
VF492996	7011	0.51	0.56
VF821946	7011	0.62	0.33
VH136222	7011	0.40	0.56
VF797171	7011	0.39	0.29
VF822023	7011	0.71	0.42
VF491970	7011	0.65	0.41
VH121642	7011	0.35	0.47
VF803293	7011	0.44	0.43
VF866290	7011	0.50	0.36
VH118579	7011	0.54	0.44
VF882800	7011	0.58	0.32
VF866230	7011	0.58	0.42
VF491976	7011	0.44	0.37
VF865678	7011	0.54	0.39
VH118513	7011	0.68	0.40

Table G11. Mathematics Grade 7 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF813502	7054	0.69	0.51
VH095465	7054	0.63	0.43
VF492966	7054	0.75	0.49
VF800136	7054	0.71	0.34
VH129959	7054	0.53	0.47
VF818181	7054	0.57	0.52
VF736954	7054	0.62	0.50
VH095405	7054	0.55	0.30
VF880250	7054	0.41	0.39
VF492708	7054	0.50	0.47
VH095484	7054	0.51	0.43
VF492578	7054	0.44	0.49
VF822880	7054	0.41	0.40
VH095274	7054	0.57	0.40
VH095287	7054	0.49	0.50
VH147509	7054	0.58	0.48
VH141616	7054	0.65	0.28
VH129916	7054	0.51	0.49
VF880897	7054	0.77	0.49
VF819294	7054	0.50	0.38
VF867395	7054	0.62	0.31
VF736963	7054	0.40	0.26
VF883138	7054	0.43	0.37
VF866491	7054	0.35	0.53
VF736931	7054	0.53	0.45
VF883264	7054	0.45	0.53
VF869623	7054	0.58	0.55
VF800078	7054	0.62	0.28
VF881807	7054	0.66	0.58
VF822986	7054	0.59	0.25
VH129853	7054	0.50	0.49
VF867256	7054	0.53	0.46
VF882715	7054	0.30	0.29
VH141553	7054	0.77	0.52
VF493067	7054	0.45	0.45
VH097710	7054	0.68	0.52
VH141570	7054	0.52	0.35
VF736938	7054	0.58	0.53
VF818347	7054	0.78	0.18
VF818184	7054	0.51	0.43
VH124615	7054	0.65	0.41

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF492915	7054	0.73	0.36
VF866531	7054	0.43	0.48
VH148173	7054	0.43	0.31
VH097803	7054	0.47	0.33
VH148935	7054	0.46	0.41
VF866386	7054	0.37	0.34
VF882739	7054	0.61	0.38
VH141596	7054	0.39	0.28
VF882691	7054	0.50	0.44
VF492589	7054	0.30	0.45
VF492992	7054	0.28	0.51
VF736947	7054	0.65	0.48
VF880171	7054	0.31	0.34
VF882910	7054	0.50	0.41
VF883244	7054	0.58	0.45
VF799837	7054	0.73	0.47
VH148915	7054	0.65	0.44
VF492295	7054	0.61	0.39

Table G12. Mathematics Grade 8 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH118733	6804	0.54	0.35
VF880798	6804	0.63	0.39
VF491923	6804	0.85	0.36
VF802924	6804	0.42	0.24
VH121865	6804	0.62	0.45
VH120033	6804	0.66	0.43
VF823307	6804	0.61	0.46
VH118897	6804	0.47	0.27
VH118056	6804	0.37	0.20
VH120009	6804	0.60	0.36
VF802927	6804	0.48	0.38
VF491915	6804	0.35	0.30
VF885510	6804	0.48	0.40
VF812743	6804	0.36	0.36
VF823444	6804	0.56	0.46
VF885577	6804	0.52	0.44
VF802937	6804	0.55	0.38
VF880849	6804	0.43	0.21
VF812762	6804	0.48	0.43
VF492856	6804	0.58	0.54
VF492216	6804	0.60	0.33
VH148416	6804	0.56	0.38
VH146719	6804	0.45	0.53
VF804260	6804	0.51	0.51
VF883648	6804	0.79	0.38
VF493034	6804	0.56	0.47
VF823848	6804	0.48	0.42
VH122227	6804	0.26	0.40
VF809838	6804	0.70	0.38
VF492199	6804	0.63	0.49
VF823784	6804	0.72	0.38
VF865996	6804	0.42	0.46
VH121958	6804	0.41	0.45
VF880646	6804	0.59	0.48
VH122429	6804	0.57	0.35
VH145979	6804	0.43	0.27
VH148455	6804	0.55	0.47
VH147216	6804	0.78	0.44
VH138993	6804	0.51	0.42
VF491975	6804	0.64	0.49
VF863290	6804	0.27	0.25

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VH137600	6804	0.50	0.57
VF492385	6804	0.52	0.34
VF880501	6804	0.63	0.43
VF883670	6804	0.53	0.43
VF866064	6804	0.47	0.46
VF492586	6804	0.46	0.32
VF492212	6804	0.58	0.33
VH139013	6804	0.60	0.40
VF865675	6804	0.52	0.35
VF809001	6804	0.42	0.52
VF492268	6804	0.38	0.23
VF804267	6804	0.46	0.31
VF822412	6804	0.56	0.50
VF812997	6804	0.54	0.46
VH148444	6804	0.70	0.36
VH137854	6804	0.48	0.41
VH137656	6804	0.46	0.44
VH139254	6804	0.52	0.32
VF880680	6804	0.60	0.41
VH122468	6804	0.45	0.41
VF822465	6804	0.53	0.49
VF492436	6804	0.74	0.40
VH122050	6804	0.56	0.37
VF810708	6804	0.74	0.43

Science

Table G13. Science Grade 4 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF287740	7477	0.49	0.32
VF287742	7477	0.52	0.39
VF430825	7477	0.73	0.39
VF287745	7477	0.50	0.25
VH102818	7477	0.77	0.35
VH102841	7477	0.59	0.49
VH102847	7477	0.77	0.35
VH102875	7477	0.54	0.38
VF387280	7477	0.33	0.35
VF387256	7477	0.62	0.43
VF431142	7477	0.67	0.44
VF283606	7477	0.86	0.38
VF431081	7477	0.83	0.31
VF287766	7477	0.71	0.46
VF287770	7477	0.91	0.25
VH123679	7477	0.73	0.43
VH123681	7477	0.64	0.53
VF430634	7477	0.61	0.44
VF430638	7477	0.40	0.34
VH123637	7477	0.63	0.49
VH123641	7477	0.82	0.46
VF430984	7477	0.67	0.40
VF430987	7477	0.42	0.36
VH129725	7477	0.53	0.40
VH129733	7477	0.63	0.54
VH129781	7477	0.51	0.36
VF431112	7477	0.62	0.42
VF431113	7477	0.69	0.33
VF269830	7477	0.48	0.28
VF269831	7477	0.50	0.38
VF406427	7477	0.57	0.48
VF407139	7477	0.58	0.39
VF393954	7477	0.78	0.43
VF393826	7477	0.72	0.47
VH146166	7477	0.69	0.52
VH146172	7477	0.70	0.42
VH126160	7477	0.72	0.50
VH126184	7477	0.79	0.43
VF393695	7477	0.84	0.43

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF393721	7477	0.59	0.39
VF386736	7477	0.83	0.49
VF386739	7477	0.73	0.44
VH149131	7477	0.59	0.44
VH149133	7477	0.36	0.33
VF431073	7477	0.63	0.51
VF431078	7477	0.49	0.42
VF430640	7477	0.55	0.46
VF282669	7477	0.55	0.45
VF269779	7477	0.74	0.46
VF269769	7477	0.55	0.37

Table G14. Science Grade 8 Classical Statistics for Operational Items

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF309025	6789	0.55	0.42
VF309015	6789	0.48	0.28
VF308924	6789	0.49	0.41
VF308925	6789	0.55	0.44
VF486687	6789	0.72	0.44
VF431198	6789	0.53	0.27
VF313292	6789	0.57	0.29
VF313289	6789	0.52	0.35
VF823953	6789	0.58	0.35
VF823970	6789	0.53	0.37
VF862725	6789	0.70	0.33
VF862740	6789	0.44	0.30
VF862757	6789	0.45	0.42
VF308708	6789	0.53	0.31
VF308712	6789	0.45	0.50
VF308713	6789	0.51	0.36
VF308932	6789	0.47	0.28
VF308933	6789	0.31	0.37
VF431368	6789	0.54	0.51
VF308876	6789	0.69	0.42
VF308882	6789	0.75	0.33
VF813827	6789	0.71	0.37
VF813887	6789	0.80	0.46
VF308906	6789	0.58	0.48
VF308910	6789	0.48	0.34
VF431549	6789	0.65	0.50
VF431554	6789	0.48	0.37
VF308941	6789	0.53	0.45
VF308944	6789	0.31	0.49
VF308886	6789	0.48	0.49
VF308887	6789	0.64	0.50
VF308890	6789	0.52	0.38
VF431602	6789	0.37	0.45
VF431598	6789	0.51	0.48
VF431188	6789	0.70	0.50
VF431193	6789	0.54	0.29
VF431624	6789	0.51	0.42
VF431626	6789	0.74	0.56
VF685187	6789	0.63	0.44
VF671386	6789	0.46	0.37

Accession Number	*N*	Average Item Score	Point Biserial Corr.
VF431609	6789	0.51	0.29
VF431608	6789	0.51	0.40
VF431704	6789	0.58	0.48
VF431703	6789	0.46	0.46
VF394547	6789	0.54	0.41
VF394561	6789	0.64	0.48
VH139756	6789	0.42	0.37
VH140079	6789	0.42	0.36
VF486149	6789	0.48	0.46
VF486146	6789	0.54	0.39

Appendix H: Rasch Difficulty, Standard Error, Fit Statistics, and *N*-counts for 2016 Field Test Items

Reading

Table H1. Reading Grade 3 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH345371	802	0.361	0.081	0.95	0.91
	VH345370	802	1.250	0.078	1.00	1.04
	VH345359	802	0.044	0.084	1.02	1.02
	VH345362	802	0.653	0.079	1.20	1.25
	VH345376	802	1.147	0.078	1.02	1.05
	VH345375	802	1.904	0.082	1.16	1.50
	VH343052	802	-0.360	0.090	1.05	1.12
	VH343057	802	2.926	0.101	1.02	1.41
	VH343076	802	0.222	0.082	0.96	0.92
	VH343072	802	-1.115	0.107	0.83	0.65
	VH343085	802	0.374	0.081	1.09	1.14
	VH343081	802	0.609	0.079	1.11	1.15
	VH345489	802	0.640	0.079	0.92	0.89
	VH345515	802	0.249	0.082	1.03	1.04
2	VH345372	774	0.404	0.082	1.00	0.99
	VH345373	774	0.806	0.079	1.06	1.07
	VH345365	774	0.806	0.079	0.91	0.89
	VH345367	774	0.329	0.083	1.06	1.08
	VH345378	774	1.206	0.079	0.97	0.99
	VH345377	774	2.480	0.091	1.13	1.48
	VH343064	774	0.141	0.085	1.11	1.17
	VH343063	774	-0.119	0.088	0.86	0.75
	VH343078	774	0.780	0.080	1.11	1.14
	VH343080	774	-0.020	0.087	0.87	0.79
	VH343087	774	2.131	0.086	1.12	1.35
	VH343088	774	1.338	0.079	1.06	1.08
	VH345525	774	-0.890	0.105	0.93	0.92
	VH345520	774	1.219	0.079	1.19	1.23

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
3	VH331034	771	0.514	0.081	1.00	1.00
	VH331037	771	-0.265	0.091	1.04	1.00
	VH331077	771	0.625	0.081	0.95	0.92
	VH331079	771	1.050	0.079	1.08	1.08
	VH331092	771	0.084	0.086	1.01	0.99
	VH331100	771	1.094	0.079	0.96	0.95
	VH324712	771	1.712	0.081	1.12	1.22
	VH324717	771	2.482	0.091	1.16	1.71
	VH324741	771	2.663	0.095	1.19	1.85
	VH324735	771	-0.751	0.101	0.90	0.76
	VH324755	771	1.791	0.082	0.99	1.09
	VH324757	771	1.213	0.079	1.20	1.30
	VH345533	771	-0.670	0.099	0.80	0.62
	VH345539	771	-0.518	0.096	0.82	0.70
4	VH331036	768	1.739	0.083	1.18	1.38
	VH331040	768	-0.255	0.090	0.97	0.89
	VH331076	768	1.164	0.080	1.16	1.21
	VH331073	768	0.542	0.081	1.00	0.97
	VH331086	768	1.043	0.080	1.06	1.09
	VH331088	768	0.980	0.080	0.83	0.80
	VH324713	768	-1.323	0.117	0.88	0.68
	VH324715	768	1.698	0.082	0.99	1.02
	VH324726	768	1.591	0.081	1.21	1.31
	VH324745	768	1.725	0.082	1.03	1.18
	VH324760	768	-0.223	0.089	0.90	0.80
	VH324754	768	1.611	0.082	1.14	1.28
	VH345612	768	-0.648	0.098	0.91	0.85
	VH345617	768	0.588	0.081	0.92	0.88
5	VH346222	775	0.667	0.080	1.07	1.08
	VH346212	775	-1.614	0.130	0.82	0.56
	VH346299	775	-1.178	0.113	0.83	0.60
	VH346290	775	2.119	0.085	1.18	1.49
	VH346342	775	0.629	0.080	1.10	1.15
	VH346337	775	1.441	0.080	1.05	1.18
	VH345196	775	2.410	0.090	1.21	1.76
	VH345197	775	1.927	0.083	1.16	1.39
	VH345201	775	0.255	0.083	0.96	0.93
	VH345200	775	2.386	0.090	1.08	1.48
	VH345206	775	0.890	0.079	0.94	0.95
	VH345203	775	0.609	0.080	1.07	1.08
	VH345608	775	1.900	0.083	1.10	1.29
	VH345614	775	-0.448	0.094	0.86	0.77

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
6	VH346227	757	1.550	0.081	1.10	1.17
	VH346215	757	1.751	0.083	0.97	1.07
	VH346286	757	2.633	0.096	1.26	2.16
	VH346233	757	0.161	0.085	1.04	1.07
	VH346333	757	0.781	0.081	1.07	1.12
	VH346331	757	1.536	0.081	1.16	1.30
	VH345198	757	-0.159	0.090	0.95	0.88
	VH345199	757	0.504	0.082	0.90	0.85
	VH345195	757	-0.018	0.088	0.89	0.84
	VH345207	757	2.708	0.097	1.18	1.83
	VH345192	757	-0.119	0.089	1.05	1.04
	VH345204	757	1.007	0.080	1.21	1.29
	VH345605	757	-1.233	0.116	0.81	0.60
	VH345579	757	-1.432	0.123	0.84	0.57
7	VH330263	758	0.526	0.082	1.20	1.30
	VH330269	758	-0.690	0.098	0.92	0.82
	VH330268	758	0.763	0.081	1.25	1.32
	VH330266	758	0.770	0.081	1.04	1.07
	VH330258	758	0.874	0.080	1.03	1.05
	VH330271	758	1.895	0.085	1.10	1.27
	VH345380	758	-0.158	0.088	0.78	0.66
	VH345379	758	0.691	0.081	1.02	1.03
	VH345385	758	0.705	0.081	1.15	1.18
	VH345384	758	-0.229	0.089	0.85	0.75
	VH345389	758	0.526	0.082	1.05	1.07
	VH345392	758	0.275	0.083	0.98	0.97
	VH345545	758	-0.671	0.098	0.80	0.62
	VH345535	758	-1.045	0.107	0.80	0.56
8	VH330275	749	1.387	0.081	1.10	1.18
	VH330270	749	-0.014	0.088	0.94	0.92
	VH330274	749	0.764	0.081	1.07	1.07
	VH330278	749	1.098	0.081	1.03	1.06
	VH330262	749	0.039	0.087	1.13	1.27
	VH330273	749	-0.719	0.101	0.85	0.67
	VH345382	749	0.365	0.084	0.84	0.76
	VH345383	749	1.501	0.082	0.99	1.01
	VH345387	749	1.209	0.081	0.99	1.03
	VH345386	749	1.327	0.081	1.06	1.15
	VH345390	749	-0.507	0.096	0.90	0.84
	VH345391	749	-0.108	0.089	0.99	0.97
	VH345567	749	-0.278	0.092	0.86	0.76
	VH345587	749	0.302	0.084	0.92	0.90

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
9	VH318252	760	-0.327	0.093	0.91	1.07
	VH318247	760	-1.658	0.133	0.79	0.47
	VH318255	760	0.443	0.083	1.07	1.13
	VH318271	760	-0.345	0.093	0.94	0.87
	VH318230	760	0.775	0.081	1.09	1.13
	VH318222	760	-1.052	0.110	0.86	0.66
	VH313589	760	-0.659	0.100	0.95	0.86
	VH313601	760	0.584	0.082	1.00	1.00
	VH313596	760	-0.293	0.092	0.96	1.02
	VH313587	760	0.146	0.086	1.01	1.00
	VH313593	760	1.413	0.081	1.12	1.21
	VH313614	760	-2.015	0.150	0.81	0.52
	VH345600	760	-1.865	0.142	0.81	0.46
	VH345595	760	-0.243	0.091	0.88	0.79
10	VH318283	766	0.301	0.084	1.01	1.06
	VH318277	766	1.241	0.080	1.07	1.08
	VH318261	766	-1.182	0.114	0.82	0.60
	VH318268	766	-0.315	0.092	0.95	0.88
	VH318285	766	2.903	0.102	1.19	2.09
	VH318281	766	-0.011	0.087	0.96	0.93
	VH313599	766	1.254	0.080	1.20	1.33
	VH313585	766	0.027	0.087	1.04	1.02
	VH313618	766	1.133	0.080	1.02	1.03
	VH313592	766	0.130	0.085	0.91	0.85
	VH313616	766	-0.418	0.094	1.00	0.97
	VH313603	766	0.691	0.081	1.08	1.09
	VH345574	766	-0.103	0.088	0.82	0.73
	VH345560	766	-1.388	0.121	0.93	1.02

Table H2. Reading Grade 4 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH330664	780	0.954	0.082	1.00	0.99
	VH330647	780	0.576	0.086	0.98	0.90
	VH330691	780	2.507	0.085	1.11	1.35
	VH330801	780	0.486	0.087	0.89	0.84
	VH330845	780	1.709	0.080	0.96	1.01
	VH330848	780	-0.268	0.100	0.85	0.77
	VH331827	780	0.050	0.094	0.94	0.94
	VH331824	780	0.187	0.091	0.98	0.87
	VH331854	780	1.007	0.082	1.02	1.06
	VH331846	780	1.120	0.081	1.13	1.21
	VH331861	780	1.093	0.081	1.07	1.12
	VH331863	780	1.901	0.080	1.08	1.21
	VH325008	780	0.705	0.084	1.04	1.02
	VH324983	780	1.224	0.080	0.92	0.89
2	VH330634	738	1.272	0.082	1.04	1.08
	VH330597	738	0.021	0.098	0.96	1.05
	VH330811	738	0.206	0.094	0.97	0.99
	VH330696	738	1.436	0.081	1.12	1.13
	VH330841	738	0.418	0.090	0.96	0.94
	VH330839	738	0.385	0.091	1.06	1.10
	VH331842	738	0.335	0.092	0.95	0.86
	VH331836	738	2.308	0.083	1.18	1.35
	VH331849	738	1.879	0.081	1.17	1.29
	VH331851	738	-0.245	0.105	0.86	0.67
	VH331865	738	1.318	0.081	1.02	1.04
	VH331866	738	1.436	0.081	1.13	1.18
	VH324993	738	1.469	0.081	1.02	1.09
	VH325010	738	0.327	0.092	0.91	0.85
3	VH345656	743	1.400	0.081	1.08	1.12
	VH345659	743	0.999	0.083	1.07	1.10
	VH345665	743	0.254	0.092	1.00	1.03
	VH345666	743	0.033	0.096	0.95	0.95
	VH345671	743	0.124	0.094	1.07	1.13
	VH345674	743	0.709	0.086	0.90	0.84
	VH346003	743	0.774	0.085	0.95	0.92
	VH346001	743	0.782	0.085	0.96	0.94
	VH346029	743	0.731	0.085	0.86	0.80
	VH345999	743	2.143	0.082	1.28	1.45
	VH346002	743	2.246	0.083	1.17	1.37
	VH345996	743	1.407	0.081	1.11	1.14
	VH325048	743	-1.137	0.133	0.86	0.60
	VH325062	743	1.068	0.082	1.03	1.04

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH345651	740	1.087	0.083	1.08	1.09
	VH345650	740	-0.063	0.098	0.90	0.77
	VH345668	740	-0.552	0.109	0.91	0.82
	VH345664	740	2.242	0.085	1.08	1.30
	VH345672	740	2.797	0.092	1.16	1.59
	VH345676	740	1.333	0.082	1.18	1.27
	VH346034	740	0.564	0.088	0.98	0.95
	VH346024	740	1.387	0.082	0.98	0.98
	VH346007	740	2.917	0.094	1.17	1.57
	VH346033	740	1.265	0.082	0.99	0.98
	VH346028	740	0.276	0.092	0.90	0.84
	VH346020	740	0.996	0.084	1.17	1.24
	VH325053	740	1.374	0.082	1.06	1.08
	VH325066	740	-0.092	0.098	0.86	0.71
5	VH311932	741	1.875	0.080	1.08	1.19
	VH311948	741	-1.838	0.178	0.95	0.83
	VH311937	741	1.219	0.081	1.00	1.00
	VH311969	741	1.317	0.080	1.05	1.07
	VH311966	741	2.058	0.081	1.15	1.26
	VH311901	741	-0.630	0.115	0.87	0.78
	VH332787	741	2.300	0.083	1.15	1.42
	VH332764	741	0.441	0.089	0.97	0.95
	VH332833	741	1.388	0.080	1.18	1.22
	VH332844	741	3.249	0.099	1.20	1.69
	VH332901	741	3.089	0.095	1.08	1.40
	VH332876	741	0.834	0.084	0.88	0.83
	VH325055	741	0.488	0.088	1.03	0.97
	VH325027	741	1.213	0.081	1.00	0.98
6	VH311955	744	0.917	0.084	1.08	1.13
	VH311934	744	1.464	0.081	1.17	1.21
	VH311943	744	0.522	0.088	0.93	0.86
	VH311941	744	-0.036	0.097	0.92	0.85
	VH312255	744	1.069	0.083	0.94	0.90
	VH311911	744	0.598	0.087	1.03	1.07
	VH332770	744	0.917	0.084	1.19	1.28
	VH332780	744	1.411	0.081	1.07	1.10
	VH332824	744	0.621	0.087	0.97	0.94
	VH332812	744	0.530	0.088	1.10	1.20
	VH332909	744	1.279	0.082	0.94	0.93
	VH332866	744	0.388	0.090	1.11	1.10
	VH325041	744	0.348	0.090	0.99	0.99
	VH325073	744	-1.562	0.152	0.87	0.79

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH346125	749	-0.112	0.099	0.94	0.86
	VH346137	749	0.837	0.084	0.98	0.96
	VH346153	749	-0.637	0.114	0.89	0.61
	VH346150	749	2.346	0.084	0.99	1.13
	VH346183	749	1.131	0.082	0.93	0.91
	VH346188	749	1.492	0.080	0.96	0.98
	VH346064	749	1.848	0.081	1.08	1.19
	VH346055	749	2.072	0.082	1.12	1.27
	VH346056	749	1.537	0.080	1.13	1.16
	VH346065	749	2.631	0.087	1.04	1.26
	VH346067	749	2.346	0.084	1.26	1.54
	VH346052	749	0.076	0.095	0.97	0.92
	VH325022	749	0.267	0.091	0.87	0.76
	VH325032	749	0.190	0.093	0.92	0.89
8	VH346120	750	1.035	0.082	0.97	0.95
	VH346134	750	0.809	0.084	0.98	0.94
	VH346149	750	0.079	0.094	0.88	0.80
	VH346155	750	1.465	0.080	0.96	0.98
	VH346185	750	0.837	0.083	1.08	1.09
	VH346184	750	1.689	0.080	1.04	1.15
	VH346062	750	0.710	0.085	0.94	0.91
	VH346059	750	1.612	0.080	1.26	1.36
	VH346061	750	1.061	0.082	1.12	1.16
	VH346068	750	0.816	0.084	1.22	1.30
	VH346073	750	1.670	0.080	1.06	1.08
	VH346043	750	1.349	0.080	1.24	1.32
	VH325019	750	0.746	0.084	0.88	0.81
	VH325003	750	-1.088	0.129	0.92	0.60
9	VH311841	752	-0.229	0.102	0.89	0.90
	VH311778	752	1.491	0.080	1.18	1.26
	VH311804	752	-0.544	0.111	0.92	0.78
	VH311800	752	0.580	0.087	1.17	1.21
	VH311846	752	0.834	0.084	1.06	1.10
	VH311772	752	0.229	0.092	0.99	1.01
	VH331783	752	0.617	0.086	0.90	0.84
	VH331761	752	1.837	0.080	0.96	1.04
	VH331767	752	-0.078	0.098	0.89	0.82
	VH331775	752	0.107	0.094	1.04	1.05
	VH331771	752	-0.208	0.101	0.90	0.82
	VH331733	752	1.549	0.080	1.15	1.25
	VH325059	752	-0.069	0.098	0.89	0.84
	VH325017	752	2.125	0.082	1.20	1.36

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH311791	743	2.345	0.084	1.22	1.44
	VH311786	743	-0.075	0.098	0.92	0.85
	VH311812	743	2.002	0.082	1.08	1.19
	VH311833	743	0.779	0.084	1.06	1.08
	VH311818	743	1.564	0.080	1.18	1.24
	VH311822	743	0.500	0.087	0.93	0.88
	VH331745	743	-0.595	0.112	0.88	0.70
	VH331785	743	0.583	0.086	1.00	0.96
	VH331778	743	1.597	0.080	1.07	1.11
	VH331790	743	1.883	0.081	1.17	1.25
	VH331772	743	1.305	0.081	1.04	1.09
	VH331792	743	1.121	0.082	1.08	1.11
	VH325023	743	-0.213	0.101	0.85	0.71
	VH325039	743	2.008	0.082	1.11	1.20

Table H3. Reading Grade 5 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH345910	794	1.239	0.081	1.07	1.08
	VH345954	794	-0.180	0.100	0.81	0.63
	VH346000	794	0.564	0.087	1.11	1.29
	VH345994	794	-0.241	0.102	0.88	0.69
	VH346066	1518	0.203	0.068	0.94	0.94
	VH346070	794	0.285	0.091	0.89	0.94
	VH332285	794	-0.211	0.101	1.16	2.00
	VH332273	794	1.318	0.081	1.33	1.47
	VH332267	794	0.326	0.091	1.04	1.17
	VH332306	794	1.893	0.080	0.92	0.95
	VH332280	794	0.617	0.087	0.93	0.88
	VH332248	794	-0.759	0.116	0.84	0.58
	VH321425	794	2.009	0.081	0.95	1.03
	VH321429	794	1.305	0.081	1.16	1.20
2	VH345940	724	1.370	0.084	1.13	1.21
	VH345974	724	1.601	0.083	0.87	0.86
	VH346031	724	0.332	0.097	1.00	0.91
	VH346006	724	0.928	0.088	1.08	1.07
	VH346066	1518	0.203	0.068	0.94	0.94
	VH346075	724	1.594	0.083	0.96	0.97
	VH332303	724	0.206	0.100	1.08	1.14
	VH332288	724	-0.314	0.114	0.85	0.74
	VH332292	724	-0.004	0.105	1.06	1.13
	VH332298	724	2.697	0.087	1.19	1.53
	VH332312	724	-0.288	0.113	0.86	0.65
	VH332296	724	0.951	0.088	1.02	0.98
	VH321434	724	-0.894	0.136	0.88	0.75
	VH321777	724	-1.419	0.163	0.83	0.39
3	VH345929	719	0.119	0.101	0.95	0.91
	VH345933	719	0.840	0.088	0.95	0.91
	VH345951	1436	2.331	0.059	1.08	1.27
	VH345943	719	3.255	0.096	1.15	1.95
	VH345966	719	2.138	0.083	1.18	1.28
	VH345914	719	1.161	0.085	1.13	1.18
	VH333338	719	1.290	0.084	1.31	1.41
	VH333335	719	1.703	0.082	1.02	1.05
	VH333563	719	-0.298	0.112	0.91	0.93
	VH333527	719	-2.190	0.218	0.87	0.66
	VH333569	719	-0.286	0.112	0.95	0.74
	VH333574	719	-0.298	0.112	0.99	1.11
	VH321628	719	1.205	0.085	1.11	1.18
	VH321690	719	-1.105	0.144	0.98	1.24

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH345939	717	1.847	0.082	1.34	1.47
	VH345951	1436	2.331	0.059	1.08	1.27
	VH345973	717	1.227	0.084	0.96	0.92
	VH345947	717	1.105	0.085	0.93	0.89
	VH345963	717	0.386	0.095	0.97	0.91
	VH345958	717	1.969	0.083	1.10	1.23
	VH333347	717	1.643	0.082	1.18	1.24
	VH333344	717	2.436	0.085	1.32	1.79
	VH333497	717	-0.511	0.118	0.87	0.69
	VH333493	717	2.071	0.083	1.12	1.25
	VH333572	717	0.457	0.094	0.99	1.03
	VH333578	717	0.275	0.097	0.99	1.05
	VH321670	717	-0.015	0.103	0.83	0.65
	VH321766	717	-0.430	0.115	0.88	0.72
5	VH332343	723	0.575	0.092	0.93	0.90
	VH332339	723	3.077	0.092	0.98	1.13
	VH332359	723	0.805	0.089	1.01	1.05
	VH332353	723	-0.417	0.116	0.92	0.65
	VH332366	723	0.506	0.093	1.01	0.94
	VH332385	723	0.047	0.103	0.91	0.83
	VH314102	723	1.901	0.082	1.03	1.05
	VH314080	723	-0.777	0.130	0.80	0.50
	VH314083	723	0.480	0.094	0.99	1.02
	VH314088	1449	2.112	0.059	1.13	1.21
	VH314106	723	2.127	0.083	1.06	1.14
	VH314090	723	0.268	0.098	1.04	1.13
	VH321688	723	-0.364	0.115	0.82	0.59
	VH321544	723	1.254	0.084	1.11	1.12
6	VH332335	726	-0.772	0.126	0.87	0.63
	VH332336	726	1.913	0.083	1.28	1.46
	VH332345	726	0.521	0.092	0.94	0.85
	VH332347	726	1.223	0.084	1.15	1.21
	VH332387	726	-0.772	0.126	0.90	0.92
	VH332378	726	0.224	0.098	1.06	1.13
	VH314085	726	0.922	0.087	1.00	0.97
	VH314104	726	2.313	0.084	1.17	1.37
	VH314078	726	0.743	0.089	0.96	0.95
	VH314101	726	0.390	0.095	0.95	0.92
	VH314088	1449	2.112	0.059	1.13	1.21
	VH314094	726	0.922	0.087	0.96	0.95
	VH321443	726	-0.440	0.115	0.96	0.87
	VH321456	726	-1.071	0.139	0.85	0.49

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH321302	727	1.974	0.082	1.15	1.23
	VH321311	727	1.481	0.083	1.06	1.09
	VH321372	727	1.590	0.082	1.09	1.12
	VH321375	727	0.819	0.089	0.95	0.94
	VH321393	727	1.536	0.083	1.13	1.19
	VH321384	727	0.390	0.095	0.98	0.98
	VH313647	727	0.047	0.103	0.95	0.85
	VH313648	727	-0.337	0.114	0.97	0.93
	VH313644	727	1.685	0.082	1.19	1.26
	VH313652	727	0.912	0.088	1.09	1.18
	VH313650	727	0.435	0.095	1.09	1.25
	VH313637	727	1.752	0.082	1.14	1.20
	VH321439	727	-0.617	0.123	0.82	0.53
	VH321468	727	0.120	0.101	1.02	1.09
8	VH321297	729	-0.165	0.107	0.84	0.71
	VH321318	729	0.202	0.098	0.89	0.79
	VH321367	729	-0.618	0.122	0.91	0.72
	VH321346	729	-0.946	0.135	0.89	0.74
	VH321380	729	0.716	0.089	1.05	1.13
	VH321386	729	1.328	0.083	1.05	1.06
	VH313641	729	1.425	0.083	1.32	1.46
	VH313645	729	-0.504	0.118	0.98	0.93
	VH313643	729	1.790	0.082	1.09	1.18
	VH313649	729	2.182	0.083	1.13	1.21
	VH313640	729	0.932	0.087	1.08	1.07
	VH313638	729	0.052	0.102	1.03	1.02
	VH321609	729	-0.358	0.113	0.82	0.57
	VH321656	729	0.635	0.090	1.04	1.07
9	VH321130	728	1.887	0.083	1.04	1.12
	VH321077	728	0.652	0.091	1.20	1.27
	VH321089	728	1.619	0.083	1.19	1.29
	VH321132	728	1.102	0.086	1.03	1.06
	VH321062	728	1.529	0.083	0.92	0.93
	VH321068	728	-0.060	0.104	0.86	0.67
	VH333040	728	-0.525	0.118	0.94	0.97
	VH332958	728	0.960	0.087	1.08	1.10
	VH333069	728	0.952	0.087	1.03	1.05
	VH333033	728	-1.220	0.146	0.93	0.90
	VH332986	728	0.167	0.099	0.88	0.82
	VH332955	728	1.190	0.085	1.14	1.17
	VH321663	728	-1.504	0.162	0.85	0.44
	VH321686	728	-0.340	0.112	0.80	0.55

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH321096	732	1.758	0.082	1.19	1.28
	VH321082	732	0.939	0.086	0.95	0.99
	VH321123	732	1.357	0.083	0.95	0.93
	VH321133	732	1.610	0.082	1.18	1.25
	VH321128	732	0.782	0.087	1.05	1.09
	VH321072	732	2.609	0.087	1.06	1.25
	VH332960	732	0.947	0.086	0.98	0.97
	VH333072	732	1.426	0.083	1.10	1.14
	VH332994	732	0.162	0.097	1.06	1.11
	VH333038	732	-0.527	0.114	1.00	1.34
	VH333083	732	2.380	0.085	1.06	1.20
	VH333056	732	0.037	0.099	0.98	0.97
	VH321471	732	0.263	0.095	0.98	0.93
	VH321676	732	-0.501	0.114	0.92	0.83

Table H4. Reading Grade 6 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH315111	780	0.337	0.093	1.07	1.04
	VH315264	780	1.284	0.082	0.95	0.91
	VH315235	780	0.090	0.098	0.96	0.94
	VH315267	780	1.995	0.081	1.18	1.24
	VH315118	780	0.239	0.095	0.94	0.89
	VH315255	780	-0.038	0.101	0.85	0.68
	VH333639	780	1.305	0.082	1.09	1.10
	VH333641	780	1.105	0.084	0.98	0.93
	VH333670	780	2.252	0.082	1.08	1.16
	VH333667	780	1.558	0.081	1.02	1.06
	VH333689	780	1.806	0.081	1.00	1.01
	VH333685	780	2.474	0.083	1.23	1.44
	VH315975	780	0.388	0.092	0.97	0.88
	VH315969	780	0.119	0.097	0.90	0.93
2	VH315258	677	2.154	0.086	1.10	1.19
	VH315243	677	-0.873	0.143	0.92	0.82
	VH315247	677	0.797	0.094	0.93	0.89
	VH315124	677	0.743	0.095	1.10	1.14
	VH315256	677	0.090	0.108	1.00	1.15
	VH315090	677	0.797	0.094	1.09	1.18
	VH333622	677	2.124	0.086	1.00	1.03
	VH333607	677	0.954	0.092	1.02	1.01
	VH333674	677	2.449	0.088	1.08	1.20
	VH333657	677	2.707	0.089	1.22	1.43
	VH333679	677	1.796	0.087	1.00	1.02
	VH333693	677	1.455	0.088	0.97	0.98
	VH315492	677	2.620	0.089	1.06	1.09
	VH315438	677	1.416	0.088	1.08	1.06
3	VH325342	678	0.635	0.097	1.03	1.04
	VH325365	678	2.599	0.089	1.22	1.51
	VH325329	678	1.395	0.088	1.14	1.24
	VH325384	678	2.193	0.087	1.21	1.35
	VH325321	678	1.364	0.088	0.98	0.95
	VH325304	678	-0.837	0.142	0.96	0.85
	VH334475	678	1.843	0.086	1.07	1.07
	VH334630	678	1.254	0.089	1.06	1.09
	VH334617	678	-0.562	0.130	1.01	1.00
	VH334624	678	0.626	0.097	0.98	0.89
	VH334389	678	0.923	0.093	0.85	0.81
	VH334393	678	-0.703	0.136	0.90	0.95
	VH315446	678	1.534	0.087	0.99	0.95
	VH315409	678	0.983	0.092	1.03	1.01

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH325336	699	0.180	0.103	1.02	1.05
	VH325360	699	1.293	0.087	1.11	1.10
	VH325368	699	0.374	0.099	0.99	0.99
	VH325300	699	0.201	0.103	1.06	1.13
	VH325308	699	1.991	0.085	1.06	1.10
	VH325318	699	1.132	0.088	1.03	1.00
	VH334398	699	0.990	0.090	0.90	0.85
	VH334395	699	-0.254	0.115	0.87	0.72
	VH334465	699	1.147	0.088	1.03	1.06
	VH334480	699	-0.553	0.125	0.88	0.69
	VH334333	699	1.769	0.085	1.11	1.16
	VH334518	699	-0.898	0.139	0.89	0.77
	VH315458	699	0.590	0.095	0.98	1.12
	VH315977	699	0.254	0.101	0.95	0.99
5	VH315342	1388	0.919	0.064	0.99	0.96
	VH315340	705	0.597	0.094	0.94	0.84
	VH315337	1388	1.828	0.060	0.97	0.97
	VH315344	705	1.616	0.085	1.16	1.20
	VH315338	705	1.055	0.089	0.91	0.85
	VH315332	705	0.130	0.103	0.84	0.67
	VH332436	1388	0.312	0.071	0.98	0.98
	VH332428	705	0.214	0.102	0.96	0.83
	VH332505	705	1.673	0.085	0.98	0.95
	VH332506	705	1.916	0.084	1.22	1.36
	VH332419	705	0.450	0.097	0.94	0.85
	VH332442	705	-0.447	0.120	0.96	0.96
	VH315467	705	0.903	0.090	1.08	1.05
	VH315429	705	1.916	0.084	1.48	1.61
6	VH315342	1388	0.919	0.064	0.99	0.96
	VH315339	683	0.213	0.105	1.14	1.55
	VH315337	1388	1.828	0.060	0.97	0.97
	VH315341	683	-0.319	0.119	1.00	1.20
	VH315333	683	2.016	0.086	1.46	1.69
	VH315331	683	2.156	0.086	1.15	1.28
	VH332436	1388	0.312	0.071	0.98	0.98
	VH332430	683	1.935	0.086	1.07	1.11
	VH332500	683	1.556	0.087	1.02	1.04
	VH332518	683	1.548	0.087	0.92	0.90
	VH332413	683	1.349	0.088	0.86	0.80
	VH332423	683	1.616	0.087	1.10	1.11
	VH315964	683	2.038	0.086	0.95	0.93
	VH315400	683	0.019	0.109	0.79	0.61

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH318937	688	2.291	0.087	0.98	1.04
	VH318930	688	1.366	0.087	1.01	1.06
	VH319047	688	1.848	0.086	1.11	1.16
	VH318974	688	0.850	0.092	0.99	0.93
	VH319070	688	-0.245	0.114	0.85	0.69
	VH319062	688	1.297	0.088	0.97	0.97
	VH315315	688	0.524	0.096	0.88	0.78
	VH315307	688	1.366	0.087	1.06	1.05
	VH315304	688	1.642	0.086	0.98	1.00
	VH315313	688	0.908	0.091	0.92	0.90
	VH315323	688	1.576	0.086	0.95	0.93
	VH315325	688	1.142	0.089	1.07	1.07
	VH315967	688	1.087	0.089	0.94	0.89
	VH315960	688	2.127	0.086	1.00	0.98
8	VH318912	698	-0.476	0.124	0.84	0.73
	VH318924	698	0.337	0.101	0.88	0.78
	VH319038	698	2.294	0.085	1.28	1.43
	VH318954	698	1.116	0.089	1.06	1.05
	VH319079	698	0.608	0.096	1.01	0.99
	VH319083	698	-0.690	0.133	0.91	0.71
	VH315310	698	1.148	0.089	1.02	1.13
	VH315308	698	2.367	0.086	1.13	1.26
	VH315327	698	-0.003	0.109	0.85	0.70
	VH315318	698	1.557	0.086	1.06	1.05
	VH315299	698	-0.555	0.127	1.02	0.98
	VH315297	698	1.378	0.087	1.04	1.05
	VH315389	698	1.789	0.085	1.05	1.10
	VH315518	698	0.599	0.096	0.95	0.86
9	VH332485	700	0.433	0.100	0.93	0.90
	VH332558	700	3.594	0.101	1.26	2.16
	VH332542	700	1.740	0.085	1.10	1.18
	VH332497	1400	0.152	0.074	0.91	0.76
	VH332502	700	1.383	0.087	1.02	1.02
	VH332530	700	0.288	0.103	0.80	0.62
	VH333943	700	1.497	0.087	1.00	1.00
	VH334052	700	3.201	0.094	1.17	1.71
	VH334048	700	0.617	0.097	1.05	1.10
	VH333809	700	2.299	0.086	1.23	1.35
	VH333802	700	0.310	0.103	0.98	1.15
	VH334043	700	-0.008	0.111	0.90	0.75
	VH315499	700	-0.933	0.146	0.87	0.55
	VH315534	700	0.579	0.097	1.11	1.24

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH332536	700	2.463	0.086	1.16	1.39
	VH332513	700	0.597	0.095	0.94	0.94
	VH332550	700	0.136	0.104	0.83	0.68
	VH332497	1400	0.152	0.074	0.91	0.76
	VH332517	700	2.661	0.088	1.32	1.57
	VH332507	700	2.294	0.085	1.14	1.26
	VH333816	700	1.608	0.085	1.01	1.02
	VH333821	700	0.449	0.097	0.92	0.82
	VH333930	700	2.199	0.085	1.03	1.09
	VH334058	700	1.082	0.089	1.17	1.37
	VH333791	700	1.305	0.087	1.02	1.05
	VH333807	700	1.350	0.087	1.09	1.14
	VH315510	700	1.388	0.086	1.08	1.13
	VH315415	700	0.439	0.098	0.89	0.80

Table H5. Reading Grade 7 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH345749	782	2.029	0.080	0.99	0.97
	VH345746	782	1.623	0.081	0.82	0.74
	VH345769	782	2.048	0.080	1.05	1.05
	VH345766	782	2.194	0.080	1.13	1.18
	VH345774	782	1.023	0.087	0.92	0.85
	VH345780	782	2.328	0.080	1.15	1.21
	VH345485	782	1.584	0.081	1.30	1.45
	VH345516	782	0.524	0.095	0.90	0.74
	VH345440	782	2.449	0.080	1.18	1.22
	VH345538	782	0.664	0.092	0.86	0.67
	VH345519	1466	1.077	0.063	1.02	1.13
	VH345564	782	1.490	0.082	1.09	1.19
	VH316374	782	0.698	0.092	0.97	0.98
	VH316736	782	3.081	0.085	1.10	1.27
2	VH345750	684	2.326	0.084	1.09	1.12
	VH345753	684	3.431	0.093	1.12	1.40
	VH345764	684	2.523	0.084	1.03	1.07
	VH345771	684	1.231	0.091	0.94	0.87
	VH345778	684	1.918	0.084	1.04	1.04
	VH345775	684	3.272	0.090	1.26	1.60
	VH345493	684	0.178	0.114	1.01	1.02
	VH345528	684	2.067	0.084	1.13	1.16
	VH345518	684	1.560	0.087	0.98	0.93
	VH345544	684	2.573	0.084	1.08	1.15
	VH345519	1466	1.077	0.063	1.02	1.13
	VH345546	684	2.474	0.084	1.20	1.25
	VH316391	684	0.479	0.105	0.96	0.92
	VH316770	684	2.680	0.085	0.94	0.94
3	VH332947	703	1.935	0.085	1.01	1.00
	VH332935	703	0.979	0.094	0.98	0.95
	VH332922	703	1.703	0.086	0.93	0.87
	VH332809	703	-0.152	0.124	0.91	0.90
	VH332915	703	0.064	0.116	0.92	0.92
	VH332891	703	0.156	0.113	1.07	1.43
	VH332641	703	1.014	0.093	1.11	1.15
	VH332639	703	0.916	0.095	1.05	1.18
	VH332676	703	1.622	0.086	1.01	0.98
	VH332617	703	0.729	0.098	1.02	1.09
	VH332647	703	0.279	0.109	0.92	0.77
	VH332663	703	0.090	0.115	0.91	0.72
	VH316266	703	0.834	0.096	0.86	0.68
	VH316281	703	1.134	0.092	0.88	0.80

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH332910	699	-0.588	0.143	0.93	0.75
	VH332940	699	1.840	0.085	1.18	1.30
	VH332943	699	2.111	0.084	1.03	1.01
	VH332929	699	1.407	0.088	0.93	0.90
	VH332846	699	0.904	0.095	0.89	0.85
	VH332799	699	1.115	0.092	0.92	0.86
	VH332602	699	0.297	0.109	0.87	0.82
	VH332599	699	0.069	0.116	1.03	1.32
	VH332637	699	1.825	0.085	1.07	1.07
	VH332672	699	0.803	0.097	1.00	0.98
	VH332593	699	0.994	0.094	1.18	1.50
	VH332671	699	1.760	0.085	1.26	1.35
	VH316274	699	0.985	0.094	1.00	0.97
	VH316403	699	3.157	0.089	1.08	1.35
5	VH333944	705	0.536	0.102	0.92	0.86
	VH333950	705	0.687	0.099	0.95	0.96
	VH333917	705	2.280	0.083	1.21	1.28
	VH333931	705	0.135	0.113	0.82	0.58
	VH333961	705	2.447	0.084	1.06	1.09
	VH333908	705	1.910	0.084	1.03	1.02
	VH345770	705	1.215	0.090	0.97	0.93
	VH345765	705	2.085	0.084	1.26	1.35
	VH345895	705	3.289	0.090	1.20	1.49
	VH345777	705	1.133	0.091	0.90	0.96
	VH345913	705	2.911	0.086	1.28	1.46
	VH345904	705	0.608	0.100	0.95	0.91
	VH316398	705	1.041	0.092	0.93	0.89
	VH316457	705	0.419	0.105	0.87	0.70
6	VH333955	681	1.439	0.089	0.96	0.92
	VH333946	681	1.227	0.092	1.08	1.05
	VH333941	681	1.227	0.092	0.95	0.90
	VH333956	681	2.086	0.085	0.96	0.95
	VH333921	681	1.955	0.086	1.00	0.99
	VH333878	681	1.261	0.091	1.09	1.22
	VH345758	681	2.830	0.087	1.25	1.45
	VH345756	681	0.845	0.097	0.96	0.98
	VH345883	681	1.725	0.087	1.21	1.26
	VH345779	681	2.216	0.085	0.92	0.92
	VH345945	681	2.028	0.085	1.19	1.33
	VH348029	681	2.238	0.085	1.08	1.10
	VH316428	681	0.585	0.103	0.77	0.63
	VH316781	681	2.411	0.085	1.15	1.21

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH333919	709	1.530	0.086	1.27	1.39
	VH333959	709	1.705	0.085	1.21	1.32
	VH333954	709	1.855	0.084	1.04	1.08
	VH333965	709	1.734	0.085	1.00	0.99
	VH333880	709	3.070	0.088	1.29	1.56
	VH333859	709	2.349	0.083	0.95	0.94
	VH344777	709	1.255	0.089	1.18	1.24
	VH344780	709	0.844	0.095	0.91	0.77
	VH344754	709	1.448	0.087	0.97	0.92
	VH344763	709	1.684	0.085	0.96	0.93
	VH344765	709	0.366	0.106	0.94	0.90
	VH344749	709	1.387	0.087	0.94	0.89
	VH316423	709	0.950	0.093	0.90	0.85
	VH316807	709	1.418	0.087	0.91	0.91
8	VH333925	691	0.876	0.095	0.97	0.92
	VH333947	691	1.219	0.090	1.16	1.26
	VH333957	691	1.643	0.086	0.98	0.98
	VH333891	691	-0.611	0.144	0.88	0.61
	VH333964	691	0.180	0.112	0.87	0.90
	VH333847	691	1.825	0.085	1.02	1.00
	VH344770	691	1.043	0.093	1.11	1.34
	VH344752	691	1.628	0.086	1.01	0.99
	VH344764	691	1.275	0.089	1.01	1.00
	VH344775	691	2.428	0.084	1.10	1.15
	VH344760	691	2.315	0.084	1.16	1.24
	VH344746	691	3.026	0.088	1.05	1.23
	VH316751	691	1.440	0.088	0.99	0.97
	VH316466	691	1.432	0.088	0.95	1.04
9	VH334238	697	2.300	0.084	1.05	1.09
	VH334213	697	2.477	0.084	1.20	1.29
	VH334227	697	1.579	0.087	1.05	1.01
	VH334200	697	0.819	0.097	0.85	0.70
	VH334204	697	2.549	0.085	1.10	1.16
	VH334199	697	0.655	0.100	0.94	0.88
	VH332675	697	1.503	0.087	1.02	1.00
	VH332634	697	2.002	0.084	1.08	1.09
	VH332636	697	2.030	0.084	1.03	1.03
	VH332631	697	1.866	0.085	1.02	1.01
	VH332622	697	2.293	0.084	1.11	1.14
	VH332613	697	1.111	0.092	0.86	0.78
	VH316459	697	2.009	0.084	1.19	1.21
	VH316687	697	2.321	0.084	1.11	1.15

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH334205	699	3.008	0.088	1.03	1.17
	VH334219	699	1.805	0.085	0.99	0.99
	VH334209	699	3.109	0.089	1.37	1.65
	VH334223	699	1.783	0.085	1.10	1.09
	VH334242	699	2.977	0.087	0.96	1.02
	VH334216	699	2.309	0.084	1.24	1.33
	VH332626	699	2.507	0.084	1.13	1.17
	VH332673	699	1.373	0.088	0.88	0.83
	VH332677	699	2.651	0.085	1.02	1.04
	VH332618	699	1.108	0.091	0.95	0.93
	VH332628	699	2.408	0.084	1.09	1.15
	VH332678	699	1.681	0.086	0.90	0.85
	VH316709	699	2.797	0.086	0.94	0.95
	VH316467	699	1.991	0.084	1.01	1.01

Table H6. Reading Grade 8 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH345147	779	0.521	0.097	0.86	0.70
	VH345177	779	1.825	0.080	1.13	1.19
	VH345151	779	1.488	0.083	0.95	0.94
	VH345172	779	0.252	0.103	0.92	0.80
	VH345124	779	1.322	0.084	0.91	0.88
	VH345121	779	1.688	0.081	0.93	0.91
	VH342979	779	1.155	0.086	0.91	0.83
	VH342994	779	3.612	0.091	1.33	1.98
	VH342980	779	2.274	0.079	1.07	1.12
	VH343009	779	2.098	0.079	1.13	1.17
	VH342990	779	1.343	0.084	0.95	0.89
	VH342964	779	1.199	0.085	0.94	0.86
	VH325149	779	0.973	0.089	0.86	0.76
	VH325228	779	3.020	0.083	1.34	1.65
2	VH345141	660	2.742	0.086	1.09	1.17
	VH345135	660	0.862	0.100	0.87	0.74
	VH345170	660	-0.298	0.139	0.85	0.59
	VH345181	660	0.801	0.102	0.83	0.70
	VH345175	660	0.503	0.109	0.88	0.76
	VH345154	660	1.827	0.087	1.00	1.00
	VH342968	660	0.329	0.114	0.98	0.92
	VH342983	660	1.380	0.091	0.98	0.93
	VH342971	660	2.786	0.086	1.28	1.42
	VH343003	660	1.445	0.090	0.93	0.91
	VH342981	660	2.301	0.085	1.09	1.12
	VH343004	660	1.027	0.097	0.92	0.85
	VH325227	660	2.853	0.086	0.97	1.00
	VH325157	660	2.071	0.085	1.12	1.15
3	VH343387	666	2.483	0.085	1.23	1.28
	VH343380	666	0.949	0.098	0.90	0.87
	VH343529	666	2.346	0.085	0.98	1.02
	VH343514	666	1.706	0.088	0.91	0.85
	VH350931	666	3.351	0.091	1.24	1.58
	VH343550	666	2.317	0.085	1.10	1.12
	VH343222	666	0.231	0.118	0.80	0.62
	VH343231	666	2.332	0.085	1.00	1.02
	VH343179	666	2.209	0.085	1.03	1.04
	VH343239	666	2.310	0.085	0.94	0.94
	VH343219	666	1.145	0.095	0.90	0.78
	VH343082	666	2.332	0.085	1.22	1.30
	VH325233	666	1.395	0.091	0.97	0.91
	VH325123	666	2.346	0.085	1.03	1.09

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH343422	675	2.800	0.086	0.98	1.07
	VH343313	675	1.202	0.094	1.03	1.11
	VH343521	675	2.904	0.086	1.39	1.62
	VH343503	675	1.229	0.093	0.85	0.73
	VH343545	675	2.713	0.085	1.01	1.03
	VH343538	675	2.590	0.085	1.10	1.17
	VH343053	675	3.039	0.087	0.98	1.07
	VH343235	675	1.999	0.086	1.17	1.22
	VH343073	675	2.396	0.085	1.03	1.04
	VH343148	675	1.439	0.090	0.88	0.83
	VH343047	675	1.631	0.088	1.01	0.99
	VH342960	675	-0.122	0.131	0.97	0.82
	VH325091	675	1.544	0.089	1.07	1.03
	VH325218	675	0.310	0.115	0.97	1.10
5	VH326366	679	0.463	0.110	0.97	0.94
	VH326386	679	1.801	0.086	1.00	0.97
	VH326377	679	2.318	0.084	1.03	1.03
	VH326381	679	0.965	0.098	1.06	1.07
	VH326380	679	2.134	0.084	0.89	0.88
	VH326372	679	1.131	0.095	0.94	0.86
	VH343939	679	3.383	0.091	1.02	1.18
	VH343960	679	2.027	0.085	1.14	1.21
	VH344147	679	2.636	0.084	1.01	1.10
	VH344131	679	2.297	0.084	1.08	1.12
	VH344167	679	3.098	0.087	1.43	1.77
	VH344176	679	0.625	0.106	0.91	0.78
	VH325164	679	2.318	0.084	1.08	1.12
	VH325144	679	1.202	0.093	1.11	1.15
6	VH326376	656	2.643	0.086	1.18	1.32
	VH326384	656	1.691	0.090	0.88	0.80
	VH326379	656	1.552	0.091	0.98	0.91
	VH326388	656	2.206	0.086	1.03	1.05
	VH326374	656	1.658	0.090	1.12	1.20
	VH326361	656	0.235	0.120	0.86	0.60
	VH343950	656	3.930	0.101	1.20	1.61
	VH343876	656	1.903	0.088	0.93	0.91
	VH344092	656	1.965	0.087	1.01	0.97
	VH344067	656	2.041	0.087	1.05	1.07
	VH344180	656	2.968	0.088	1.06	1.18
	VH344153	656	3.069	0.089	1.19	1.41
	VH325199	656	0.848	0.103	0.84	0.69
	VH325171	656	1.618	0.090	0.98	0.94

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH345018	653	3.143	0.089	1.11	1.27
	VH345022	653	0.826	0.103	0.96	0.92
	VH345004	653	0.323	0.117	0.83	0.61
	VH345012	653	0.530	0.111	0.83	0.69
	VH345026	653	1.678	0.089	1.23	1.33
	VH345028	653	-0.079	0.133	0.86	0.57
	VH343135	653	1.467	0.091	1.06	1.03
	VH343132	653	2.585	0.085	0.97	0.98
	VH343191	653	1.161	0.096	0.98	0.96
	VH343196	653	1.501	0.091	0.95	0.89
	VH343126	653	0.336	0.117	0.87	0.67
	VH343129	653	1.278	0.094	1.29	1.76
	VH325186	653	3.120	0.089	0.97	1.04
	VH325211	653	2.052	0.086	1.10	1.13
8	VH344994	679	0.128	0.120	0.93	0.88
	VH344998	679	0.382	0.112	0.97	0.88
	VH345031	679	1.279	0.091	1.08	1.13
	VH345020	679	0.266	0.115	0.89	0.68
	VH345016	679	1.865	0.085	0.98	0.93
	VH344978	679	1.930	0.085	1.23	1.32
	VH343193	679	1.930	0.085	1.00	0.98
	VH343184	679	1.763	0.086	1.01	1.00
	VH343176	679	1.337	0.091	0.97	0.95
	VH343186	679	2.528	0.084	1.11	1.22
	VH343124	679	1.836	0.085	1.04	1.04
	VH343115	679	0.913	0.098	0.95	0.80
	VH325231	679	3.188	0.088	0.97	1.03
	VH325195	679	1.394	0.090	0.93	0.90
9	VH345688	667	1.539	0.089	0.86	0.79
	VH345726	667	-0.601	0.154	0.92	0.63
	VH345730	667	1.750	0.087	0.97	0.93
	VH345706	667	2.331	0.085	1.04	1.07
	VH345694	667	0.712	0.104	1.02	1.08
	VH345734	667	1.900	0.086	0.90	0.86
	VH343247	667	1.795	0.087	0.96	0.94
	VH343240	667	2.331	0.085	0.99	1.02
	VH343260	667	1.602	0.089	0.93	0.88
	VH343230	667	0.887	0.100	0.90	0.80
	VH343234	667	1.696	0.088	1.01	1.01
	VH343263	667	1.750	0.087	1.19	1.33
	VH325216	667	3.975	0.103	1.27	1.94
	VH325130	667	0.498	0.109	0.87	0.72

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH345686	679	2.985	0.087	1.06	1.20
	VH345723	679	1.667	0.087	1.10	1.13
	VH345735	679	2.383	0.084	1.06	1.12
	VH345718	679	3.735	0.098	1.14	1.61
	VH345703	679	3.114	0.088	1.13	1.24
	VH345733	679	2.657	0.084	1.04	1.08
	VH343254	679	1.077	0.094	1.01	0.95
	VH343257	679	-0.225	0.132	0.86	0.77
	VH343232	679	1.429	0.089	1.06	1.07
	VH343250	679	0.987	0.096	0.89	0.82
	VH343237	679	1.445	0.089	1.00	1.02
	VH343224	679	1.667	0.087	1.28	1.45
	VH325220	679	1.902	0.085	1.00	0.99
	VH325191	679	1.113	0.094	0.94	0.87

Mathematics

Table H7. Mathematics Grade 3 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH327773	836	1.245	0.080	1.01	1.08
	VH331639	836	2.571	0.099	1.10	1.67
	VH327891	836	1.005	0.079	0.99	1.01
	VH326762	836	0.992	0.079	0.95	0.98
	VH326729	836	0.659	0.078	1.05	1.06
	VH337124	836	0.202	0.079	1.17	1.23
	VH327898	836	0.629	0.078	1.04	1.10
	VH318564	836	1.323	0.081	1.02	1.09
	VH336404	836	2.078	0.089	1.43	2.04
	VH318704	836	1.131	0.080	1.16	1.22
	VH303943	836	-0.593	0.086	0.84	0.71
	VH303635	836	-0.706	0.088	1.04	1.14
2	VH318837	775	1.336	0.082	1.02	1.07
	VH303570	775	0.887	0.081	1.31	1.43
	VH327878	775	1.390	0.083	1.16	1.34
	VH336409	775	-1.506	0.113	0.99	0.95
	VH326721	775	0.998	0.081	1.15	1.20
	VH336574	775	1.989	0.089	1.21	1.41
	VH273044	775	1.752	0.086	1.12	1.41
	VH331638	775	-0.423	0.088	1.00	1.05
	VH303611	775	0.361	0.081	0.97	0.96
	VH273115	775	0.745	0.080	1.08	1.06
	VH327000	775	2.699	0.104	1.01	1.16
	VH303869	775	0.972	0.081	1.09	1.15
3	VH318818	781	0.226	0.082	0.93	0.85
	VH310953	781	2.917	0.109	1.00	1.50
	VH303648	781	-0.547	0.090	0.98	1.01
	VH326759	781	1.194	0.081	1.05	1.13
	VH310966	781	0.952	0.081	1.07	1.10
	VH326993	781	0.945	0.081	1.05	1.05
	VH327895	781	2.154	0.091	1.16	1.49
	VH318584	781	0.412	0.081	0.99	0.96
	VH303879	781	1.906	0.088	1.29	1.73
	VH303664	781	1.148	0.081	1.01	1.07
	VH310884	781	0.485	0.081	1.10	1.13
	VH326786	781	1.506	0.083	1.38	1.62

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH273074	776	0.122	0.082	1.12	1.14
	VH336532	776	2.019	0.090	1.46	2.07
	VH327797	776	1.205	0.082	1.11	1.18
	VH303523	776	0.858	0.080	1.05	1.08
	VH331640	776	1.066	0.081	1.01	1.04
	VH327871	776	-0.021	0.083	0.95	0.91
	VH326997	776	-0.167	0.084	1.26	1.38
	VH336381	776	0.314	0.081	1.33	1.49
	VH273083	776	1.599	0.085	1.29	1.53
	VH318718	776	0.386	0.081	0.93	0.91
	VH304079	776	1.570	0.085	1.13	1.43
	VH310932	776	1.185	0.082	1.02	1.07
5	VH303507	769	-0.735	0.093	0.86	0.74
	VH310964	769	2.215	0.094	0.98	1.10
	VH327801	769	0.033	0.083	1.09	1.06
	VH335375	769	0.971	0.081	1.42	1.58
	VH310863	769	0.984	0.081	1.05	1.07
	VH318631	769	0.231	0.082	0.89	0.81
	VH327913	769	1.206	0.082	1.12	1.19
	VH327886	769	0.898	0.081	1.12	1.17
	VH336373	769	2.026	0.091	1.22	1.69
	VH327903	769	0.392	0.081	0.85	0.85
	VH326797	769	0.806	0.081	1.15	1.19
	VH303698	769	1.240	0.083	1.29	1.51
6	VH327839	772	1.279	0.083	1.03	1.11
	VH310911	772	2.413	0.098	1.39	2.01
	VH327815	772	0.655	0.081	1.04	1.10
	VH336427	772	0.088	0.083	1.23	1.48
	VH327922	772	1.095	0.082	1.13	1.22
	VH331717	772	0.060	0.084	0.97	0.91
	VH326974	772	1.748	0.087	0.83	0.76
	VH318645	772	0.542	0.082	0.86	0.81
	VH303657	772	0.281	0.082	1.13	1.12
	VH273148	772	0.039	0.084	0.96	0.93
	VH326790	772	-0.784	0.094	1.08	1.11
	VH310857	772	2.376	0.097	1.24	1.47

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH327819	757	-0.599	0.092	0.88	0.73
	VH310958	757	2.694	0.104	1.22	2.11
	VH303596	757	2.342	0.097	1.36	2.06
	VH335365	757	-1.166	0.103	1.03	1.17
	VH326990	757	0.612	0.083	0.90	0.86
	VH273027	757	1.518	0.086	1.10	1.27
	VH303935	757	2.380	0.097	0.94	0.95
	VH273095	757	2.188	0.094	1.26	2.10
	VH310949	757	-1.870	0.126	1.04	1.27
	VH326794	757	1.309	0.084	1.28	1.47
	VH303950	757	-0.283	0.088	0.83	0.71
	VH336376	757	1.416	0.085	1.12	1.22
8	VH273033	736	2.587	0.104	1.11	1.59
	VH303643	736	0.336	0.085	0.89	0.83
	VH327923	736	0.393	0.084	0.96	0.94
	VH303560	736	1.202	0.085	1.31	1.48
	VH318657	736	1.122	0.085	1.01	1.10
	VH272617	736	2.523	0.102	1.10	1.50
	VH310937	736	0.507	0.084	1.04	1.11
	VH272624	736	-0.343	0.090	0.92	0.91
	VH326761	736	2.361	0.099	1.13	1.69
	VH273103	736	0.817	0.084	1.10	1.14
	VH303940	736	-0.161	0.088	0.97	0.87
	VH304127	736	0.754	0.084	1.31	1.38
9	VH318768	741	1.379	0.085	1.07	1.19
	VH326465	741	1.802	0.089	1.06	1.33
	VH303624	741	-0.021	0.087	0.97	0.94
	VH335356	741	1.930	0.090	1.30	1.63
	VH318803	741	0.362	0.085	1.12	1.12
	VH303949	741	-0.679	0.096	0.94	0.83
	VH272510	741	2.978	0.111	0.91	1.05
	VH318734	741	0.114	0.086	1.05	1.24
	VH303681	741	0.617	0.084	1.24	1.34
	VH318609	741	0.246	0.085	0.98	1.02
	VH326964	741	1.971	0.091	1.13	1.42
	VH310934	741	-1.137	0.106	0.90	0.93

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH318777	735	0.592	0.083	1.01	1.00
	VH303651	735	-0.405	0.090	0.93	0.90
	VH303554	735	0.895	0.083	1.14	1.24
	VH335372	735	-0.018	0.086	1.30	1.57
	VH326984	735	0.186	0.085	1.00	0.95
	VH318334	735	-2.212	0.143	0.92	0.95
	VH272517	735	-0.246	0.088	0.98	0.95
	VH310928	735	0.099	0.085	1.19	1.34
	VH273136	735	0.453	0.083	0.90	0.86
	VH273090	735	1.479	0.086	1.57	2.17
	VH303946	735	-0.293	0.089	0.95	0.94
	VH303629	735	0.026	0.086	0.91	0.84

Table H8. Mathematics Grade 4 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH327010	820	2.037	0.080	0.92	0.91
	VH331635	820	2.395	0.082	1.17	1.31
	VH313887	820	3.033	0.090	1.16	1.55
	VH305767	820	1.980	0.079	1.16	1.28
	VH305882	820	1.418	0.078	0.99	1.01
	VH313961	820	2.268	0.081	1.12	1.23
	VH326696	820	2.673	0.085	1.13	1.31
	VH327015	820	3.219	0.094	1.14	1.64
	VH335023	820	2.623	0.084	1.30	1.70
	VH326687	820	2.139	0.080	1.09	1.13
	VH329506	820	1.614	0.078	0.93	0.91
	VH317610	820	0.979	0.080	0.91	0.93
2	VH305896	754	1.967	0.082	0.84	0.80
	VH305749	754	4.695	0.143	1.14	2.88
	VH313944	754	2.162	0.082	1.07	1.11
	VH331636	754	0.716	0.087	0.96	0.92
	VH334941	754	2.183	0.083	1.04	1.06
	VH334950	754	2.272	0.083	1.10	1.17
	VH326661	754	1.085	0.083	1.32	1.51
	VH317460	754	0.769	0.086	0.92	0.83
	VH329325	754	0.671	0.088	0.89	0.80
	VH317521	754	1.324	0.082	1.21	1.40
	VH306121	754	3.178	0.094	1.33	1.73
	VH326676	754	2.169	0.083	1.05	1.11
3	VH335221	746	0.782	0.087	1.10	1.21
	VH331152	746	3.044	0.093	1.22	1.69
	VH313929	746	1.821	0.082	1.05	1.06
	VH327013	746	2.418	0.085	0.94	1.00
	VH306097	746	1.139	0.084	0.86	0.79
	VH313822	746	2.052	0.083	0.97	0.98
	VH326691	746	2.368	0.085	1.27	1.41
	VH317762	746	4.137	0.119	1.05	1.40
	VH313766	746	1.720	0.082	1.08	1.14
	VH319829	746	0.960	0.085	0.95	0.95
	VH335019	746	3.104	0.094	1.44	2.10
	VH317589	746	1.061	0.084	0.92	0.91

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH331156	742	2.526	0.086	1.19	1.37
	VH326717	742	3.182	0.096	1.03	1.27
	VH326999	742	2.886	0.091	0.97	1.19
	VH313814	742	3.219	0.097	1.19	1.61
	VH305890	742	1.883	0.082	1.18	1.26
	VH326671	742	1.768	0.082	1.18	1.25
	VH327038	742	2.886	0.091	1.12	1.42
	VH317816	742	1.384	0.082	0.90	0.86
	VH305918	742	2.054	0.083	1.14	1.22
	VH313880	742	2.609	0.087	1.02	1.15
	VH319812	742	1.694	0.082	1.13	1.20
	VH317601	742	1.661	0.082	0.88	0.85
5	VH313730	733	2.860	0.090	1.21	1.47
	VH313920	733	1.956	0.083	1.01	1.00
	VH331634	733	2.542	0.086	0.94	1.03
	VH334637	733	2.749	0.089	1.25	1.53
	VH317504	733	0.568	0.090	0.86	0.80
	VH317471	733	2.550	0.086	1.08	1.17
	VH326668	733	3.591	0.104	1.08	1.39
	VH317809	733	2.107	0.083	1.22	1.27
	VH306130	733	1.481	0.083	1.11	1.13
	VH329260	733	1.610	0.082	0.82	0.77
	VH334982	733	1.563	0.082	1.11	1.12
	VH313889	733	1.569	0.082	1.12	1.15
6	VH334652	738	0.631	0.090	1.05	1.21
	VH326714	738	1.832	0.083	1.09	1.09
	VH305785	738	1.304	0.084	1.10	1.17
	VH327006	738	2.240	0.085	0.93	0.96
	VH305906	738	1.009	0.086	0.93	0.88
	VH317694	738	3.545	0.102	1.00	1.14
	VH327067	738	1.853	0.083	0.94	0.91
	VH317724	738	1.569	0.083	0.88	0.80
	VH306127	738	0.759	0.088	0.86	0.78
	VH317577	738	0.957	0.086	0.90	0.84
	VH329431	738	-0.691	0.123	0.88	0.62
	VH313898	738	1.147	0.085	1.07	1.22

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH313955	743	2.466	0.086	1.12	1.21
	VH305760	743	1.602	0.082	1.01	0.99
	VH326711	743	2.512	0.087	1.16	1.34
	VH327002	743	0.057	0.097	0.96	0.79
	VH317685	743	1.615	0.082	0.94	0.91
	VH317539	743	0.140	0.095	0.94	0.82
	VH326680	743	2.504	0.087	1.23	1.40
	VH327022	743	1.683	0.082	1.20	1.30
	VH319890	743	-1.111	0.136	0.95	0.85
	VH327046	743	1.582	0.082	0.96	0.96
	VH313751	743	1.029	0.084	0.91	0.87
	VH317629	743	0.681	0.087	0.98	0.94
8	VH331637	746	0.924	0.085	1.01	1.01
	VH326998	746	1.392	0.083	1.01	1.00
	VH326709	746	2.417	0.086	0.96	1.01
	VH335041	746	3.455	0.103	1.29	1.90
	VH305790	746	2.357	0.086	0.99	0.96
	VH317703	746	0.651	0.087	1.01	1.08
	VH317782	746	1.821	0.083	0.83	0.79
	VH327054	746	1.965	0.083	0.94	0.93
	VH313740	746	1.616	0.082	1.08	1.10
	VH306155	746	2.816	0.091	1.12	1.34
	VH317741	746	1.473	0.082	1.04	1.00
	VH317801	746	2.833	0.091	1.00	1.07
9	VH313837	735	2.579	0.087	1.33	1.57
	VH331160	735	2.228	0.084	1.25	1.46
	VH313967	735	2.813	0.090	1.12	1.32
	VH335038	735	2.045	0.083	1.21	1.29
	VH327062	735	2.271	0.085	0.92	0.93
	VH305909	735	2.038	0.083	0.98	1.02
	VH317710	735	1.564	0.083	0.95	0.95
	VH319860	735	0.043	0.100	0.83	0.68
	VH306115	735	0.898	0.086	0.87	0.79
	VH335031	735	2.686	0.088	1.21	1.43
	VH319897	735	-0.515	0.115	0.96	0.85
	VH317772	735	1.674	0.083	1.12	1.15

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH335034	715	1.772	0.083	1.23	1.31
	VH313801	715	3.142	0.096	1.22	1.63
	VH305755	715	1.537	0.083	0.92	0.89
	VH331141	715	2.689	0.089	1.18	1.45
	VH327030	715	1.862	0.083	1.06	1.09
	VH317690	715	0.768	0.088	0.96	0.90
	VH317513	715	1.779	0.083	1.00	1.01
	VH305784	715	3.068	0.095	1.27	1.70
	VH306147	715	2.778	0.090	1.28	1.62
	VH329239	715	-0.508	0.117	0.95	0.92
	VH317729	715	2.229	0.085	1.04	1.09
	VH314029	715	2.877	0.092	1.17	1.44

Table H9. Mathematics Grade 5 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH319261	813	3.136	0.088	1.23	1.48
	VH335026	813	2.768	0.084	1.29	1.47
	VH332183	813	1.898	0.079	1.30	1.37
	VH333962	813	2.882	0.085	1.05	1.16
	VH314306	813	3.480	0.093	1.16	1.62
	VH326916	813	1.728	0.079	0.94	0.91
	VH336215	813	3.596	0.095	1.25	1.78
	VH326780	813	1.595	0.080	0.92	0.92
	VH334660	813	2.896	0.085	1.18	1.33
	VH319641	813	2.081	0.080	1.19	1.28
	VH092639	813	1.195	0.081	0.92	0.84
	VH335003	813	4.211	0.110	1.30	2.32
2	VH326904	728	-0.085	0.116	1.01	1.21
	VH336104	728	2.286	0.085	1.26	1.52
	VH332140	728	1.497	0.087	0.90	0.84
	VH317267	728	2.466	0.085	0.98	0.97
	VH334657	728	3.536	0.096	1.08	1.28
	VH334292	728	2.488	0.085	1.14	1.26
	VH314260	728	0.675	0.097	1.14	1.40
	VH319556	728	2.221	0.085	0.89	0.92
	VH326751	728	0.291	0.105	0.98	1.19
	VH334986	728	3.018	0.089	1.43	1.68
	VH332121	728	2.350	0.085	1.05	1.11
	VH314191	728	1.913	0.085	0.97	0.93
3	VH317107	712	2.739	0.086	1.12	1.14
	VH334073	712	3.688	0.098	1.24	1.63
	VH335036	712	2.563	0.085	1.17	1.22
	VH334937	712	2.078	0.084	0.94	0.92
	VH314297	712	0.887	0.093	1.06	1.10
	VH314361	712	3.046	0.089	1.31	1.49
	VH317292	712	2.376	0.085	1.17	1.20
	VH326804	712	3.910	0.103	1.38	2.21
	VH095032	712	2.071	0.084	0.95	0.92
	VH319663	712	2.498	0.085	1.30	1.42
	VH336288	712	4.156	0.109	1.01	1.39
	VH334942	712	2.952	0.088	1.11	1.21

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH314353	712	2.640	0.086	1.17	1.25
	VH092839	712	2.067	0.084	1.17	1.27
	VH334324	712	3.814	0.102	1.22	1.91
	VH326935	712	4.065	0.107	1.23	1.90
	VH319572	712	0.926	0.093	1.09	1.10
	VH326775	712	0.458	0.102	0.86	0.79
	VH317339	712	2.201	0.084	1.20	1.29
	VH336229	712	2.159	0.084	1.00	1.01
	VH332193	712	2.351	0.085	1.11	1.20
	VH336353	712	2.685	0.086	1.34	1.47
	VH335051	712	1.889	0.084	1.03	0.99
	VH333915	712	3.150	0.091	1.14	1.23
5	VH332144	716	2.751	0.087	1.10	1.20
	VH319266	716	2.805	0.088	1.17	1.29
	VH317013	716	0.325	0.102	1.01	1.21
	VH334396	716	3.341	0.094	1.27	1.60
	VH326943	716	2.287	0.085	1.13	1.18
	VH314337	716	3.368	0.095	1.25	1.63
	VH319583	716	1.647	0.085	1.07	1.14
	VH335048	716	1.226	0.088	1.03	1.03
	VH333922	716	4.728	0.131	1.31	3.27
	VH336339	716	3.184	0.092	1.05	1.23
	VH326796	716	1.918	0.084	0.99	1.01
	VH314178	716	0.989	0.090	1.01	1.04
6	VH332143	745	1.738	0.083	0.82	0.80
	VH319433	745	3.646	0.097	1.21	1.86
	VH334079	745	3.723	0.099	1.31	2.01
	VH314214	745	0.686	0.094	1.03	1.46
	VH314320	745	3.267	0.091	1.39	1.82
	VH326965	745	1.827	0.083	1.10	1.09
	VH317122	745	2.012	0.083	1.05	1.07
	VH326785	745	3.004	0.088	1.24	1.46
	VH336088	745	2.981	0.087	1.07	1.16
	VH319630	745	2.182	0.083	1.19	1.30
	VH336271	745	3.544	0.095	0.92	1.05
	VH335044	745	2.794	0.086	1.02	1.02

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH332139	729	2.109	0.083	0.88	0.87
	VH317020	729	1.733	0.084	1.00	1.09
	VH326987	729	2.193	0.084	1.28	1.40
	VH333949	729	3.204	0.091	1.14	1.37
	VH314314	729	3.563	0.097	1.40	1.95
	VH326748	729	1.612	0.084	1.11	1.21
	VH335040	729	1.789	0.084	1.13	1.14
	VH319654	729	3.658	0.098	1.00	1.11
	VH317145	729	1.977	0.083	0.92	0.88
	VH326811	729	3.463	0.095	1.19	1.50
	VH099656	729	2.172	0.083	1.10	1.12
	VH334282	729	3.535	0.096	1.06	1.27
8	VH334959	722	3.656	0.098	1.03	1.27
	VH319473	722	2.423	0.084	1.32	1.48
	VH317046	722	3.686	0.099	1.19	1.51
	VH334354	722	3.686	0.099	0.92	0.99
	VH314286	722	3.085	0.089	1.21	1.34
	VH336361	722	1.996	0.083	1.18	1.26
	VH317322	722	3.053	0.089	1.16	1.33
	VH332141	722	2.472	0.084	1.00	0.99
	VH334313	722	2.133	0.083	1.07	1.08
	VH319532	722	2.291	0.083	1.04	1.05
	VH333894	722	2.153	0.083	1.17	1.19
	VH326813	722	5.274	0.156	1.03	1.77
9	VH334973	722	1.627	0.086	1.12	1.20
	VH317141	722	1.152	0.090	1.00	0.92
	VH334626	722	4.018	0.105	1.01	0.98
	VH314274	722	1.708	0.086	1.02	1.00
	VH332207	722	0.850	0.094	0.91	0.84
	VH326774	722	1.286	0.088	0.92	0.88
	VH317303	722	3.099	0.090	1.07	1.23
	VH104496	722	3.239	0.092	1.43	1.78
	VH332126	722	1.918	0.085	0.85	0.80
	VH319545	722	3.942	0.104	1.05	1.33
	VH326899	722	1.255	0.089	1.00	1.05
	VH314204	722	3.091	0.090	1.01	1.08

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH334348	719	2.864	0.088	1.09	1.20
	VH335010	719	1.333	0.087	0.93	0.90
	VH317138	719	2.872	0.088	1.08	1.14
	VH314281	719	2.007	0.084	1.03	1.05
	VH334373	719	3.480	0.096	1.22	1.69
	VH335006	719	2.989	0.089	1.09	1.22
	VH336248	719	4.050	0.107	1.17	1.63
	VH326722	719	1.052	0.090	1.03	1.11
	VH333824	719	1.195	0.088	1.34	1.67
	VH326814	719	1.272	0.088	0.95	0.92
	VH332138	719	2.163	0.084	0.89	0.86
	VH319509	719	2.958	0.089	1.31	1.55

Table H10. Mathematics Grade 6 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH135854	824	2.624	0.078	1.22	1.30
	VF862767	824	1.572	0.080	1.00	1.00
	VH316009	824	3.295	0.084	1.06	1.15
	VF882940	824	2.163	0.077	0.96	0.95
	VH331472	824	2.916	0.080	1.15	1.21
	VH329926	824	0.910	0.088	0.94	0.87
	VH333341	824	2.509	0.078	1.34	1.47
	VH314512	824	2.353	0.077	0.92	0.90
	VH331938	824	3.091	0.081	1.15	1.27
	VH329962	824	3.309	0.084	0.98	1.11
	VH330025	824	3.747	0.090	1.35	1.82
	VH332000	824	2.747	0.079	1.14	1.20
2	VH316001	682	1.977	0.087	0.92	0.92
	VH319189	682	2.752	0.086	0.99	0.98
	VH320665	682	2.334	0.086	1.03	1.04
	VH326082	682	2.414	0.086	1.04	1.04
	VH333594	682	2.976	0.087	0.93	0.89
	VH314593	682	3.501	0.092	1.18	1.38
	VH329993	682	2.097	0.086	0.89	0.83
	VH330017	682	2.260	0.086	1.08	1.09
	VH136338	682	2.356	0.086	1.17	1.26
	VH331623	682	2.082	0.086	1.03	1.03
	VH314476	682	3.442	0.091	0.94	1.06
	VH329597	682	2.976	0.087	0.90	0.93
3	VH320741	684	0.830	0.103	0.93	0.87
	VH316017	684	2.319	0.085	0.94	0.90
	VH325587	684	2.876	0.086	0.93	0.93
	VH319122	684	2.363	0.085	0.82	0.77
	VH330042	684	3.011	0.087	1.35	1.50
	VH331487	684	2.385	0.085	1.04	1.06
	VH332053	684	4.147	0.103	1.21	1.53
	VH331630	684	2.501	0.085	1.40	1.66
	VH329607	684	1.775	0.088	1.01	1.01
	VH331620	684	2.515	0.085	1.01	0.99
	VH331633	684	3.189	0.089	1.19	1.28
	VH314495	684	2.816	0.086	1.00	1.00

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH326139	681	2.494	0.084	1.15	1.21
	VH319131	681	2.852	0.084	0.89	0.90
	VH316021	681	4.232	0.104	1.37	2.20
	VF741705	681	1.509	0.090	0.99	0.98
	VH329539	681	1.676	0.088	0.88	0.85
	VH314392	681	4.356	0.108	0.99	1.27
	VH332012	681	2.641	0.084	1.02	1.04
	VH331627	681	3.909	0.097	1.27	1.55
	VH331936	681	4.167	0.103	1.35	1.92
	VH329981	681	2.171	0.084	0.96	0.97
	VH314542	681	3.054	0.086	1.19	1.28
	VH332001	681	4.054	0.100	1.15	1.39
5	VH325593	688	4.177	0.105	1.12	1.45
	VH319203	688	1.957	0.085	1.15	1.25
	VF741703	688	2.659	0.084	1.11	1.14
	VH325591	688	2.926	0.086	1.11	1.16
	VH331624	688	5.500	0.157	1.24	2.36
	VH314559	688	3.237	0.089	1.20	1.39
	VH331547	688	3.365	0.090	1.09	1.11
	VH331985	688	2.911	0.086	1.04	1.06
	VH331946	688	3.106	0.087	1.06	1.10
	VH331521	688	3.023	0.087	1.18	1.25
	VH329584	688	2.510	0.084	0.97	1.00
	VH314487	688	3.175	0.088	1.02	1.06
6	VH121079	702	1.329	0.091	1.14	1.32
	VH118695	702	3.055	0.085	1.07	1.09
	VH326102	702	3.287	0.088	1.19	1.32
	VH320813	702	3.026	0.085	1.03	1.07
	VH314584	702	2.437	0.083	0.96	0.94
	VH329936	702	3.396	0.089	1.16	1.30
	VH331988	702	3.557	0.091	1.08	1.20
	VH331980	702	3.181	0.087	1.05	1.13
	VH331626	702	3.419	0.089	1.18	1.27
	VH333348	702	4.535	0.112	1.13	1.38
	VH314517	702	2.819	0.084	1.16	1.20
	VH330051	702	2.847	0.084	0.98	1.00

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH326121	673	3.532	0.092	1.19	1.33
	VH319180	673	2.925	0.086	1.04	1.07
	VH325579	673	2.531	0.085	1.06	1.08
	VH121052	673	3.277	0.089	1.09	1.17
	VH331968	673	3.333	0.090	1.30	1.48
	VH329789	673	1.992	0.086	1.00	1.06
	VH333625	673	4.083	0.102	1.08	1.37
	VH331594	673	3.854	0.097	1.16	1.53
	VH331619	673	3.206	0.088	1.02	1.04
	VH332010	673	3.671	0.094	1.08	1.27
	VH314506	673	2.523	0.085	0.91	0.87
	VH330012	673	2.444	0.085	1.12	1.13
8	VH316165	686	2.818	0.085	0.98	0.99
	VH316030	686	2.890	0.085	1.11	1.16
	VH320829	686	2.905	0.085	0.96	0.94
	VH319153	686	1.856	0.086	0.95	0.91
	VH141195	686	3.684	0.093	1.34	1.67
	VH333343	686	3.845	0.096	0.96	1.07
	VH331992	686	5.024	0.131	1.16	1.91
	VH314590	686	2.912	0.085	1.07	1.11
	VH331625	686	3.701	0.093	1.32	1.51
	VH332080	686	3.632	0.092	1.24	1.46
	VH314597	686	3.474	0.090	1.08	1.17
	VH331585	686	2.919	0.085	1.11	1.16
9	VF862661	687	1.954	0.085	0.94	0.88
	VH316181	687	2.784	0.084	1.00	1.01
	VH326109	687	3.602	0.092	1.20	1.38
	VH316141	687	3.289	0.088	1.17	1.26
	VH314579	687	3.776	0.095	1.13	1.34
	VH331610	687	3.731	0.094	1.01	1.09
	VH333616	687	3.049	0.086	1.05	1.08
	VH331939	687	1.724	0.087	0.99	1.09
	VH331929	687	3.049	0.086	1.33	1.47
	VH333334	687	3.093	0.086	0.98	0.99
	VH331516	687	2.504	0.083	1.09	1.10
	VH332031	687	2.650	0.084	1.01	1.01

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH316125	700	3.041	0.086	1.01	1.05
	VH320721	700	1.939	0.085	1.03	1.07
	VH316016	700	1.308	0.092	0.87	0.80
	VH325575	700	0.584	0.107	0.98	0.93
	VH332100	700	2.873	0.085	1.10	1.13
	VH331448	700	3.222	0.088	1.28	1.46
	VH333293	700	3.619	0.093	1.03	1.07
	VH331951	700	2.378	0.084	1.17	1.24
	VH331621	700	3.585	0.092	1.12	1.26
	VH332098	700	3.214	0.088	1.03	1.05
	VH329531	700	2.923	0.085	1.15	1.25
	VH331495	700	3.355	0.089	1.18	1.30

Table H11. Mathematics Grade 7 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH306429	820	3.510	0.081	0.99	1.02
	VH303120	820	3.850	0.085	1.05	1.11
	VH306193	820	4.167	0.090	1.19	1.44
	VH325443	820	3.152	0.078	0.99	0.97
	VH312445	820	3.696	0.083	1.08	1.13
	VH330863	820	4.050	0.088	1.27	1.46
	VH328433	820	4.012	0.087	0.97	1.06
	VH333671	820	2.859	0.077	1.16	1.20
	VH335315	820	4.457	0.095	1.32	1.66
	VH333706	820	2.421	0.077	0.98	1.08
	VH328319	820	2.770	0.077	0.97	0.95
	VF492957	820	2.564	0.077	1.13	1.17
2	VH325549	681	4.012	0.093	1.10	1.17
	VH325400	681	3.403	0.086	1.16	1.23
	VH306275	681	2.761	0.084	0.98	0.95
	VH306418	681	4.304	0.098	1.25	1.60
	VH331221	681	2.139	0.087	0.96	0.90
	VH328916	681	3.373	0.086	0.98	1.00
	VH335579	681	2.986	0.084	0.94	0.93
	VH328327	681	3.014	0.084	1.01	1.00
	VH331013	681	4.154	0.095	1.14	1.29
	VH312469	681	3.731	0.089	1.35	1.49
	VH335465	681	4.422	0.101	1.21	1.42
	VH331383	681	3.163	0.085	1.10	1.13
3	VH325435	683	3.650	0.088	1.15	1.28
	VH303346	683	3.681	0.088	1.36	1.54
	VH306390	683	3.189	0.085	0.86	0.83
	VH325467	683	3.349	0.086	0.89	0.87
	VH312723	683	3.816	0.090	1.10	1.20
	VH328394	683	4.146	0.095	0.95	1.05
	VH333655	683	3.089	0.084	0.96	0.94
	VH335583	683	3.712	0.089	1.01	1.05
	VH331398	683	3.505	0.087	1.15	1.25
	VH330880	683	4.584	0.103	1.15	1.32
	VH335784	683	3.046	0.084	1.20	1.25
	VH312436	683	4.693	0.106	1.29	1.71

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH325544	700	4.696	0.104	1.12	1.33
	VF492587	700	4.520	0.100	1.15	1.33
	VH306332	700	1.652	0.094	0.87	0.70
	VH306208	700	2.756	0.084	0.84	0.79
	VH331271	700	2.032	0.088	0.89	0.80
	VH312742	700	2.945	0.084	1.18	1.22
	VH328920	700	4.158	0.093	1.02	1.11
	VH100013	700	4.193	0.094	0.96	1.00
	VH335476	700	3.384	0.085	1.05	1.06
	VH333687	700	3.508	0.086	1.07	1.12
	VH331403	700	3.057	0.084	1.00	0.98
	VH330884	700	3.156	0.084	1.01	1.00
5	VH325421	699	2.975	0.083	1.20	1.23
	VH325475	699	6.085	0.159	0.92	1.50
	VH303234	699	3.367	0.085	1.09	1.18
	VH306239	699	3.526	0.086	1.10	1.15
	VH330956	699	3.211	0.084	0.89	0.90
	VH312497	699	5.332	0.124	1.05	1.45
	VH331263	699	3.461	0.085	1.14	1.20
	VH335467	699	3.615	0.086	1.29	1.42
	VH328943	699	4.263	0.095	1.19	1.32
	VH328326	699	2.968	0.083	1.01	1.01
	VH331269	699	2.145	0.087	0.94	0.86
	VH333637	699	4.718	0.105	1.26	1.68
6	VH306229	671	3.500	0.088	0.93	0.90
	VH307249	671	3.401	0.087	1.00	0.99
	VH325528	671	4.385	0.100	1.23	1.54
	VH303324	671	4.056	0.094	1.29	1.50
	VH330883	671	2.460	0.086	0.91	0.86
	VH331280	671	2.370	0.087	0.98	0.96
	VH328421	671	4.047	0.094	1.06	1.18
	VH312744	671	4.038	0.094	1.14	1.21
	VH333644	671	4.499	0.103	1.33	1.73
	VH312448	671	2.211	0.088	1.11	1.32
	VH336013	671	1.572	0.097	0.91	0.86
	VH328884	671	3.515	0.088	1.06	1.10

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH303131	720	3.110	0.083	1.14	1.18
	VH325538	720	2.623	0.082	0.97	0.98
	VH325396	720	3.680	0.087	0.98	0.99
	VH306236	720	1.837	0.088	0.93	0.87
	VH328317	720	2.731	0.082	0.99	0.97
	VH328875	720	4.121	0.093	1.20	1.32
	VH331207	720	3.680	0.087	1.04	1.10
	VH335778	720	3.531	0.085	1.05	1.10
	VH312895	720	4.411	0.098	1.19	1.45
	VF866380	720	5.194	0.119	1.23	1.96
	VH335481	720	2.650	0.082	0.93	0.90
	VH330965	720	3.233	0.083	0.93	0.92
8	VH325533	683	5.125	0.117	1.36	2.12
	VH306422	683	3.524	0.087	0.99	1.02
	VH303297	683	4.057	0.092	1.07	1.15
	VH303147	683	3.614	0.087	1.00	1.05
	VH335614	683	2.966	0.084	1.00	0.98
	VH331003	683	3.302	0.085	1.02	1.04
	VH328888	683	4.126	0.093	1.15	1.35
	VH335799	683	3.216	0.085	1.13	1.18
	VH328414	683	3.841	0.090	1.03	1.05
	VH331240	683	4.187	0.094	1.28	1.51
	VH312475	683	4.334	0.097	1.08	1.14
	VH312631	683	2.617	0.085	0.97	0.93
9	VH306233	692	2.751	0.084	0.99	1.00
	VH095539	692	4.933	0.111	1.22	1.69
	VH303186	692	3.852	0.090	1.32	1.45
	VH325449	692	2.391	0.086	0.91	0.90
	VH331267	692	2.850	0.084	1.06	1.05
	VF882725	692	3.538	0.087	1.15	1.22
	VH312486	692	3.106	0.085	1.11	1.12
	VH335314	692	2.399	0.085	0.97	0.98
	VH333677	692	3.478	0.087	1.13	1.15
	VH328932	692	3.092	0.085	0.93	0.92
	VH335317	692	3.876	0.090	1.00	1.01
	VH312734	692	3.708	0.089	0.98	1.05

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH306378	702	2.375	0.085	0.85	0.80
	VH306201	702	4.004	0.091	1.04	1.12
	VH303156	702	4.428	0.098	1.23	1.46
	VH325459	702	3.017	0.083	1.07	1.07
	VH336010	702	3.108	0.083	1.00	1.02
	VH330990	702	2.983	0.083	0.86	0.83
	VH328322	702	3.630	0.086	1.10	1.16
	VH312539	702	4.566	0.101	1.34	1.93
	VH331234	702	4.680	0.103	1.10	1.35
	VH333640	702	4.130	0.092	1.31	1.49
	VH335824	702	4.139	0.093	1.15	1.32
	VH312439	702	3.789	0.088	1.06	1.13

Table H12. Mathematics Grade 8 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH273319	779	2.849	0.078	0.93	0.94
	VH273219	779	4.328	0.091	1.25	1.49
	VF812655	779	2.952	0.078	1.04	1.07
	VF823461	779	3.032	0.078	1.03	1.03
	VH327154	779	3.667	0.082	1.15	1.18
	VH335264	779	3.489	0.080	0.96	0.95
	VH326522	779	2.641	0.078	0.95	0.91
	VH326397	779	2.555	0.079	0.90	0.86
	VH327295	779	3.896	0.085	1.02	1.14
	VH279222	779	3.399	0.080	1.18	1.21
	VH328152	779	3.154	0.079	0.95	0.92
	VH275351	779	2.873	0.078	0.92	0.88
2	VH273345	662	3.913	0.089	0.96	0.97
	VH273404	662	4.191	0.092	1.07	1.11
	VH273241	662	3.594	0.086	1.02	1.01
	VH278285	662	4.074	0.091	1.11	1.16
	VH279127	662	3.706	0.087	1.22	1.29
	VH326444	662	3.969	0.089	1.05	1.11
	VH277021	662	2.876	0.085	0.98	0.95
	VH335216	662	2.761	0.085	1.00	0.98
	VH335113	662	3.147	0.084	1.06	1.07
	VH327161	662	4.340	0.095	1.05	1.07
	VH327423	662	2.659	0.086	0.87	0.82
	VH328223	662	3.304	0.085	0.99	0.97
3	VH273284	660	3.183	0.085	0.87	0.82
	VH277940	660	2.576	0.088	1.05	1.04
	VH273342	660	3.830	0.088	0.90	0.87
	VH273400	660	3.948	0.089	1.04	1.09
	VH275357	660	3.044	0.085	1.03	1.09
	VH326529	660	5.541	0.125	1.15	2.00
	VH335279	660	2.979	0.085	1.17	1.29
	VH327122	660	4.423	0.096	1.02	1.12
	VH279238	660	3.616	0.087	1.28	1.35
	VH328168	660	3.139	0.085	0.93	0.89
	VH327173	660	2.957	0.086	0.96	0.93
	VH326508	660	3.586	0.086	1.13	1.15

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH120095	674	3.930	0.088	1.09	1.17
	VH273378	674	2.637	0.085	1.01	1.00
	VH273231	674	4.499	0.097	1.17	1.33
	VH275661	674	2.989	0.083	1.01	1.02
	VH327094	674	3.163	0.083	1.01	1.00
	VH335194	674	4.231	0.092	0.99	1.05
	VH326525	674	4.033	0.089	1.01	1.02
	VH326402	674	4.748	0.103	1.17	1.29
	VH335226	674	3.337	0.084	0.99	0.97
	VH328252	674	2.440	0.086	0.98	0.98
	VH279210	674	3.309	0.084	0.99	1.00
	VH327547	674	2.892	0.084	1.06	1.10
5	VH278724	673	3.679	0.087	1.16	1.21
	VH279104	673	4.675	0.102	1.24	1.47
	VH273412	673	4.373	0.096	1.11	1.21
	VH273267	673	2.103	0.091	0.87	0.77
	VH328109	673	3.295	0.085	1.02	1.00
	VH327481	673	4.143	0.092	1.04	1.10
	VH326462	673	3.087	0.085	1.09	1.12
	VH275352	673	2.649	0.086	1.02	1.02
	VH327109	673	2.582	0.086	0.91	0.85
	VH335237	673	4.194	0.093	1.13	1.26
	VH326518	673	3.849	0.089	0.94	0.96
	VH335277	673	2.176	0.090	0.95	1.00
6	VH275564	650	3.126	0.086	0.99	0.96
	VH273383	650	4.073	0.092	0.97	0.99
	VH273236	650	5.026	0.110	1.13	1.43
	VH279008	650	3.924	0.090	1.05	1.11
	VH279185	650	3.500	0.087	1.14	1.15
	VH335272	650	3.827	0.089	0.97	0.99
	VH326399	650	4.450	0.098	1.13	1.26
	VH327545	650	2.505	0.089	1.06	1.16
	VH335121	650	3.327	0.086	1.08	1.09
	VH335242	650	4.908	0.107	1.21	1.64
	VH327164	650	3.119	0.086	0.86	0.81
	VH328117	650	2.505	0.089	0.88	0.83

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH277469	662	3.717	0.087	1.12	1.16
	VH275779	662	3.596	0.086	0.98	1.04
	VH273280	662	3.167	0.085	1.00	1.00
	VH273369	662	3.687	0.087	0.91	0.90
	VH328183	662	3.664	0.087	1.06	1.05
	VH327150	662	2.786	0.085	0.94	0.93
	VH335206	662	3.304	0.085	0.92	0.89
	VH326505	662	3.840	0.088	1.19	1.24
	VH327439	662	4.115	0.092	0.97	1.00
	VH327101	662	4.040	0.091	1.04	1.12
	VH276482	662	3.903	0.089	0.87	0.84
	VH326449	662	3.694	0.087	1.39	1.49
8	VH275743	708	2.291	0.087	0.89	0.83
	VH275759	708	3.089	0.082	1.04	1.05
	VH273295	708	4.402	0.093	1.18	1.37
	VH273335	708	3.408	0.083	1.07	1.08
	VH335208	708	3.554	0.084	1.03	1.03
	VH335262	708	2.431	0.085	0.95	0.91
	VH326498	708	4.160	0.089	1.13	1.23
	VH327176	708	2.624	0.084	0.96	0.96
	VH328125	708	2.790	0.083	0.99	0.98
	VH326531	708	4.410	0.093	1.05	1.14
	VH327542	708	3.666	0.084	0.99	1.00
	VH275489	708	4.826	0.102	1.01	1.16
9	VH273262	670	2.993	0.084	1.10	1.18
	VH279045	670	3.570	0.086	1.10	1.15
	VH273388	670	4.627	0.100	1.05	1.18
	VH277658	670	4.034	0.090	1.02	1.07
	VH327429	670	3.953	0.089	1.18	1.29
	VH326511	670	4.271	0.094	1.10	1.20
	VH328260	670	2.785	0.085	0.89	0.83
	VH327115	670	2.964	0.085	0.86	0.84
	VH335243	670	2.528	0.086	0.99	0.98
	VH327335	670	3.293	0.085	1.06	1.05
	VH275483	670	4.452	0.097	0.96	1.09
	VH327497	670	3.300	0.085	1.02	1.03

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH273397	666	4.321	0.094	1.18	1.29
	VH273366	666	4.700	0.101	0.94	1.03
	VH273257	666	3.156	0.085	0.83	0.79
	VH278508	666	3.797	0.088	1.00	1.00
	VH327449	666	4.091	0.091	0.99	1.01
	VH328245	666	2.925	0.085	1.03	1.01
	VH327551	666	1.695	0.100	0.92	0.82
	VH275440	666	3.257	0.085	1.15	1.16
	VH279229	666	4.629	0.100	1.14	1.33
	VH335202	666	3.541	0.086	1.09	1.12
	VH335214	666	3.315	0.085	1.12	1.15
	VH326513	666	4.741	0.102	1.23	1.54

Science

Table H13. Science Grade 4 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH260643	826	1.691	0.082	1.23	1.45
	VH260644	826	0.808	0.076	0.90	0.90
	VH260636	826	2.449	0.096	1.38	2.50
	VH260662	826	1.774	0.084	1.19	1.51
	VH260514	826	0.628	0.076	1.06	1.06
	VH097892	826	-1.617	0.111	0.91	0.81
	VH097820	826	-0.088	0.080	1.25	1.33
	VH256003	826	0.287	0.077	1.01	1.02
	VH256810	826	-0.100	0.080	0.95	0.92
	VH255995	826	1.545	0.081	0.99	1.02
	VH256791	826	2.306	0.093	1.24	1.81
	VH256803	826	1.969	0.087	1.16	1.38
2	VH260516	762	-0.029	0.083	1.03	1.05
	VH260654	762	1.007	0.080	0.94	0.95
	VH260512	762	1.099	0.081	1.13	1.19
	VH264414	762	1.705	0.086	1.20	1.42
	VH260651	762	1.318	0.082	1.14	1.20
	VH097815	762	-0.680	0.092	0.96	0.91
	VH097878	762	-2.434	0.158	0.87	0.51
	VH256019	762	0.067	0.082	0.89	0.85
	VH256022	762	-0.968	0.098	1.07	1.15
	VH256027	762	0.201	0.081	1.03	1.03
	VH256821	762	-0.383	0.087	0.96	1.02
	VH256852	762	1.942	0.089	1.16	1.39
3	VH320328	741	1.478	0.084	1.00	1.01
	VH320263	741	2.018	0.092	1.19	1.34
	VH320261	741	0.903	0.081	1.09	1.15
	VH320349	741	0.136	0.082	0.99	0.95
	VH320281	741	2.212	0.096	1.36	1.91
	VH311274	741	1.646	0.086	1.07	1.10
	VH311280	741	-1.226	0.105	0.82	0.60
	VH264410	741	-0.505	0.089	1.01	0.98
	VH264364	741	0.395	0.081	0.94	0.91
	VH264406	741	1.331	0.083	1.18	1.26
	VH264348	741	-0.296	0.086	1.00	1.03
	VH264404	741	1.194	0.082	1.02	1.07

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH320338	733	0.481	0.082	1.08	1.09
	VH320242	733	0.926	0.082	1.16	1.23
	VH320273	733	1.257	0.083	1.16	1.22
	VH320256	733	1.369	0.084	1.05	1.09
	VH320333	733	-0.363	0.088	1.04	1.05
	VH311283	733	1.769	0.088	1.18	1.33
	VH311275	733	-0.034	0.085	0.95	0.89
	VH253167	733	0.879	0.082	1.07	1.10
	VH253123	733	0.374	0.082	0.96	0.97
	VH253274	733	0.023	0.084	1.10	1.10
	VH253209	733	0.713	0.081	1.12	1.16
	VH253228	733	0.115	0.083	1.01	0.97
5	VH264369	732	-0.238	0.087	0.93	0.90
	VH264373	732	-0.567	0.092	0.97	0.93
	VH264376	732	1.172	0.083	1.12	1.17
	VH264397	732	2.277	0.097	1.11	1.34
	VH264412	732	-0.323	0.088	0.96	0.97
	VH320346	732	0.702	0.082	1.22	1.28
	VH320294	732	0.183	0.083	1.02	1.04
	VH253127	732	0.127	0.084	0.98	0.96
	VH253180	732	1.317	0.084	1.14	1.22
	VH253201	732	-0.394	0.089	1.12	1.22
	VH253194	732	0.210	0.083	1.10	1.10
	VH253271	732	0.889	0.082	0.99	1.03
6	VH311558	753	-2.244	0.144	0.96	1.04
	VH311552	753	0.069	0.083	1.03	1.08
	VH311529	753	0.471	0.081	1.20	1.24
	VH311553	753	1.187	0.082	1.15	1.22
	VH311510	753	0.157	0.082	1.11	1.09
	VH311264	753	0.875	0.081	1.02	1.03
	VH311255	753	0.654	0.081	1.00	0.98
	VH260801	753	0.279	0.082	0.93	0.92
	VH260767	753	1.477	0.084	1.11	1.22
	VH260807	753	0.205	0.082	0.97	0.95
	VH260785	753	0.764	0.081	0.90	0.88
	VH260824	753	1.477	0.084	1.20	1.39

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH311495	746	-0.454	0.088	0.96	0.92
	VH311498	746	1.961	0.091	1.17	1.41
	VH311543	746	2.305	0.098	1.15	1.53
	VH311501	746	1.428	0.084	1.35	1.47
	VH311526	746	0.696	0.080	1.24	1.28
	VH256383	746	1.310	0.083	1.18	1.40
	VH256373	746	0.632	0.080	1.11	1.14
	VH267013	746	0.625	0.080	1.04	1.07
	VH267016	746	0.850	0.080	0.95	0.95
	VH267026	746	-0.766	0.094	0.88	0.73
	VH266958	746	0.394	0.080	0.83	0.79
	VH266924	746	1.310	0.083	0.99	1.01
8	VH267032	739	-0.489	0.091	0.87	0.74
	VH266994	739	-0.481	0.091	1.04	0.99
	VH267007	739	0.812	0.081	1.08	1.07
	VH266971	739	0.923	0.081	1.07	1.08
	VH267052	739	2.167	0.093	1.20	1.55
	VH256405	739	-1.055	0.104	0.85	0.66
	VH256416	739	-0.226	0.087	0.96	0.96
	VH264469	739	1.906	0.088	1.20	1.42
	VH264455	739	0.490	0.081	1.07	1.07
	VH264464	739	2.652	0.103	1.38	2.32
	VH264443	739	1.431	0.083	1.07	1.10
	VH264458	739	2.245	0.094	1.23	1.81
9	VH264488	732	0.484	0.081	1.03	1.01
	VH264484	732	2.295	0.097	1.30	1.93
	VH264486	732	2.139	0.094	1.22	1.60
	VH264461	732	1.735	0.088	1.23	1.34
	VH264474	732	0.669	0.081	1.02	1.01
	VH256376	732	0.643	0.081	1.10	1.13
	VH256424	732	1.007	0.082	1.04	1.04
	VH320167	732	0.404	0.082	0.99	0.99
	VH320192	732	-0.660	0.093	0.92	0.83
	VH320146	732	0.788	0.081	0.97	0.96
	VH320130	732	1.926	0.091	1.25	1.62
	VH320182	732	0.715	0.081	0.97	0.96

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH267198	713	0.918	0.082	1.04	1.08
	VH267368	713	1.135	0.083	1.14	1.21
	VH267359	713	1.162	0.083	1.28	1.44
	VH267376	713	-1.530	0.119	0.91	0.68
	VH267389	713	0.958	0.082	1.09	1.16
	VH256391	713	-0.190	0.087	1.03	1.04
	VH256369	713	0.313	0.083	0.99	0.97
	VH320124	713	0.223	0.083	0.94	0.92
	VH320119	713	0.985	0.082	0.95	0.95
	VH320173	713	-0.519	0.092	0.85	0.75
	VH320190	713	1.954	0.092	1.30	1.61
	VH325817	713	0.952	0.082	1.04	1.05

Table H14. Science Grade 8 IRT Statistics for Field Test Items

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
1	VH272770	774	-1.962	0.105	0.88	0.70
	VH272762	774	-2.150	0.112	0.87	0.63
	VH272778	774	-0.765	0.082	0.98	0.94
	VH272766	774	-0.711	0.082	1.03	1.04
	VH272782	774	-1.521	0.094	0.87	0.75
	VH260222	774	1.135	0.087	1.14	1.26
	VH260216	774	0.527	0.080	1.13	1.16
	VH260290	774	0.501	0.080	1.16	1.20
	VH260307	774	-0.088	0.079	1.02	1.03
	VH260339	774	0.995	0.085	1.10	1.23
	VH260355	774	0.135	0.079	1.15	1.19
	VH260370	774	-1.602	0.096	0.94	0.91
2	VH260318	658	-0.318	0.087	1.09	1.11
	VH260326	658	-0.527	0.089	1.08	1.09
	VH260329	658	-1.376	0.103	0.85	0.69
	VH260347	658	0.722	0.087	1.15	1.21
	VH260361	658	-0.857	0.093	1.13	1.16
	VH260229	658	0.768	0.088	1.17	1.22
	VH260244	658	0.481	0.086	1.08	1.10
	VH272750	658	-1.441	0.104	0.87	0.81
	VH272759	658	-1.253	0.100	0.92	0.95
	VH272765	658	-1.165	0.098	0.99	1.01
	VH272761	658	-1.708	0.112	0.96	0.92
	VH272772	658	-0.071	0.086	0.93	0.92
3	VH304307	664	-0.384	0.087	1.02	1.05
	VH304378	664	0.033	0.085	0.93	0.91
	VH304364	664	-1.256	0.100	1.05	1.19
	VH304374	664	-1.130	0.097	0.90	0.83
	VH304354	664	0.442	0.085	1.08	1.09
	VH260236	664	0.987	0.089	1.28	1.43
	VH260257	664	-0.679	0.090	1.04	1.06
	VH267619	664	1.082	0.090	1.15	1.25
	VH267562	664	-0.853	0.092	0.99	0.98
	VH267471	664	-0.206	0.086	1.20	1.31
	VH267634	664	1.315	0.093	1.13	1.27
	VH267596	664	1.673	0.099	1.08	1.20

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
4	VH267601	670	1.129	0.089	1.31	1.43
	VH267491	670	-0.170	0.085	1.04	1.06
	VH267352	670	0.927	0.087	1.16	1.22
	VH267437	670	1.194	0.090	1.08	1.19
	VH267625	670	-0.752	0.090	0.89	0.81
	VH260247	670	0.973	0.087	0.99	1.06
	VH260266	670	0.406	0.084	1.14	1.16
	VH304390	670	0.590	0.085	1.14	1.20
	VH304370	670	-0.206	0.085	0.98	0.97
	VH304400	670	0.734	0.085	1.04	1.04
	VH304392	670	0.049	0.084	0.95	0.95
	VH304395	670	1.326	0.092	1.15	1.31
5	VH268068	685	-0.174	0.084	1.00	1.00
	VH268093	685	-0.817	0.090	0.94	0.91
	VH268157	685	1.783	0.101	1.28	1.79
	VH313550	685	0.832	0.086	1.24	1.27
	VH268128	685	-1.051	0.093	1.01	1.10
	VH260259	685	1.007	0.088	1.26	1.38
	VH260262	685	-1.025	0.093	0.99	1.12
	VH310799	685	-1.299	0.098	0.87	0.72
	VH310801	685	-0.433	0.086	0.91	0.86
	VH310773	685	-0.698	0.088	1.01	1.07
	VH310788	685	0.149	0.083	0.96	0.95
	VH310812	685	0.991	0.088	1.08	1.18
6	VH310771	650	0.268	0.086	1.02	1.02
	VH310779	650	0.401	0.086	1.04	1.05
	VH310782	650	1.429	0.097	1.10	1.24
	VH310816	650	1.170	0.093	1.18	1.34
	VH310822	650	0.342	0.086	1.03	1.02
	VH260417	650	1.111	0.092	1.06	1.16
	VH260427	650	1.644	0.101	1.30	1.66
	VH268075	650	-0.151	0.086	0.92	0.94
	VH268148	650	-0.129	0.086	1.03	1.03
	VH268153	650	-0.316	0.087	0.94	0.93
	VH268119	650	0.268	0.086	1.20	1.27
	VH268155	650	-0.400	0.088	0.90	0.84

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
7	VH305912	652	-0.457	0.088	0.98	0.99
	VH305886	652	-1.239	0.100	0.99	1.09
	VH305898	652	-0.296	0.087	1.03	1.07
	VH305911	652	1.518	0.097	1.05	1.22
	VH305917	652	1.164	0.091	1.27	1.42
	VH260431	652	0.675	0.087	1.15	1.21
	VH260477	652	-0.488	0.089	0.95	0.95
	VH271293	652	-0.664	0.090	0.96	0.92
	VH271303	652	-0.840	0.093	0.91	0.84
	VH271336	652	0.140	0.085	1.21	1.26
	VH271325	652	-0.154	0.086	0.98	0.96
	VH271363	652	-0.465	0.088	0.90	0.86
8	VH271321	706	-0.443	0.084	1.04	1.07
	VH271329	706	1.794	0.100	1.17	1.43
	VH271311	706	0.094	0.082	1.10	1.12
	VH271285	706	-0.219	0.083	0.93	0.89
	VH271366	706	-0.130	0.083	0.93	0.88
	VH260423	706	0.047	0.082	0.98	0.97
	VH260469	706	-0.844	0.089	1.00	1.08
	VH305902	706	-0.096	0.083	1.12	1.19
	VH305880	706	-0.330	0.084	0.94	0.94
	VH305894	706	1.117	0.088	1.21	1.36
	VH305908	706	0.310	0.082	1.01	1.03
	VH305913	706	0.135	0.082	0.89	0.86
9	VH095960	656	-1.413	0.104	0.96	0.90
	VH095948	656	0.278	0.085	1.12	1.17
	VH095976	656	1.595	0.098	1.22	1.44
	VH096248	656	-0.674	0.090	0.95	0.89
	VH096235	656	-0.892	0.093	0.90	0.85
	VH260453	656	-0.340	0.087	0.88	0.83
	VH260440	656	0.767	0.087	1.13	1.21
	VH305813	656	-1.026	0.096	0.92	0.85
	VH305598	656	0.501	0.085	0.99	0.99
	VH305787	656	0.835	0.087	1.02	1.03
	VH305819	656	1.168	0.091	1.25	1.38
	VH305822	656	-0.235	0.086	0.94	0.91

Form Number	Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
10	VH305589	674	-0.032	0.084	1.03	1.04
	VH305772	674	0.964	0.089	1.21	1.31
	VH305792	674	1.510	0.097	1.02	1.13
	VH305609	674	-0.336	0.085	0.88	0.86
	VH305821	674	1.084	0.090	1.07	1.17
	VH260475	674	-0.137	0.084	1.03	1.02
	VH260438	674	0.240	0.084	0.99	0.97
	VH096228	674	0.247	0.084	1.23	1.30
	VH096002	674	-1.964	0.118	0.93	0.76
	VH096008	674	-0.496	0.086	0.91	0.86
	VH095997	674	0.902	0.088	1.17	1.21
	VH095991	674	-0.018	0.084	1.05	1.04

Appendix I: Rasch Difficulty, Standard Error, and Fit Statistics for 2016 Operational Items

Reading

Table I1. Reading Grade 3 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF394057	7672	-0.031	0.028	1.13	1.19
VF394053	7672	-0.479	0.030	0.94	0.88
VF394041	7672	-0.047	0.028	1.10	1.15
VF394045	7672	0.348	0.026	1.02	1.01
VF394046	7672	-1.056	0.034	0.85	0.68
VF394050	7672	-0.915	0.033	1.02	1.09
VF394049	7672	-0.712	0.032	1.03	1.07
VF394051	7672	-0.445	0.030	1.19	1.45
VF389477	7672	0.287	0.026	0.90	0.84
VF389620	7672	0.451	0.026	0.98	0.97
VF389446	7672	0.818	0.025	1.13	1.18
VF389470	7672	0.057	0.027	0.91	0.85
VF389473	7672	0.648	0.026	0.97	0.95
VF389457	7672	1.323	0.025	1.09	1.16
VF389165	7672	0.059	0.027	0.88	0.83
VF821120	7672	-0.352	0.029	0.93	0.83
VF821218	7672	-0.545	0.030	0.94	0.86
VF821123	7672	-0.286	0.029	1.11	1.22
VF821312	7672	0.446	0.026	1.01	1.00
VF821272	7672	-1.603	0.041	0.82	0.51
VF821338	7672	0.108	0.027	0.98	0.94
VF821362	7672	0.369	0.026	0.98	0.97
VF883330	7672	1.038	0.025	0.96	0.97
VF882884	7672	1.086	0.025	0.99	1.03
VF883326	7672	0.234	0.027	1.11	1.17
VF883549	7672	-0.582	0.031	0.99	0.87
VF883561	7672	0.825	0.025	1.14	1.17
VF883364	7672	-1.203	0.036	0.85	0.67
VF883614	7672	-0.517	0.030	0.96	0.94
VF883619	7672	0.419	0.026	0.95	0.91
VF883622	7672	0.161	0.027	1.07	1.15
VH143172	7672	0.466	0.026	0.95	0.94
VH143162	7672	-1.293	0.037	0.92	0.76

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH143197	7672	0.924	0.025	1.12	1.17
VH143131	7672	-0.312	0.029	1.07	1.14
VH143274	7672	-0.028	0.028	1.07	1.17
VH143252	7672	0.567	0.026	0.97	0.97
VF497759	7672	1.144	0.025	0.95	0.97
VF497725	7672	0.137	0.027	0.84	0.75
VF497758	7672	0.547	0.026	0.99	0.98
VF497767	7672	1.094	0.025	1.02	1.04
VF497766	7672	1.615	0.026	1.15	1.34
VF497718	7672	-0.013	0.028	1.11	1.17
VH152948	7672	0.543	0.026	1.06	1.06
VH153021	7672	-0.540	0.030	1.00	1.06
VH153042	7672	-0.829	0.032	0.88	0.85
VF885214	7672	-0.004	0.027	0.98	0.96
VF885379	7672	0.906	0.025	1.08	1.11
VF885192	7672	-1.120	0.035	0.78	0.55
VF885434	7672	0.280	0.026	0.92	0.88

Table I2. Reading Grade 4 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF880215	7478	0.151	0.030	0.95	0.89
VF880204	7478	1.825	0.026	0.96	1.00
VF880210	7478	1.347	0.026	1.10	1.14
VF880326	7478	0.810	0.027	1.05	1.08
VF880311	7478	1.227	0.026	1.00	1.02
VF880321	7478	1.454	0.026	1.08	1.14
VF880314	7478	2.447	0.027	1.11	1.54
VF880343	7478	0.996	0.026	1.03	1.06
VF880345	7478	1.134	0.026	1.00	1.02
VF880350	7478	0.693	0.027	1.02	1.03
VH145672	7478	0.772	0.027	1.14	1.18
VH145677	7478	1.897	0.026	1.04	1.11
VH145638	7478	0.427	0.028	0.93	0.86
VH145684	7478	2.472	0.027	1.07	1.34
VH145660	7478	1.405	0.026	1.10	1.16
VH145702	7478	1.393	0.026	1.10	1.15
VH145595	7478	0.678	0.027	0.96	0.90
VH150628	7478	0.321	0.029	1.02	0.99
VH150639	7478	-0.630	0.036	0.92	0.80
VH150646	7478	0.630	0.027	1.03	1.11
VH150612	7478	0.295	0.029	1.04	0.99
VH150620	7478	1.087	0.026	1.10	1.12
VH150605	7478	-0.599	0.036	0.91	0.75
VH150637	7478	-0.489	0.034	0.92	0.78
VF884843	7478	-0.924	0.039	1.00	0.71
VF884830	7478	0.433	0.028	1.05	1.07
VF884836	7478	0.750	0.027	0.93	0.89
VF884910	7478	-0.606	0.036	1.16	1.20
VF884900	7478	-0.982	0.040	1.08	0.87
VF884925	7478	0.264	0.029	0.99	0.99
VF884913	7478	1.631	0.026	1.01	1.06
VF497338	7478	1.011	0.026	1.29	1.46
VF497297	7478	-1.320	0.044	0.88	0.64
VF497314	7478	0.378	0.029	0.82	0.72
VF497322	7478	0.560	0.028	0.94	0.87
VF497303	7478	-0.289	0.033	0.84	0.72
VF497330	7478	0.039	0.030	0.90	0.79

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF497326	7478	0.898	0.027	0.89	0.83
VF497327	7478	0.278	0.029	1.05	1.03
VF407243	7478	0.157	0.030	1.02	1.08
VF407287	7478	-0.780	0.037	0.87	0.71
VF407232	7478	-0.568	0.035	0.84	0.64
VF407295	7478	-1.199	0.043	0.83	0.69
VF407297	7478	0.341	0.029	0.98	0.97
VF407235	7478	1.251	0.026	1.18	1.25
VF407298	7478	0.659	0.027	1.08	1.12
VF407282	7478	-0.962	0.039	0.75	0.53
VF885226	7478	-0.966	0.040	0.98	0.96
VF885195	7478	-0.453	0.034	0.89	0.80
VH152031	7478	1.373	0.026	1.06	1.08

Table I3. Reading Grade 5 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH134512	7296	-0.485	0.037	0.93	0.78
VH134422	7296	0.170	0.031	0.84	0.72
VH134412	7296	1.303	0.027	1.14	1.20
VH134521	7296	-0.258	0.035	1.12	1.65
VH134580	7296	0.612	0.029	1.09	1.23
VH143307	7296	0.486	0.029	0.99	0.99
VH143304	7296	1.269	0.027	0.94	0.91
VH143328	7296	1.576	0.026	1.04	1.06
VH143318	7296	-0.325	0.035	1.00	1.21
VH143331	7296	-0.123	0.033	1.00	1.03
VF407319	7296	0.839	0.028	1.07	1.10
VF407388	7296	0.500	0.029	1.08	1.02
VF407329	7296	0.058	0.032	0.95	0.81
VF407332	7296	0.257	0.031	0.94	0.90
VF407360	7296	-0.984	0.043	0.89	0.76
VF407355	7296	-0.247	0.034	0.93	0.93
VH144472	7296	-0.655	0.039	1.04	1.29
VH144483	7296	0.350	0.030	1.01	0.99
VH144474	7296	0.959	0.028	0.94	0.90
VH144509	7296	0.950	0.028	1.15	1.23
VH144512	7296	-0.083	0.033	0.98	0.99
VH144593	7296	-0.838	0.041	0.90	0.86
VF496865	7296	0.765	0.028	0.93	0.88
VF496879	7296	0.953	0.028	1.03	1.02
VF496213	7296	0.564	0.029	1.10	1.16
VF496206	7296	0.943	0.028	1.06	1.04
VF496209	7296	0.126	0.032	0.96	0.89
VF496212	7296	-0.047	0.033	0.91	0.81
VF496221	7296	-1.319	0.048	0.86	0.54
VF496200	7296	-0.149	0.034	0.98	1.06
VF880864	7296	-0.801	0.040	0.89	0.75
VF882769	7296	0.497	0.029	0.99	0.98
VF882762	7296	1.611	0.026	1.09	1.15
VF882790	7296	1.641	0.026	1.04	1.09
VF909893	7296	1.206	0.027	1.13	1.20
VF882786	7296	0.587	0.029	1.16	1.26
VH136637	7296	0.010	0.032	1.00	0.85
VH136631	7296	-0.205	0.034	0.94	0.86
VH136633	7296	0.298	0.030	1.03	0.99
VH136634	7296	0.862	0.028	1.00	0.97

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH136607	7296	-1.146	0.045	0.95	0.98
VF497025	7296	0.777	0.028	0.91	0.84
VF497030	7296	0.809	0.028	1.00	0.96
VF497037	7296	-0.421	0.036	0.84	0.63
VF497012	7296	0.727	0.028	0.92	0.90
VF497016	7296	0.319	0.030	0.91	0.78
VF885197	7296	-0.293	0.035	0.75	0.57
VF885217	7296	0.296	0.031	0.87	0.77
VF885221	7296	1.333	0.027	1.14	1.23
VH150300	7296	1.092	0.027	1.05	1.08
VH150408	7296	0.647	0.029	0.98	0.95
VH150398	7296	2.035	0.026	0.99	1.06
VH150312	7296	0.818	0.028	0.90	0.81
VH150267	7296	1.489	0.026	1.17	1.24

Table I4. Reading Grade 6 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF497047	6989	0.436	0.031	1.10	1.22
VF497049	6989	0.746	0.030	1.18	1.27
VF497035	6989	0.418	0.031	0.87	0.81
VF497034	6989	-0.012	0.034	0.99	0.93
VF497041	6989	-0.118	0.035	0.85	0.70
VF496873	6989	0.982	0.029	1.03	1.07
VF496204	6989	1.223	0.028	0.95	0.89
VF496208	6989	1.243	0.028	1.01	0.99
VF496863	6989	1.228	0.028	0.95	0.91
VF496191	6989	1.544	0.027	0.94	0.93
VF496867	6989	0.891	0.029	0.96	0.92
VF496415	6989	1.063	0.028	1.09	1.19
VF496172	6989	1.100	0.028	1.00	0.97
VF496055	6989	1.917	0.027	1.00	1.02
VF496083	6989	-0.933	0.045	1.26	1.06
VF496097	6989	0.666	0.030	0.91	0.81
VF496065	6989	0.907	0.029	0.95	0.94
VF496071	6989	0.040	0.034	1.28	1.25
VF496100	6989	-0.199	0.036	1.45	1.38
VH151373	6989	0.885	0.029	1.00	1.03
VH151417	6989	1.358	0.028	1.11	1.16
VH151420	6989	0.599	0.030	0.86	0.73
VH151386	6989	0.851	0.029	0.96	0.94
VH151408	6989	0.983	0.029	1.00	0.97
VH153064	6989	1.498	0.027	1.09	1.13
VH153081	6989	1.485	0.027	1.08	1.15
VH153209	6989	-0.242	0.036	0.88	0.75
VH153197	6989	1.125	0.028	0.94	0.90
VH153196	6989	0.472	0.031	0.91	0.82
VF495908	6989	0.075	0.034	0.96	0.89
VF495961	6989	1.263	0.028	1.07	1.17
VF495968	6989	2.096	0.027	1.05	1.11
VF495918	6989	0.720	0.030	1.02	0.99
VF495945	6989	1.279	0.028	1.13	1.19
VF495925	6989	0.520	0.031	0.99	1.00
VF806599	6989	-0.225	0.036	0.86	0.65
VF806591	6989	0.677	0.030	0.91	0.82
VF806593	6989	1.247	0.028	1.02	0.99
VF806588	6989	1.210	0.028	1.18	1.36
VF806587	6989	1.620	0.027	1.14	1.16

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH129551	6989	0.179	0.033	0.91	0.89
VH129569	6989	-0.682	0.041	0.82	0.54
VH129567	6989	0.080	0.034	0.92	0.88
VH129555	6989	0.623	0.030	0.93	0.86
VH129558	6989	1.641	0.027	1.07	1.10
VH129579	6989	1.436	0.027	1.06	1.08
VH129572	6989	-0.048	0.035	0.90	0.79
VF884733	6989	1.115	0.028	1.06	1.10
VF884772	6989	2.398	0.027	1.06	1.15
VF884844	6989	1.142	0.028	1.02	1.01
VF884880	6989	0.589	0.030	1.01	0.95
VH152232	6989	1.858	0.027	1.02	1.04
VF884630	6989	1.439	0.027	1.15	1.22
VF884988	6989	1.011	0.029	0.91	0.84
VF884628	6989	1.866	0.027	1.16	1.24
VH152242	6989	-0.509	0.039	0.87	0.67

Table I5. Reading Grade 7 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF496949	7043	0.314	0.034	1.02	1.06
VF496944	7043	1.396	0.028	1.10	1.12
VF496901	7043	1.296	0.028	1.21	1.37
VF496913	7043	-0.053	0.038	0.89	0.67
VF496908	7043	-0.628	0.045	0.92	0.76
VF496903	7043	0.381	0.033	0.91	0.76
VF496895	7043	2.163	0.027	0.95	0.95
VF496900	7043	-0.139	0.039	0.87	0.64
VF497956	7043	1.255	0.028	1.17	1.35
VF497955	7043	0.309	0.034	0.89	0.73
VF497958	7043	1.120	0.029	0.91	0.85
VF497951	7043	1.180	0.029	1.00	0.98
VF497961	7043	1.678	0.027	0.94	0.90
VF497978	7043	1.978	0.027	1.01	1.00
VF497974	7043	1.786	0.027	1.04	1.03
VF865426	7043	0.301	0.034	0.93	0.84
VF865413	7043	0.786	0.031	0.97	0.94
VF865473	7043	2.047	0.027	0.95	0.93
VF865494	7043	1.981	0.027	1.21	1.31
VF865477	7043	0.590	0.032	0.98	0.97
VF906623	7043	2.380	0.027	1.09	1.13
VF865627	7043	1.062	0.029	0.89	0.84
VF820260	7043	1.931	0.027	0.94	0.92
VF820251	7043	1.442	0.028	1.09	1.19
VF820333	7043	1.966	0.027	1.21	1.30
VF820315	7043	1.403	0.028	0.98	0.95
VF820301	7043	0.707	0.031	0.94	0.90
VF820351	7043	1.968	0.027	1.07	1.08
VF820210	7043	1.414	0.028	1.00	1.01
VF498023	7043	0.297	0.034	0.83	0.67
VF497995	7043	1.049	0.029	0.81	0.70
VF498018	7043	0.659	0.031	0.84	0.72
VF497980	7043	2.339	0.027	1.00	1.02
VF498062	7043	1.824	0.027	1.04	1.06
VF498011	7043	1.855	0.027	1.11	1.15
VF498051	7043	1.630	0.027	1.02	1.02
VF498054	7043	1.766	0.027	0.94	0.92
VF498057	7043	1.508	0.028	1.00	1.03
VF498063	7043	2.236	0.027	1.08	1.10
VF498034	7043	1.402	0.028	0.96	0.90

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF498052	7043	2.530	0.027	1.12	1.19
VF820464	7043	1.509	0.028	1.12	1.14
VF820412	7043	0.605	0.032	1.04	1.09
VF820449	7043	1.021	0.029	1.04	1.13
VF820435	7043	2.681	0.027	1.10	1.19
VF820391	7043	1.404	0.028	1.03	1.05
VF820404	7043	2.429	0.027	0.96	0.96
VF864902	7043	1.133	0.029	0.90	0.82
VF864898	7043	0.193	0.035	0.98	1.00
VF865057	7043	1.052	0.029	0.83	0.72
VF865072	7043	0.902	0.030	0.89	0.78
VF865094	7043	3.089	0.028	1.15	1.30
VF865104	7043	2.031	0.027	0.88	0.85
VF885398	7043	1.461	0.028	1.02	1.04
VF885820	7043	2.193	0.027	0.89	0.88
VH150842	7043	0.815	0.030	0.98	0.97

Table I6. Reading Grade 8 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH152367	6790	1.879	0.027	0.92	0.91
VH152400	6790	0.964	0.031	1.05	1.13
VH152415	6790	1.501	0.028	1.19	1.28
VH152422	6790	0.452	0.035	0.91	0.78
VH152362	6790	0.374	0.035	0.93	0.86
VH152391	6790	0.422	0.035	0.84	0.68
VH152434	6790	1.054	0.030	0.95	0.88
VF497207	6790	0.052	0.039	0.89	0.75
VF497196	6790	0.636	0.033	0.98	0.97
VF497178	6790	1.907	0.027	1.12	1.14
VF497213	6790	0.401	0.035	0.95	0.93
VF497203	6790	1.517	0.028	0.93	0.95
VF497209	6790	1.230	0.029	1.07	1.11
VF497173	6790	2.515	0.027	1.08	1.14
VF497229	6790	1.356	0.029	1.11	1.15
VF497244	6790	0.593	0.033	0.97	0.96
VF497242	6790	2.384	0.027	1.05	1.09
VF497255	6790	0.068	0.039	0.96	1.03
VF497271	6790	0.676	0.033	0.88	0.74
VF867197	6790	1.021	0.030	0.91	0.84
VF867239	6790	1.544	0.028	1.00	0.99
VF867326	6790	1.056	0.030	0.87	0.80
VF867274	6790	0.866	0.031	0.99	1.00
VF867355	6790	3.041	0.028	1.04	1.15
VF497096	6790	1.532	0.028	0.96	0.89
VF497103	6790	0.838	0.032	0.93	0.92
VF497098	6790	2.307	0.027	1.10	1.17
VF497114	6790	1.032	0.030	0.94	0.94
VF497115	6790	0.791	0.032	1.09	1.05
VF819971	6790	2.645	0.027	1.15	1.24
VF820011	6790	1.798	0.027	1.21	1.29
VF820174	6790	0.812	0.032	1.05	0.95
VF820025	6790	0.883	0.031	1.01	0.95
VF820170	6790	2.555	0.027	1.02	1.07
VF820236	6790	2.594	0.027	1.04	1.11
VF820159	6790	0.907	0.031	1.21	1.35
VF497329	6790	0.770	0.032	0.91	0.83
VF497328	6790	1.578	0.028	1.11	1.17
VF497325	6790	1.635	0.028	0.89	0.86
VF497349	6790	-0.111	0.041	0.90	0.81

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF497370	6790	1.539	0.028	1.10	1.14
VF497363	6790	2.126	0.027	1.05	1.07
VF820734	6790	0.026	0.039	0.90	0.87
VF820781	6790	1.559	0.028	0.94	0.90
VF820727	6790	1.962	0.027	0.88	0.85
VF820720	6790	2.051	0.027	1.03	1.04
VF820786	6790	1.598	0.028	1.23	1.37
VF820792	6790	1.211	0.029	1.04	1.00
VF820796	6790	2.428	0.027	1.12	1.19
VF820801	6790	-0.100	0.041	0.83	0.63
VH151102	6790	1.294	0.029	0.95	0.92
VH151057	6790	0.115	0.038	0.91	0.83
VH151086	6790	0.173	0.037	0.84	0.64
VF883716	6790	0.547	0.034	0.95	0.99
VF883823	6790	0.805	0.032	0.94	0.87
VF883653	6790	-0.092	0.040	0.88	0.73

Mathematics

Table I7. Mathematics Grade 3 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH095532	7651	-1.127	0.032	0.98	0.92
VH103654	7651	-1.072	0.032	0.95	0.89
VF393788	7651	-0.644	0.029	0.98	0.96
VF821723	7651	0.765	0.026	1.06	1.09
VF803172	7651	0.597	0.026	1.04	1.06
VF406331	7651	-0.686	0.030	0.89	0.83
VF387500	7651	-0.175	0.027	0.97	0.92
VF821698	7651	0.661	0.026	1.05	1.08
VF493136	7651	-0.641	0.029	0.85	0.77
VF492342	7651	-0.947	0.031	0.86	0.74
VF406297	7651	0.707	0.026	1.21	1.30
VF865405	7651	0.041	0.027	1.03	1.01
VH103554	7651	-0.267	0.028	0.98	0.97
VF866235	7651	-0.044	0.027	0.91	0.84
VH095414	7651	1.152	0.026	1.10	1.18
VF737752	7651	0.535	0.026	0.86	0.83
VF740960	7651	0.602	0.026	1.06	1.09
VF866360	7651	-1.834	0.039	0.83	0.50
VF866898	7651	-0.864	0.031	0.93	0.91
VF740890	7651	0.765	0.026	1.07	1.11
VH095347	7651	-0.045	0.027	0.90	0.85
VH095406	7651	0.252	0.026	1.08	1.09
VH103449	7651	-0.891	0.031	1.15	1.33
VF865570	7651	0.053	0.027	0.87	0.79
VF821652	7651	-0.061	0.027	0.98	0.94
VF866354	7651	-0.946	0.031	0.81	0.67
VH098018	7651	0.299	0.026	1.40	1.61
VF803266	7651	0.384	0.026	0.84	0.78
VF865389	7651	-0.020	0.027	0.87	0.81
VF819676	7651	-1.622	0.037	1.00	1.01
VF494693	7651	0.391	0.026	0.97	0.95
VF394378	7651	1.220	0.027	1.30	1.47
VH103497	7651	0.472	0.026	0.89	0.83
VF867001	7651	-0.130	0.027	0.93	0.86
VF803121	7651	0.428	0.026	1.14	1.24
VF393751	7651	-0.947	0.031	0.92	0.83
VF865316	7651	-1.904	0.040	0.84	0.53
VF493127	7651	0.349	0.026	0.99	0.98

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH094951	7651	0.814	0.026	0.93	0.91
VF821767	7651	1.050	0.026	1.15	1.25
VF394232	7651	-1.203	0.033	0.80	0.72
VF394239	7651	0.185	0.027	0.82	0.77
VF493461	7651	1.057	0.026	1.20	1.36
VF867224	7651	-0.940	0.031	0.98	0.91
VF394358	7651	0.233	0.027	0.90	0.84
VF493364	7651	0.451	0.026	1.06	1.08
VF866264	7651	-0.566	0.029	0.91	0.84
VF493387	7651	-1.025	0.032	1.10	1.10
VH098050	7651	0.623	0.026	1.05	1.07
VH095303	7651	1.513	0.027	0.99	1.06

Table I8. Mathematics Grade 4 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH097534	7457	-0.222	0.033	1.02	1.23
VH097139	7457	0.666	0.028	0.84	0.73
VF492371	7457	-1.079	0.043	0.91	1.10
VH118288	7457	1.579	0.026	1.13	1.19
VF866677	7457	0.195	0.030	0.94	0.84
VF822854	7457	-1.209	0.045	0.83	0.62
VF800697	7457	0.615	0.028	1.35	1.81
VF492358	7457	-0.893	0.040	0.95	1.07
VF815962	7457	0.535	0.028	1.00	1.04
VH103556	7457	1.497	0.026	1.17	1.24
VH097502	7457	-0.165	0.033	0.97	1.03
VF867084	7457	1.869	0.026	0.92	0.90
VF823145	7457	-0.140	0.033	0.95	0.87
VF801214	7457	1.825	0.026	0.95	0.94
VH097287	7457	1.791	0.026	0.82	0.79
VF823138	7457	0.775	0.028	0.88	0.82
VF497391	7457	0.507	0.029	0.89	0.83
VF864145	7457	1.818	0.026	1.04	1.07
VF492392	7457	0.706	0.028	1.04	1.20
VF493144	7457	1.009	0.027	0.89	0.82
VF493149	7457	-0.814	0.039	0.97	0.92
VF492339	7457	-0.822	0.039	0.87	0.69
VF493228	7457	1.144	0.027	0.81	0.74
VF393714	7457	1.487	0.026	1.06	1.09
VF866857	7457	1.132	0.027	0.95	0.88
VF741948	7457	2.209	0.027	1.18	1.28
VF492376	7457	0.104	0.031	0.99	1.11
VF864149	7457	1.660	0.026	1.00	0.99
VF493219	7457	0.184	0.030	1.05	1.19
VF863975	7457	-0.520	0.036	1.09	1.05
VF880294	7457	1.063	0.027	0.85	0.78
VF741944	7457	0.010	0.031	1.17	1.48
VF816041	7457	0.899	0.027	1.01	1.04
VF801835	7457	0.715	0.028	1.16	1.15
VH103619	7457	2.192	0.027	1.12	1.21
VH097472	7457	2.028	0.026	1.05	1.12
VH097255	7457	-0.131	0.032	0.90	0.82
VF815957	7457	0.785	0.028	1.01	0.97
VF801227	7457	2.106	0.027	1.02	1.04
VF492338	7457	0.612	0.028	1.01	1.02

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF866381	7457	1.115	0.027	0.86	0.79
VF492354	7457	-0.012	0.032	0.94	0.92
VF493301	7457	1.780	0.026	1.13	1.19
VH097368	7457	1.906	0.026	0.90	0.87
VH097423	7457	-0.764	0.039	1.06	1.20
VF393648	7457	2.429	0.027	1.11	1.22
VF823371	7457	0.562	0.028	1.00	1.02
VF864100	7457	0.357	0.029	0.94	0.96
VF492364	7457	0.102	0.031	0.98	0.94
VF816057	7457	0.664	0.028	0.94	0.90
VF867088	7457	2.012	0.026	0.98	1.01
VF866870	7457	2.120	0.027	1.10	1.19
VH097521	7457	0.607	0.028	0.90	0.85
VF493262	7457	0.822	0.027	0.98	0.95
VF497402	7457	1.531	0.026	0.97	0.97
VF880252	7457	0.133	0.031	0.86	0.70
VF493361	7457	1.612	0.026	1.27	1.37
VF816151	7457	-0.235	0.033	0.94	0.92
VF815909	7457	-0.385	0.035	0.87	0.70

Table I9. Mathematics Grade 5 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF491888	7294	-0.071	0.036	0.88	0.86
VF492269	7294	0.966	0.029	0.94	0.88
VF819900	7294	0.079	0.034	1.09	1.21
VF491924	7294	1.042	0.028	0.96	0.95
VF492083	7294	2.681	0.027	0.92	0.91
VF864604	7294	1.732	0.027	0.98	0.99
VF741081	7294	1.423	0.027	1.06	1.05
VF491942	7294	0.782	0.030	0.95	0.89
VF815846	7294	2.019	0.027	1.14	1.24
VF797963	7294	1.550	0.027	0.94	0.91
VF491626	7294	0.814	0.029	0.85	0.73
VH094283	7294	2.659	0.027	1.19	1.30
VF492313	7294	0.792	0.030	0.85	0.73
VF741371	7294	0.182	0.033	0.95	0.85
VF823759	7294	2.768	0.028	1.03	1.11
VH095125	7294	2.940	0.028	1.16	1.36
VF819994	7294	1.759	0.027	1.06	1.05
VF866061	7294	1.746	0.027	0.98	0.96
VH095132	7294	2.487	0.027	0.92	0.92
VF492296	7294	2.442	0.027	1.23	1.32
VF740936	7294	0.290	0.033	0.87	0.93
VF741539	7294	0.245	0.033	0.90	0.82
VH099648	7294	3.216	0.029	1.07	1.22
VF492255	7294	1.197	0.028	0.86	0.76
VH103584	7294	1.073	0.028	0.84	0.72
VF802894	7294	1.710	0.027	1.04	1.03
VF492113	7294	1.154	0.028	0.98	0.93
VF491914	7294	1.417	0.027	1.15	1.19
VF492077	7294	1.357	0.028	0.89	0.82
VF491635	7294	2.072	0.027	0.90	0.87
VH093957	7294	1.577	0.027	1.01	0.98
VF491992	7294	1.267	0.028	1.05	1.13
VF491755	7294	2.442	0.027	0.96	0.96
VF492099	7294	1.623	0.027	0.89	0.86
VH095017	7294	1.586	0.027	0.92	0.88
VF798083	7294	1.337	0.028	0.91	0.86
VH094425	7294	1.847	0.027	0.91	0.89
VH094296	7294	1.513	0.027	1.28	1.46
VF491804	7294	1.621	0.027	0.77	0.68
VF492374	7294	0.787	0.030	1.00	1.01

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF823764	7294	0.230	0.033	0.94	1.08
VF491932	7294	1.615	0.027	0.91	0.85
VF491630	7294	2.084	0.027	1.09	1.13
VF492397	7294	2.203	0.027	1.01	1.01
VF492304	7294	0.596	0.031	0.89	0.77
VF491905	7294	0.997	0.029	0.93	0.89
VF880813	7294	1.774	0.027	0.97	0.94
VF492435	7294	1.887	0.027	1.27	1.45
VH103721	7294	2.132	0.027	0.97	0.95
VF864609	7294	2.095	0.027	0.89	0.86
VF802778	7294	3.128	0.029	1.14	1.35
VF492001	7294	1.065	0.028	1.17	1.27
VF866103	7294	1.540	0.027	1.09	1.11
VF491753	7294	1.280	0.028	1.13	1.23
VF819989	7294	2.561	0.027	1.09	1.16
VF491761	7294	1.310	0.028	0.99	0.98
VF815982	7294	2.690	0.027	0.98	1.02
VF802014	7294	1.528	0.027	0.84	0.77
VF797938	7294	1.129	0.028	0.88	0.83

Table I10. Mathematics Grade 6 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH135098	6995	0.901	0.032	0.95	0.87
VF862804	6995	0.585	0.034	1.05	1.27
VF862699	6995	1.399	0.029	0.93	0.95
VF811515	6995	2.913	0.027	1.03	1.04
VF492647	6995	1.089	0.030	0.90	0.78
VF492066	6995	2.404	0.027	0.96	0.96
VF882963	6995	2.572	0.027	1.22	1.31
VF862786	6995	1.912	0.027	0.88	0.85
VF809839	6995	1.609	0.028	0.87	0.77
VH135857	6995	1.185	0.030	1.00	1.00
VF492192	6995	2.273	0.027	1.07	1.07
VF741723	6995	1.886	0.027	1.04	1.07
VF492759	6995	2.776	0.027	1.09	1.15
VF810696	6995	2.511	0.027	0.94	0.93
VF492030	6995	1.639	0.028	0.99	0.98
VF812185	6995	1.911	0.027	1.03	1.02
VF492093	6995	2.524	0.027	0.92	0.92
VF492160	6995	2.745	0.027	1.06	1.09
VH118676	6995	2.107	0.027	0.94	0.91
VH140963	6995	2.509	0.027	0.95	0.95
VF492388	6995	1.786	0.027	0.98	0.99
VF741557	6995	2.887	0.027	0.99	1.01
VH121098	6995	1.894	0.027	1.16	1.26
VF809034	6995	1.160	0.030	0.99	0.97
VF741572	6995	2.306	0.027	1.01	1.02
VF741928	6995	0.473	0.035	0.92	0.80
VF492722	6995	2.913	0.027	1.03	1.05
VF865682	6995	3.549	0.029	1.08	1.19
VF866265	6995	1.137	0.030	1.05	1.16
VF491935	6995	1.348	0.029	0.86	0.77
VF866278	6995	2.611	0.027	1.10	1.16
VF865635	6995	2.125	0.027	0.98	0.98
VF797964	6995	3.122	0.028	1.22	1.34
VF822007	6995	2.848	0.027	1.00	1.03
VF883067	6995	1.796	0.027	0.98	0.98
VF491931	6995	-0.006	0.041	0.94	0.59
VH121115	6995	1.160	0.030	0.86	0.76
VF741533	6995	2.806	0.027	1.00	1.02
VH121032	6995	2.010	0.027	0.85	0.78
VH136168	6995	1.529	0.028	0.91	0.83

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF866221	6995	2.342	0.027	1.02	1.03
VF821920	6995	1.891	0.027	1.08	1.11
VF797977	6995	0.955	0.031	0.96	1.00
VF821998	6995	0.803	0.032	0.92	0.86
VF492996	6995	2.329	0.027	0.87	0.82
VF821946	6995	1.902	0.027	1.10	1.19
VH136222	6995	2.997	0.027	0.86	0.84
VF797171	6995	3.076	0.027	1.16	1.25
VF822023	6995	1.362	0.029	0.98	0.97
VF491970	6995	1.728	0.028	1.00	1.00
VH121642	6995	3.295	0.028	0.94	0.99
VF803293	6995	2.640	0.027	1.00	1.01
VF866290	6995	2.519	0.027	1.09	1.11
VH118579	6995	2.309	0.027	1.00	1.00
VF882800	6995	2.215	0.027	1.12	1.14
VF866230	6995	2.200	0.027	1.01	1.03
VF491976	6995	2.891	0.027	1.08	1.12
VF865678	6995	2.276	0.027	1.05	1.05
VH118513	6995	1.553	0.028	1.00	1.03

Table I11. Mathematics Grade 7 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF813502	7042	1.980	0.028	0.88	0.80
VH095465	7042	2.304	0.027	0.98	0.95
VF492966	7042	1.612	0.030	0.88	0.75
VF800136	7042	1.819	0.029	1.06	1.06
VH129959	7042	2.783	0.026	0.96	0.94
VF818181	7042	3.416	0.027	1.08	1.08
VF736954	7042	2.318	0.027	0.91	0.85
VH095405	7042	2.720	0.026	1.14	1.17
VF880250	7042	3.221	0.027	1.04	1.05
VF492708	7042	2.883	0.026	0.97	0.96
VH095484	7042	2.897	0.026	1.01	1.01
VF492578	7042	3.537	0.027	1.00	1.02
VF822880	7042	3.405	0.027	1.05	1.08
VH095274	7042	2.603	0.027	1.02	1.03
VH095287	7042	2.971	0.027	0.93	0.90
VH147509	7042	2.538	0.027	0.94	0.92
VH141616	7042	2.170	0.027	1.12	1.29
VH129916	7042	2.882	0.026	0.94	0.93
VF880897	7042	1.443	0.031	0.88	0.74
VF819294	7042	2.958	0.026	1.06	1.08
VF867395	7042	2.338	0.027	1.10	1.25
VF736963	7042	3.467	0.027	1.20	1.27
VF883138	7042	3.196	0.027	1.06	1.07
VF866491	7042	3.828	0.028	0.91	0.92
VF736931	7042	2.803	0.026	0.98	0.96
VF883264	7042	3.206	0.027	0.90	0.88
VF869623	7042	2.468	0.027	0.87	0.83
VF800078	7042	2.447	0.027	1.12	1.25
VF881807	7042	2.110	0.028	0.81	0.72
VF822986	7042	2.481	0.027	1.18	1.25
VH129853	7042	2.926	0.026	0.94	0.92
VF867256	7042	2.691	0.027	0.98	0.96
VF882715	7042	3.812	0.028	1.09	1.19
VH141553	7042	1.474	0.030	0.83	0.68
VF493067	7042	3.052	0.027	0.98	0.98
VH097710	7042	2.026	0.028	0.87	0.79
VH141570	7042	2.831	0.026	1.09	1.14
VF736938	7042	2.540	0.027	0.89	0.83
VF818347	7042	1.557	0.030	1.09	1.53
VF818184	7042	2.717	0.027	1.01	1.06

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH124615	7042	2.212	0.027	0.99	1.00
VF492915	7042	1.694	0.029	1.01	1.06
VF866531	7042	3.290	0.027	0.95	0.94
VH148173	7042	3.299	0.027	1.14	1.18
VH097803	7042	3.103	0.027	1.12	1.15
VH148935	7042	3.166	0.027	1.03	1.04
VF866386	7042	3.603	0.028	1.10	1.15
VF882739	7042	2.396	0.027	1.04	1.02
VH141596	7042	3.511	0.027	1.18	1.26
VF882691	7042	2.794	0.026	1.00	0.97
VF492589	7042	4.026	0.029	0.96	1.04
VF492992	7042	4.145	0.030	0.88	0.88
VF736947	7042	2.182	0.027	0.92	0.86
VF880171	7042	3.994	0.029	1.12	1.18
VF882910	7042	2.946	0.026	1.03	1.05
VF883244	7042	2.586	0.027	0.97	0.95
VF799837	7042	1.599	0.030	0.96	0.90
VH148915	7042	2.213	0.027	0.97	0.93
VF492295	7042	2.383	0.027	1.02	1.06

Table I12. Mathematics Grade 8 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VH118733	6801	2.952	0.027	1.06	1.06
VF880798	6801	2.507	0.027	0.99	1.02
VF491923	6801	0.962	0.038	1.06	0.91
VF802924	6801	3.653	0.027	1.19	1.26
VH121865	6801	2.561	0.027	0.94	0.90
VH120033	6801	2.353	0.028	0.94	0.92
VF823307	6801	2.630	0.027	0.93	0.91
VH118897	6801	3.293	0.027	1.15	1.18
VH118056	6801	3.811	0.028	1.21	1.31
VH120009	6801	2.687	0.027	1.03	1.04
VF802927	6801	3.264	0.027	1.03	1.03
VF491915	6801	3.935	0.028	1.10	1.17
VF885510	6801	3.219	0.027	1.01	1.01
VF812743	6801	3.675	0.027	1.01	1.03
VF823444	6801	2.942	0.027	0.95	0.93
VF885577	6801	3.110	0.027	0.97	0.97
VF802937	6801	2.956	0.027	1.03	1.03
VF880849	6801	3.135	0.027	1.21	1.37
VF812762	6801	3.135	0.027	0.98	0.97
VF492856	6801	2.801	0.027	0.86	0.81
VF492216	6801	2.694	0.027	1.06	1.07
VH148416	6801	2.873	0.027	1.02	1.04
VH146719	6801	3.394	0.027	0.89	0.87
VF804260	6801	3.097	0.027	0.91	0.87
VF883648	6801	1.565	0.032	0.95	0.93
VF493034	6801	2.907	0.027	0.93	0.90
VF823848	6801	3.282	0.027	0.99	0.98
VH122227	6801	4.455	0.031	0.97	1.06
VF809838	6801	2.225	0.028	0.96	0.95
VF492199	6801	2.527	0.027	0.90	0.85
VF823784	6801	2.214	0.028	0.94	0.89
VF865996	6801	3.567	0.027	0.96	0.96
VH121958	6801	3.623	0.027	0.96	0.96
VF880646	6801	2.815	0.027	0.91	0.87
VH122429	6801	2.834	0.027	1.05	1.06
VH145979	6801	3.514	0.027	1.16	1.21
VH148455	6801	2.907	0.027	0.94	0.90
VH147216	6801	1.665	0.031	0.89	0.78
VH138993	6801	3.106	0.027	0.99	1.01
VF491975	6801	2.461	0.027	0.90	0.86

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF863290	6801	4.342	0.030	1.15	1.24
VH137600	6801	3.160	0.027	0.84	0.81
VF492385	6801	3.044	0.027	1.07	1.08
VF880501	6801	2.511	0.027	0.95	0.92
VF883670	6801	3.037	0.027	0.97	0.97
VF866064	6801	3.383	0.027	0.96	0.97
VF492586	6801	3.394	0.027	1.10	1.12
VF492212	6801	2.742	0.027	1.06	1.09
VH139013	6801	2.678	0.027	1.00	0.98
VF865675	6801	3.124	0.027	1.06	1.04
VF809001	6801	3.590	0.027	0.89	0.88
VF492268	6801	3.761	0.028	1.19	1.25
VF804267	6801	3.320	0.027	1.10	1.13
VF822412	6801	2.755	0.027	0.91	0.87
VF812997	6801	3.116	0.027	0.95	0.92
VH148444	6801	2.141	0.028	0.99	1.01
VH137854	6801	3.246	0.027	1.01	1.01
VH137656	6801	3.370	0.027	0.97	0.97
VH139254	6801	3.082	0.027	1.09	1.10
VF880680	6801	2.759	0.027	0.97	1.02
VH122468	6801	3.422	0.027	1.00	1.01
VF822465	6801	3.086	0.027	0.91	0.89
VF492436	6801	1.963	0.029	0.93	0.90
VH122050	6801	2.888	0.027	1.04	1.02
VF810708	6801	1.915	0.030	0.91	0.84

Science

Table I13. Science Grade 4 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF287740	7464	0.672	0.026	1.12	1.17
VF287742	7464	0.549	0.026	1.05	1.08
VF430825	7464	-0.477	0.028	1.01	0.98
VF287745	7464	0.718	0.026	1.19	1.25
VH102818	7464	-0.715	0.030	1.02	1.09
VH102841	7464	0.286	0.026	0.94	0.91
VH102847	7464	-0.729	0.030	1.02	1.12
VH102875	7464	0.532	0.026	1.06	1.09
VF387280	7464	1.648	0.027	1.07	1.15
VF387256	7464	0.219	0.026	0.98	1.02
VF431142	7464	-0.089	0.027	0.97	0.96
VF283606	7464	-1.498	0.036	0.99	0.89
VF431081	7464	-0.974	0.031	0.91	1.08
VF287766	7464	-0.384	0.028	0.94	0.89
VF287770	7464	-1.987	0.042	1.01	1.22
VH123679	7464	-0.473	0.028	0.97	0.92
VH123681	7464	0.051	0.027	0.89	0.83
VF430634	7464	0.179	0.026	0.99	0.98
VF430638	7464	1.214	0.026	1.09	1.13
VH123637	7464	0.073	0.026	0.93	0.89
VH123641	7464	-1.116	0.032	0.89	0.76
VF430984	7464	-0.201	0.027	1.05	1.06
VF430987	7464	1.088	0.026	1.07	1.12
VH129725	7464	0.571	0.026	1.03	1.04
VH129733	7464	0.056	0.027	0.88	0.81
VH129781	7464	0.659	0.026	1.08	1.12
VF431112	7464	0.036	0.027	1.02	1.03
VF431113	7464	-0.249	0.027	1.07	1.18
VF269830	7464	0.775	0.026	1.17	1.21
VF269831	7464	0.623	0.026	1.06	1.07
VF406427	7464	0.586	0.026	0.95	0.92
VF407139	7464	0.343	0.026	1.05	1.03
VF393954	7464	-0.648	0.029	0.90	0.84
VF393826	7464	-0.533	0.029	0.97	0.95
VH146166	7464	-0.242	0.027	0.89	0.83
VH146172	7464	-0.323	0.028	0.98	1.00
VH126160	7464	-0.395	0.028	0.90	0.84
VH126184	7464	-0.852	0.031	0.94	0.90

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF393695	7464	-1.271	0.034	0.91	0.86
VF393721	7464	0.237	0.026	1.05	1.05
VF386736	7464	-1.177	0.033	0.86	0.72
VF386739	7464	-0.419	0.028	0.93	0.89
VH149131	7464	0.282	0.026	0.99	0.98
VH149133	7464	1.447	0.027	1.07	1.18
VF431073	7464	0.090	0.026	0.91	0.87
VF431078	7464	0.797	0.026	0.99	1.01
VF430640	7464	0.504	0.026	0.97	0.98
VF282669	7464	0.471	0.026	0.98	0.96
VF269779	7464	-0.555	0.029	0.93	0.88
VF269769	7464	0.493	0.026	1.06	1.09

Table I14. Science Grade 8 IRT Statistics for Operational Items

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF309025	6789	0.007	0.027	0.99	1.00
VF309015	6789	0.304	0.027	1.13	1.17
VF308924	6789	0.250	0.027	1.00	1.01
VF308925	6789	-0.041	0.027	0.97	0.96
VF486687	6789	-0.916	0.029	0.93	0.88
VF431198	6789	0.042	0.027	1.14	1.19
VF313292	6789	-0.135	0.027	1.11	1.18
VF313289	6789	0.118	0.027	1.06	1.07
VF823953	6789	-0.213	0.027	1.05	1.08
VF823970	6789	0.032	0.027	1.05	1.06
VF862725	6789	-0.839	0.029	1.04	1.07
VF862740	6789	0.500	0.027	1.11	1.13
VF862757	6789	0.415	0.027	0.99	0.99
VF308708	6789	0.032	0.027	1.10	1.13
VF308712	6789	0.443	0.027	0.91	0.90
VF308713	6789	0.165	0.027	1.06	1.07
VF308932	6789	0.312	0.027	1.13	1.17
VF308933	6789	1.126	0.029	1.01	1.06
VF431368	6789	-0.020	0.027	0.90	0.88
VF308876	6789	-0.733	0.028	0.96	0.93
VF308882	6789	-0.875	0.029	0.96	0.94
VF813827	6789	-0.909	0.029	1.01	1.00
VF813887	6789	-1.414	0.032	0.88	0.75
VF308906	6789	-0.184	0.027	0.93	0.89
VF308910	6789	0.267	0.027	1.08	1.08
VF431549	6789	-0.560	0.028	0.90	0.84
VF431554	6789	0.297	0.027	1.04	1.06
VF308941	6789	0.078	0.027	0.96	0.94
VF308944	6789	0.991	0.028	0.86	0.84
VF308886	6789	0.299	0.027	0.93	0.91
VF308887	6789	-0.518	0.028	0.89	0.84
VF308890	6789	0.084	0.027	1.03	1.04
VF431602	6789	0.667	0.027	0.93	0.92
VF431598	6789	0.298	0.027	0.94	0.92
VF431188	6789	-0.837	0.029	0.88	0.79
VF431193	6789	-0.014	0.027	1.12	1.19
VF431624	6789	0.243	0.027	1.00	0.99
VF431626	6789	-1.023	0.030	0.82	0.68
VF685187	6789	-0.465	0.027	0.95	0.94
VF671386	6789	0.379	0.027	1.04	1.06

Accession Number	*N*	Rasch Difficulty	Rasch *SE*	Infit	Outfit
VF431609	6789	0.061	0.027	1.13	1.20
VF431608	6789	-0.082	0.027	1.03	1.06
VF431704	6789	-0.204	0.027	0.93	0.91
VF431703	6789	0.312	0.027	0.94	0.94
VF394547	6789	0.003	0.027	1.00	0.99
VF394561	6789	-0.507	0.028	0.92	0.87
VH139756	6789	0.589	0.027	1.04	1.06
VH140079	6789	0.601	0.027	1.06	1.08
VF486149	6789	0.511	0.027	0.97	0.98
VF486146	6789	0.112	0.027	1.02	1.00

Appendix J: Frequency of Individual Accommodations for 2016 PAWS Tests

Students received the same accommodations for all subjects (reading, mathematics, and science). The only exceptions are for those accommodations shaded in yellow that were not allowed for the Reading test.

Reading

Table J1. Frequency of IEP Student's Standard Accommodations for Reading: Presentation

Code	Accommodation	3	4	5	6	7	8
1	Student uses a Braille Special Test Form.	2	0	0	0	0	0
2	Student uses a Large Print Special Test Form.	8	5	6	6	1	1
3	Student uses an Audio Special Test Form.	0	0	0	0	0	0
4	Student uses magnification devices.	5	2	0	2	2	2
5	Student uses color overlays to reduce glare or enhance text.	9	10	3	2	6	1
6	Student uses templates to reduce the amount of visible print.	20	6	11	6	8	8
7	Student uses tactile graphics.	1	0	0	0	0	0
8	Sign language interpreter signs directions in all content areas and/or signs test questions as written in all content areas EXCEPT Reading. The interpreter may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one interpreter be provided for each individual student.	1	2	1	2	0	0
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	2	2	6	0	1	2
10	A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.	144	160	162	73	110	41
11	Student asks for clarification of directions (NOT test questions or answer choices).	248	317	341	266	202	221
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	5	5	8	3	2	3
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	4	1	3	3	14	4

Table J2. Frequency of IEP Student's Standard Accommodations for Reading: Response

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.	21	23	20	20	6	6
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.	2	1	9	9	23	9
16	Student uses speech-to-text conversion or voice recognition in all content areas. The margins for this document should match as closely as possible the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	1	3	7	3	0
17	Student uses a Brailler. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	0	0	0	0	0
18	Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	0	0	0	1	1
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	73	69	49	54	62	19
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.	45	50	50	55	53	16

Table J3. Frequency of IEP Student's Standard Accommodations for Reading: Setting

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.	528	604	671	627	603	505

Table J4. Frequency of IEP Student's Standard Accommodations for Reading: Timing and Scheduling

Code	Accommodation	3	4	5	6	7	8
22	Student is provided with extended time to complete the assessment.	342	426	458	399	368	326
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	335	423	427	328	247	183
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	100	101	85	54	59	25

Table J5. Frequency of ELLs Standard Accommodations for Reading: Presentation

Code	Accommodation	3	4	5	6	7	8
25	A certified staff member or access assistant translates written directions to the student.	13	7	3	6	2	3
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.	64	28	25	20	4	14
27	A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one reader be provided for each individual student.	28	9	8	8	3	3
28	Student uses a bilingual dictionary provided by the school.	11	3	4	9	3	13

Table J6. Frequency of ELLs Standard Accommodations for Reading: Setting

Code	Accommodation	3	4	5	6	7	8
29	Student takes test in a different or individual location, or in a small group.	88	43	34	22	6	20

Table J7. Frequency of ELLs Standard Accommodations for Reading: Timing and Scheduling

Code	Accommodation	3	4	5	6	7	8
30	Student is provided with multiple, individual breaks as needed.	63	25	21	16	1	8
31	Student is allowed to complete the test over multiple days.	31	14	13	2	0	2

Mathematics

Table J8. Frequency of IEP Student's Standard Accommodations for Mathematics: Presentation

Code	Accommodation	3	4	5	6	7	8
1	Student uses a Braille Special Test Form.	1	0	0	0	0	0
2	Student uses a Large Print Special Test Form.	7	4	5	6	1	1
3	Student uses an Audio Special Test Form.	15	32	47	30	36	42
4	Student uses magnification devices.	4	2	0	2	1	2
5	Student uses color overlays to reduce glare or enhance text.	7	6	2	2	6	2
6	Student uses templates to reduce the amount of visible print.	15	4	6	10	10	11
7	Student uses tactile graphics.	1	0	0	0	1	0
8	Sign language interpreter signs directions in all content areas and/or signs test questions as written in all content areas EXCEPT Reading. The interpreter may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one interpreter be provided for each individual student.	1	3	1	3	0	0
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	1	6	4	2	4	7
10	A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.	417	473	478	395	337	266
11	Student asks for clarification of directions (NOT test questions or answer choices).	247	342	377	269	232	222
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	4	7	9	5	2	3
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	8	15	10	13	18	7

Table J9. Frequency of IEP Student's Standard Accommodations for Mathematics: Response

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.	22	20	22	18	7	11
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.	2	2	9	9	27	25
16	Student uses speech-to-text conversion or voice recognition in all content areas. The margins for this document should match as closely as possible the same space as is allowed in the Student Test and Answer Book. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	1	3	7	2	0
17	Student uses a Brailler. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	0	0	0	0	0
18	Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	0	0	3	1	1
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	78	75	53	53	62	20
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.	100	118	105	136	119	36

Table J10. Frequency of IEP Student's Standard Accommodations for Mathematics: Setting

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.	534	608	659	638	612	502

Table J11. Frequency of IEP Student's Standard Accommodations for Mathematics: Timing and Scheduling

Code	Accommodation	3	4	5	6	7	8
22	Student is provided with extended time to complete the assessment.	314	432	471	396	377	332
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	310	409	439	318	260	176
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	96	90	85	54	58	20

Table J12. Frequency of ELLs Standard Accommodations for Mathematics: Presentation

Code	Accommodation	3	4	5	6	7	8
25	A certified staff member or access assistant translates written directions to the student.	13	10	3	13	16	9
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.	64	31	23	27	15	15
27	A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one reader be provided for each individual student.	154	77	37	27	17	12
28	Student uses a bilingual dictionary provided by the school.	66	8	12	16	16	13

Table J13. Frequency of ELLs Standard Accommodations for Mathematics: Setting

Code	Accommodation	3	4	5	6	7	8
29	Student takes test in a different or individual location, or in a small group.	155	87	45	31	20	29

Table J14. Frequency of ELLs Standard Accommodations for Mathematics: Timing and Scheduling

Code	Accommodation	3	4	5	6	7	8
30	Student is provided with multiple, individual breaks as needed.	58	69	20	19	11	13
31	Student is allowed to complete the test over multiple days.	53	16	22	4	0	1

Science

Table J15. Frequency of IEP Student's Standard Accommodations for Science: Presentation

Code	Accommodation	4	8
1	Student uses a Braille Special Test Form.	0	0
2	Student uses a Large Print Special Test Form.	5	1
3	Student uses an Audio Special Test Form.	27	42
4	Student uses magnification devices.	3	2
5	Student uses color overlays to reduce glare or enhance text.	7	1
6	Student uses templates to reduce the amount of visible print.	2	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. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one interpreter be provided for each individual student.	2	0
9	A certified staff member or access assistant provides visual cues to students who are deaf or hard of hearing.	2	2
10	A certified staff member or access assistant reads directions word-for-word as written in all content areas and/or reads or re-reads test questions word-for-word as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. It is recommended that one reader be provided for each individual student.	487	273
11	Student asks for clarification of directions (NOT test questions or answer choices).	334	232
12	Student uses audio amplification devices, including and/or in addition to hearing aids to increase clarity.	4	9
13	Student uses text-to-speech software in all content areas EXCEPT Reading.	11	8

Table J16. Frequency of IEP Student's Standard Accommodations for Science: Response

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.	22	9
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.	1	25
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.	3	0
17	Student uses a Brailler. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	0	0
18	Student uses a tape recorder to record test responses rather than writing on a paper. A certified staff member or access assistant transcribes verbatim the student's work into the Student Test and Answer Book on the pages that the student's response is to be written.	1	1
19	A certified staff member or access assistant monitors the placement of student responses on the Student Test and Answer Book.	68	20
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.	75	37

Table J17. Frequency of IEP Student's Standard Accommodations for Science: Setting

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.	609	516

Table J18. Frequency of IEP Student's Standard Accommodations for Science: Timing and Scheduling

Code	Accommodation	4	8
22	Student is provided with extended time to complete the assessment.	433	318
23	Student is provided with multiple, individual breaks as needed, monitored by a teacher or access assistant.	412	174
24	Student takes the tests at the time of day when he or she is most likely to demonstrate peak performance.	90	25

Table J19. Frequency of ELLs Standard Accommodations for Science: Presentation

Code	Accommodation	4	8
25	A certified staff member or access assistant translates written directions to the student.	10	9
26	A certified staff member or access assistant re-reads, simplifies, or clarifies directions in English or in the student's primary language (NOT test questions or answer choices) without clueing correct responses.	29	7
27	A certified staff member or access assistant reads and/or re-reads test questions in English, word-for-word, exactly as written in all content areas EXCEPT Reading. Raters may not clarify, interpret, define word meanings, elaborate, or provide assistance to students. Raters need to be familiar with the terminology and symbols specific to the content. It is recommended that one reader be provided for each individual student.	78	11
28	Student uses a bilingual dictionary provided by the school.	15	12

Table J20. Frequency of ELLs Standard Accommodations for Science: Setting

Code	Accommodation	4	8
29	Student takes test in a different or individual location, or in a small group.	85	28

Table J21. Frequency of ELLs Standard Accommodations for Science: Timing and Scheduling

Code	Accommodation	4	8
30	Student is provided with multiple, individual breaks as needed.	58	16
31	Student is allowed to complete the test over multiple days.	15	1

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

Reading

Table K1. Reading Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	*N*	Mean Scale Score	*SD* Scale Score
3	Yes	1066	562.5	47.7
	No	6617	604.7	43.5
4	Yes	979	579.3	48.3
	No	6503	627.1	43.6
5	Yes	990	584.0	49.4
	No	6333	638.0	46.1
6	Yes	952	585.1	44.0
	No	6061	645.5	47.4
7	Yes	870	603.5	37.1
	No	6182	657.6	44.8
8	Yes	778	603.8	38.5
	No	6015	661.7	41.1

Mathematics

Table K2. Mathematics Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	*N*	Mean Scale Score	*SD* Scale Score
3	Yes	1059	572.8	51.6
	No	6624	609.1	52.9
4	Yes	981	612.8	48.6
	No	6504	652.3	50.6
5	Yes	985	625.3	46.9
	No	6340	672.0	53.7
6	Yes	949	640.8	40.7
	No	6062	686.5	46.4
7	Yes	867	661.4	33.7
	No	6188	705.3	46.4
8	Yes	781	670.2	27.3
	No	6023	714.5	42.5

Science

Table K3. Science Mean Scale Scores, Counts, and Scale Score Standard Deviations for IEP Students by Accommodation Status

Grade	Accommodated	*N*	Mean Scale Score	*SD* Scale Score
4	Yes	982	644.6	46.7
	No	6495	677.8	51.0
8	Yes	779	607.4	34.5
	No	6010	652.8	46.0

Appendix L: Scale Score Descriptive Statistics by Demographic Subgroup

Table L1. Summary Statistics of Reading, Mathematics, and Science Scale Score by Grade

Grade	*N*	Mean	*SD*
		Reading	
3	7683	598.9	46.5
4	7482	620.8	47.1
5	7323	630.7	50.1
6	7013	637.3	51.3
7	7052	650.9	47.4
8	6793	655.0	44.8
		Mathematics	
3	7683	604.1	54.2
4	7485	647.2	52.1
5	7325	665.8	55.2
6	7011	680.4	48.3
7	7055	699.9	47.3
8	6804	709.4	43.4
		Science	
4	7477	673.4	51.7
8	6789	647.6	47.1

Reading

Table L2. Summary Statistics of Reading Grade 3 Scale Score

Group	*N*	Mean	*SD*
Total	7683	598.9	46.5
Male	3984	596.2	47.2
Female	3696	601.7	45.4
Unknown	3	659.7	99.9
Amer. Indian	295	563.2	40.6
Asian	55	605.7	40.7
African American	89	588.4	48.3
Haw./Pac. Islander	11	577.1	48.1
Hispanic/Latino	997	578.4	44.8
White	6025	604.3	45.3
Multiracial	192	595.6	43.6
Unknown	19	585.8	61.7
Free/Reduced Lunch	2893	586.2	44.9
Not Free/Reduced Lunch	4790	606.5	45.8
Special Education	1066	562.5	47.7
Not Special Education	6617	604.7	43.5
English Language Learner	338	558.5	35.5
Not English Language Learner	7345	600.7	46.1

Table L3. Summary Statistics of Reading Grade 4 Scale Score

Group	*N*	Mean	*SD*
Total	7482	620.8	47.1
Male	3876	617.8	47.2
Female	3603	624.1	46.8
Unknown	3	595.3	64.6
Amer. Indian	287	581.9	46.4
Asian	60	627.9	40.7
African American	82	605.9	47.6
Haw./Pac. Islander	8	617.9	53.7
Hispanic/Latino	1089	604.3	45.4
White	5811	626.1	45.8
Multiracial	136	618.9	47.6
Unknown	9	600.1	52.7
Free/Reduced Lunch	2717	607.7	45.7
Not Free/Reduced Lunch	4765	628.3	46.2
Special Education	979	579.3	48.3
Not Special Education	6503	627.1	43.6
English Language Learner	187	563.5	35.8
Not English Language Learner	7295	622.3	46.4

Table L4. Summary Statistics of Reading Grade 5 Scale Score

Group	*N*	Mean	*SD*
Total	7323	630.7	50.1
Male	3686	627.0	50.4
Female	3633	634.5	49.5
Unknown	4	586.8	35.8
Amer. Indian	285	592.2	47.1
Asian	55	634.0	46.4
African American	77	618.2	45.3
Haw./Pac. Islander	15	597.8	48.7
Hispanic/Latino	1126	612.4	47.5
White	5613	636.7	49.1
Multiracial	144	625.5	46.9
Unknown	8	607.5	49.3
Free/Reduced Lunch	2624	616.4	48.5
Not Free/Reduced Lunch	4699	638.7	49.2
Special Education	990	584.0	49.4
Not Special Education	6333	638.0	46.1
English Language Learner	150	560.5	34.5
Not English Language Learner	7173	632.1	49.3

Table L5. Summary Statistics of Reading Grade 6 Scale Score

Group	*N*	Mean	*SD*
Total	7013	637.3	51.3
Male	3671	633.1	52.1
Female	3341	641.9	50.0
Unknown	1	615.0	.
Amer. Indian	263	598.3	45.8
Asian	66	653.1	60.2
African American	95	626.5	45.5
Haw./Pac. Islander	9	615.7	60.0
Hispanic/Latino	1000	618.1	47.8
White	5437	642.6	50.5
Multiracial	141	641.6	49.2
Unknown	2	623.0	11.3
Free/Reduced Lunch	2472	621.1	48.8
Not Free/Reduced Lunch	4541	646.1	50.4
Special Education	952	585.1	44.0
Not Special Education	6061	645.5	47.4
English Language Learner	106	570.7	31.9
Not English Language Learner	6907	638.3	50.9

Table L6. Summary Statistics of Reading Grade 7 Scale Score

Group	*N*	Mean	*SD*
Total	7052	650.9	47.4
Male	3638	646.7	47.6
Female	3413	655.4	46.8
Unknown	1	749.0	.
Amer. Indian	262	612.3	38.5
Asian	78	659.5	48.4
African American	73	637.7	52.2
Haw./Pac. Islander	10	634.3	51.3
Hispanic/Latino	963	633.4	42.6
White	5515	656.0	46.9
Multiracial	145	646.4	47.1
Unknown	6	655.2	70.5
Free/Reduced Lunch	2451	634.6	44.2
Not Free/Reduced Lunch	4601	659.6	46.8
Special Education	870	603.5	37.1
Not Special Education	6182	657.6	44.8
English Language Learner	109	596.7	30.6
Not English Language Learner	6943	651.8	47.1

Table L7. Summary Statistics of Reading Grade 8 Scale Score

Group	*N*	Mean	*SD*
Total	6793	655.0	44.8
Male	3529	648.5	45.0
Female	3264	662.1	43.5
Unknown	.	.	.
Amer. Indian	254	629.6	41.5
Asian	62	669.1	50.0
African American	66	632.4	54.6
Haw./Pac. Islander	10	630.8	35.0
Hispanic/Latino	931	642.0	42.9
White	5329	658.6	44.2
Multiracial	137	656.6	41.9
Unknown	4	651.0	75.6
Free/Reduced Lunch	2236	641.9	43.9
Not Free/Reduced Lunch	4557	661.5	43.8
Special Education	778	603.8	38.5
Not Special Education	6015	661.7	41.1
English Language Learner	137	605.0	35.3
Not English Language Learner	6656	656.1	44.4

Mathematics

Table L8. Summary Statistics of Mathematics Grade 3 Scale Score

Group	*N*	Mean	*SD*
Total	7683	604.1	54.2
Male	3982	607.8	56.4
Female	3700	600.2	51.5
Unknown	1	674.0	.
Amer. Indian	293	564.5	43.6
Asian	60	629.0	56.2
African American	90	583.1	51.5
Haw./Pac. Islander	11	589.0	63.3
Hispanic/Latino	1004	583.4	51.8
White	6016	610.0	53.3
Multiracial	191	596.2	54.3
Unknown	18	559.9	39.3
Free/Reduced Lunch	2898	590.8	51.2
Not Free/Reduced Lunch	4785	612.2	54.4
Special Education	1059	572.8	51.6
Not Special Education	6624	609.1	52.9
English Language Learner	346	568.4	46.4
Not English Language Learner	7337	605.8	54.0

Table L9. Summary Statistics of Mathematics Grade 4 Scale Score

Group	*N*	Mean	*SD*
Total	7485	647.2	52.1
Male	3879	651.2	54.5
Female	3603	642.8	48.9
Unknown	3	638.7	110.0
Amer. Indian	287	607.3	46.6
Asian	62	648.8	47.3
African American	83	623.2	46.5
Haw./Pac. Islander	8	622.4	59.2
Hispanic/Latino	1095	630.3	46.9
White	5801	653.0	51.7
Multiracial	137	638.7	49.0
Unknown	12	604.1	61.7
Free/Reduced Lunch	2726	635.4	49.2
Not Free/Reduced Lunch	4759	653.9	52.5
Special Education	981	612.8	48.6
Not Special Education	6504	652.3	50.6
English Language Learner	188	601.0	31.9
Not English Language Learner	7297	648.3	52.0

Table L10. Summary Statistics of Mathematics Grade 5 Scale Score

Group	*N*	Mean	*SD*
Total	7325	665.8	55.2
Male	3688	665.9	56.6
Female	3633	665.7	53.8
Unknown	4	627.0	58.9
Amer. Indian	284	626.9	44.5
Asian	58	687.0	62.5
African American	79	646.9	51.7
Haw./Pac. Islander	15	648.9	60.3
Hispanic/Latino	1126	648.1	49.1
White	5611	671.5	55.2
Multiracial	144	660.8	52.7
Unknown	8	634.6	57.8
Free/Reduced Lunch	2624	652.7	51.2
Not Free/Reduced Lunch	4701	673.1	56.0
Special Education	985	625.3	46.9
Not Special Education	6340	672.0	53.7
English Language Learner	157	612.9	36.2
Not English Language Learner	7168	666.9	55.0

Table L11. Summary Statistics of Mathematics Grade 6 Scale Score

Group	*N*	Mean	*SD*
Total	7011	680.4	48.3
Male	3668	678.9	49.3
Female	3342	682.0	47.2
Unknown	1	661.0	.
Amer. Indian	261	644.5	39.0
Asian	66	699.7	54.3
African American	95	669.9	41.0
Haw./Pac. Islander	9	656.6	45.1
Hispanic/Latino	1004	664.9	43.0
White	5433	684.9	48.3
Multiracial	141	681.5	48.6
Unknown	2	645.5	21.9
Free/Reduced Lunch	2473	666.4	42.7
Not Free/Reduced Lunch	4538	688.0	49.5
Special Education	949	640.8	40.7
Not Special Education	6062	686.5	46.4
English Language Learner	109	634.2	40.5
Not English Language Learner	6902	681.1	48.1

Table L12. Summary Statistics of Mathematics Grade 7 Scale Score

Group	*N*	Mean	*SD*
Total	7055	699.9	47.3
Male	3638	699.4	47.9
Female	3416	700.3	46.6
Unknown	1	843.0	.
Amer. Indian	262	667.7	35.8
Asian	80	723.1	52.8
African American	72	684.9	46.8
Haw./Pac. Islander	11	690.7	51.5
Hispanic/Latino	973	684.8	41.5
White	5506	704.1	47.4
Multiracial	145	692.8	46.4
Unknown	6	732.2	69.2
Free/Reduced Lunch	2449	686.0	41.8
Not Free/Reduced Lunch	4606	707.2	48.4
Special Education	867	661.4	33.7
Not Special Education	6188	705.3	46.4
English Language Learner	114	655.9	28.1
Not English Language Learner	6941	700.6	47.2

Table L13. Summary Statistics of Mathematics Grade 8 Scale Score

Group	*N*	Mean	*SD*
Total	6804	709.4	43.4
Male	3535	708.3	43.9
Female	3269	710.6	42.9
Unknown	-	-	-
Amer. Indian	254	682.6	32.4
Asian	65	726.7	49.9
African American	66	690.3	42.3
Haw./Pac. Islander	10	673.0	22.4
Hispanic/Latino	939	697.0	38.1
White	5329	713.1	43.9
Multiracial	137	704.8	37.0
Unknown	4	695.5	42.8
Free/Reduced Lunch	2238	697.7	38.8
Not Free/Reduced Lunch	4566	715.1	44.4
Special Education	781	670.2	27.3
Not Special Education	6023	714.5	42.5
English Language Learner	140	673.9	26.7
Not English Language Learner	6664	710.2	43.4

Science

Table L14. Summary Statistics of Science Grade 4 Scale Score

Group	*N*	Mean	*SD*
Total	7477	673.4	51.7
Male	3877	674.2	52.7
Female	3597	672.5	50.6
Unknown	3	691.3	41.5
Amer. Indian	283	624.8	42.5
Asian	62	671.1	42.8
African American	83	651.3	48.0
Haw./Pac. Islander	8	665.9	59.3
Hispanic/Latino	1093	652.2	47.4
White	5802	680.3	50.5
Multiracial	136	667.1	51.7
Unknown	10	664.8	37.5
Free/Reduced Lunch	2713	659.2	47.9
Not Free/Reduced Lunch	4764	681.5	52.0
Special Education	982	644.6	46.7
Not Special Education	6495	677.8	51.0
English Language Learner	185	618.2	31.8
Not English Language Learner	7292	674.8	51.3

Table L15. Summary Statistics of Science Grade 8 Scale Score

Group	*N*	Mean	*SD*
Total	6789	647.6	47.1
Male	3524	648.0	48.2
Female	3265	647.1	45.8
Unknown	-	-	-
Amer. Indian	253	617.2	40.6
Asian	65	652.2	53.7
African American	66	625.9	43.3
Haw./Pac. Islander	10	613.5	35.8
Hispanic/Latino	936	628.9	41.6
White	5320	652.6	46.9
Multiracial	135	644.7	42.9
Unknown	4	655.8	44.1
Free/Reduced Lunch	2231	633.1	43.0
Not Free/Reduced Lunch	4558	654.6	47.4
Special Education	779	607.4	34.5
Not Special Education	6010	652.8	46.0
English Language Learner	140	598.2	24.6
Not English Language Learner	6649	648.6	46.9

Appendix M: Raw Score to Scale Score Tables

Reading

Table M1. Reading Grade 3 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-4.780	375	63	1
1	-4.070	375	45	1
2	-3.344	406	32	1
3	-2.905	426	27	1
4	-2.583	440	23	1
5	-2.326	451	21	1
6	-2.109	461	20	1
7	-1.920	469	19	1
8	-1.752	476	18	1
9	-1.598	483	17	1
10	-1.457	489	16	1
11	-1.325	495	16	1
12	-1.201	500	15	1
13	-1.084	506	15	1
14	-0.971	511	15	1
15	-0.863	515	14	1
16	-0.759	520	14	1
17	-0.658	524	14	1
18	-0.560	529	14	1
19	-0.464	533	14	1
20	-0.370	537	13	1
21	-0.277	541	13	1
22	-0.185	545	13	1
23	-0.095	549	13	1
24	-0.004	553	13	2
25	0.085	557	13	2
26	0.175	561	13	2
27	0.265	565	13	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
28	0.356	569	13	2
29	0.447	573	13	2
30	0.539	577	13	2
31	0.633	581	13	2
32	0.728	585	14	2
33	0.826	589	14	2
34	0.926	594	14	3
35	1.029	598	14	3
36	1.136	603	14	3
37	1.247	608	15	3
38	1.363	613	15	3
39	1.485	618	16	3
40	1.616	624	16	3
41	1.755	630	17	3
42	1.907	637	17	3
43	2.073	644	18	4
44	2.260	652	20	4
45	2.475	662	21	4
46	2.730	673	23	4
47	3.049	687	27	4
48	3.485	706	32	4
49	4.209	738	45	4
50	4.917	769	63	4

Table M2. Reading Grade 4 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-4.549	400	63	1
1	-3.837	400	45	1
2	-3.106	417	32	1
3	-2.662	436	27	1
4	-2.336	451	24	1
5	-2.073	462	22	1
6	-1.851	472	20	1
7	-1.656	480	19	1
8	-1.482	488	18	1
9	-1.323	495	17	1
10	-1.176	502	17	1
11	-1.038	508	16	1
12	-0.908	513	16	1
13	-0.784	519	15	1
14	-0.666	524	15	1
15	-0.552	529	15	1
16	-0.442	534	14	1
17	-0.335	538	14	1
18	-0.230	543	14	1
19	-0.128	548	14	1
20	-0.028	552	14	1
21	0.071	556	14	1
22	0.169	561	14	1
23	0.266	565	14	1
24	0.362	569	14	2
25	0.458	573	14	2
26	0.553	577	14	2
27	0.649	582	14	2
28	0.746	586	14	2
29	0.843	590	14	2
30	0.941	594	14	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
31	1.041	599	14	2
32	1.142	603	14	2
33	1.246	608	14	3
34	1.352	613	14	3
35	1.461	617	15	3
36	1.574	622	15	3
37	1.691	627	15	3
38	1.814	633	16	3
39	1.943	638	16	3
40	2.080	644	17	3
41	2.227	651	17	3
42	2.385	658	18	3
43	2.559	665	19	4
44	2.754	674	20	4
45	2.976	684	22	4
46	3.240	695	24	4
47	3.567	710	27	4
48	4.013	729	32	4
49	4.746	761	45	4
50	5.460	793	63	4

Table M3. Reading Grade 5 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-4.561	425	63	1
1	-3.852	425	45	1
2	-3.127	425	32	1
3	-2.690	435	27	1
4	-2.370	449	23	1
5	-2.115	460	21	1
6	-1.900	470	20	1
7	-1.713	478	18	1
8	-1.546	485	17	1
9	-1.395	492	17	1
10	-1.256	498	16	1
11	-1.127	504	16	1
12	-1.005	509	15	1
13	-0.890	514	15	1
14	-0.781	519	14	1
15	-0.676	523	14	1
16	-0.575	528	14	1
17	-0.477	532	14	1
18	-0.382	536	13	1
19	-0.290	540	13	1
20	-0.199	544	13	1
21	-0.110	548	13	1
22	-0.023	552	13	1
23	0.063	556	13	1
24	0.148	560	13	1
25	0.233	563	13	1
26	0.317	567	13	1
27	0.401	571	13	1
28	0.485	574	13	1
29	0.569	578	13	2
30	0.653	582	13	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
31	0.738	586	13	2
32	0.824	589	13	2
33	0.910	593	13	2
34	0.999	597	13	2
35	1.088	601	13	2
36	1.180	605	13	2
37	1.274	609	14	2
38	1.371	613	14	2
39	1.471	618	14	2
40	1.575	622	14	3
41	1.683	627	15	3
42	1.797	632	15	3
43	1.917	637	15	3
44	2.045	643	16	3
45	2.183	649	17	3
46	2.332	656	17	3
47	2.497	663	18	3
48	2.682	671	19	4
49	2.895	680	21	4
50	3.148	691	23	4
51	3.466	705	26	4
52	3.901	724	32	4
53	4.624	756	45	4
54	5.331	787	63	4

Table M4. Reading Grade 6 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-4.138	450	63	1
1	-3.429	450	45	1
2	-2.706	450	32	1
3	-2.270	454	26	1
4	-1.951	468	23	1
5	-1.697	479	21	1
6	-1.484	488	20	1
7	-1.299	496	18	1
8	-1.134	503	17	1
9	-0.985	510	17	1
10	-0.847	516	16	1
11	-0.720	522	15	1
12	-0.600	527	15	1
13	-0.487	532	15	1
14	-0.380	536	14	1
15	-0.277	541	14	1
16	-0.178	545	14	1
17	-0.083	550	13	1
18	0.009	554	13	1
19	0.099	558	13	1
20	0.187	561	13	1
21	0.273	565	13	1
22	0.357	569	13	1
23	0.440	572	13	1
24	0.522	576	13	1
25	0.604	580	12	1
26	0.684	583	12	1
27	0.764	587	12	1
28	0.844	590	12	2
29	0.924	594	12	2
30	1.004	597	12	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
31	1.084	601	12	2
32	1.165	604	13	2
33	1.246	608	13	2
34	1.328	611	13	2
35	1.412	615	13	2
36	1.497	619	13	2
37	1.583	623	13	2
38	1.672	627	13	2
39	1.763	631	13	3
40	1.857	635	14	3
41	1.954	639	14	3
42	2.055	643	14	3
43	2.160	648	14	3
44	2.271	653	15	3
45	2.389	658	15	3
46	2.514	664	16	3
47	2.649	669	16	3
48	2.795	676	17	3
49	2.957	683	18	4
50	3.140	691	19	4
51	3.350	700	21	4
52	3.601	711	23	4
53	3.916	725	26	4
54	4.348	744	32	4
55	5.068	776	44	4
56	5.774	807	63	4

Table M5. Reading Grade 7 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-3.663	475	63	1
1	-2.954	475	45	1
2	-2.228	475	32	1
3	-1.790	475	27	1
4	-1.470	489	23	1
5	-1.215	500	21	1
6	-1.000	509	20	1
7	-0.813	517	18	1
8	-0.647	525	17	1
9	-0.496	531	17	1
10	-0.357	537	16	1
11	-0.228	543	15	1
12	-0.107	548	15	1
13	0.007	553	15	1
14	0.116	558	14	1
15	0.220	563	14	1
16	0.320	567	14	1
17	0.416	571	14	1
18	0.510	576	13	1
19	0.601	580	13	1
20	0.690	583	13	1
21	0.777	587	13	1
22	0.863	591	13	1
23	0.948	595	13	1
24	1.031	598	13	1
25	1.113	602	13	1
26	1.195	606	13	2
27	1.277	609	13	2
28	1.358	613	13	2
29	1.439	616	13	2
30	1.521	620	13	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
31	1.602	623	13	2
32	1.684	627	13	2
33	1.767	631	13	2
34	1.851	634	13	2
35	1.936	638	13	2
36	2.022	642	13	3
37	2.111	646	13	3
38	2.201	650	13	3
39	2.293	654	13	3
40	2.389	658	14	3
41	2.487	662	14	3
42	2.590	667	14	3
43	2.697	672	15	3
44	2.809	676	15	3
45	2.928	682	15	3
46	3.055	687	16	3
47	3.192	693	17	4
48	3.340	700	17	4
49	3.504	707	18	4
50	3.688	715	19	4
51	3.900	724	21	4
52	4.152	735	23	4
53	4.469	749	26	4
54	4.904	768	32	4
55	5.625	800	45	4
56	6.333	831	63	4

Table M6. Reading Grade 8 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-3.793	500	63	1
1	-3.086	500	45	1
2	-2.365	500	32	1
3	-1.932	500	26	1
4	-1.615	500	23	1
5	-1.363	500	21	1
6	-1.151	503	19	1
7	-0.967	511	18	1
8	-0.803	518	17	1
9	-0.654	524	17	1
10	-0.518	530	16	1
11	-0.391	536	15	1
12	-0.271	541	15	1
13	-0.158	546	15	1
14	-0.051	551	14	1
15	0.052	555	14	1
16	0.152	560	14	1
17	0.248	564	14	1
18	0.341	568	13	1
19	0.432	572	13	1
20	0.521	576	13	1
21	0.608	580	13	1
22	0.694	584	13	1
23	0.779	587	13	1
24	0.862	591	13	1
25	0.945	595	13	1
26	1.028	598	13	1
27	1.110	602	13	1
28	1.192	605	13	1
29	1.275	609	13	1
30	1.357	613	13	1

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
31	1.440	616	13	2
32	1.524	620	13	2
33	1.608	624	13	2
34	1.693	627	13	2
35	1.780	631	13	2
36	1.869	635	13	2
37	1.959	639	13	2
38	2.051	643	13	2
39	2.146	647	14	2
40	2.243	652	14	2
41	2.345	656	14	3
42	2.450	661	14	3
43	2.559	665	15	3
44	2.675	671	15	3
45	2.796	676	16	3
46	2.926	682	16	3
47	3.066	688	17	3
48	3.217	694	17	3
49	3.384	702	18	3
50	3.572	710	20	3
51	3.787	719	21	4
52	4.043	731	23	4
53	4.363	745	27	4
54	4.801	764	32	4
55	5.526	796	45	4
56	6.236	827	63	4

Mathematics

Table M7. Mathematics Grade 3 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-5.002	375	62	1
1	-4.290	384	44	1
2	-3.559	416	32	1
3	-3.115	435	26	1
4	-2.789	449	23	1
5	-2.528	461	21	1
6	-2.307	470	20	1
7	-2.114	479	18	1
8	-1.941	486	18	1
9	-1.784	493	17	1
10	-1.638	499	16	1
11	-1.503	505	16	1
12	-1.375	511	15	1
13	-1.254	516	15	1
14	-1.138	521	15	1
15	-1.026	526	14	1
16	-0.919	531	14	1
17	-0.815	535	14	1
18	-0.713	539	14	1
19	-0.614	544	14	1
20	-0.516	548	13	1
21	-0.420	552	13	2
22	-0.325	556	13	2
23	-0.232	560	13	2
24	-0.139	564	13	2
25	-0.046	568	13	2
26	0.047	572	13	2
27	0.139	576	13	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
28	0.232	580	13	2
29	0.326	585	13	2
30	0.421	589	13	2
31	0.517	593	13	2
32	0.614	597	14	2
33	0.714	601	14	3
34	0.816	606	14	3
35	0.921	610	14	3
36	1.029	615	14	3
37	1.142	620	15	3
38	1.260	625	15	3
39	1.384	630	15	3
40	1.516	636	16	3
41	1.657	642	17	3
42	1.810	649	17	3
43	1.978	656	18	3
44	2.166	664	19	4
45	2.382	674	21	4
46	2.638	685	23	4
47	2.958	699	26	4
48	3.395	718	32	4
49	4.120	749	44	4
50	4.828	780	62	4

Table M8. Mathematics Grade 4 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-4.448	400	62	1
1	-3.738	408	44	1
2	-3.011	440	32	1
3	-2.571	459	26	1
4	-2.249	473	23	1
5	-1.991	484	21	1
6	-1.774	493	19	1
7	-1.585	502	18	1
8	-1.416	509	17	1
9	-1.263	516	17	1
10	-1.122	522	16	1
11	-0.990	527	15	1
12	-0.867	533	15	1
13	-0.750	538	15	1
14	-0.639	543	14	1
15	-0.532	547	14	1
16	-0.430	552	14	1
17	-0.331	556	14	1
18	-0.235	560	13	1
19	-0.141	564	13	1
20	-0.050	568	13	1
21	0.039	572	13	1
22	0.127	576	13	1
23	0.213	580	13	1
24	0.299	583	13	1
25	0.383	587	13	2
26	0.466	591	13	2
27	0.549	594	12	2
28	0.631	598	12	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
29	0.713	601	12	2
30	0.795	605	12	2
31	0.877	608	12	2
32	0.960	612	12	2
33	1.042	616	13	2
34	1.125	619	13	2
35	1.209	623	13	2
36	1.294	627	13	2
37	1.379	630	13	2
38	1.467	634	13	2
39	1.555	638	13	3
40	1.646	642	13	3
41	1.738	646	13	3
42	1.833	650	13	3
43	1.931	654	14	3
44	2.032	659	14	3
45	2.137	663	14	3
46	2.247	668	15	3
47	2.362	673	15	3
48	2.483	678	15	3
49	2.612	684	16	3
50	2.751	690	16	3
51	2.902	696	17	3
52	3.068	704	18	4
53	3.254	712	19	4
54	3.468	721	21	4
55	3.723	732	23	4
56	4.041	746	26	4
57	4.477	765	32	4
58	5.201	796	44	4

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
59	5.909	827	62	4

Table M9. Mathematics Grade 5 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-3.489	425	62	1
1	-2.782	450	44	1
2	-2.060	481	32	1
3	-1.626	500	26	1
4	-1.309	514	23	1
5	-1.057	525	21	1
6	-0.846	534	19	1
7	-0.662	542	18	1
8	-0.499	549	17	1
9	-0.351	555	16	1
10	-0.215	561	16	1
11	-0.090	567	15	1
12	0.028	572	15	1
13	0.140	576	14	1
14	0.246	581	14	1
15	0.347	585	14	1
16	0.444	590	13	1
17	0.537	594	13	1
18	0.628	598	13	1
19	0.716	601	13	1
20	0.802	605	13	1
21	0.886	609	13	2
22	0.969	612	12	2
23	1.050	616	12	2
24	1.130	619	12	2
25	1.209	623	12	2
26	1.287	626	12	2
27	1.365	630	12	2
28	1.442	633	12	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
29	1.519	636	12	2
30	1.596	640	12	2
31	1.673	643	12	2
32	1.751	646	12	2
33	1.828	650	12	2
34	1.907	653	12	3
35	1.986	657	12	3
36	2.066	660	12	3
37	2.147	664	12	3
38	2.230	667	13	3
39	2.314	671	13	3
40	2.400	675	13	3
41	2.488	678	13	3
42	2.579	682	13	3
43	2.673	686	13	3
44	2.771	691	14	3
45	2.872	695	14	3
46	2.978	700	14	3
47	3.090	705	15	3
48	3.208	710	15	3
49	3.334	715	16	3
50	3.470	721	16	3
51	3.618	727	17	4
52	3.781	735	18	4
53	3.965	743	19	4
54	4.177	752	21	4
55	4.429	763	23	4
56	4.746	776	26	4
57	5.180	795	32	4
58	5.902	827	44	4

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
59	6.610	857	62	4

Table M10. Mathematics Grade 6 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-3.029	450	62	1
1	-2.320	470	44	1
2	-1.595	501	32	1
3	-1.159	520	26	1
4	-0.839	534	23	1
5	-0.585	545	21	1
6	-0.371	554	19	1
7	-0.186	562	18	1
8	-0.021	569	17	1
9	0.129	576	16	1
10	0.266	582	16	1
11	0.393	587	15	1
12	0.512	593	15	1
13	0.625	598	14	1
14	0.732	602	14	1
15	0.834	607	14	1
16	0.932	611	13	1
17	1.026	615	13	1
18	1.117	619	13	1
19	1.206	623	13	1
20	1.293	627	13	1
21	1.377	630	13	2
22	1.460	634	12	2
23	1.541	637	12	2
24	1.622	641	12	2
25	1.701	644	12	2
26	1.779	648	12	2
27	1.857	651	12	2
28	1.934	654	12	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
29	2.011	658	12	2
30	2.088	661	12	2
31	2.164	664	12	2
32	2.241	668	12	2
33	2.318	671	12	2
34	2.396	674	12	2
35	2.475	678	12	3
36	2.554	681	12	3
37	2.634	685	12	3
38	2.716	688	12	3
39	2.799	692	13	3
40	2.884	696	13	3
41	2.971	699	13	3
42	3.060	703	13	3
43	3.153	707	13	3
44	3.248	711	14	3
45	3.348	716	14	3
46	3.452	720	14	3
47	3.562	725	15	3
48	3.678	730	15	3
49	3.802	735	16	3
50	3.936	741	16	3
51	4.081	748	17	4
52	4.242	755	18	4
53	4.424	762	19	4
54	4.633	772	21	4
55	4.882	782	23	4
56	5.196	796	26	4
57	5.627	815	31	4
58	6.346	846	44	4

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
59	7.051	876	62	4

Table M11. Mathematics Grade 7 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-2.234	475	62	1
1	-1.529	504	44	1
2	-0.811	535	31	1
3	-0.380	554	26	1
4	-0.066	568	23	1
5	0.183	578	21	1
6	0.391	587	19	1
7	0.572	595	18	1
8	0.733	602	17	1
9	0.878	609	16	1
10	1.011	614	15	1
11	1.134	620	15	1
12	1.249	625	15	1
13	1.358	629	14	1
14	1.461	634	14	1
15	1.560	638	13	1
16	1.655	642	13	1
17	1.746	646	13	1
18	1.835	650	13	1
19	1.920	654	13	2
20	2.004	657	13	2
21	2.086	661	12	2
22	2.166	664	12	2
23	2.245	668	12	2
24	2.323	671	12	2
25	2.400	675	12	2
26	2.476	678	12	2
27	2.551	681	12	2
28	2.626	684	12	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
29	2.701	688	12	2
30	2.776	691	12	2
31	2.851	694	12	2
32	2.926	697	12	3
33	3.001	701	12	3
34	3.077	704	12	3
35	3.154	707	12	3
36	3.232	711	12	3
37	3.311	714	12	3
38	3.391	718	12	3
39	3.473	721	13	3
40	3.556	725	13	3
41	3.642	728	13	3
42	3.731	732	13	3
43	3.822	736	13	3
44	3.917	740	13	3
45	4.015	745	14	3
46	4.119	749	14	3
47	4.228	754	15	4
48	4.343	759	15	4
49	4.466	764	16	4
50	4.599	770	16	4
51	4.744	776	17	4
52	4.905	783	18	4
53	5.086	791	19	4
54	5.294	800	21	4
55	5.544	811	23	4
56	5.857	825	26	4
57	6.288	843	31	4
58	7.007	875	44	4

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
59	7.712	905	62	4

Table M12. Mathematics Grade 8 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-2.095	500	62	1
1	-1.389	510	44	1
2	-0.670	541	31	1
3	-0.238	560	26	1
4	0.075	574	23	1
5	0.324	584	21	1
6	0.532	594	19	1
7	0.711	601	18	1
8	0.870	608	17	1
9	1.013	614	16	1
10	1.144	620	15	1
11	1.265	625	15	1
12	1.377	630	14	1
13	1.483	635	14	1
14	1.583	639	14	1
15	1.679	643	13	1
16	1.770	647	13	1
17	1.857	651	13	1
18	1.941	655	13	1
19	2.023	658	12	1
20	2.102	662	12	1
21	2.179	665	12	2
22	2.254	668	12	2
23	2.327	671	12	2
24	2.400	675	12	2
25	2.470	678	12	2
26	2.540	681	11	2
27	2.609	684	11	2
28	2.677	687	11	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
29	2.745	690	11	2
30	2.812	692	11	2
31	2.879	695	11	2
32	2.946	698	11	2
33	3.012	701	11	2
34	3.079	704	11	2
35	3.145	707	11	3
36	3.212	710	11	3
37	3.280	713	11	3
38	3.347	716	11	3
39	3.416	719	11	3
40	3.485	722	11	3
41	3.556	725	12	3
42	3.627	728	12	3
43	3.700	731	12	3
44	3.774	734	12	3
45	3.850	738	12	3
46	3.928	741	12	3
47	4.009	744	12	3
48	4.092	748	13	3
49	4.178	752	13	3
50	4.268	756	13	3
51	4.361	760	13	3
52	4.460	764	14	4
53	4.564	769	14	4
54	4.675	773	15	4
55	4.794	779	15	4
56	4.922	784	16	4
57	5.063	790	17	4
58	5.219	797	18	4

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
59	5.396	805	19	4
60	5.600	813	20	4
61	5.845	824	23	4
62	6.155	838	26	4
63	6.582	856	31	4
64	7.297	887	44	4
65	8.001	918	62	4

Science

Table M13. Science Grade 4 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-4.887	402	69	1
1	-4.176	436	49	1
2	-3.446	471	35	1
3	-3.005	493	29	1
4	-2.681	508	26	1
5	-2.422	521	23	1
6	-2.203	531	22	1
7	-2.013	540	20	1
8	-1.842	549	19	1
9	-1.688	556	19	1
10	-1.545	563	18	1
11	-1.413	569	17	1
12	-1.288	575	17	1
13	-1.169	581	16	1
14	-1.056	587	16	1
15	-0.948	592	16	1
16	-0.843	597	15	1
17	-0.742	602	15	1
18	-0.643	607	15	1
19	-0.547	611	15	1
20	-0.452	616	15	2
21	-0.359	620	15	2
22	-0.267	625	15	2
23	-0.176	629	15	2
24	-0.085	633	15	2
25	0.005	638	14	2
26	0.095	642	14	2
27	0.186	646	15	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
28	0.276	651	15	2
29	0.368	655	15	2
30	0.461	660	15	2
31	0.555	664	15	2
32	0.650	669	15	3
33	0.748	674	15	3
34	0.849	678	15	3
35	0.952	683	16	3
36	1.059	689	16	3
37	1.171	694	16	3
38	1.287	700	17	3
39	1.410	705	17	3
40	1.541	712	18	3
41	1.681	719	18	3
42	1.833	726	19	4
43	2.001	734	20	4
44	2.188	743	22	4
45	2.404	753	23	4
46	2.660	766	26	4
47	2.980	781	29	4
48	3.417	802	35	4
49	4.142	837	49	4
50	4.851	871	69	4

Table M14. Science Grade 8 Raw Score to Scale Score

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
0	-4.760	408	69	1
1	-4.053	442	49	1
2	-3.332	477	35	1
3	-2.898	498	29	1
4	-2.582	513	25	1
5	-2.330	525	23	1
6	-2.118	535	21	1
7	-1.934	544	20	1
8	-1.770	552	19	1
9	-1.622	559	18	1
10	-1.485	566	17	1
11	-1.358	572	17	1
12	-1.239	578	16	1
13	-1.126	583	16	1
14	-1.019	588	16	1
15	-0.916	593	15	1
16	-0.816	598	15	1
17	-0.720	603	15	1
18	-0.627	607	15	2
19	-0.535	612	15	2
20	-0.446	616	14	2
21	-0.357	620	14	2
22	-0.270	624	14	2
23	-0.184	629	14	2
24	-0.099	633	14	2
25	-0.013	637	14	2
26	0.072	641	14	2
27	0.157	645	14	2
28	0.243	649	14	2

Raw Score	Theta	Rounded Scale Score	Standard Error	Performance Level
29	0.330	653	14	2
30	0.417	658	14	3
31	0.506	662	14	3
32	0.597	666	15	3
33	0.690	671	15	3
34	0.785	675	15	3
35	0.884	680	15	3
36	0.986	685	16	3
37	1.092	690	16	3
38	1.203	695	16	3
39	1.321	701	17	3
40	1.446	707	17	3
41	1.581	714	18	4
42	1.728	721	19	4
43	1.890	729	20	4
44	2.072	737	21	4
45	2.281	747	23	4
46	2.531	760	25	4
47	2.845	775	29	4
48	3.276	795	35	4
49	3.995	830	49	4
50	4.701	864	69	4

Appendix N: Performance Level Percentages by Demographic Subgroup

Table N1.Performance Levels of Reading, Mathematics, and Science by Grade Revised

Grade	Below Basic	Basic	Proficient	Advanced
		Reading		
3	15.5	26.9	39.4	18.2
4	12.8	22.0	46.5	18.8
5	14.1	25.2	38.1	22.6
6	18.0	24.3	36.7	20.9
7	18.1	22.2	38.8	20.9
8	17.6	29.0	45.4	8.0
		Mathematics		
3	16.9	30.5	37.3	15.3
4	11.0	33.9	39.6	15.5
5	13.9	30.1	41.7	14.4
6	13.4	37.1	39.8	9.8
7	14.8	36.2	35.1	13.9
8	13.1	39.4	34.7	12.8
		Science		
4	12.9	32.7	37.1	17.2
8	19.6	39.1	30.8	10.4

Reading

Table N2. Performance Levels of Reading by Grade 3 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	15.5	26.9	39.4	18.2
Male	17.3	27.0	38.4	17.3
Female	13.5	26.7	40.6	19.2
Unknown	0.0	33.3	33.3	33.3
Amer. Indian	39.7	36.3	20.0	4.1
Asian	12.7	21.8	45.5	20.0
African American	19.1	33.7	36.0	11.2
Haw./Pac. Islander	18.2	54.5	9.1	18.2
Hispanic/Latino	28.6	32.0	30.6	8.8
White	12.0	25.3	41.9	20.7
Multiracial	15.1	30.7	40.1	14.1
Unknown	31.6	26.3	31.6	10.5
Free/Reduced Lunch	21.7	32.1	35.2	11.0
Not Free/Reduced Lunch	11.7	23.7	42.0	22.6
Special Education	43.5	30.1	19.5	6.8
Not Special Education	10.9	26.3	42.6	20.1
English Language Learner	43.8	36.4	18.9	0.9
Not English Language Learner	14.2	26.4	40.4	19.0

Table N3. Performance Levels of Reading by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	12.8	22.0	46.5	18.8
Male	14.4	22.7	45.6	17.3
Female	11.0	21.1	47.5	20.3
Unknown	33.3	33.3	33.3	0.0
Amer. Indian	36.6	34.1	24.4	4.9
Asian	10.0	18.3	51.7	20.0
African American	18.3	31.7	36.6	13.4
Haw./Pac. Islander	25.0	12.5	25.0	37.5
Hispanic/Latino	20.2	29.9	40.0	9.8
White	10.1	19.8	49.0	21.2
Multiracial	15.4	22.8	44.1	17.6
Unknown	22.2	33.3	44.4	0.0
Free/Reduced Lunch	19.1	26.1	43.7	11.1
Not Free/Reduced Lunch	9.2	19.6	48.1	23.1
Special Education	41.9	29.6	22.8	5.7
Not Special Education	8.4	20.8	50.1	20.7
English Language Learner	54.5	31.6	13.9	0.0
Not English Language Learner	11.7	21.7	47.3	19.2

Table N4. Performance Levels of Reading by Grade 5 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	14.1	25.2	38.1	22.6
Male	16.1	26.3	37.1	20.5
Female	12.0	24.1	39.1	24.7
Unknown	50.0	25.0	25.0	0.0
Amer. Indian	39.3	34.7	19.6	6.3
Asian	7.3	30.9	38.2	23.6
African American	19.5	33.8	36.4	10.4
Haw./Pac. Islander	26.7	33.3	33.3	6.7
Hispanic/Latino	21.6	33.7	32.8	12.0
White	11.3	22.8	39.9	26.0
Multiracial	13.9	26.4	46.5	13.2
Unknown	37.5	25.0	12.5	25.0
Free/Reduced Lunch	20.9	30.3	34.0	14.8
Not Free/Reduced Lunch	10.3	22.4	40.3	26.9
Special Education	47.0	29.9	17.1	6.1
Not Special Education	9.0	24.5	41.4	25.2
English Language Learner	68.0	26.7	5.3	0.0
Not English Language Learner	13.0	25.2	38.8	23.1

Table N5. Performance Levels of Reading by Grade 6 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	18.0	24.3	36.7	20.9
Male	20.9	25.1	34.6	19.4
Female	14.9	23.4	39.0	22.7
Unknown	0.0	100.0	0.0	0.0
Amer. Indian	44.1	31.2	18.3	6.5
Asian	16.7	15.2	33.3	34.8
African American	21.1	34.7	30.5	13.7
Haw./Pac. Islander	33.3	22.2	22.2	22.2
Hispanic/Latino	28.9	29.0	31.7	10.4
White	14.8	23.0	38.7	23.5
Multiracial	13.5	28.4	36.9	21.3
Unknown	0.0	50.0	50.0	0.0
Free/Reduced Lunch	26.5	29.0	33.1	11.4
Not Free/Reduced Lunch	13.4	21.8	38.6	26.2
Special Education	58.3	25.8	12.2	3.7
Not Special Education	11.7	24.1	40.5	23.7
English Language Learner	71.7	20.8	7.5	0.0
Not English Language Learner	17.2	24.4	37.1	21.3

Table N6. Performance Levels of Reading by Grade 7 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	18.1	22.2	38.8	20.9
Male	21.1	22.8	37.4	18.8
Female	14.9	21.7	40.2	23.2
Unknown	0.0	0.0	0.0	100.0
Amer. Indian	47.3	30.2	19.1	3.4
Asian	16.7	16.7	37.2	29.5
African American	32.9	20.5	30.1	16.4
Haw./Pac. Islander	10.0	60.0	10.0	20.0
Hispanic/Latino	27.3	29.6	33.6	9.4
White	14.9	20.5	40.7	23.9
Multiracial	17.9	25.5	42.8	13.8
Unknown	33.3	16.7	16.7	33.3
Free/Reduced Lunch	26.9	28.2	33.6	11.3
Not Free/Reduced Lunch	13.4	19.1	41.5	26.1
Special Education	57.2	25.7	14.6	2.4
Not Special Education	12.6	21.7	42.2	23.6
English Language Learner	65.1	25.7	8.3	0.9
Not English Language Learner	17.3	22.2	39.2	21.3

Table N7. Performance Levels of Reading by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	17.6	29.0	45.4	8.0
Male	21.9	30.4	41.9	5.8
Female	13.1	27.5	49.1	10.4
Unknown	0.0	0.0	0.0	0.0
Amer. Indian	37.4	35.4	24.4	2.8
Asian	8.1	32.3	41.9	17.7
African American	40.9	16.7	39.4	3.0
Haw./Pac. Islander	20.0	60.0	20.0	0.0
Hispanic/Latino	25.8	33.5	36.8	3.9
White	15.2	28.0	47.9	8.9
Multiracial	14.6	29.9	48.2	7.3
Unknown	25.0	0.0	75.0	0.0
Free/Reduced Lunch	25.5	33.2	37.2	4.1
Not Free/Reduced Lunch	13.8	26.9	49.4	9.9
Special Education	61.2	27.8	10.7	0.4
Not Special Education	12.0	29.2	49.9	9.0
English Language Learner	63.5	25.5	10.9	0.0
Not English Language Learner	16.7	29.1	46.1	8.1

Mathematics

Table N8. Performance Levels of Mathematics by Grade 3 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	16.9	30.5	37.3	15.3
Male	16.4	28.2	37.5	17.9
Female	17.5	33.0	37.0	12.5
Unknown	0.0	0.0	0.0	100.0
Amer. Indian	40.6	38.6	17.4	3.4
Asian	10.0	21.7	36.7	31.7
African American	30.0	33.3	31.1	5.6
Haw./Pac. Islander	27.3	45.5	18.2	9.1
Hispanic/Latino	28.1	35.5	27.9	8.6
White	13.5	29.3	40.0	17.2
Multiracial	20.9	30.9	36.6	11.5
Unknown	44.4	33.3	22.2	0.0
Free/Reduced Lunch	22.8	35.0	32.3	9.9
Not Free/Reduced Lunch	13.4	27.8	40.3	18.6
Special Education	38.6	32.6	22.0	6.8
Not Special Education	13.5	30.2	39.7	16.7
English Language Learner	36.7	38.7	20.5	4.0
Not English Language Learner	16.0	30.1	38.0	15.8

Table N9. Performance Levels of Mathematics by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	11.0	33.9	39.6	15.5
Male	11.0	31.2	39.1	18.7
Female	11.1	36.9	40.0	12.0
Unknown	33.3	33.3	0.0	33.3
Amer. Indian	37.6	38.7	19.2	4.5
Asian	6.5	37.1	40.3	16.1
African American	18.1	49.4	26.5	6.0
Haw./Pac. Islander	25.0	37.5	25.0	12.5
Hispanic/Latino	16.7	43.7	30.5	9.0
White	8.3	31.6	42.7	17.4
Multiracial	16.8	35.0	34.3	13.9
Unknown	50.0	33.3	8.3	8.3
Free/Reduced Lunch	15.6	39.1	34.9	10.4
Not Free/Reduced Lunch	8.4	31.0	42.2	18.4
Special Education	32.5	40.6	21.4	5.5
Not Special Education	7.8	32.9	42.3	17.0
English Language Learner	35.1	54.8	9.0	1.1
Not English Language Learner	10.4	33.4	40.3	15.8

Table N10. Performance Levels of Mathematics by Grade 5 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	13.9	30.1	41.7	14.4
Male	14.9	29.4	41.0	14.8
Female	12.8	30.8	42.4	14.0
Unknown	75.0	0.0	25.0	0.0
Amer. Indian	38.7	35.6	22.5	3.2
Asian	8.6	20.7	39.7	31.0
African American	22.8	35.4	35.4	6.3
Haw./Pac. Islander	26.7	33.3	33.3	6.7
Hispanic/Latino	21.0	37.2	33.8	7.9
White	11.0	28.3	44.4	16.3
Multiracial	15.3	31.9	40.3	12.5
Unknown	50.0	12.5	37.5	0.0
Free/Reduced Lunch	19.2	35.0	36.8	9.0
Not Free/Reduced Lunch	10.9	27.3	44.4	17.4
Special Education	41.1	33.9	22.1	2.8
Not Special Education	9.7	29.5	44.7	16.2
English Language Learner	52.2	35.7	10.8	1.3
Not English Language Learner	13.1	29.9	42.3	14.7

Table N11. Performance Levels of Mathematics by Grade 6 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	13.4	37.1	39.8	9.8
Male	14.8	37.3	37.9	10.0
Female	11.8	36.8	41.9	9.5
Unknown	0.0	100.0	0.0	0.0
Amer. Indian	36.8	43.7	17.2	2.3
Asian	10.6	27.3	42.4	19.7
African American	13.7	47.4	34.7	4.2
Haw./Pac. Islander	33.3	44.4	11.1	11.1
Hispanic/Latino	20.7	42.3	32.8	4.2
White	11.0	35.6	42.3	11.1
Multiracial	10.6	38.3	41.1	9.9
Unknown	0.0	100.0	0.0	0.0
Free/Reduced Lunch	19.0	44.1	32.1	4.8
Not Free/Reduced Lunch	10.3	33.2	44.0	12.5
Special Education	44.4	39.1	14.6	1.9
Not Special Education	8.5	36.7	43.7	11.0
English Language Learner	51.4	33.9	13.8	0.9
Not English Language Learner	12.8	37.1	40.2	9.9

Table N12. Performance Levels of Mathematics by Grade 7 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	14.8	36.2	35.1	13.9
Male	16.1	34.9	35.1	13.9
Female	13.4	37.6	35.1	13.9
Unknown	0.0	0.0	0.0	100.0
Amer. Indian	36.3	45.8	14.1	3.8
Asian	12.5	20.0	36.3	31.3
African American	27.8	36.1	30.6	5.6
Haw./Pac. Islander	18.2	45.5	18.2	18.2
Hispanic/Latino	21.3	45.7	25.4	7.6
White	12.4	34.3	37.9	15.4
Multiracial	17.9	37.9	34.5	9.7
Unknown	0.0	33.3	16.7	50.0
Free/Reduced Lunch	20.3	44.4	27.7	7.6
Not Free/Reduced Lunch	11.9	31.8	39.0	17.2
Special Education	44.2	42.0	12.2	1.6
Not Special Education	10.7	35.4	38.3	15.6
English Language Learner	55.3	34.2	9.6	0.9
Not English Language Learner	14.1	36.2	35.5	14.1

Table N13. Performance Levels of Mathematics by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	13.1	39.4	34.7	12.8
Male	14.5	38.9	34.5	12.1
Female	11.6	39.9	35.0	13.6
Unknown	0.0	0.0	0.0	0.0
Amer. Indian	31.9	50.8	14.2	3.1
Asian	6.2	32.3	38.5	23.1
African American	28.8	39.4	27.3	4.5
Haw./Pac. Islander	50.0	40.0	10.0	0.0
Hispanic/Latino	18.0	47.0	27.7	7.3
White	11.2	37.5	37.0	14.3
Multiracial	11.7	43.1	36.5	8.8
Unknown	25.0	25.0	50.0	0.0
Free/Reduced Lunch	18.4	45.9	28.4	7.3
Not Free/Reduced Lunch	10.5	36.2	37.8	15.5
Special Education	44.0	47.2	7.8	0.9
Not Special Education	9.1	38.4	38.2	14.3
English Language Learner	39.3	50.7	9.3	0.7
Not English Language Learner	12.5	39.1	35.3	13.1

Science

Table N14. Performance Levels of Science by Grade 4 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	12.9	32.7	37.1	17.2
Male	13.2	32.0	36.8	18.0
Female	12.6	33.5	37.6	16.4
Unknown	0.0	33.3	33.3	33.3
Amer. Indian	43.8	39.2	15.2	1.8
Asian	8.1	40.3	38.7	12.9
African American	22.9	45.8	22.9	8.4
Haw./Pac. Islander	12.5	37.5	25.0	25.0
Hispanic/Latino	22.5	42.9	26.9	7.7
White	9.4	30.1	40.4	20.0
Multiracial	14.7	37.5	33.8	14.0
Unknown	10.0	40.0	40.0	10.0
Free/Reduced Lunch	18.3	38.9	33.6	9.2
Not Free/Reduced Lunch	9.8	29.2	39.2	21.8
Special Education	27.1	44.3	22.6	6.0
Not Special Education	10.7	31.0	39.3	18.9
English Language Learner	48.6	46.5	4.9	0.0
Not English Language Learner	12.0	32.4	38.0	17.7

Table N15. Performance Levels of Science by Grade 8 Demographic Subgroup

Group	Below Basic	Basic	Proficient	Advanced
Total	19.6	39.1	30.8	10.4
Male	21.0	36.8	31.2	11.0
Female	18.2	41.7	30.3	9.8
Unknown	0.0	0.0	0.0	0.0
Amer. Indian	47.0	35.6	14.2	3.2
Asian	20.0	36.9	26.2	16.9
African American	42.4	31.8	22.7	3.0
Haw./Pac. Islander	60.0	30.0	10.0	0.0
Hispanic/Latino	31.0	43.9	21.3	3.8
White	16.0	38.5	33.5	12.0
Multiracial	18.5	43.7	29.6	8.1
Unknown	25.0	25.0	50.0	0.0
Free/Reduced Lunch	28.1	43.4	23.4	5.1
Not Free/Reduced Lunch	15.5	37.0	34.4	13.1
Special Education	52.9	38.6	7.2	1.3
Not Special Education	15.3	39.2	33.8	11.6
English Language Learner	67.1	29.3	3.6	0.0
Not English Language Learner	18.6	39.3	31.4	10.7