WISCONSIN KNOWLEDGE AND CONCEPTS EXAMINATIONS

FALL 2007 WKCE TECHNICAL MANUAL

JUNE 2008

CTB/MCGRAW-HILL
20 RYAN RANCH ROAD
MONTEREY, CALIFORNIA 93940

Copyright

Developed and published under contract with the Wisconsin Department of Public Instruction by CTB/McGraw-Hill LLC, a subsidiary of The McGraw-Hill Companies, Inc., 20 Ryan Ranch Road, Monterey, California 93940-5703. Copyright © 2008 by the Wisconsin Department of Public Instruction. All rights reserved. Only State of Wisconsin educators and citizens may copy, download and/or print the document, located online at http://www.dpi.state.wi.us. Any other use or reproduction of this document, in whole or in part, requires written permission of the Wisconsin Department of Public Instruction.

Foreword

The technical information herein is intended for use by those who evaluate tests, interpret scores, or use test results in making educational decisions. It is assumed that the reader has technical knowledge of test construction and measurement procedures, as stated in *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999).

Copyright © 2008 by the Wisconsin Department of Public Instruction

Acknowledgments

The following CTB/McGraw-Hill and DPI staff members are primarily responsible for the content and statistical quality of this report:

CTB/McGraw-Hill:

Thakur Karkee,
Research Scientist

Greg Friedman,
Statistical Analyst

Kevin Fatica,
Research Associate

Margie Tully,
Publishing Director

Robin Hughes,
Scoring Project Manager

Leslie Miller,
Program Director

Wisconsin Department of Public Instruction – Office of Educational Accountability:

Lynette Russell,
Director

Viji Somasundaram,
Education Consultant

Duane Dorn,
Education Consultant

Alison Colby,
Education Consultant

Melissa Olson,
Education Consultant

Copyright © 2008 by the Wisconsin Department of Public Instruction

Acknowledgments, Cont'd

Jason Engle,
Education Consultant

Amy Marsman,
Education Consultant

Copyright © 2008 by the Wisconsin Department of Public Instruction

TABLE OF CONTENTS

Copyright ii

Foreword iii

Acknowledgments iv

Acknowledgments, Cont'd v

Part 1: Overview 1

Part 2: Test Design and Item Development 6

2.1 Content Framework and Assessment Limits 6
2.2 Test Blueprint 7
2.3 Reading Passage Selection 8
2.4 Item Development and Editing 8
2.5 Content/Bias Review and Item Alignment 10

Part 3: Test Form Development 12

3.1 Overall Test Book Development Process 12
3.1.1 WKCE Fall 2007 Form Selection 12
3.1.2 WKCE Field Test Item Selection 14
3.1.3 Quality Reviews 14
3.2 Description of the WKCE 2007 Tests 15
3.2.1 Reading 15
3.2.2 Mathematics 16
3.2.3 Language Arts 16
3.2.4 Social Studies 17
3.2.5 Science 17
3.3 Customer Approvals 17
3.3.1 Item Content and Bias Review 17
3.3.2 Item Selection Approval 17
3.3.3 Manuscript Approvals 18
3.3.4 Second Pages Approvals 18
3.3.5 Final Pages Sign-off 18

Part 4: Test Administration 19

4.1 Accommodations 19
4.2 Reporting Results of Assessments Taken with Accommodations 20
4.3 Test Security 20
4.4 Test Administration 22

Part 5: Scoring 24

5.1 Scoring of Multiple-Choice Items 24
5.1.1 Scanning and Calibration of Scanners 25
5.2 Scoring of Open-Ended or Constructed-Response Items 26
5.2.1 Description of Scoring Rubrics 26
5.2.2 Handscoring Process 27
5.2.3 Electronic Handscoring System 27
5.2.4 Anchor Papers and Training Papers 27
5.2.5 Scoring Personnel and Qualifications 28
5.2.6 Reader Training 28
5.2.7 Inter-Rater Reliability 29
5.3 Distribution of Constructed-Response (CR) Item Scores 30

Part 6: Characteristics of Sample Data 32

Copyright © 2008 by the Wisconsin Department of Public Instruction

6.1 Calibration Sample Data 32
6.1.1 Demographic Comparison of Sample and Population Data, by the Five NCLB Groupings 32

Part 7: Calibration, Equating, and Deriving Scale Scores 35

7.1 Calibration and Equating Methods 35
7.1.1 Calibration Models 36
7.1.2 Calibration Software 37
7.2 Deriving Scale Scores in the WKCE 37
7.3 Calibration Results 43
7.3.1 IRT Item Parameters 43
7.3.2 IRT Item Fit 44
7.3.3 Standard Error of Measurement 45
7.3.4 LOSS and HOSS 46
7.3.5 Test Characteristic Curves 46

Part 8: Test Results 47

8.1 Classical Item Analysis: Item Level Statistics 47
Reading 49
Mathematics 50
Language Arts 51
Social Studies 51
Science 51
Classical Item Analysis Summary 51
8.1.1 Speededness 52
8.2 Raw Score Results 52
Reading 54
Mathematics 54
Language Arts 55
Social Studies 55
Science 55
8.3 Summary Statistics for Scale Scores 56
Reading 56
Mathematics 57
Language Arts 57
Social Studies 58
Science 58
8.4 Cut Scores and Performance Level Classifications 58
Reading 59
Mathematics 60
Language Arts 60
Social Studies 61
Science 61
8.5 Standard Performance Indicator (SPI) for Content Standard 61
Reading 62
Mathematics 63
Language Arts 63
Social Studies 63
Science 63
Summary of Student Achievement Results 64

Part 9: Reliability 65

9.1 Measures of Internal Consistency and SEM 66
9.2 Classification Consistency and Accuracy 68
9.2.1 Kolen and Kim's Method for Pattern Scoring 69
9.3 Inter-rater Reliability for CR Items and Writing Prompts 72

Copyright © 2008 by the Wisconsin Department of Public Instruction

Reading 75
Mathematics 75
Writing 75

Part 10: Validity 77

10.1 Differential Item Functioning 79
10.2 Construct Validity 84
10.3 Erasure Analysis 86

Part 11: Summary and Recommendations 87

References 88

Tables and Figures 91

Appendix 1 361

Copyright © 2008 by the Wisconsin Department of Public Instruction

TABLE OF TABLES

PART 2

Table 2-1 Target Reading Test Blueprint: Grades 3–8, 10 91
Table 2-2 Target Mathematics Test Blueprint: Grades 3–8, 10 92
Table 2-3 Target Language Arts Test Blueprint: Grades 4, 8, 10 93
Table 2-4 Target Science Test Blueprint: Grades 4, 8, 10 93
Table 2-5 Target Social Studies Test Blueprint: Grades 4, 8, 10 93
Table 2-6 Actual Reading Test Blueprint: Grades 3–8, 10 94
Table 2-6 Cont'd Actual Reading Test Blueprint: Grades 3–8, 10 95
Table 2-7 Actual Mathematics Test Blueprint: Grades 3–8, 10 96
Table 2-8 Actual Language Arts Test Blueprint: Grades 4, 8, 10 97
Table 2-9 Actual Science Test Blueprint: Grades 4, 8, 10 97
Table 2-10 Actual Social Studies Test Blueprint: Grades 4, 8, 10 97
Table 2-11 Reading: 2007 Item Development Plan 98
Table 2-12 Mathematics: 2007 Item Development Plan 98
Table 2-14 Mathematics 2007 Item Development 99
Table 2-16 Mathematics: 2007 Item Development by Reporting Category and Item Format 100
Table 2-17 Item Development Each Year and Total to Date 101
Table 2-18 Reading: 2007 Item Selection Review Results 102
Table 2-19 Mathematics: 2007 Item Selection Review Results 106

PART 3

Table 3-1 Fall 2007 Test Configuration for Operational (OP) Items 111
Table 3-2 Reading: 2007 Embedded Field Test Items 112
Table 3-3 Mathematics: 2007 Embedded Field Test Items 113
Table 3-4 Unique Items Field Tested Each Year and Total to Date 114

PART 4

Table 4-1 Test Accommodations 115

PART 5

Table 5-1 Reading Rubric, Grades 3–8 and 10 126
Table 5-2 Mathematics Rubric, Grades 3–8 and 10 127
Table 5-3 Writing Rubric, Conventions of Written English, Grade 4 128
Table 5-4 Writing Rubric, Composing, Grade 4 129
Table 5-5 Writing Rubric, Conventions of Written English, Grade 8 130
Table 5-6 Writing Rubric, Composing, Grade 8 131
Table 5-7 Writing Rubric, Conventions of Written English, Grade 10 133
Table 5-8 Writing Rubric, Composing, Grade 10 134
Table 5-9 Score Distribution for Reading CR Items 136
Table 5-10 Score Distribution for Mathematics CR Items 138
Table 5-11 Score Distribution for Grades 4, 8, and 10 Writing: Composing Rubric 141
Table 5-12 Pct. Distribution of Scores, Grades 4, 8, and 10 Writing: Composing Rubric 141
Table 5-13 Score Distribution, Grades 4, 8, and 10 Writing: Convention Rubric 142
Table 5-14 Pct. Distribution of Scores for Grades 4, 8, and 10 Writing: Convention Rubric 142
Table 5-15 Score Distribution for Grades 4, 8, and 10 Writing: Total Score 143
Table 5-16 Pct. for Grades 4, 8, and 10 Writing: Total Score 143

Copyright © 2008 by the Wisconsin Department of Public Instruction

TABLE OF TABLES CONT'D

PART 6

Table 6-1 The Current 13 Calibration Districts 144
Table 6-2 Total Number of Students for Census and CD 13 145
Table 6-3 Number and Percent of Census and 13 CD Students, by Gender 146
Table 6-4 Number and Percent of Census and 13 CD Students, by Ethnicity 147
Table 6-5 Number and Percent of Census and 13 CD Students, by Socioeconomic Status 148
Table 6-6 Number and Percent of Census and 13 CD Students, by Disability 149
Table 6-7 Number and Percent of Census and 13 CD Students, by ELP 150

PART 7

Table 7-1 Item Flagged Based on Yen's Q_1 151
Table 7-2 Scoring Table for Reading Grade 3 152
Table 7-3 Scoring Table for Reading Grade 4 153
Table 7-4 Scoring Table for Reading Grade 5 154
Table 7-5 Scoring Table for Reading Grade 6 155
Table 7-6 Scoring Table for Reading Grade 7 156
Table 7-7 Scoring Table for Reading Grade 8 157
Table 7-8 Scoring Table for Reading Grade 10 158
Table 7-9 Scoring Table for Mathematics Grade 3 159
Table 7-10 Scoring Table for Mathematics Grade 4 160
Table 7-11 Scoring Table for Mathematics Grade 5 161
Table 7-12 Scoring Table for Mathematics Grade 6 162
Table 7-13 Scoring Table for Mathematics Grade 7 163
Table 7-14 Scoring Table for Mathematics Grade 8 164
Table 7-15 Scoring Table for Mathematics Grade 10 165
Table 7-16 Scoring Table for Language Arts Grade 4 166
Table 7-17 Scoring Table for Language Arts Grade 8 167
Table 7-18 Scoring Table for Language Arts Grade 10 168
Table 7-19 Scoring Table for Social Studies Grade 4 169
Table 7-20 Scoring Table for Social Studies Grade 8 170
Table 7-21 Scoring Table for Social Studies Grade 10 171
Table 7-22 Scoring Table for Science Grade 4 172
Table 7-23 Scoring Table for Science Grade 8 173
Table 7-24 Scoring Table for Science Grade 10 174
Table 7-25 The Number of Students and Percents at LOSS and HOSS 175

PART 8

Table 8-1 Item Analysis Grade 3 Reading 176
Table 8-2 Item Analysis Grade 4 Reading 180
Table 8-3 Item Analysis Grade 5 Reading 184
Table 8-4 Item Analysis Grade 6 Reading 188
Table 8-5 Item Analysis Grade 7 Reading 192
Table 8-6 Item Analysis Grade 8 Reading 196
Table 8-7 Item Analysis Grade 10 Reading 200
Table 8-8 Item Analysis Grade 3 Mathematics 203
Table 8-9 Item Analysis Grade 4 Mathematics 207
Table 8-10 Item Analysis Grade 5 Mathematics 211
Table 8-11 Item Analysis Grade 6 Mathematics 216
Table 8-12 Item Analysis Grade 7 Mathematics 220
Table 8-13 Item Analysis Grade 8 Mathematics 224

Copyright © 2008 by the Wisconsin Department of Public Instruction

TABLE OF TABLES CONT'D

PART 8 CONT'D

Table 8-14 Item Analysis Grade 10 Mathematics 228
Table 8-15 Item Analysis Grade 4 Language Arts 231
Table 8-16 Item Analysis Grade 8 Language Arts 233
Table 8-17 Item Analysis Grade 10 Language Arts 235
Table 8-18 Item Analysis Grade 4 Social Studies 237
Table 8-19 Item Analysis Grade 8 Social Studies 239
Table 8-20 Item Analysis Grade 10 Social Studies 241
Table 8-21 Item Analysis Grade 4 Science 243
Table 8-22 Item Analysis Grade 8 Science 245
Table 8-23 Item Analysis Grade 10 Science 247
Table 8-24 The Number of Items Flagged 249
Table 8-25 Raw Score Descriptive Statistics based on Census Data 250
Table 8-26 Raw Score Descriptive Statistics based on 13 Districts 251
Table 8-27 Raw Score Descriptive Statistics by Gender 252
Table 8-28 Raw Score Descriptive Statistics for Reading by Ethnicity 253
Table 8-29 Raw Score Descriptive Statistics for Mathematics by Ethnicity 254
Table 8-30 Raw Score Descriptive Statistics for Language Arts by Ethnicity 255
Table 8-31 Raw Score Descriptive Statistics for Social Studies by Ethnicity 255
Table 8-32 Raw Score Descriptive Statistics for Science by Ethnicity 256
Table 8-33 Raw Score Descriptive Statistics by Socioeconomic Status 257
Table 8-34 Raw Score Descriptive Statistics by Disability 258
Table 8-35 Raw Score Descriptive Statistics by English Language Proficiency 259
Table 8-36 Scale Score Mean and Standard Deviation for Census and 13 CD Results 260
Table 8-37 Scale Score Descriptive Statistics based on 13 Districts 261
Table 8-38 Scale Score Descriptive Statistics based on Census Data 262
Table 8-39 Scale Score Descriptive Statistics by Gender 263
Table 8-40 Scale Score Descriptive Statistics for Reading by Ethnicity 264
Table 8-41 Scale Score Descriptive Statistics for Mathematics by Ethnicity 265
Table 8-42 Scale Score Descriptive Statistics for Language Arts by Ethnicity 266
Table 8-43 Scale Score Descriptive Statistics for Social Studies by Ethnicity 267
Table 8-44 Scale Score Descriptive Statistics for Science by Ethnicity 267
Table 8-45 Scale Score Descriptive Statistics by Socioeconomic Status 268
Table 8-46 Scale Score Descriptive Statistics by Disability 269
Table 8-47 Scale Score Descriptive Statistics by English Language Proficiency 270
Table 8-48 Performance Level Cut Scores for all Contents 271
Table 8-49 Pct. of Students in Each Performance Level by Sub-Group (Reading) 272
Table 8-50 Pct. of Students in Each Performance Level by Sub-Group (Mathematics) 275
Table 8-51 Pct. of Students in Each Performance Level by Sub-Group (Language Arts) 278
Table 8-52 Pct. of Students in Each Performance Level by Sub-Group (Social Studies) 279
Table 8-53 Percentage of Students in Each Performance Level by Sub-Group (Science) 280
Table 8-54 Cut scores and Associated Impact Data for WKCE-CRT Reading 281
Table 8-55 Cut scores and Associated Impact Data for WKCE-CRT Mathematics 281
Table 8-56 Cut scores and Associated Impact Data for WKCE-CRT Language Arts 282
Table 8-57 Cut scores and Associated Impact Data for WKCE-CRT Social Studies 282
Table 8-58 Cut scores and Associated Impact Data for WKCE-CRT Science 283
Table 8-59 Summary Statistics for Reading Content Standards Raw and SPI Scores 284
Table 8-60 Summary Statistics for Mathematics Content Standards Raw and SPI Scores 286
Table 8-61 Summary Statistics for Language Arts Content Standards Raw and SPI Scores 288
Table 8-62 Summary Statistics for Social Studies Content Standards Raw and SPI Scores 289
Table 8-63 Summary Statistics for Science Content Standards Raw and SPI Scores 290

Copyright © 2008 by the Wisconsin Department of Public Instruction

TABLE OF TABLES CONT'D

PART 9

Table 9-1 Reliability for Total Group and Subgroups Using Cronbach's Alpha 291
Table 9-2 Standard Error of Measurement for Total Group and Subgroups 292
Table 9-3 Cronbach Alpha Reliability Coefficients for Content Standards 293
Table 9-4 Standard Error of Measurement per Content Standard 294
Table 9-5 Classif. Consistency and Classification Accuracy for Reading Grade 3 295
Table 9-6 Classif. Consistency and Classification Accuracy for Reading Grade 4 296
Table 9-7 Classif. Consistency and Classification Accuracy for Reading Grade 5 297
Table 9-8 Classif. Consistency and Classification Accuracy for Reading Grade 6 298
Table 9-9 Classif. Consistency and Classification Accuracy for Reading Grade 7 299
Table 9-10 Classif. Consistency and Classification Accuracy for Reading Grade 8 300
Table 9-11 Classif. Consistency and Classification Accuracy for Reading Grade 10 301
Table 9-12 Classif. Consistency and Classification Accuracy for Mathematics Grade 3 302
Table 9-13 Classif. Consistency and Classification Accuracy for Mathematics Grade 4 303
Table 9-14 Classif. Consistency and Classification Accuracy for Mathematics Grade 5 304
Table 9-15 Classif. Consistency and Classification Accuracy for Mathematics Grade 6 305
Table 9-16 Classif. Consistency and Classification Accuracy for Mathematics Grade 7 306
Table 9-17 Classif. Consistency and Classification Accuracy for Mathematics Grade 8 307
Table 9-18 Classif. Consistency and Classification Accuracy for Mathematics Grade 10 308
Table 9-19 Classif. Consistency and Classification Accuracy for Language Arts Grade 4 309
Table 9-20 Classif. Consistency and Classification Accuracy for Language Arts Grade 8 310
Table 9-21 Classif. Consistency and Classification Accuracy for Language Arts Grade 10 311
Table 9-22 Classif. Consistency and Classification Accuracy for Social Studies Grade 4 312
Table 9-23 Classif. Consistency and Classification Accuracy for Social Studies Grade 8 313
Table 9-24 Classif. Consistency and Classification Accuracy for Social Studies Grade 10 314
Table 9-25 Classif. Consistency and Classification Accuracy for Science Grade 4 315
Table 9-26 Classif. Consistency and Classification Accuracy for Science Grade 8 316
Table 9-27 Classif. Consistency and Classification Accuracy for Science Grade 10 317
Table 9-28 Inter-Rater Reliability Reading 318
Table 9-29 Inter-Rater Reliability Mathematics 321
Table 9-30 Inter-Rater Reliability, Writing Prompts 329

PART 10

Table 10-1 Items Flagged for DIF, By Gender 330
Table 10-2 Items Flagged for DIF, By Ethnicity, African American 331
Table 10-3 Items Flagged for DIF, By Ethnicity, Hispanic 332
Table 10-4 Items Flagged for DIF, By Ethnicity, Asian 333
Table 10-5 Items Flagged for DIF, By Ethnicity, American Indian 336
Table 10-6 Items Flagged for DIF, By English Language Proficiency 337
Table 10-7 Items Flagged for DIF, By Socioeconomic Status 339
Table 10-8 Items Flagged for DIF, By Disability Status 340
Table 10-9 Correlations among Reading Objectives 341
Table 10-10 Correlations among Mathematics Objectives 342
Table 10-11 Correlations among Language Arts Objectives 343
Table 10-12 Correlations among Social Studies Objectives 343
Table 10-13 Correlations among Science Objectives 344
Table 10-14 Factor Analysis 345

Copyright © 2008 by the Wisconsin Department of Public Instruction

TABLE OF FIGURES

PART 7

Figure 7-1 SEM Curves, Reading Grades 3-6 346
Figure 7-2 SEM Curves, Mathematics Grades 3-6 348
Figure 7-3 SEM Curves, Language Arts Grades 4, 8, 10 350
Figure 7-4 SEM Curves, Social Studies Grades 4, 8, 10 351
Figure 7-5 SEM Curves, Science Grades 4, 8, 10 352
Figure 7-6 TCC Curve for Reading Grades 3-8, 10 353
Figure 7-7 TCC Curve for Mathematics Grades 3-8, 10 354
Figure 7-8 TCC Curve for Language Arts Grades 4, 8, 10 355
Figure 7-9 TCC Curve for Social Studies Grades 4, 8, 10 356
Figure 7-10 TCC Curve for Science Grades 4, 8, 10 357

PART 9

Figure 9-1 Reading Indices for Classification Consistency and Classification Accuracy 358
Figure 9-2 Mathematics Indices for Classification Consistency and Classification Accuracy 358
Figure 9-3 Language Arts Indices for Classification Consistency and Classification Accuracy 359
Figure 9-4 Social Studies Indices for Classification Consistency and Classification Accuracy 359
Figure 9-5 Science Indices for Classification Consistency and Classification Accuracy 360

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 1: Overview

The Fall 2007 WKCE Technical Report documents the processes and procedures applied in the Fall 2007 WKCE, and the results. This report also shows how the processes, procedures, and results of the Fall 2007 administration relate to the issues of validity and reliability, and to the *Standards for Educational and Psychological Testing* (American Education Research Association, American Psychological Association National Council on Measurement in Education, 1999). This report demonstrates that the Fall 2007 WKCE adhered to the appropriate standards and practices of educational assessment. Ultimately, this report serves to document evidence that valid inferences about Wisconsin student performance can be derived from the Fall 2007 WKCE.

The *Improving America's Schools Act* of 1994 requires that states establish challenging academic standards as well as aligned annual assessments. The *Goals 2000: Educate America Act and the Elementary and Secondary Education Act* spell out additional requirements to ensure that citizens receive coherent information about whether and to what degree students are meeting rigorous academic standards. This Technical Report is an important part of meeting those requirements.

Wisconsin students in grades 4, 8, and 10 began taking the Wisconsin Knowledge and Concepts (WKCE) norm-referenced assessments in the 1997 school year. The assessments used at that time were TerraNova tests developed by CTB/McGraw-Hill. The selection of those tests was partly predicated on an awareness of standards being developed. In January 1998, the *Wisconsin Model Academic Standard*s were adopted. These new standards were the work of the Governor's *Commission on Model Academic Standards*, chaired by then-current Lieutenant Governor McCallum and the Wisconsin Department of Public Instruction (DPI). The Model Academic Standards would measure student performance in the same subjects as the TerraNova tests.

Beginning in the 2005–2006 school year, the federal *No Child Left Behind Act* (NCLB) required all states to test all students in Reading and Mathematics in grades 3 through 8 and once in high school (in grade 10 under Wisconsin law s. 118.30). Based on the NCLB legislation, student performance, reported in terms of proficiency categories, is used to determine the adequate yearly progress of students at the school, district, and state levels.

Specifically, under *No Child Left Behind*, each state has been required, beginning with the 2005-06 school year, to measure every child's progress in Reading and Mathematics in each of grades 3 through 8 and at least once during grades 10 through 12. Beginning with school year 2007-2008, states were also required administer Science assessments at least once during grades 3-5, grades 6-9, and grades 10-12. Students in grades 4, 8 and 10 are, and will continue to be assessed in Language Arts, Science and Social Studies as required by state law (s. 118.30 Wisconsin Statutes).

It is within this policy context that the WKCE was constructed, as a criterion-referenced test, for the Fall 2005 administration, replacing the previously existing norm-referenced WKCE Reading and Mathematics tests. The new criterion-referenced WKCE was designed specifically

Copyright © 2008 by the Wisconsin Department of Public Instruction

for Wisconsin students, and specifically to measure their performance on the Wisconsin Model Academic Standards adopted by the state. The new criterion-referenced WKCE assessments were designed to evaluate students' knowledge and to measure achievement in the basic skills taught in schools at grades 3-8 and 10. The Fall 2007 WKCE is the third administration of this new criterion-referenced assessment.

The WKCE tests consist of criterion-referenced items written by Wisconsin teachers and items from CTB/McGraw-Hill's norm-referenced test, TerraNova, The Second Edition® (TerraNova, CTB/McGraw-Hill, 2001). The WKCE tests include Reading and Mathematics at grades 3-8 and 10 and Science, Social Studies and Language Arts (including Writing) at grades 4, 8, and 10.

Based on the input of Wisconsin educators and the Wisconsin Model Academic Standards, a design was derived, developed, administered, and scored. The present Technical Report documents all aspects of the testing cycle in the subsequent chapters. The structure of the present Technical Report mirrors the testing cycle. A brief content summary of the report is provided below.

Design

- Part 2 of this report describes test design, item development process, and some aspects of the content-related validity of the WKCE tests.
- More specifically Part 2 describes how CTB, DPI, and Wisconsin educators collaborated through a series of test development processes to ensure that the appropriate content was included in the WKCE, and to ensure that the test items adequately sampled the domain of content knowledge necessary to make legitimate inferences about student performance.
- Wisconsin Model Academic Standards were translated into grade-level content frameworks, which in turn, formed the basis for test blueprints and item specifications.
- Wisconsin educators were involved in design at every step to ensure the appropriateness of the test to the standards.
- Test design started in August 2003 with the convention of approximately 35 educators per content area for grades 3–8 and 10 to establish the grade-level content frameworks based on the Wisconsin Model Academic Standards, establish assessment limits, create the test blueprint, and to review reading passage and page specifications. The test specifications documents created and later approved by DPI serve as a foundation for item and test development through 2008-2009 and will continue through 2009-2010.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Item and Form Development

- Part 3 discusses key development tasks and issues related to creating the Fall 2007 test forms.
- Item development was based on the approved test blueprints, with a sufficient quantity of items written across years to develop multiple operational test forms.
- Part 3 discusses the process of selecting operational test items, the content distribution of field test items, and the process of obtaining DPI approvals.
- A total of 4,482 items (4,042 multiple-choice items and 440 constructed-response items) were field-tested for the WKCE by the end of the Fall 2007 administration.
- Selection of the Fall 2007 operational forms was done using the ITEMWIN software similar to previous administrations for all grades and content areas.

Administration

- Part 4 briefly describes test administration and accommodations.
- The test administration window was October 22 - November 27, 2007.
- Delivery of materials was handled through the district and school assessment coordinators.
- Two test books were used for the grades 4, 8 and 10. In past administrations, only one book was used per grade. The change here related to manufacturing recommendations for the maximum number of pages per book. Two books also coincided nicely with the NCLB content areas.

Scoring

- Part 5 documents how the multiple-choice and constructed-response items were scanned and scored.
- The machine scanning process and the hand scoring process, including the development and review of the scoring rubrics, anchor (sample) papers, and writing prompts, as well as the training of scoring personnel, ongoing quality assurance, the application of an inter-rater reliability assessment, and a systematic review of the resulting score distributions, all supported the development of reliable scores.
- The scoring rubrics used in handscoring are presented in detail for all content areas with handscored items.

Characteristics of Sample Data

- As detailed in Part 6, scoring and all norms for the Fall 2007 WKCE tests were based on a sample of student response data. The school districts used to obtain the

Copyright © 2008 by the Wisconsin Department of Public Instruction

calibration sample are identified in the report. These districts were successfully used for the same purposes in previous administrations.

- Part 6 describes demographic characteristics of the calibration sample data and provides a comparison of sample data and the census data. The demographic comparison references each of the five NCLB groupings (gender, ethnicity, socioeconomic status, disability status, and English Language Proficiency). The sample data was demonstrated to be a sufficiently representative sample of the state student population.
- Mean scale scores and standard deviations were also derived from the calibration district data, compared to the results from the entire student population, and this comparison further demonstrated the validity of using the sample data.

Calibration and Scaling

- Part 7 reviews calibration, equating, scoring methods, and calibration results. Evaluation of the calibration results includes model-to-data fit and the standard error of measurement.
- Part 7 also explains how a student's scale score is derived from the raw score. Examples of a very low-performing student, a very high-performing student, and several students with a 50% correct raw score are provided. Several students with the same 50% correct raw score are provided in order to illustrate how students with the same raw score can have different scale scores.
- The Fall 2007 WKCE was calibrated and scaled using two different item response theory (IRT) models, one for constructed-response items and one for multiple-choice items, which are the basis of most large-scale standardized testing.
- Calibration and scaling results and scoring tables which include standard error of measurement are also presented.
- Item-pattern scoring was applied to the Fall 2007 WKCE. As discussed in Part 7, item-pattern scoring is generally recommended over number-correct scoring because it produces more accurate scores for individual students.

Test Results

- Part 8 summarizes item analysis, raw scores, scale scores, performance levels, and a standard performance indicator score (SPI) for content standards.
- Evidence in support of reliability of the test is established through an item analysis which includes Cronbach's alpha, standard error of measurement (SEM), and the omit rate.
- Summary descriptive statistics for all scores (raw scores, scale score, SPI scores and performance levels) are reported for all students and the five NCLB groupings.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Reliability

- Part 9 elaborates on the reliability of the test based on results presented in previous parts of the report.
- Standard error of measurement (SEM) was assessed for raw score and scale score.
- Inter-rater reliability was estimated for all of the constructed-response items.
- Internal consistency was assessed for all multiple-choice and constructed-response items using Chronbach's alpha.
- Classification consistency and accuracy were estimated for performance classification.

Validity

- Part 10 reviews the main validity issues discussed in all prior chapters and provides additional validity evidence supporting of the WKCE tests.
- Factor analysis and correlations among content standards are presented as construct validity.
- An analysis of differential item functioning (DIF) is presented.
- Erasure analysis, a procedure used to identify high erasure rates, is also discussed.

Summary Recommendations

- Key findings of the 2007 administration are presented in the body of the report. However, some items of a more technical nature, which stand out as key recommendations and summary statements that should be considered in subsequent administrations are presented in Part 11.
- Recommendations based on the Fall 2007 administration cover three different phases of the testing cycle: item development, scoring, and psychometric, or measurement-based, research and evaluation.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 2: Test Design and Item Development

Part 2 of the Technical Report describes how CTB, DPI, and Wisconsin educators collaborated through a series of test development processes to ensure that the appropriate content was included in the WKCE, and to ensure that the test items adequately sampled the domain of content knowledge necessary to make legitimate inferences about student performance. Part 2 documents the test development process for the Fall 2007 test administration and the development of new items to be field tested in fall 2008.

As described below, the Wisconsin Model Academic Standards were central to the entire test design process. Part 2 of the Technical Report demonstrates the adherence of the WKCE program to the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999), and specifically to standards 1.2, 1.6, 3.1, 3.2, 3.3, 3.11, 6.4, 6.15, 13.3, and 13.5.

The test development activities during 2007 adhered to the test specifications documents developed in previous years. The 2006 Technical Report (Parts 2, 3, and 4) provides a detailed account of the development of the test specifications documents during previous years. The assessment frameworks, test design, test blueprints, reading passage specifications, item specifications, art specifications, and style guide were all developed in 2003, the first year of the WKCE program. The role of Wisconsin educators was an essential component of the development of the WKCE, because of their professional expertise and judgment when providing content-related validity evidence in test development.

During the first year of the contract, August 2003 to August 2004, the test specifications documents were developed through an extended, collaborative process with DPI and based on the contributions of Wisconsin educators during meetings conducted in 2003 (see the Fall 2006 Technical Report, p. 6). Test specifications include the test blueprint, passage specifications, item specifications, page specifications, and style guide.

According to the most recent edition of the *Standards* (AERA, APA, NCME, 1999), "Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of test scores" (p. 9). Much of the content-related validity evidence is produced during the test development process. The content-related evidence supports inferences from a sample of observations (the test) to a domain of observations (the content area). A substantial source of content-related validity evidence is the expert judgment that the test items are an adequate and representative sample of the domain being measured. Content-related validity evidence can support interpretations of test scores in terms of performance over a performance domain. If the content domain is specified clearly, and a representative sample of performance tasks is drawn from the domain, then inferences about expected performance over the domain based on observed performances should be legitimate.

2.1 Content Framework and Assessment Limits

The Assessment Framework documents created by DPI provide information about the content measured at each grade level and explain the relationships among the Model Academic

Copyright © 2008 by the Wisconsin Department of Public Instruction

Standards, the Assessment Framework, and classroom instruction. The Framework documents are located on DPI's website at http://dpi.wi.gov/oea/wkce.html. The Fall 2006 Technical Report, section 3.1.1, explains the structure and development of the Assessment Frameworks.

The Assessment Frameworks specify the broad categories within the content area at which test sub-scores may be reported, for example, “Number Operations and Relations,” or “Measurement” for Mathematics, and “Understands Text” and “Analyzes Text” for Reading. These broad categories are further delineated into subskills. For example, Number Operations and Relations is further sub-divided into “Reading, Writing, and Representing numbers,” and “Ordering and Comparing Numbers,” and so forth. Assessment limits are bulleted statements which identify the specific content that is eligible for testing for each subskill and may clarify how the content could be assessed. For example, in Mathematics, the size of numbers or the types of plane and solid geometric figures that are appropriate at each grade level would be specified in the assessment limits. For Reading, the assessment limits clarify which prefixes or suffixes or which literary devices are appropriate to assess at each grade level. For the grade 4, 8, and 10 Science assessments, the Model Academic Standards served as the foundation for the creation of the Science Assessment Frameworks. The Model Academic Standards for Language Arts and Social Studies provide the content framework for these content area tests at grades 4, 8, and 10.

The Assessment Frameworks and Item Specifications documents were used by the professional test item writers and by CTB assessment editors when reviewing, editing, and preparing the items for the January 2008 Item Selection Review meeting. The Wisconsin educators who attended the Item Selection review meeting used the assessment frameworks to identify the objective and subskill measured by each item, confirming that test items are aligned to the Wisconsin Model Academic Standards.

2.2 Test Blueprint

The test blueprints specify the number of multiple-choice (MC) and constructed-response (CR) items for each reporting category and subskill. The process used for developing the blueprints was described in detail in the Fall 2006 Technical Report (Parts 2 and 3). Tables 2-1 through 2-5 present the target blueprints for the Fall 2007 test. Tables 2-6 through 2-10 present the actual test blueprints showing how the items selected for the Fall 2007 forms were distributed by reporting category and subskill for each item type.

In 2007, some changes were made to the blueprints for Mathematics, Science, and Language Arts grade 8. In addition, Depth of Knowledge (DOK) requirements were incorporated into the Reading and Mathematics blueprints to indicate the number of items needed at each DOK level for each reporting category.

The Mathematics blueprints were modified to reflect the inclusion of a two-point constructed-response item and the subsequent reduction of the three-point constructed-response items from four to three. In addition, the number of multiple-choice items for each reporting category was adjusted to reflect the use of multiple-choice items for reporting category A.

Copyright © 2008 by the Wisconsin Department of Public Instruction

The Science blueprints were modified slightly to show a shift in emphasis among reporting categories A and B and among reporting categories G and H.

The Language Arts grade 8 blueprint changes involved shifting two multiple-choice items from reporting category D to reporting category B. This change was made in response to Wisconsin educators' concerns expressed at the 2005 content review that the language test should not require excessive reading. When selecting test forms for 2005 and 2006, CTB made the effort to minimize the number of item sets that use a common stimulus, such as a brief essay or letter. However, when selecting the 2008 form, the use of two lengthy stimuli would have been necessary in order to meet the blueprint. CTB brought this concern to the attention of DPI and suggested that two items be shifted from category D to category B. DPI approved this change to the blueprint on March 9, 2007.

2.3 Reading Passage Selection

Reading passages on the 2007 operational* forms were selected, reviewed, and approved between 2003–2005. The process of selecting, reviewing, and approving reading passages was detailed in the Fall 2006 Technical Report (see section 3.1.3).

For the field test items embedded in the Fall 2007 forms, item development for Reading focused on developing additional items for existing and previously used reading passages; therefore, there were no new reading passages selected and approved.

Item development during 2007 focused on developing additional items for existing reading passages as well as for some new reading passages at each grade level. CTB submitted ten commissioned and 13 permissioned passages to DPI for consideration. DPI conducted a passage review meeting and informed CTB of which ten passages were approved for item development. Items were also developed for 24 passages that were previously used to augment existing items for these passages. The summary report for the January 2008 Item Selection Review Meeting provides additional information about the passage selection. This document is available at http://www.dpi.state.wi.us./

2.4 Item Development and Editing

This section describes the process and results of developing test items during 2007 for field testing in Fall 2008. The development of items included as embedded field test items on the Fall 2007 forms is described in the 2006 Technical Report, section 4.2.

In August 2007, CTB editors developed plans for new item development for items to be field tested on the Fall 2008 forms. In addition to developing new items to meet DOK needs, other goals for item development included creating additional items for certain subskills to

* Operational items are those items that contribute to student scores. Operational items are contrasted with field-test items, which do not contribute to student scores. Operational items are abbreviated in this report as OP, and field-test items are abbreviated as FT.

Copyright © 2008 by the Wisconsin Department of Public Instruction

increase the item pool, providing flexibility in meeting the blueprint, and increasing overall flexibility in selecting items for forms.

An alignment study was conducted by Dr. Norman Webb of the University of Wisconsin-Madison using the Fall 2005 test forms. The results of the alignment study were reviewed by DPI's Technical Advisory Committee (TAC) in their December 2005 meeting. It was determined that the DOK level of the items should be reviewed for Reading, Math and Science. The TAC recommended, based on the alignment study, test forms should have 65% or more items at or above the DOK level of the objective to have a fully aligned test. As a result of the item alignment study, item development in 2006 and 2007 focused on creating items at higher levels of DOK for targeted objectives and subskills. The 2006 Technical Report, section 4.2.3, describes the implications of the alignment study on item development in greater detail.

The Mathematics and Reading item development plans for grades 3-8 were determined by content editors and supervisors who analyzed the Mathematics and Reading item banks. Careful evaluation of the banks was conducted to determine if a sufficient number of items existed to meet test development criteria for each grade and content area. Criteria included meeting blueprints in terms of content diversity and the DOK requirement for each objective. Content editors also evaluated item banks relative to Reading and Mathematics blueprint changes requested by DPI in 2007 (see section 2.2).

The Reading item development plan for grades 3–8 was based on creating additional items for existing passages in order to have more items measuring higher levels of DOK. Additional items for existing passages were warranted in order to expand the item set for the passage, and to allow greater flexibility when selecting operational forms.

The plan for Mathematics focused primarily on developing DOK level 3 items for statistics and probability, and on creating multiple-choice items that measure objective A, "Mathematical Processes." This was done in response to DPI's request to include multiple-choice items for this standard on the test blueprint. Additional items for remaining objectives were developed to broaden content diversity and flexibility of the item bank.

The development plans were presented to DPI in August 2007. Tables 2-11 and 2-12 present the Reading and Mathematics item development plans for the items to be field tested in Fall 2008; these plans represent the minimum number of items to be developed in 2007. Tables 2-13 and 2-14 show the number of items CTB developed prior to the Item Selection Review meeting (January 9–11, 2008), the number of items written by the committees during the review meeting, and the total number of items reviewed. Tables 2-15 and 2-16 show the number of items developed for Reading and Mathematics by grade level, reporting category, and item format. The number of items developed exceeded the number of items proposed in the plans. Increased development was a result of continued evaluation of the item banks by the content development team. Table 2-17 shows how many multiple-choice, constructed-response and total items have been written to date.

A staff of professional item writers, many of them experienced teachers, wrote the WKCE test items developed in 2007. Item writers adhered to the item specifications as they

Copyright © 2008 by the Wisconsin Department of Public Instruction

drafted and revised items. CTB assessment editors also used the item specifications during editorial reviews and revisions of the items. The item specifications provide detailed information regarding the following:

- item type
- content strand, standard, objective, subskills to be measured
- clarification statement of the task students will perform when answering each item type
- assessment limits
- stimulus attributes (stems, graphics, narratives)
- response attributes (general, correct response, acceptable distractors, unacceptable distractors)
- scoring rubric attributes (general or item/task specific)
- sample items

Throughout the item development and review process, the alignment between the item and the content standard/subskill/assessment limit was checked during each editing phase. All test items were carefully reviewed for content and style by test development specialists, Wisconsin educators, and the content specialists from DPI. All test items developed in 2007 were reviewed internally by CTB supervisors familiar with the Wisconsin content frameworks and item specifications. During all item reviews, careful attention was paid to verifying that each item measured the intended objective, subskill, and assessment limit. If there any misalignment was found, the item was either rejected, edited to achieve greater alignment, or a different subskill or assessment limit was assigned.

2.5 Content/Bias Review and Item Alignment

Following the internal editorial reviews, the new items written in 2007 were reviewed by committees of Wisconsin educators at the Item Selection Review meeting held January 9–11, 2008. The list of educators participating in the review committees was included in the January 2008 Item Selection Review Meeting Summary Report (http://www.dpi.state.wi.us./).

A primary purpose and emphasis of the item review meeting was to verify the alignment of each item to an objective, subskill, and assessment limit of the Wisconsin Assessment Framework and to a DOK level. CTB developed the items to target specific objectives, subskills, and DOK levels, and documented the alignment of the items to the Framework. However, in order to simulate an external content alignment study, as per the advice of the TAC, participants were asked to identify the objective and subskill to which each item best aligned. Therefore, the content and DOK alignment information was not included on the hardcopy item cards in the review books, on the review forms, nor on the item templates projected on screen. The committee participants' independent alignment of the items to the content objective, subskill, and DOK level were important contributions to evidence that the items are valid measures of the assessment frameworks.

The committee participants were provided with a handout describing each of the four levels of the Depth of Knowledge framework, which is based on the framework developed by

Copyright © 2008 by the Wisconsin Department of Public Instruction

Norman Webb of the University of Wisconsin-Madison. The presentation on the first day provided an explanation of the concept and a few sample items showing the distinction between depth of knowledge and item difficulty.

As participants reviewed each item prior to discussion, participants individually identified the objective, subskill, and the DOK level to which they thought the item aligned. Participants recorded the content objective, and subskill, and DOK level on their review form and then, in turn, verbally reported the information to the DPI recorder. In this manner, DPI collected data regarding the consensus alignment. As discussion or editing ensued for each item, participants could revise their judgment regarding the item's content and DOK alignment. Committee discussions and editing resulted in the writing of a few Reading and Mathematics items for each grade during the Item Selection Review meeting.

Tables 2-18 and 2-19 present the results of the Item Selection Review meeting for Reading and Mathematics by grade level, reporting category, subskill, item format, depth of knowledge, and the recommendations of the committees.

The degree of concurrence between CTB editors and committee participants in DOK assignment was variable, as it was at the 2007 content review meeting. Agreement ranged from 43% to 67%, with an overall agreement of 55% for Reading and 63% for Mathematics. The overall concurrence for Reading and Mathematics was 59%. In addition to calculating the percent concurrence, a Pearson correlation coefficient was calculated to determine the relationship between the CTB-assigned DOK level and the committee's DOK assignment. The correlation coefficient for Reading was 0.58, and the coefficient for Mathematics was 0.53. The overall correlation coefficient was 0.52.

Agreement between CTB and committee DOK ratings is highest for DOK levels 1 and 3. Based on the evaluation survey, participants reported having a slightly easier time assigning a DOK level to items than assigning a content objective/subskill. Additional information about the Item Selection Review meeting, including detailed results of the item alignment, and review, and the evaluation survey, is available in the January 9-11, 2008 Item Selection Review Summary Report (http://www.dpi.state.wi.us./).

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 3: Test Form Development

Part 3 of the Technical Report focuses on key development tasks and issues related to creating the Fall 2007 operational test forms and the items that appeared as embedded field test items in the Fall 2007 test books. The test specifications and item development activities described in Part 2 explain how specific development processes provided evidence to support test validity, primarily through the use of expert professional judgment from Wisconsin educators and from CTB test development specialists. The foundation test specifications documents—assessment framework, assessment limits, passage specifications, item specifications, test blueprints, art and page specifications, and style guide—developed and approved during the initial phases of the project continued to serve as critical guides for the ongoing development and field testing of items. These documents contribute to ensuring that each form of the test accurately measures the content in consistent and stable ways, thus providing evidence of the test's validity. Information is provided in Part 3 relating to the following topics:

- a general discussion of CTB's test book creation and editing process;
- the process of selecting operational test items;
- the content distribution of field test items; and
- the process of obtaining DPI approvals.

A comprehensive, multi-segmented development process guides the development of assessment materials. The following section outlines this process in general terms. The remainder of Part 3 provides details of how these processes were implemented in Wisconsin. This section of the Technical Report addresses the following AERA/APA/NCME standards: 1.6, 3.1, 3.5, 3.6, 3.7, 3.9, 3.11, 3.16, 6.4, 6.15, 7.3, 7.4, 7.7, 13.3, and 13.5.

3.1 Overall Test Book Development Process

The creation of test book materials involved the expertise of multiple CTB departments, DPI, and Wisconsin educators. The activities that contributed to the creation of the test book materials are described below.

3.1.1 WKCE Fall 2007 Form Selection

The WKCE operational test forms for all content areas and grade levels use a common item linking design in order to equate the forms from year to year. The minimum number of linking items per content area is as follows:

- Reading: 16 items
- Mathematics: 19 items
- Language Arts: 13 items
- Social Studies: 15 items
- Science: 15 items

Copyright © 2008 by the Wisconsin Department of Public Instruction

CTB assessment editors selected items for the 2007 operational forms using multiple selection criteria, including:

- blueprint
- diversity of content represented by individual items
- classical item statistics (avoiding multiple-choice items with *p*-values below 0.30, constructed-response items with *p*-values below 0.20; or items with positive point-biserial on distractors)
- IRT statistics (range of scale score values, maximizing information)
- bias flags (avoiding items with C flag)
- poor fit flags

CTB assessment editors used CTB's proprietary software ITEMWIN to select items for the Fall 2007 operational test forms for all content areas and grade levels. The ITEMWIN software (Burket, 2000) allows the content editor to make informed decisions regarding item selection. This software monitors the impact of each decision made during the item selection process and offers a variety of options for grouping, classifying, sorting, and ranking items to highlight key information as it is needed.

The ITEMWIN program has three parts. The first part is used to select a working item pool of manageable size from the larger tryout pool; items clearly inappropriate to the target grade range are eliminated. There is information about each item in the pool, including the item format to which the item is assigned, a descriptive phrase about the item, the association of the item with a stimulus, the item parameters, a fit rating indicating how well the item fits the expectations based on the IRT model used, and a "bias" rating indicating whether item performance reflects extraneous or construct-irrelevant information making the item unfairly difficult for one of the NCLB subgroups. This "bias" concept is discussed further in Part 10 in the context of differential item functioning, or "DIF."

The second part of the ITEMWIN program uses the working item pool created in the first step to perform the actual test selection. Typically, the developer begins by specifying the number of items to be included in the test and a target number of items for each item format. The program can then be prompted to automatically select a test that represents the best possible statistical combination of items. These automatic selections can then be used as a reference set to which other selections are compared. Successive selections are plotted on a graphic display that shows the test characteristic curve for each set of selected items.

In the third part of the program, a table shows both expected number correct and standard error of measurement as functions of scale score, as well as statistical and graphical summaries of bias, fit, and the average standard error of the test as selected. Any fault in the selection, whether the test is too easy or too difficult for the target grade, contains biased items, or does not adequately cover part of the range, becomes apparent as the final statistics are generated. Content editors and CTB Research Department staff examined these statistics for each of the WKCE selections to confirm that each form had an appropriate scale score range and when the test characteristic curves for all grades were compared side-by-side, that there was an appropriate

Copyright © 2008 by the Wisconsin Department of Public Instruction

progression in difficulty. In addition, CTB content editors reviewed each selection for content diversity to ensure no two items were similar in content.

CTB assessment editors prepared a document for each selected form that summarized the test and item characteristics, submitted their selections to the content supervisor for review, and then to the publishing project manager. Appendix 1 shows the Form Selection Summary Document. The supervisor and manager requested changes to the selections, as necessary, in order to improve the test characteristic curve (TCC) or standard error (SE) curve. Form selections were then submitted to the CTB Research Scientist for review. Additional revisions may have been requested at this stage. For the Reading and Mathematics selections, it was important to ensure the test characteristic curves for all grade levels formed a progression. CTB Research Scientists reviewed the form selections to ensure the test characteristic curves for the 2007 forms were as similar as possible to the 2006 forms, and that curves for the linking items were as similar as possible to curves for the complete forms.

Upon approval of the selections by the Research Department, the CTB editor submitted the selections to DPI for review. For some selections, DPI requested revisions for content, difficulty, or statistical reasons. Upon making the requested changes and submitting revised selection summary forms, all operational forms were approved by DPI. Table 3-1 shows the structure of operational test forms in the Fall 2007 WKCE.

3.1.2 WKCE Field Test Item Selection

In addition to the operational items, a set of new field test items were included in the Fall 2007 test books for Mathematics and Reading. Tables 3-2 and 3-3 present the number of field test items embedded on the Fall 2007 forms by grade level, reporting category, and item format for Reading and Mathematics respectively. The items embedded on the Fall 2007 forms were developed in 2006 and reviewed in January 2007. The 2006 Technical Report (Parts 2, 3, and 4) documents the process of developing these items and the procedures used at the 2007 Item Selection Review meeting and the results of that meeting.

Mathematics and Reading items approved at the January 2007 Item Review were all candidates for placement on the 2007 field test forms. Items that increased the DOK levels of operational items, or which were needed to meet the blueprint changes, were given first priority in field test sessions. Remaining field test slots were filled using items that increased the flexibility and diversity of the operational item banks. Table 3-4 shows the number of unique field test items tested through Fall 2007.

3.1.3 Quality Reviews

A smooth test administration requires that all test materials, including test books, manipulatives, and test administration manuals align with each other. All items, page numbers, and administration times must be accurate in all components of the test program. When materials are not in alignment, not only can rework and additional costs be incurred, but there is also the possibility of jeopardizing the validity of test results and creating poor publicity. Therefore, to

Copyright © 2008 by the Wisconsin Department of Public Instruction

help ensure all documents required for the administration of a test are in alignment with each other, a materials integration review (MIR) is conducted prior to moving the materials on to the Quality Assurance (QA) Department within CTB.

During the MIR, a proctor simulated the test experience by administering the test to two test takers for each grade and content area using the WKCE examiner's manual. The purpose of this review is twofold: to ensure the test materials are in alignment with each other and to verify the answer keys are correct.

In addition a Quality Assurance (QA) review was conducted on each test book and all ancillary materials. The purpose of the QA review is to ensure all publishable products meet the standards and expectations of DPI. The QA review includes, but is not limited to, the review for: page number location/order, header/footer information, "go on" and "stop" signs, item sequence numbering, accuracy of directions, vertical and horizontal alignment, conventions of written English, clarity/accuracy of art, accuracy of cross-references, and that there is only one correct answer to each item. This QA review occurred at the end of the page production cycles and prior to releasing the materials to Manufacturing.

In addition to the MIR and QA review steps, the WKCE test books were also reviewed by CTB's Technology department to verify the scannable test books were constructed to meet CTB's scanning and scoring specifications. With each round of page production, CTB production staff viewed the position of answer choice bubbles to confirm they were "on grid" and readable by CTB scanners. In addition, at the second pages stage, all test books were reviewed by Technology specialists to verify that bubbles were on grid, that there was no "bubble back-up" that would interfere with accurate scanning, and that other scannable page elements were properly placed.

3.2 Description of the WKCE 2007 Tests

The 2007 test books contained Reading and Mathematics content in a single test booklet at each grade for grades 3, 5, 6 and 7. Test content for grades 4, 8, and 10 was included in two unique test booklets. Reading, Mathematics, and Science content for grades 4, 8, 10 was comprised in Book 1; Language Arts, Writing, and Social Studies comprised Book 2.

The Reading and Mathematics tests for grades 3–8 and 10 consist of custom items developed specifically for the WKCE. Language Arts, Science, and Social Studies at grades 4 and 8 consist primarily of TerraNova items; but a few custom multiple-choice items were added to address content standards not adequately covered by the TerraNova items. The grade 10 Language Arts, Science, and Social Studies tests consist of custom items previously developed for Wisconsin.

3.2.1 Reading

Table 3-1 presents the configuration of the operational tests. The Reading test for grades 3–8 had one operation passage for each of the six types of passages: short literary, long literary,

Copyright © 2008 by the Wisconsin Department of Public Instruction

short informational, long informational, poetry, and everyday text. The embedded field test session had one or two passages, which could be any combination of the six types of passages. Table 3-2 presents the number of Reading embedded field test items by grade, form, objective, and item type.

There were four test sessions: three containing operational items and the fourth containing the field test items. Each grade had at least one set of paired reading passages with a few items that required analyzing or synthesizing ideas from the passages. Each of the three sessions with operational items had approximately 18 multiple-choice items. Two of the three operational sessions included a constructed-response item. One of the constructed-response items was for the reporting category Analyzing Text, while the other was for the reporting category Evaluate and Extend Text. Each session was allotted 40 minutes of testing time. The field test session for each grade was allotted 30 minutes. The grade 10 test consisted of three sessions: Sessions 1 and 2 were 35 minutes and Session 3 was 40 minutes.

For grades 3–8, there were four different forms. The operational items in all forms were the same, but the embedded field test items differed by form. Grade 10 had one form and did not contain any embedded field test items for Reading.

3.2.2 Mathematics

Table 3-1 also shows the operational Mathematics test structure. The Mathematics test for grades 3, 4, and 5 had three sessions with operational items and one session for field test items. Grades 6, 7, and 8 had five sessions—four with operational items and one with field test items. The grade 10 test had four operational sessions. Table 3-3 presents the number of Mathematics embedded field test items by grade, form, objective, and item type.

The first session at each grade and the first part of the field test session at grades 3–8 was a “non-calculator” session. Grades 3 and 4 do not permit the use of calculators for any session. For these grades, if a student is provided an accommodation that allows the use of a calculator, the calculator may not be used to answer the items in Session 1 or the first part of the field test session.

For grades 3–8, there were four different forms. The operational items in all forms were the same, but the embedded field test items differed by form. Grade 10 had one form and did not contain any embedded field test items for Mathematics.

3.2.3 Language Arts

The operational test configurations of Language Arts tests for grades 4, 8, and 10 are presented in Table 3-1 as well. The grade 4 and 8 Language Arts tests consisted of 24 *TerraNova* multiple-choice items and six custom multiple-choice items that measure Content Standard F, Research and Inquiry. The session was allotted 30 minutes of testing time. There was a writing session in grades 4 and 8 that presented an operational writing prompt. This session was allotted

Copyright © 2008 by the Wisconsin Department of Public Instruction

30 minutes. The grade 10 test consisted entirely of custom items developed for Wisconsin. The test was administered in two sessions; the first session contained the 30 MC items, and the second session contained the writing prompt.

3.2.4 Social Studies

Table 3-1 also presents the operational Social Studies test structure. The Social Studies test at grades 4 and 8 consisted almost entirely of *TerraNova* items, but also included a few custom items previously developed for the WKCE. There was one test session at these grades. The grade 10 test consisted of 50 custom multiple-choice items developed for Wisconsin. The test was administered in two sessions. Each session was timed at 30 minutes.

3.2.5 Science

Table 3-1 presents the operational Science test structure as well. The Science test at grades 4 and 8 consisted almost entirely of *TerraNova* items, but also included a few custom items previously developed for the WKCE. There was one test session at these grades, which was allotted 40 minutes. The grade 10 test consisted entirely of custom items developed for Wisconsin. The test was administered in two sessions; each session was allotted 40 minutes.

3.3 Customer Approvals

The development phases where DPI approval was obtained included:

- item content and bias review
- item selection for the Fall 2006 operational forms
- manuscript
- second pages
- final pages (prior to release to manufacturing)

3.3.1 Item Content and Bias Review

Following the review of items, CTB and DPI staff reviewed the edits recommended by the educator committees. DPI gave final approval of educator recommendations. DPI and CTB each kept a copy of the item review book with the edits marked.

3.3.2 Item Selection Approval

ITEMWIN selection summary reports were submitted to DPI, which included graphics of test characteristic and standard error curves, lists of items selected, and summary test statistics. DPI approval was obtained using a sign-off form.

Copyright © 2008 by the Wisconsin Department of Public Instruction

3.3.3 Manuscript Approvals

CTB content editors submitted a copy of the test book manuscript to the CTB Production team. The manuscripts show the items as sequenced within test sessions. The manuscripts for the test administration manuals were also submitted to DPI for review, and many content changes were addressed at this stage. DPI approval was obtained using a sign-off form.

Production returned the test book pages to CTB style editors as first pages. CTB style editors reviewed first pages to ensure pages followed the proper format. CTB content editors reviewed first pages for format and content issues. Content editors marked first pages to indicate content changes requested by DPI on the manuscript sign-off form. CTB content editors submitted a copy of first pages with correction markup to Production, and the edits were incorporated in the InDesign file. CTB editors reviewed the corrected pages before submitting them to DPI for review. If an edit was not incorporated correctly, it was re-marked for correction.

3.3.4 Second Pages Approvals

The second pages represent DPI's first review of the composite test books or test administration manual pages. By this point, all content issues had been resolved. That is, the focus of the approval was on format and presentation issues, rather than content. DPI approval was obtained using a sign-off form.

3.3.5 Final Pages Sign-off

The final pages represent DPI's last opportunity to review test book and test administration manual pages prior to releasing the materials to CTB's Manufacturing team. At this stage, the materials had been through CTB's quality assurance process and all queries had been resolved. The focus of this review was to verify that previously-requested edits had been made, and that there were no errors in content or conventions of standard written English. DPI approval was obtained using a sign-off form.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 4: Test Administration

In the Fall of 2007, Wisconsin administered assessments in Reading and Mathematics grades 3-8, and 10, as well as assessments in Language Arts, Social Studies, and Science grades 4, 8, and 10. The test administration window was October 22 - November 27, 2007. Part 4 of the Technical Report describes a set of standardized procedures and policies applied to administer WKCE assessments. The issue of test security in test administration has important implications for the integrity of the results and thus the validity of WKCE scores. Documentation citing the written procedures provided to test administrators and school personnel, in order to standardize the administration of the test, are also provided here. The following AERA/ APA/ NCME standards are addressed here: 1.13, 3.3, 3.19, 3.20, 3.21, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 6.11, 6.15, 9.1, 10.1, and 10.2.

The Wisconsin Department of Public Instruction (DPI) is committed to the proposition that all schools, and all students within schools, will be held accountable to a common set of high academic content standards. Students, who have an Individualized Education Plan (IEP) - a 504 plan (under Section 504 of the Rehabilitation Act of 1973) - or are identified as Limited English Proficient (LEP), or Formerly Limited English Proficient (FLEP) may be eligible to receive testing accommodations. Accommodations are changes in the routine conditions under which a student takes an assessment in order to provide the student equal opportunity to demonstrate their knowledge. The types of accommodations and guidelines for test administration conditions are described below.

4.1 Accommodations

Accommodations were allowed for eligible individual students participating in the WKCE. Accommodations provided to a student must be documented in a current IEP and used during routine instruction. IEP teams were directed to refer to the WKCE accommodations policy (http://www.dpi.wi.gov/oea/pdf/assessmatrix07.pdf). Test administrators indicated which accommodations were used by each student by completing the Student Assessment Report located on the back cover of the student Answer Document. The following accommodation information was collected from the Student Assessment Report:

Type of Accommodation:

- Used translation
- Signed test questions and content to student
- Used Braille
- Used assistive device (e.g., text-talker, adaptive keyboard, picture symbols)
- Used objects or manipulatives
- Used another DPI-approved accommodation
- Used a non-allowed accommodation, resulting in the invalidation of test results.

Copyright © 2008 by the Wisconsin Department of Public Instruction

In 2007 the State of Wisconsin developed Spanish and Hmong translation scripts for the WKCE. The aim of these scripts is to better help students demonstrate what they know on the WKCE without the interference of language. Students whose native language is Spanish or Hmong will be given the choice to use all or parts of the translation accommodation, which includes a bilingual glossary of commonly used testing terms, translation of the test directions, and a written translation script of Mathematics, Science and Social Studies test items. DPI recommends that educators also consult the list of allowable accommodations in order to create the most appropriate testing situation for their students.

DPI recognizes that 15% of the Wisconsin ELL population speaks a language other than Spanish or Hmong. Districts who serve students who speak languages other than Spanish or Hmong may use qualified translators to provide oral translation support to students. However, the use of translation supports is restricted to Mathematics, Science, and Social Studies Tests.

Table 4-1 provides the list of standard accommodations made available for the Fall 2007 WKCE assessments, and the number and percent of students provided these accommodations. Table 4-1 provides a summary view of the accommodations provided, based on all students. The table is split across pages by accommodation, with one accommodation per page. Additional accommodation tables were also delivered to DPI which detailed the accommodations provided for subgroup populations of interest, including the five NCLB groupings discussed in this report, (gender, ethnicity, socioeconomic status, disability status, and English Language Proficiency) as well as migrant status.

4.2 Reporting Results of Assessments Taken with Accommodations

Scores of assessments taken with accommodations were included with the results of students who took these tests under standard conditions and presented at the school, district, and state level.

4.3 Test Security

The primary goal of test security is to protect the integrity of the examinations. To ensure that trends in achievement results can be calculated across years, and in order to provide longitudinal data, a certain number of test questions must be repeated from year to year. If any of these questions are made public, the validity of the test may be compromised. Access to test materials was limited to those educators who required access. DPI ensured that all who had access to test materials understood the critical need for test security, presented during the 2006 and 2007 Pre-Test Workshops, and outlined the acceptable and unacceptable test preparation and administration practices (Do's/Don'ts sheet provided in the Test Coordinator Kits). All WKCE tests were administered under secure testing conditions established by the DPI.

Wisconsin Student Assessment Security Warning: The following statement was directed by DPI to appear on every test booklet beginning with the 2004-2005 school year:

Copyright © 2008 by the Wisconsin Department of Public Instruction

Test Security

All passages, stimuli, and questions used in the *Wisconsin Knowledge and Concepts Examinations-Criterion-Referenced Test* are CONFIDENTIAL and must be kept SECURE at all times. Unauthorized use, duplication or reproduction of ANY or ALL portions of the test materials is prohibited. Violation of security can result in district disciplinary action, prosecution, and/or penalties by the Department of Public Instruction or CTB/McGraw-Hill.

Other security measures for WKCE test administrations are described below.

ELL students and students with disabilities were allowed to use highlighters. Test administrators were instructed to carefully supervise the use of highlighters because they may cause smudging of pencil marks and bubbles, which could affect reliability of scanning and scoring. If highlighters were used, the following guidelines were provided:

Guidelines for Highlighters:

1. Do not allow the highlighting of track marks, litho codes, skunk lines, barcodes, preslugged bubbles or any carbon black printing. The highlighters cause these black inks to blur and bleed.

2. Do not allow the highlighting of pencil marks of any kind, whether bubbles or handwriting. The highlighters cause pencil marks to blur and bleed.

3. Use only a highlighter from the following list, which were tested and found to have minimal problems:

   - Avery Hi-liter
   - Avery Hi-liter, thin-tipped
   - Bic Brite-Liner
   - Sanford Major Accent
   - Sanford Pocket Accent, thin-tipped

Use of restroom and wireless communications:

Test security must be maintained during all breaks within a testing session. To lessen the risk of a security breach occurring during these breaks, students requiring the use of restroom facilities must be escorted by either the proctor or a test examiner. In addition, students must not be allowed to use any form of wireless communication during these breaks.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Parameters for marking test booklet with No. 2 Pencil:

- Do not mark in the bubble answer positions.
- Do not mark in the student Pre-ID Barcode on barcode label.
- Do not mark in the timing tracks (the parallel lines along the side of the test booklet).
- Do not mark in the skunk lines (the little squares and rectangles across the bottom of each page of the test booklet).
- Do not mark in the Litho codes (the squares and numbers across the bottom of the document on the first and last page of the test booklet).
- Do not mark more than one answer bubble as the scanner cannot determine a response.

4.4 Test Administration

In order to ensure standardized testing administration for all students, a Guide for District Assessment Coordinators and School Assessment Coordinators was made available to all test coordinators (DPI, 2007-2008). The Guide included the following topics:

- Test Security
- Schedule of Important Dates
- Responsibilities of District Assessment Coordinators (DACs)
- Responsibilities of School Assessment Coordinators (SACs)
- Scheduling Test Administration
- Preparations for Testing
- Students to be Tested
- Standard Accommodations
- DPI Guidelines to Facilitate Participation of Students with Special Needs
- Data Collections and Reporting
- Word Processors, Scribes, Tape Recorders, and Large Print and Braille Tests
- Test Materials
- Receiving Test Materials
- Inventorying Test Materials
- Procedures During Test Administration
- Procedures Following Test Administration
- Returning Materials to CTB/McGraw-Hill

In addition, Test Administration Manuals were made available to all test administrators. They included the following:

- Test Materials
- Test Security
- Testing Schedules
- Test Administration Guidelines
- Guidelines for Preparing Students to Take the Test

Copyright © 2008 by the Wisconsin Department of Public Instruction

- Use of Appropriate Testing Procedures
- Student Identification Information
- Explanation of Symbols
- Detailed Scripts for Administration of Each Part of Each Test
- Procedures Following Test Administration

For specific information related to test administration, refer to the Test Coordinator's Manual and/or the Test Administration Manuals available at (http://www.dpi.wi.gov/oea/publications.html).

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 5: Scoring

The fundamental purpose of Part 5 is to demonstrate adherence to AERA/APA/NCME standards for scoring, including 3.22, 3.23, 3.24, 5.8 and 5.9. Part 5 describes:

- Scoring process of multiple-choice items
  - Scanning and calibration of scanners

- Scoring of open-ended or constructed-response items
  - Scoring rubrics
  - Handscoring process
  - Electronic handscoring system
  - Selection of readers or Scoring personnel
  - Selection of anchor papers
  - Distribution of constructed-response item scores

5.1 Scoring of Multiple-Choice Items

The Fall 2007 WKCE Student Test documents were returned to CTB's scoring facility by the districts at the conclusion of test administration. Test materials were tracked throughout the entire process, from retrieval to document retention.

CTB's Scoring Operations processes were organized into Lean Processing Scanning Cells. Each workcell was a self-contained, cross-functional team made up of the stations, equipment and personnel skill-sets necessary to efficiently and accurately complete the operational processing cycle for student test documents.

The process, through which the answer documents traveled, in order, was as follows:

Receiving – Answer documents were tracked from retrieval to receipt at CTB, checked for damage in shipping, verified for full box counts, registered into an internal tracking system called On-Hold Tracking System (OHTS), and passed along to Login.

Login – Answer documents were removed from the boxes, the pre-work verified for district accuracy, and stacks aligned and cut for scanning.

Scanning – Stacks were fed through the scanners (see detail below for scanner information and calibration statistics) and problems monitored and rectified.

Updates – Raw scoring and editing of scanned student data was performed using a sophisticated system of edits to determine the integrity of each batch scanned and produce an error list to be resolved by trained editors using pre-defined guidelines.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Document Retention – Documents were moved to a staging area where they were caged, warehoused and retained for retrieval during the specified retention period. At the end of the specified retention period for the WKCE contract, these documents will be securely destroyed according to contract guidelines.

Within the workcell, documents were staged, then image-scanned using customized scanning software and high-speed image scanners, and then processed through CTB's proprietary Winscore editing system. Documents were moved directly from process to process, or sat momentarily in mini-queues. Once complete, documents were prepared for secure document retention.

The scanners efficiently captured all multiple-choice, ancillary and student demographic data. The scanning production systems separated the constructed-response items for handscoring which were later re-merged with the scored multiple-choice data file for correction, analysis and reports. The following section describes scanning of student responses and quality assurance process that goes into it.

5.1.1 Scanning and Calibration of Scanners

This section provides description of the capacity of scanners and the scanner calibration process. CTB's scanning software utilized the speed of optical scanners to capture document images and bubbled data. The scanners were able to run at a rated speed with no interruptions except for problems with the physical documents. Technology termed Optical Mark Recognition (OMR) detected all pencil marks in the answer section of the scanned document. The scanners scanned in OMR mode, capturing student response images called TIFFs, as well as reading scannable barcodes. CTB's proprietary Winscore scanning program evaluated detectable marks on both sides of each page, recording the intensity and coordinates of solid marks for resolution in the raw scoring step. Completed bubbles were turned into characters of data representing test item responses, or other information. The scanner reported intensities in the range 0 (lightest) to 15 (darkest). The form identification (i.e. 'skunk marks') determined the type of the document and the headers determined customer identification, district, school and class.

At the beginning of each shift and every 5000 sheets thereafter, a diagnostic sheet was run to assess camera functionality. CTB cell leads cleaned the scanners at the end of their shift and ran the quick check utility to confirm the equipment was ready for the next shift. If the scanner did not pass quick check or a diagnostic check, a field engineer was called in to address the problem. If the camera was adjusted in any way, the scanner was re-calibrated and the quick check utility was run again. When both passed, the scanner was released for scoring. All scanners were calibrated as scheduled.

The scanners detected all pencil marks in the answer section of the scanned document. Winscore scored the darkest mark for each question as the intended response. Additional (faint) marks on the answer document can be evaluated in erasure analyses, an option that was not produced for the WKCE.

Copyright © 2008 by the Wisconsin Department of Public Instruction

All images were captured during scanning using hi-resolution technology, also called gray-scale. Any item determined to be 'unclear' was electronically retrieved in gray-scale in the Electronic Handscoring System (EHS).

The integrity of student responses was ensured by, but not limited to, the following checks:

Reliability check - This checked that low scores, either among groups or at the individual student level, were within reasonable ranges.

Biographical data – This checked for obvious errors such as duplicate barcode numbers, and that district and school data matched school headers sheets.

Student counts – Actual book counts generated by the scanners were compared to counts provided by the district on the School Group List and School Header Sheet. In 2007, 281 such incidents were identified and resolved.

School name/number – Pre-assigned school numbers and names were verified against data provided by DPI.

Fall 2007 WKCE books were scanned using these scanning machines. No re-calibration was necessary during the WKCE Fall 2007 administration. Processing metrics obtained were:

- Number of sheets scanned: 26,283,588
- Number of books scanned/processed: 631,761

These scored responses were later retrieved by Research and Technology for statistical analyses and production of reports.

5.2 Scoring of Open-Ended or Constructed-Response Items

This section documents handscoring processes used for WKCE for open-ended or constructed-response items as a part of demonstrating validity evidence of handscoring process. The processes include description of scoring rubrics, handscoring process, selection of sample (anchor) model paper for reader (scorer) training, the process of selecting readers, inter-rater reliability, and distributions of scores.

5.2.1 Description of Scoring Rubrics

For the 2007 operational test administration, the Reading and Mathematics forms at grades 3–8 had constructed-response items. In addition, a Writing prompt was administered at grades 4, 8, and 10. The Writing prompts at grades 4, 8, and 10 were scored using two holistic rubrics: a 3-point rubric for conventions of written English, and a 6-point rubric for composing. Tables 5-1 to 5-8 present the scoring rubrics.

Copyright © 2008 by the Wisconsin Department of Public Instruction

5.2.2 Handscoring Process

Handscoring readers were trained using customer-approved training material (rubrics, anchors, horizontal training, and two validation rounds). Once qualified, readers were required to maintain accuracy standards throughout the project. These requirements were assessed at the item level, primarily through each reader's daily checkset performance, as well as agreement rates with other readers on the second reads, and targeted read-behinds with team leaders. Data monitors generated reports daily that flagged any readers falling below the established quality standards for any item, providing insight on reader scoring trends (such as difficulty with any particular score point), and shared these reports with handscoring supervisors. Those readers identified in the reports received additional coaching such as additional training or review, targeted read-behinds or additional checksets. Readers who did not meet standards with these initial corrective actions were administered another validation (recalibration) round. Failure to recalibrate resulted in dismissal from the scoring assignment. This process was in place throughout the entire handscoring window.

5.2.3 Electronic Handscoring System

For WKCE programs, the Electronic Handscoring System was used to score constructed-response items. EHS presented images of scanned test books to trained readers, who assigned scores for constructed-response items. Scanned output was viewed on high-quality 19-inch workstation monitors. Images of each student's responses were automatically routed to two or more readers when required, and images of specific subsets of test items routed to designated groups of readers trained to score these items.

5.2.4 Anchor Papers and Training Papers

Anchor papers (model papers written by WKCE students) were selected to train raters for scoring. These papers play an important role in deciding which level of writing should receive which score. Prior to the actual scoring, the CTB Scoring Center created training materials. CTB randomly sampled student answer documents to ensure a representative sample of the possible responses. Range-finding meetings were held with DPI staff and educators to select sample papers of each score point. Sample papers were used to construct scoring guides and training papers. CTB's Scoring Team collaborated with DPI to make necessary revisions to the rubrics and in the selection of scoring guides and training papers.

This process included several presorting steps and subsequent iterative/consensus processes in order to achieve agreement and precision through a kind of "round robin" scoring, followed by discussion and selection.

Model responses/papers were selected for anchor training, horizontal training, and qualifying, and were consolidated into training formats. Once approved by DPI, the Scoring Guides (consisting of rubrics, anchors, and annotations) served as a constant guide, setting the course for all subsequent training and scoring.

Copyright © 2008 by the Wisconsin Department of Public Instruction

5.2.5 Scoring Personnel and Qualifications

CTB recruited, trained, and managed staff to complete all handscoring operations within the timelines of the contract. This involved extensive consultation between CTB Scoring, CTB Publishing, Wisconsin educators, and DPI to review scoring rubrics, to develop anchor papers and other reader training materials, and to provide analyses of student responses to tryout forms. The characteristics of the readers, team leaders, and scoring supervisors are described below.

Readers

Many CTB readers had many years of classroom teaching experience. The CTB reader pool included editors, published authors, and individuals with advanced degrees. The minimum qualification for all readers was a Bachelor's degree. Evaluator staff was comprised of individuals from many walks of life—from retired or current educators, to engineers, possessing Bachelor's to Doctorate degrees. Readers were required to participate in training and successfully pass at least one of two qualification rounds. Once qualified, readers could start scoring, but throughout the scoring processes, their scoring performance was assessed by a supervisor and data monitoring staff through the use of checksets and read-behinds and the review of inter-rater reliability statistics.

Team Leaders

Team leaders were selected on the basis of their ability to maintain a high degree of scoring accuracy and consistency, often across multiple subjects and grades. Team leaders were also required to possess good interpersonal and leadership skills in order to be effective when training and counseling readers. Team leaders were responsible for a small team of readers, and in addition to scoring, which includes performing read-behinds on readers, performing counseling and coaching of readers when required by data monitoring or supervisor staff. Team leaders working on the writing component also resolved discrepant scores.

Scoring Supervisors

Scoring supervisors were the core group at CTB who directed and organized the assessment process and trained team leaders and readers. Scoring supervisors had extensive experience as team leaders prior to their qualification and selection. Scoring supervisors were subject area experts in the content areas they supervised and trained. They oversaw all team leaders and readers.

5.2.6 Reader Training

Validation was a critical task in the assessment training process, and the final determinant of reader readiness. All readers, including team leaders, were required to achieve standard

Copyright © 2008 by the Wisconsin Department of Public Instruction

percent agreement on the qualifying round that followed training. Standard percent agreement varies with the score point range of an item (0 – 1 point, 95%; 0 – 2 points, 90% etc.). Those readers not validating on the first attempt received further training prior to taking an additional qualifying round. Only those who were successfully validated qualified as readers to score tests. Team leaders were required to complete two validation rounds with 80 percent exact agreement in each round.

5.2.7 Inter-Rater Reliability

Checksets

Throughout the course of the handscoring process, calibration sets of pre-scored papers (checksets) were administered daily to the team leaders as well as to the readers. These were used to monitor scoring accuracy and to maintain a consistent focus on the established rubric and guidelines. This monitoring occurred without reader knowledge. Readers whose checkset scores fell below the qualifying level were flagged for additional coaching (training review, targeted read behinds, etc.). Those readers who remained below standard were given another validation (recalibration) round. Readers unable to recalibrate were dismissed.

Read-Behinds

The "read-behind" was another valuable intra-rater reliability monitoring technique used. Each team leader was able to read a random selection of a reader's scored items, targeting at the item and score point level, or at the Rater Item Block level, where (potentially) more than one item was being scored. The scores were compared, and if they agreed, the team leader was able to offer feedback, which enhanced the reader's confidence and ability to score quickly and accurately. However, if a reader strayed from the standard established in the training and validation samples, the aberrant scoring was detected, and the team leader was able to offer guidance necessary to refocus the reader's effort. Readers whose scoring was inconsistent were read behind more frequently by their team leaders, thus correcting any scoring variations.

Double Reads

Each Writing constructed-response was scored by two readers independently and inter-rater reliability was monitored throughout the scoring process. For the 6-point rubric, if the scores of the two assigned readers differed by one point, the student received the higher of the two scores. If the scores of the two readers differed by more than one point, a third rating was provided by an expert rater, who resolved the discrepancy and assigned a final score. For the 3-point rubric, if the scores of the two assigned readers differed by one point, a third rating was provided by an expert rater, who resolved the adjacency and assigned the final score. For Reading and Mathematics items, the first score assigned is the final score, however five percent were double read (2nd read) for statistical purposes.

Copyright © 2008 by the Wisconsin Department of Public Instruction

5.3 Distribution of Constructed-Response (CR) Item Scores

Tables 5-9 to 5-16 show the percentage of examinees at each score level (1 point, 2 points, etc.) for each CR item. These distributions of examinee responses across score levels were provided to check the reasonability of CR items. The score distributions and the condition code counts yield several indications of reasonability. For example, a score distribution with a very high proportion of students scoring at either the highest or the lowest score level could indicate a need to investigate the reasonability of the item further.

The tables also use four condition codes. Condition code "A" denotes no response or no attempt, "B" represents illegible, "C" means another language, and "D" denotes off-topic. Responses receiving condition codes were converted to zero score points. All responses for operational items* were scored, but only a portion of the responses for field test items was scored. Operational and field test items can be identified based on the number of students (N) in the tables.

Tables 5-9 and 5-10 show the percentage of students who received different score points in Reading and Mathematics constructed-response items. These tables show the score distribution of the first read. All Reading items have one part, and a maximum score of 3 points. As can be seen in Table 5-9 for Reading, most items split scores between 0, 1, or 2 score points, and relatively few students scored at the 3-point level on most items.

All Mathematics items in grades 3-8 have two parts. Part A has one score point, and Part B has two score points. In grade 10, the Mathematics items have one part. Looking across the Part A CR items in Table 5-10, one may observe that approximately 22% to 89% of students obtained the single score point for these items. This suggests a wide range of difficulty levels across these CR items in Part A. For the Part B items, the score distributions per item spread across the range of available score points. The most common condition code in Reading and Mathematics was "A," indicating that these students did not attempt the CR items.

Tables 5-11 to 5-16 present the score distributions for the Writing Prompts. The tables for Writing are split between the 6-point Composing rubric and the 3-point Convention rubric. There is one writing prompt in each grade 4, 8, and 10. While only the score from the first read is used for Reading and Mathematics, the average of two raw scores is used for Writing. As such, the similarity of first rater's score and the second rater's score is important. To evaluate this, scores from the first read and second read were reported, and the difference between them was also presented. As can be observed in the tables, the distributions of the two reads were similar across all score levels for grades 4, 8, and 10. These tables are split between counts and percentages.

As can be seen in Tables 5-11 and 5-12, most scores in the Composing rubric were in the middle of the 6-point range, and relatively few students were at the low and the high extremes. The Convention rubric showed similar results. As can be seen in Tables 5-13 and 5-14, a high

* As indicated earlier, operational items are those items that contribute to student scores. Operational items are contrasted with field-test items, which do not contribute to student scores. Operational items may be abbreviated in this report as OP, and field-test items are abbreviated as FT.

Copyright © 2008 by the Wisconsin Department of Public Instruction

proportion of students scored in the middle level of the 3-point range for the Convention rubric, and relatively few students scored at either 1 point or 3 points.

Tables 5-15 and 5-16 show the total score, combining both Composing and Convention. The combined scores for most students were in the middle or upper-middle range of the 9 point total, from 4 points to 6 points. The highest and lowest levels of scoring were less common, but in every grade, a small proportion of students obtained 0 score points, and a small proportion obtained the highest possible score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 6: Characteristics of Sample Data

The calibration, equating, and scoring of the Fall 2007 WKCE were based on a pre-selected sample of student data from the state. This arrangement was chosen in order to expedite the return of score reports to districts. In accordance with AERA/APA/NCME standards 1.5, 1.13, 2.4, 4.7, and 6.1, Part 6 provides a description of how the sample data was selected and a comparison of the sample data and the population data in terms of demographic characteristics. Part 6 serves to demonstrate that the sample data were sufficiently representative of the Wisconsin student population for the purposes of calibration. This documentation forms part of the validity evidence supporting the WKCE program.

6.1 Calibration Sample Data

Table 6-1 lists the 13 school districts from which the sample student data were obtained. These districts are referred to as the calibration districts or as the calibration sample. These districts have been used to obtain a calibration sample every year of the WKCE since the Fall 2005 administration. This year, one district, Ashland, was dropped from the list for logistic reasons. Prior to the Fall 2007 administration, fourteen calibration districts were used. The calibration sample was selected to represent the state student population in terms of demographic composition and student performance, based on recommendations from the Wisconsin Department of Public Instruction (DPI) and, more fundamentally, on an analysis of the demographic and performance profiles of these districts.

The reader may wish to observe that the sampling unit in the sample data is the district. All of the schools and students within these districts were part of the sample. Only invalid records were excluded. Readers interested in how records were determined to be invalid can refer to the Fall 2006 WKCE Technical Report, section 7.2. Readers may also wish to note that the census data includes students from the 13 districts. In other words, the 13 districts were a subset of the census data. Table 6-2 lists the number of students in the sample data and the census data, for each grade level.

6.1.1 Demographic Comparison of Sample and Population Data, by the Five NCLB Groupings

The calibration districts were determined to be a representative sample of all Wisconsin students based on two analyses. First, the demographic composition of the sample data was examined and compared to the demographic composition of the entire student population. This included a focus on the five *No Child Left Behind* (NCLB) groupings (gender, ethnicity, socioeconomic status, disability status, and English Language Proficiency). Then, the sample data and the population data were compared based on the mean scale score. The results of this latter comparison are provided in Part 8, where results of the Fall 2007 administration are discussed. As detailed below and in Part 8, the results of both of these analyses indicate that the 13 calibration districts formed a sufficiently representative sample of all Wisconsin students. In the following paragraphs readers will find that the demographic profiles of two data sets are

Copyright © 2008 by the Wisconsin Department of Public Instruction

similar, but some differences exist. However, because WKCE applies item response theory to student response data, given a sufficient number of students at each point along the ability range, these discrepancies will not have any measurable impact on calibration results.

Side-by-side comparisons of the two student profiles, as derived from the calibration sample and the WI census data, are presented below. The two data sources are compared first in terms of gender, followed by ethnicity, socioeconomic status (SES), disability status, and lastly by English Language Proficiency (ELP), in Tables 6-3 to Table 6-7.

As shown in Table 6-3, there were more male students in both the census and the calibration sample data. In most grades, males slightly outnumbered females. The difference between the sample and the census data was small. The two data sources present the same pattern of gender distribution in the WI student population.

Table 6-4 presents ethnicity characteristics. Overall, the two data sources show the same pattern: White students were numerically predominant, by far, in all grades. African American students were the next largest group, followed by Hispanic, Asian, and American Indian students. However, compared to the census data, the calibration sample contained a smaller proportion of White students, and a larger proportion of students of every other ethnicity. In the census data, White students were 75% to 80% of the student population across grades. In the calibration sample, White students were 63% to 70% of the population, in each grade level.

Table 6-5 provides two profiles of the student population in terms of socio-economic status (SES) based on the two data sources. Students were identified as “Economically Disadvantaged” or “Not Economically Disadvantaged.” In the table, “Yes” means “Economically Disadvantaged.” The census and the sample data again show similar profiles: a large proportion of the student population was identified as Economically Disadvantaged in both data sources. However, students identified as Economically Disadvantaged formed a smaller proportion of the total student population in the census than in the sample. The census data shows that 32% to 37% of students in grades 3-8 were identified as Economically Disadvantaged, and 26% of students in grade 10 were identified as Economically Disadvantaged. The proportion declines in each grade, from grade 3 to grade 10. The calibration sample data shows that 36% to 40% of students were identified as Economically Disadvantaged in grades 3-8, and in grade 10, the proportion drops to 30%. Both the sample data and the census data reflect a general decline in the proportion of students identified as Economically Disadvantaged from grade 3 to grade 10.

Table 6-6 presents the data on disability status. In this table, “Yes” denotes “Disabled.” The census data and the calibration sample data show very similar patterns. In both cases, approximately 1 in 7 students were identified as Disabled. The calibration sample shows only a slightly larger proportion of the population as Disabled than the census data. The difference between the census and the sample was not more than 2% in any grade level.

Table 6-7 shows the percentage of students that were “Proficient in the English Language” and “Not Proficient in the English Language” based on both the census data and the sample data. In the table, “No” indicates students “Not Proficient in English Language.” The

Copyright © 2008 by the Wisconsin Department of Public Instruction

census data and the calibration sample data show a similar pattern. Most students were Proficient in the English Language in both data sources. However, the population that was Not Proficient in the English Language was larger in the calibration sample than in the census. The census data shows that from 4% to 7% of students were Not Proficient in the English Language. The calibration sample shows that 10% to 15% of students were Not Proficient in the English Language. Both the census and the sample data show more students in the lower grades and fewer students in the higher grades as Not Proficient in the English Language.

In summary, the demographic profiles of the two data sets are very similar, and while some differences exist, the sample data were sufficiently representative of the WI student population for the purposes of calibration. The quality of the sample data and its comparability to the entire student population of WI forms an important part of the validity evidence supporting the WKCE program.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 7: Calibration, Equating, and Deriving Scale Scores

The WKCE program reports scores based on item response theory (IRT). These scores are established through the processes of calibration, equating, and item-pattern scoring. Part 7 of the Technical Report serves to describe these processes as they were applied in the Fall 2007 administration, as well as the results. This portion of the Technical report addresses AERA/ APA/NCME standards 1.13, 4.1, 4.2, 4.3, 4.10, 4.11, 7.1, 7.2, and 7.10.

Readers should note that calibration, equating, and scoring using IRT are mathematically complex processes and a full understanding of these topics requires a background in psychometrics. However, in order to make these processes more accessible and transparent to a wider range of audiences, a brief non-technical explanation of how scale scores are derived from raw scores is provided in section 7.3. Additional references to supplemental materials are also provided there.

At the outset here, non-technical readers may note that, in psychometric terms, calibration is a process of estimating item parameters. Sections 7.1, 7.1.1, and 7.1.2 largely serve to explain how item parameters were estimated for the WKCE. Part 7 begins with a description of the calibration and equating methods used in the Fall 2007 WKCE, the calibration models are then discussed, and the software used is also reviewed. Then the derivation of scale scores from raw scores is explained, with a focus on non-technical audiences. Part 7 then moves on to review the results of the calibration process, using model-to-data fit statistics, and the standard error of measurement.

7.1 Calibration and Equating Methods

IRT provides theoretical models of the item response process. Multiple-choice (MC) and constructed-response (CR) items have different characteristics; as a result, IRT provides different models for each item type. In the Fall 2007 WKCE, the three-parameter logistic (3PL) IRT model (Lord & Novick, 1968; Lord, 1980) was used for MC items, and a two-parameter partial credit (2PPC) model (Muraki, 1992; Yen, 1993) was used for CR items. All Language Arts, Social Studies, and Science, items were calibrated using the 3PL model, because all of the items in these content areas were MC items. Since the Reading and Mathematics tests consisted of both MC and CR items, a simultaneous calibration with the 3PL and 2PPC models (Muraki, 1992; Yen, 1993) was implemented. Simultaneous calibration was chosen for the mixed format tests, in part, because a single scale that reflects the trait assessed by the two item types is theoretically attractive and technically feasible. The single scale also communicates the instructionally sound idea that the skills to be assessed relate to the same underlying qualities and characteristics, and that they can be taught and measured using a variety of assessment modes. In considering the simultaneous calibration process, it is also important to recall the position of Thissen, Wainer, and Wang (1992) that items of diverse types can be scaled together provided that the different types of items assess the same primary characteristics. A simultaneous calibration was also applied to the Writing Prompt in grade 10. The 3PL and 2PPC models are described in the next section.

Copyright © 2008 by the Wisconsin Department of Public Instruction

By design, there was a special set of items in each content area that were common to both the current administration and a prior administration. This arrangement is called a *common item non-equivalent group* design. The purpose of this design is to link the current item parameters to a base scale, using the common items. Linking the current test forms to the previously established scales is necessary in order to obtain results that are comparable across administrations. The equating process also mitigates differences in test difficulty between forms from the current and the previous year, which are built to be similar in difficulty and content (Kolen & Brennan, 1995). The items from previous administrations are called *anchor items*. In each grade and content area, each anchor set was a miniature version of the total test, which adequately represented the test content coverage in terms of item difficulty and the test blueprint. The Stocking and Lord (1983) procedure was used to place the estimated parameters on the scale from which the anchor items were drawn. This procedure estimates the linear transformation constants by minimizing the distance between the characteristic curves for the total test and anchor set. Field test items were calibrated together with *operational* items (the items that contributed to the student scores), and put in the operational scale of the Fall 2007 WKCE using the item parameters of the operational items as anchor items.

For additional information on the scaling methods and procedures used in the WKCE, readers can refer to Part 8 and Part 11 of the WKCE Technical Report from the Fall 2005 administration.

7.1.1 Calibration Models

The three-parameter logistic model defines a MC item in terms of three characteristics, or *item parameters*: the item difficulty (or its location on a scale of difficulty/ability), the item discrimination (or item differences on discrimination), and the level of guessing. The two-parameter partial credit model defines a CR item in terms of an item discrimination parameter and a difficulty parameter for each score point.

In the 3PL model (Lord & Novick, 1968; Lord, 1980), the probability that a student with scale score θ responds correctly to item *i* is:

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

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

The 2PPC model (Muraki, 1992; Yen, 1993) is a special case of Bock's (1972) nominal model. Bock's model states that the probability of an examinee with ability θ having a score at the *k*-th level of the *j*-th item is:

Copyright © 2008 by the Wisconsin Department of Public Instruction

$$P_{jk}(\theta) = P(x_j = k-1|\theta) = \frac{\exp Z_{jk}}{\sum_{i=1}^{m_j} \exp Z_{ji}}, \; k = 1,...m_j,$$

where $Z_{jk} = A_{jk}\theta + C_{jk}$.

For the special case of the 2PPC model used here, the following constraints were used:

$$A_{jk} = \alpha_j(k-1), \text{ and } C_{jk} = -\sum_{i=0}^{k-1} \gamma_{ji}, \text{ where } \gamma_{j0} = 0,$$

where α_j and γ_{ji} are parameters freely estimated from the data. The first constraint implies that higher item scores reflect higher ability levels and items can vary in their discriminations. The 2PPC model estimates a total of m_j independent item parameters; for each item there are m_j–1 independent γ_{ji} parameters and one α_j parameter.

The item calibration process is a process of estimating item parameters. Parameters are estimated in an *iterative* process using a computer software program called PARDUX (discussed below). The PARDUX program operates by estimating person parameters (ability) and item parameters (e.g., difficulty) through a series of iterations, until the change in parameter estimates between iterations is reduced to a given threshold.

7.1.2 Calibration Software

The IRT models and the student response data from the Fall 2007 administration were used together to estimate item parameters for each test. The IRT models were implemented using CTB's PARDUX software (Burket, 1991). PARDUX estimates parameters simultaneously for MC and CR items using marginal maximum likelihood procedures implemented with the expected maximum (EM) algorithm (Bock & Aitkin, 1981; Thissen, 1982). PARSCALE, MULTILOG, and BIGSTEPS are among the most widely known and used IRT programs. Extensive simulation studies and comparisons between PARDUX and MULTILOG (Thissen, 1990), a program widely used for research purposes, have shown that PARDUX provides precise parameter and ability estimates, and it performs more efficiently than MULTILOG (Fitzpatrick, 1991). Simulation studies have also compared PARDUX with PARSCALE (Muraki & Bock, 1991), and with BIGSTEPS (Wright & Linacre, 1992). Fitzpatrick and Julian (1996) found that PARDUX provided precise parameter and ability estimates, and performed more efficiently than the other programs. Extensive research with simulation data has also shown that the IRT procedures used here produce accurate vertical scaling (Yen & Burket, 1997).

7.2 Deriving Scale Scores in the WKCE

Simply put, a scale score in the WKCE can be interpreted as a highly probable estimate of a student's ability in a given content area. These scores are based on the student's responses to

Copyright © 2008 by the Wisconsin Department of Public Instruction

all of the items on a given test, and the scores account for the characteristics of the items that are in the test (such as how difficult the items are).

Scale scores in the WKCE are based on the theoretical models of the item response process described above and elaborated upon below. The basic idea represented by these models is that the probability of a correct response to a given item is a function of examinee ability, and the characteristics of the item, such as how difficult it is. Item Response Theory assumes that, generally speaking, we can expect that as examinee ability increases, the probability of a correct response to a given item also increases given certain conditions and assumptions. This description applies specifically to multiple-choice items, and constructed-response items are handled only slightly differently, with logic that is essentially the same.

Whether looking at an individual item, or at a group of items that make up a complete test, IRT uses probability models to describe the relationship between a student's ability and his/her observed scores. To review, the 3PL model described above is used to estimate the probability of a correct response for each of the multiple-choice items. The model is provided here because its components are reviewed in the following paragraphs.

$$P(u_i = 1 \mid \theta) = c_i + \frac{1 - c_i}{1 + e^{-1.7 a_i (\theta - b_i)}} \qquad (1)$$

In this model, θ denotes a measured ability (e.g., Language Arts ability), and u_i represents an observed score on a particular item. For MC items, the observed score u_i is either 0 or 1, indicating either an incorrect response or a correct response, respectively speaking. For a MC item, the probability model can be denoted as P (u_i=1|θ), i.e., P is an estimation of the probability that a student with an ability value θ would answer the item-*i* correctly.

The terms on the right hand side of the equation above (a_i, b_i, c_i) represent the parameters in this model: *discrimination*, *difficulty*, and a *pseudo-guessing factor*, as indicated above. Simply put, discrimination refers to how well an item sorts students by ability level, difficulty represents the difficulty of the item or its location in an ability continuum, and the pseudo-guessing factor represents the probability of a low ability student guessing the correct response.

Given any particular response pattern: $u_1 u_2 \cdots u_n$ on a test with some number of items (*n*-items), the "likelihood function" or the probability of a student with a given ability value (θ) would produce this particular response pattern is given by:

$$P(u_1 u_2 \cdots u_n \mid \theta) = \prod_{i=1}^{n} P(u_i \mid \theta) \qquad (2)$$

Copyright © 2008 by the Wisconsin Department of Public Instruction

The formula here indicates that the "estimated maximum likelihood" IRT item-pattern scoring method searches for that ability estimate (θ_0) which maximizes the probability function in (2) and it assigns an ability estimate (θ_0) as the test score for the student with the response pattern $u_1 u_2 \cdots u_n$. In other words, the scale score is the most likely or most probable estimate of student ability - produced in a context where item parameters are known, and based on all of the items in a given test.

As the text above and equation (1) indicate, the item-pattern scoring method takes into account not only a student's total raw score but also the psychometric characteristics of all items the student responded to, including the items the student responded to incorrectly.

Consider the following example: suppose six examinees in the fourth grade take a 30 item multiple-choice test in Language Arts. Suppose further that the properties, or parameters, of the items on that test are as follows:

Table A. Item Parameters for a Test

Item No.	Discrimination (a)	Location (b)	Guessing (c)
1	0.0341	318.75	0.16
2	0.0342	244.62	0.20
3	0.0234	257.56	0.20
4	0.0306	235.00	0.20
5	0.0125	342.39	0.17
6	0.0305	261.51	0.16
7	0.0316	296.93	0.19
8	0.0228	252.70	0.20
9	0.0383	266.28	0.20
10	0.0229	308.84	0.11
11	0.0536	259.00	0.21
12	0.0478	245.19	0.20
13	0.0418	276.25	0.28
14	0.0377	287.60	0.23
15	0.0177	316.08	0.24
16	0.0398	286.13	0.13
17	0.0523	290.65	0.26
18	0.0387	280.23	0.14
19	0.0329	315.71	0.21
20	0.0370	287.88	0.25
21	0.0387	280.25	0.18
22	0.0321	285.86	0.17
23	0.0219	302.52	0.13
24	0.0551	301.11	0.26
25	0.0165	324.24	0.19
26	0.0279	297.19	0.11
27	0.0423	296.06	0.28

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table A. Item Parameters for a Test, Cont'd.

Item No.	Discrimination (a)	Location (b)	Guessing (c)
28	0.0658	324.76	0.21
29	0.0488	281.56	0.32
30	0.0237	345.32	0.37

Now suppose the student response patterns for these six examinees are as follows, where 0 represents an incorrect response, and 1 represents a correct response to a multiple-choice item:

Table B. Item Response Pattern

Student	Response Pattern $(u_1 u_2 \cdots u_n)$	Raw Score	Item-pattern Score
Pam	100001100101000000000000000101	7	140
Craig	101010101010101010101010101010	15	246
Vicki	010101010101010101010101010101	15	266
Tom	001100110011001100110011001101	15	259
Evan	110011001100110011001100110010	15	265
Dan	111111111111111111111111011111	29	379

The first student, Pam, answered only seven out of 30 items correctly and obtained a scale score of 140, which is equal to the lowest point on the score range, called the "lowest obtainable scale score" or LOSS. The next four students each answered 15 out of 30 items correctly, but the response pattern of each of these students is different. The *raw score* of each of these students is 15. However, the maximum likelihood item-pattern scoring method produced a different scale score for each examinee. Scale scores were 246 for Craig, 266 for Vicki, 259 for Tom, and 265 for Evan. These scores can be accounted for by considering the pattern of the student responses on the test together with the properties (or parameters) of the items, as shown in Table A. By referring to Table A, the reader can observe that Vicki and Evan answered some difficult and highly discriminating items correctly, (whereas Craig and Tom did not), and the ability estimates for Vicki and Evan therefore include "credit" for answering these more difficult and more discriminating items, (whereas the estimates for Craig and Tom do not). The remaining student, Dan, scored 29 out of the 30 items correctly and obtained a scale score of 379, which is near the upper limit of the scale score range, called the "highest obtainable scale score" or HOSS.

Figure 1 below shows the probability of each ability level estimate (or scale score) for each of the examinees above. The total scale score range for Language Arts is plotted on the horizontal axis. As indicated by the two vertical lines in the plot, the lower and upper limits of the range are 140 and 420 respectively. The likelihood or probability of all possible ability

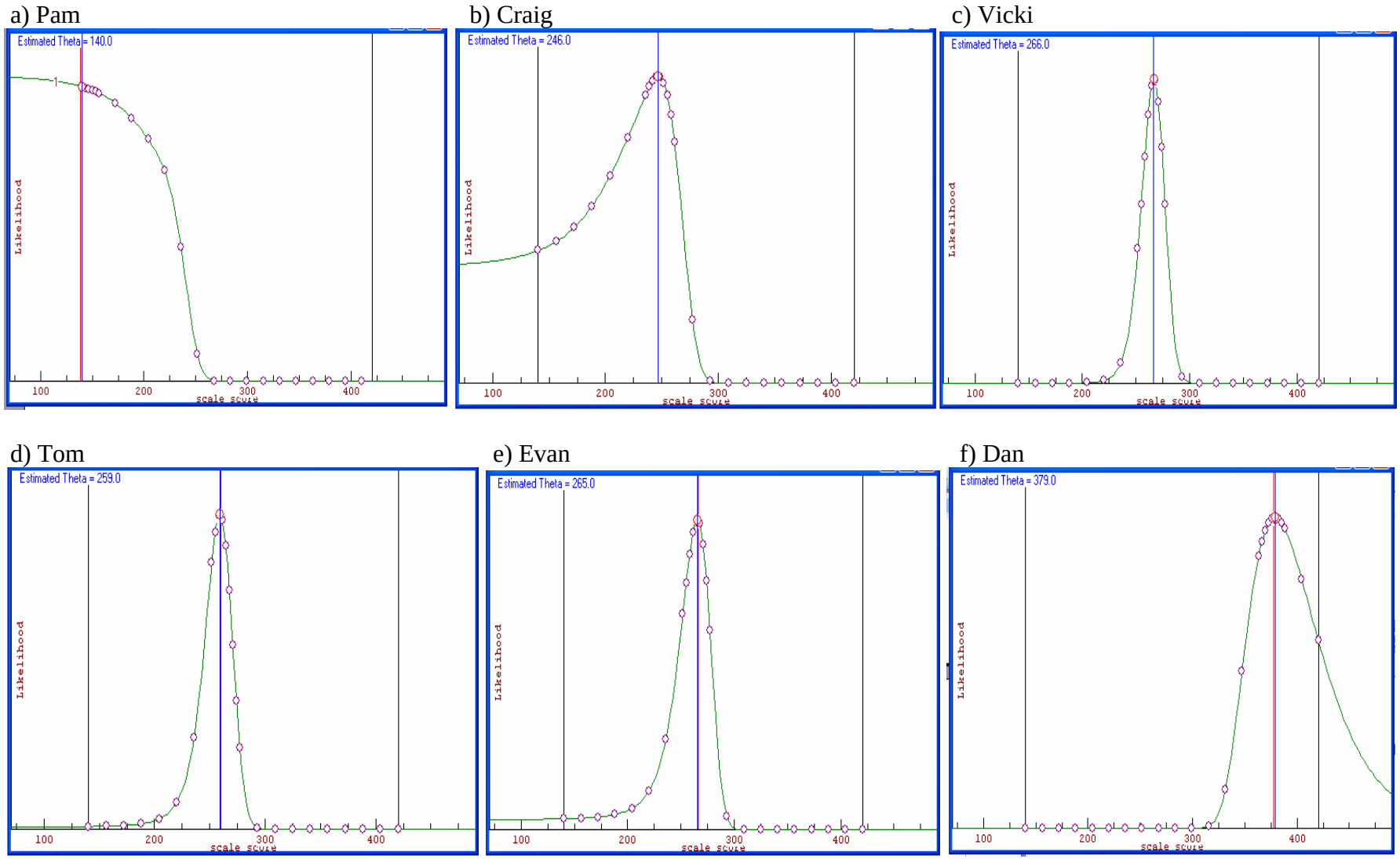
Copyright © 2008 by the Wisconsin Department of Public Instruction

estimates for each examinee is plotted on the vertical axis, and ranges from 0 to 1.0. The higher the likelihood, the more probable the estimate actually reflects the examinee's ability level.

As indicated above, scale scores are the most likely, or the "maximum likelihood," estimates of examinee ability. As can be observed for Vicki, Tom, and Evan, scores which are plus or minus only a few scale score points are markedly less likely estimates of examinee ability. In the case of Craig, as can be seen in the plot, it is highly unlikely that his true ability level is even a few scale score points higher, though some slightly lower estimates are almost as likely as the "maximum likelihood" estimate reported. In the case of Pam, some slightly higher scores were almost as likely as likely as the "maximum likelihood" estimate reported. Those scores which appear to be more likely than the reported score are outside of the scale score range of the test (below the LOSS). In the case of Dan, a few slightly higher or lower scores were almost as likely as the "maximum likelihood" estimate reported.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 1. “Likelihood Functions,” or the probability of each ability level estimate (or scale score)*



*The circular dots in the likelihood functions indicate that the software program used is searching for a maximum likelihood estimate (scale score) for the student.

Copyright © 2008 by the Wisconsin Department of Public Instruction

There are two IRT-based scoring methods generally used for large scale assessments: number-correct scoring and item-pattern scoring. Item-pattern scoring may be recommended over number-correct scoring for several reasons. Two reasons are pertinent for present purposes: accuracy and reliability.

Item-pattern scoring produces more accurate scores for individual students. Specifically, it produces a smaller standard error of measurement (SEM) across the scale score range for a given test compared to number-correct scoring. The smaller the SEM, the more confidence test users can have in the accuracy of the test results. The increase in accuracy is equivalent, on the average, to approximately a 15% to 20% increase in test length (Yen, 1984; Yen & Candell, 1991).

Second, reliability tends to be higher using item-pattern scoring, which means a) fewer items are needed to achieve a given level of reliability, and that b) a given test with a given number of items will have higher reliability than when using number correct scoring. Yen (1984) has demonstrated that an equivalent level of reliability for a 20-item test scored by the number-correct scoring method could be obtained with a 16- or 17-item test scored by the item-pattern scoring method.

Several supplements to this simplified outline of IRT are available. Introductory discussions of IRT can be found in Educational Measurement (Linn, 1989), or Chapter 11 in Introduction to Measurement Theory (Allen & Yen, 1979). More advanced discussions of partial credit models may be found in Muraki (1990, 1992), Yen (1993), and van der Linden and Hambleton (1997). The procedures applied here are similar to those followed in the development of the *TerraNova* test (CTB/McGraw-Hill, 1997), *TerraNova* 2nd Edition (CTB/McGraw-Hill, 2000), and the prior Wisconsin Knowledge and Concept Exam (WKCE) (CTB/McGraw-Hill, 1997-2004). For additional information on the technical details of the item-pattern scoring, readers can also refer to Yen & Candell (1991) and to *TerraNova* 2nd Edition (CTB/McGraw-Hill, 2000).

7.3 Calibration Results

Now that the calibration, equating, and scoring methods have been discussed, the following sections move on to describe the calibration results, in terms of the estimation of item parameters, model-to-data fit, and the standard error of measurement of the scale scores across content areas and grades.

7.3.1 IRT Item Parameters

All operational items “converged,” meaning parameters were successfully estimated for every item. There were three field-test (FT) items that did not converge, or for which parameters could not be estimated, during calibration. Those items were Mathematics grade 4 Form A item 56, Reading grade 6 Form A item 63, and Reading grade 8 Form B item 66.

Copyright © 2008 by the Wisconsin Department of Public Instruction

There were nine CR FT items (three in Reading and six in Mathematics) dropped from the calibration process based on recommendations from a joint item review between CTB and DPI. The list of these items was shared between CTB and DPI at that time. None of the non-converged items or those dropped during the item review will be used for any future testing without revising and re-field testing them.

The item parameters for both operational and field-test items derived from the procedures described above were provided to CTB's Content Development and to DPI for the purposes of replenishing the item pool for future test selection.

7.3.2 IRT Item Fit

A statistical procedure was used to identify items that did not fit the IRT model. Item model fit information was obtained for each item using a *Z*-statistic. The *Z*-statistic is a transformation of the chi-square (Q_1) statistic that takes into account differing numbers of score levels as well as sample size:

$$Z_j = \frac{(Q_{1j} - DF_j)}{\sqrt{2DF_j}}$$

where Q_{1j} is the item chi-square statistic, *j* is an item, and DF is the degrees of freedom for a given item *j*.

The *Z*-statistic is an index of the degree to which obtained proportions of students with each item score match the proportions that would be predicted by the estimated student ability and item parameters. These values, along with the associated chi-squares (Q_1), are computed for ten intervals corresponding to deciles of the ability distribution (Yen, 1984). Because the value of *Z* increases as the sample size increases, with other things being equal, the critical values for *Z* were established using the following equation (Yen, 1991a):

$$Z_{crit,j} = \frac{4N_j}{1500}$$

where $Z_{crit,j}$ is the critical value of *Z* for item *j*, and N_j is the number of students who responded to item *j*.

Table 7-1 presents items that were flagged for less than optimal fit based on the *Z* statistics above. The table specifies content area, grade level, test book form, the test booklet item number, the part name for CR items (A or B), the status of the item (operational or field test), item type (MC or CR), N size (the number of students), and *Z* and critical *Z*, as described above. To take an example from the table then, the Reading grade 3 field test CR item number 64 was flagged because its *Z* value of 16.87 is larger than the critical *Z* value of 6.8 based on a sample size of 2,555. These flagged items should be avoided from future selection unless there is a compelling reason that they should be included, such as meeting the test blueprint.

Copyright © 2008 by the Wisconsin Department of Public Instruction

The flagged items were further investigated using the item characteristic curves (ICC) in PARDUX. These ICCs simultaneously plot the characteristics of an item (e.g., item difficulty, item discrimination, the level of guessing) based on both the expectations of the IRT model, and based on the actual student responses. When the expected characteristics and the observed characteristics are sufficiently similar, the model "fits" the data. Where differences between what the model expects and what is actually observed reach certain thresholds, "misfit" is said to exist.

The flagging of an item for less than optimal fit is one of many criteria for providing item information to content experts for future test selection. The main issue in item fit is where the misfit happens. Misfit can readily occur where there are too few students. If the misfit happens around the lower or higher ability range, where there are not many students, this could be a smaller issue. However, if the misfit happens around the middle of the ability range, where there are many students, this could provide cause for more concern about the item, and possibly lead to the item being dropped from the test. For CR items, there are, in general, a small number of students at the lower and higher score levels, and with small sample sizes, misfit is thereby easily introduced. In a similar manner, misfit for MC items often happens at the lower ability range or at the higher ability range, where there are fewer students.

7.3.3 Standard Error of Measurement

The reliability of a reported test score can be characterized by the standard error associated with the score. An observed score should not be regarded as an absolute value, but as a point within a range that with a certain degree of probability includes a student's true score. The standard error of measurement (SEM) is used to obtain a range within which a student's true score is likely to fall, that is, with a certain degree of probability. It is expected that 68% of the time a student's score obtained from a single testing would fall within one SEM of that student's true score and that 95% of the time the obtained score would fall within two standard errors of true score.

Standard errors of measurement (SEM) for the Fall 2007 WKCE scale scores, obtained from item-pattern scoring, are displayed graphically for each grade and content area in Figures 7-1 through 7-5. Each curve is plotted as a function of the scale scores. These figures show the scale score range within which measurement is most accurate. The figures also show that extreme scale scores have more measurement error than moderate scores. The forms lose accuracy of measurement for scale scores near the high or low extremes because there are fewer students at these score ranges. Note that for convenience, the lower and upper limits of the scale, referred to as the lowest and highest obtainable scale score (LOSS and HOSS), were used as the starting scale score and the last scale score here. LOSS and HOSS are further discussed in the next section.

Because of the nature of item-pattern scoring, a scoring table showing a simple, direct conversion from raw to scale score cannot be generated for the 2007 WKCE. However, scoring tables showing a rough relationship among raw score, scale score, and standard error of measurement (SEM) can be produced, and that data is presented in Tables 7-2 through 7-24.

Copyright © 2008 by the Wisconsin Department of Public Instruction

7.3.4 LOSS and HOSS

As has been established, a scale score is a “maximum likelihood” ability estimate. The maximum likelihood procedure cannot produce scale score estimates for students with perfect scores or scores below the level expected by guessing. Although maximum likelihood estimates are available for students with extreme scores other than zero or perfect, these estimates generally have large standard errors of measurement. Therefore, scores are established for these extreme highs and lows based on a rational but necessarily non-maximum likelihood procedure. These values, which are set separately by grade, are called the lowest obtainable scale score (LOSS) and the highest obtainable scale score (HOSS).

Table 7-25 shows the number and percent of students at the lowest obtainable scale score (LOSS) and the highest obtainable scale score (HOSS). In general, there should not be many students clustered at the LOSS or HOSS. An accumulation of a higher proportion of students in the LOSS or HOSS may indicate a floor or ceiling effect. In grade 4 Social Studies, 2.10% of students were at the HOSS. This percentage of students at the HOSS for grade 4 indicates that the 2008 WKCE may need some more difficult items for grade 4 in order to reduce the proportion of students scoring at the HOSS.

In all other content areas and grades, the percentages at LOSS and HOSS were relatively small and showed no indication of floor or ceiling effects. For Reading, the percent at the LOSS ranged from 0.86% to 1.43% across grades. The percent at HOSS for Reading ranged from 0.00% to 0.12% across grades. For Mathematics, the percent at the LOSS ranged from 0.14% to 1.41% and the percent at the HOSS ranged from 0.05% to 0.29%. The percent at the LOSS in Language Arts ranged from 0.37% to 1.12%, and the percent at the HOSS ranged from 0.01% to 1.50%. For Social Studies, the percent at the LOSS ranged from 0.28% to 0.54%, and the percent at the HOSS ranged from 0.19% to 2.10%, as indicated above. In Science, the percent at the LOSS ranged from 0.22% to 0.69%, and the percent at the HOSS ranged from 0.36% to 1.26%.

7.3.5 Test Characteristic Curves

Test characteristic curves are provided in Figures 7-6 to 7-10. These curves model the relationship between student ability and expected scoring outcomes at the test level. By following the plotted line for any grade level and content area, one can observe the estimated scoring outcome (the estimated proportion of the maximum correct score) plotted as a function of examinee ability, based on the IRT models, methods, and scaling processes described above. The vertical relationship across grade levels that can be observed the in TCCs reflects the typical growth pattern: as grade level increases, ability levels are also expected to increase, across the ability range.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 8: Test Results

Part 8 presents a classical item analysis, and several summaries of the scoring results of the Fall 2007 administration. The summary results discussed are raw scores, scale scores, performance level results, and scores based on each of the content standards within each content area called *standardized performance indicator* (SPI) scores. Combined, the classical item analysis and the four forms of scores provided offer the reader several vantage points from which to understand and evaluate the WKCE testing program and the results derived. The AERA/APA/NCME standards addressed in Part 8 include: 1.5, 3.18, 4.3, 4.5, 4.6, 4.7, 4.19, 7.1, 7.10, 13.15, and 13.19.

8.1 Classical Item Analysis: Item Level Statistics

Three statistics are frequently used in item analysis at the item level: the proportion correct (*p*-value), a correlation index of item discrimination, and omit rates.

The *p*-value is an indication of the difficulty of an item. The *p*-value for a MC item represents the proportion of students who answered the item correctly. If all students answered a given MC item correctly, its *p*-value would be 1.0. If only 30% of students answered the question correctly, the *p*-value would be 0.30. So, the lower the *p*-value, the more difficult the item is. Item *p*-value is a good indication of difficulty; it takes student performance into account, and it makes comparing items in terms of a common statistic very simple. A test made up of items well distributed across the range of item difficulty levels is desirable, because it supports assessment of students at all ability levels.

The *p*-value for a CR item represents the mean proportion of possible raw score points that students actually obtained for the item. A *p*-value of 0.33 for a given CR item would indicate that, on average, students obtained one-third of the possible points for the item. If the *p*-value were 0.75, this would indicate a much easier item, where, on average, students scored 75% of the maximum possible points for the item. As such, for CR items as well, *p*-value indicates difficulty and the lower the *p*-value, the more difficult the item is.

The point-biserial correlation is a correlation between the score on a particular item and the total score on the test (which is calculated by summing up the scores on the remaining items). The value of the point-biserial correlation ranges from -1.0 to +1.0. A large positive point-biserial value (say 0.80) indicates that students with high scores in the test are also getting the item right and students with low scores on the overall test are getting the item wrong. A low point-biserial value (say 0.10) implies that students who answered the item correctly tend to do poorly on the overall test and vice versa. The point-biserial correlation is often described as an item-to-total score correlation.

Copyright © 2008 by the Wisconsin Department of Public Instruction

The point-biserial correlation is only appropriate for dichotomous level data (yes/no, right/wrong), so for the CR items, a Pearson correlation between the item score and the total score on the test was computed. This Pearson correlation can be interpreted the same way.[*]

The point-biserial and Pearson correlation values for operational items (items that counted in student scores) were based on operational items only. The point-biserial values for field-test items were based on the operational and field-test items per form.

A formula similar to the point-biserial correlation was applied to compute the correlation between each distractor and the total score. In general, negative correlations are expected for all distractors when an item is good. However, a small positive correlation for a distractor can often mean that the distractor is very attractive for low-performing students.

The omit rates refer to the percentage of students that did not respond to an item. A high omit rate can indicate an especially difficult item, or if located near the end of the test, it can indicate a speeded test, where students lack time.

For the Fall 2007 administration, items were flagged for further investigation when certain thresholds were reached. The *p*-value was flagged when the statistic fell below 0.30 for MC items. This would indicate an especially difficult item, where fewer than 30% of students obtained the correct answer. The point-biserial correlation was flagged where the coefficient was below 0.15. A low point-biserial value may indicates that the item is not performing appropriately as expected, for example, the item was not discriminating between students of different ability levels, or that most students were guessing on the item. The omit-rate was flagged when it was above 5%. Distractors were flagged when they had a positive correlation with the total test score.

Tables 8-1 through 8-23 present an item-level item analysis for all grades and contents. The flagged items listed across all grades and contents are also provided in a single summary table, Table 8-24. In addition to providing results based on the statistics above, the item analysis tables differentiate between those items used for scoring the WKCE (operational items) and those used to replenish the item pool (field test items). The tables also provide the test book form and test book item number, which can be used to understand the location of test items as students actually encountered them in test booklets. The item analysis tables also indicate item type (MC or CR). Readers should note that because item analysis was performed by form, and FT items can occur in more than one form, the same FT item can appear as flagged more than once. Operational items appear only once.

Readers may also wish to note that the results presented below may differ slightly from testing results presented on the Wisconsin Department of Public Instruction website due to slight differences in the way the two sources define which students should be included or excluded

---

[*] For both the point-biserial and the Pearson correlation, the studied item was excluded from the computation of the total score so as to not artificially inflate the correlation statistic. This effect would be most noticeable for CR items worth several points.

Copyright © 2008 by the Wisconsin Department of Public Instruction

from the summary results. Official final results are based on application of detailed school-level rules for including or excluding students, based on factors such as whether the student moved, or how long they were in one school or another over the course of the year.

Detailed item analysis results for each grade and content area are provided below. There were 1,130 items in the Fall 2007 WKCE assessments. Readers are encouraged to view the summary table (Table 8-24) and to consult the brief summary provided below the detailed results presented here in order to obtain an overall view of the item analysis results. As explained in the summary and detailed below, overall, a small number of items was flagged, and field test items were flagged more often than operational items. Both of these findings may be expected in large scale assessments such as the WKCE.

Reading

Speaking now with reference to operational items and all flagging criteria, in Reading grade 3, three items were flagged. One item was flagged for a distractor, one was flagged for the omit rate, and one was flagged for *p*-value. In grade 4, two items were flagged, both for a distractor. In grade 5, five items were flagged, four for a positive point-biserial for a distractor and one for the point-biserial for the correct choice being less than 0.15. In grade 6, two items were flagged for a positive point-biserial for a distractor. In grade 7, three items were flagged for a positive point-biserial for a distractor, of these three items one was also flagged for a point-biserial for the correct choice less than 0.15. In grade 8, six items were flagged, five of them for a positive point-biserial for a distractor and one for a point-biserial for the correct choice less than 0.15. One of the items flagged for a positive point-biserial for a distractor was also flagged for a point-biserial for the correct choice less than 0.15. For grade 10, two items were flagged, both for a positive point-biserial for a distractor.

Among field test items, in Reading grade 3, looking across all criteria, five items were flagged. Two items were flagged for both the point-biserial for the correct choice being less than 0.15 and a positive point-biserial for a distractor, one item was flagged for a positive point-biserial for a distractor only, and two items were flagged for *p*-value less than 0.30. In grade 4, six items were flagged. One item was flagged for a point-biserial for the correct choice less than 0.15, a positive point-biserial for a distracter, and p-value less than 0.30. All other items were flagged for a positive point-biserial for a distractor. In grade 5, three items were flagged: two for a positive point-biserial for a distractor, and one for the point-biserial for the correct choice less than 0.15, and a positive point-biserial for a distracter. In grade 6, six items were flagged. One was flagged for the point-biserial less than 0.15 and a positive point-biserial for a distractor, one was flagged for the point-biserial correlation less than 0.15 only, three items were flagged for a positive point-biserial for a distractor only, and one was flagged for *p*-value less than 0.30. In grade 7, six items were flagged. Two items were flagged for the point-biserial correlation less than 0.15, a positive point-biserial for a distractor, and *p*-value less than 0.30. Two items were flagged for *p*-value less than 0.30 only. One item was flagged for both the point-biserial correlation less than 0.15 and a positive point-biserial for a distractor. One item was also flagged for a positive point-biserial correlation for a distractor. In grade 8, five FT items were flagged. All of these items were flagged for a positive point-biserial for a distractor, but one of these

Copyright © 2008 by the Wisconsin Department of Public Instruction

items was also flagged for a point-biserial correlation less than 0.15. In grade 10, there were no FT items.

Mathematics

Among operational items, in grade 3, four items were flagged, all for *p*-value. In grade 4, four items were flagged: one for a point-biserial for the correct choice less than 0.15, two for *p*-value less than 0.30, and one for an omit rate larger than 5%. Three items in grade 5 were flagged: one for *p*-value less than 0.30 and two for a positive point-biserial for a distractor. There were four items flagged in grade 6: two for a point-biserial correlation less than 0.15, one for a positive point-biserial for a distractor, and one for *p*-value less than 0.30. Five operational items were flagged in grade 7, of which, three were flagged for a positive point-biserial for a distractor, one was flagged for *p*-value less than 0.30, and one was flagged for a point-biserial correlation less than 0.15 and a positive point-biserial for a distractor. In grade 8, six items were flagged. Four of these items were flagged for *p*-value less than 0.30 and two were flagged for a positive point-biserial for a distractor. In grade 10, one item was flagged for *p*-value less than 0.30 and two for a positive point-biserial for a distractor.

Among FT items, in grade 3, five items were flagged. One of these items was flagged for the point-biserial correlation, a distractor, and for *p*-value. One item was flagged for an omit rate larger than 5%, one item was flagged for a point-biserial less than 0.15, a positive point-biserial for a distracter, and for *p*-value less than 0.30. One item was flagged for a positive point-biserial for a distractor and p-value less than 0.30. One item was flagged for a positive point-biserial for a distractor only, and one item was flagged for p-value less than 0.30 only. In grade 4, five FT items were flagged: two for a positive point-biserial for a distractor and p-value less than 0.30, one for a point-biserial for the correct choice less than 0.15, one for a point-biserial for the correct choice less than 0.15, a positive point-biserial for a distractor, and p-value less than 0.30. One item was also flagged for a positive point-biserial for a distractor only. In grade 5, two items were flagged for a point-biserial correlation less than 0.15, a positive point-biserial for a distractor, and for *p*-value less than 0.30. One item was flagged for a point-biserial less than 0.15 and p-value less than 0.30, and one item was flagged for p-value less than 0.30 only. In grade 6, five FT items were flagged. Two were flagged for both a point-biserial correlation less than 0.15 and for a positive point-biserial for a distractor, one item was flagged for a positive point-biserial for a distractor and *p*-value less than 0.30, and one item was flagged for a point-biserial correlation less than 0.15, a positive point-biserial for a distractor, and *p*-value less than 0.30. One item was also flagged for p-value less than 0.30 only. In grade 7, 10 FT items were flagged: one for a point-biserial for a distractor and *p*-value less than 0.30, one for the point-biserial correlation less than 0.15, a positive point-biserial for a distractor, and *p*-value less than 0.30, two were flagged for *p*-value less than 0.30 only, two were flagged for the point-biserial correlation less than 0.15 and for a positive point-biserial for a distractor, two item were flagged for a positive point-biserial for a distractor only, and two for a point-biserial for the correct choice less than 0.15. In grade 8, two items were flagged for a positive point-biserial for a distractor and *p*-value less than 0.30, three items were flagged for the point-biserial correlation less than 0.15, a positive point-biserial for a distractor, and *p*-value less than 0.30, one item was flagged for a positive point-biserial for a distractor only, one item was flagged for a point-biserial

Copyright © 2008 by the Wisconsin Department of Public Instruction

less than 0.15 and a positive point-biserial for a distractor, and three items were flagged for *p*-value less than 0.30 only. There were no FT items in grade 10.

Language Arts

In Language Arts, speaking with reference to operational items, in grade 4, two operational items were flagged, both for a positive point-biserial for a distractor, of which, one was also flagged for a point-biserial less than 0.15. In grade 8, two items were flagged: one for a point-biserial less than 0.15 and one for a point-biserial less than 0.15 and a positive point-biserial for a distractor. In grade 10, three items were flagged, all for a positive point-biserial for a distractor. One was also flagged for p-value less than 0.30. There were no FT items in Language Arts.

Social Studies

In Social Studies, looking at operational items, one operational item was flagged in grade 4 and 8, and both of were flagged for a positive point-biserial for a distractor. In grade 10, three operational items were flagged. One of these items was flagged for the point-biserial correlation less than 0.15, a positive point-biserial for a distractor, and for *p*-value less than 0.30. One item was flagged for a positive point-biserial for a distractor and p-value less than 0.30. One item was flagged for a positive point-biserial for a distractor only. There were no FT items in Social Studies.

Science

Among operational items, in Science, three items in grade 4 were flagged. Two were flagged for a point-biserial less than 0.15 and a positive point-biserial for a distractor, and one was flagged for a point-biserial less than 0.15 only. In grade 8, three items were flagged: one for an omit rate higher than 5%, one for a positive point-biserial for a distractor and an omit rate higher than 5%, and one for a positive point-biserial for a distractor only. In grade 10, three items were flagged: one for the point-biserial correlation less than 0.15 and a positive point-biserial for a distractor, and two for a positive point-biserial for a distractor only. There were no FT items in Science.

Classical Item Analysis Summary

In summary, out of a total of 1,130 items in the 2007 WKCE assessments, a small proportion was flagged. This summary view is apparent in Table 8-24, which, as indicated above, summarizes all flagging for both operational and field-test items. As is evident in the table, flagging was more common among field-test items than among operational items.

Copyright © 2008 by the Wisconsin Department of Public Instruction

The item analysis also indicated that the *p*-values of the items in the operational tests were well distributed throughout the range of difficulty levels, with point-biserial correlations reasonably high for most items. Field-test items were flagged for out of range for *p*-values and point-biserial correlations more often than operational items.

Readers may wish to note that the flagging of an item is just one aspect of a total evaluation of an item. In the case of *p*-value for example, items flagged for high *p*-value (an especially easy item) or low *p*-value (an especially difficult item) may be necessary in order to measure student ability at the extreme high and low ranges, and to represent a given content standard. Flagged items do not necessarily indicate a challenge to test validity; rather they highlight notable item characteristics that should be taken into account with care when evaluating the items in a test, either individually or as a whole.

These classical statistics were provided to CTB's Content Development and to DPI for their review and evaluation, which may inform decisions regarding including or not including particular items in the item pool for future test selections.

8.1.1 Speededness

The degree to which a test is speeded can be evaluated by examining the percentage of students who fail to respond to the last items on a test, or the last items in a timed section. WKCE tests were established to have a sufficient amount of time for all students. Preliminary research on omit rates did not indicate speededness was an issue in WKCE tests.

8.2 Raw Score Results

A raw score corresponds to the number of items correctly responded to by a student in a test. Some audiences prefer raw scores over standardized scores for ease of understanding. For example, it is easy to explain to parents that their child correctly scored 50 items out of 60 items in a test rather than that he/she received a score of 479. However, in order to be useful, a raw score has to be interpreted with reference to the number of items in the test and the maximum raw score possible. In the raw score tables drawn upon below, readers will find appropriate references to the test configuration, however, readers may also find it useful to refer to Table 3-1, which reflects the test structure for all grades and contents.

The raw score table (Table 8-25) includes the number of students, the average number of items correct, test difficulty, the standard deviation (SD) of raw scores, skewness of the raw score distribution, kurtosis, the minimum observed score, the maximum observed score, maximum possible score, reliability (Cronbach's alpha), and the standard error of measurement for raw scores. Summary statistics for raw scores are based on all students that took the Fall 2007 assessment.

An equivalent table is also provided (Table 8-26) based on the sample of students used for calibration, for comparison and reference. Readers will recall that the calibration sample was

Copyright © 2008 by the Wisconsin Department of Public Instruction

determined to be representative of the entire student population based on both the demographic composition of the sample (as has been discussed in detail) and the performance profile. The raw score information provided here is part of the documentation provided in support of the validity of the WKCE. The similarity of the performance profiles of the sample data and the census data is discussed in more detail in the next section, which addresses scale scores.

In terms of the measurements applied in the raw score table, note first that the mean number of items correct, or the mean raw score, should be understood by grade and content area, and specifically in the context of the maximum possible score points, which varies by grade and content area. In Mathematics for example, the maximum possible raw score ranges from 57 to 62 and in Reading it ranges from 56 to 60.

Raw score results include an index of relative difficulty of the tests. Interested readers can refer to previous tests for comparison. Test difficulty is computed as mean raw score divided by maximum possible score points, and ranges from 0 to 1.0. A larger test difficulty statistic indicates a mean raw score which is closer to the maximum possible score, and it indicates, therefore, an easier test. A smaller test difficulty statistic indicates a mean raw score which is further from the maximum possible score, and it indicates therefore, a more difficult test. Consider an example: the test difficulty statistic would be 0.90 if a mean score of 45 were obtained on a test with a maximum possible score of 50. This could be considered an easier test. On the other hand, test difficulty would be 0.50 if a mean raw score of 25 were obtained on the same test. This could then be considered a more difficult test. In Reading grade 3, the test difficulty statistic (0.66) was obtained by taking the mean raw score of 39.48 and dividing it by 60.

Like the standard deviation (SD), skewness and kurtosis also describe the shape of a distribution. When a distribution is perfectly normal, skewness is 0. A negative skew indicates the presence of some extreme low scores and (because the mean is sensitive to extreme scores) a corresponding increase in the number of student scores above the mean. A positive skew indicates a distribution with some extreme high scores and a corresponding increase in the number of scores below the mean. Kurtosis describes a distribution in terms of its degree of peakedness. When a distribution is perfectly normal, Kurtosis is 0. A negative Kurtosis statistic indicates a distribution which is flatter than a perfectly normal curve, and a positive Kurtosis statistic indicates a distribution which is more peaked than a perfectly normal curve.

Where any student failed all items on the test, the minimum observed score is 0. Where any student obtained the full scores for all items, the maximum observed score is equal to the maximum number of points possible on the test. For example, as can be observed in Table 8-25, in Reading grade 3, there is at least one student who failed all items, but no student obtained a perfect raw score (the maximum obtained score was 59, and the maximum possible score was 60).

A reliable test is one with high reliability as represented by statistics such as Cronbach's alpha and a low standard error of measurement (SEM). Test length (number of items and score points) is one of the important factors that influence reliability statistics and SEM. These concepts are described further in Part 10: Reliability. However, for present purposes, the reader

Copyright © 2008 by the Wisconsin Department of Public Instruction

should note that measurement error is associated with every test score. A student's true score is the hypothetical average score that would result if the test could be administered repeatedly without the effects of practice or fatigue. Obtained scores should not be regarded as absolute, but as one point within a range that, with a certain degree of probability, includes a student's true score. Results are summarized and discussed below. The raw score results for the five NCLB groupings are drawn from tables 8-27 to 8-35.

Reading

Test difficulty in Reading ranged from 0.64 to 0.68 across grades, indicating that each test was of approximately equal difficulty for each grade. Standard deviations, in all cases, indicated a moderate degree of dispersion in scores. For every grade, alpha was high (0.90 to 0.93), and SEM ranged from 3.04 to 3.22. In terms of gender, female students, as a group, had a slightly higher score than male students, as a group, in every grade. The difference was small, but consistent. There was also slightly less variation, or dispersion, in female scores, as indicated by the smaller standard deviation in each grade. Grouped and compared by ethnicity, mean scores varied more widely. In every grade level, White students had the highest score, followed by Asian students, American Indian students, Hispanic students, and African American students. Taken as a group, in every grade, those students who were Not Economically Disadvantaged scored higher than students who were Economically Disadvantaged. The mean Reading score was higher among Not Disabled students than among Disabled students in every grade. Students who were Proficient in English had a higher mean score, as a group, than those students who were Not Proficient in English.

Mathematics

Looking across grades, test difficulty ranged from 0.57 in grades 8 and 10, to 0.70 and 0.71 in grades 3 and 4. Alpha was high for every grade (0.90 to 0.93) except for grade 4 where alpha was marginally lower, at 0.89. SEM ranged from 2.94 to 3.39. Split and compared by gender, male students had a higher mean raw score in every grade, but that difference in mean score was very small, and the two mean scores were nearly equivalent in each grade. Scores varied more widely by ethnicity, where, in every grade, White students, as a group, had the highest mean score, followed by Asian students, American Indian students, Hispanic students, and African American students. Grouped and compared by SES, the mean score was higher among those Not Economically Disadvantaged than among those who were Economically Disadvantaged. The mean score among those students who were Not Disabled was higher than the mean score among students who were Disabled. Those students who were Proficient in English had a higher mean score, as a group, than students who were Not Proficient in English, in every grade level.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Language Arts

There were some small differences in test difficulty across grades in Language Arts. Test difficulty ranged from 0.62 to 0.72. Note that there were 30 possible points in grades 4 and 8, but 39 in grade 10, because in grade 10, a 9-point writing prompt is part of the Language Arts score. Looking across all grades, alpha ranged from 0.82 to 0.85. SEM ranged from 2.15 to 2.52. As a group, female students scored higher than male students, by a small margin, in each grade. Taken as a group, White students had the highest mean score, followed by Asian students, American Indian students, Hispanic students, and African American students. In every grade, the mean score was higher among those students who were Not Economically Disadvantaged than among students who were Economically Disadvantaged. Students who were Not Disabled had a higher mean raw score than students who were Disabled, in each grade. In every grade, the mean score was higher among students who were Proficient in English than among those students who were Not Proficient.

Social Studies

In Social Studies, there were some differences in test difficulty by grade level. Test difficulty ranged from 0.64 in grade 10 to 0.76 in grade 4. Across all grades, alpha ranged from 0.87 to 0.90 and SEM ranged from 2.28 to 2.91. Mean raw scores in Social Studies showed only a minute difference between genders. The differences were less than one raw score point, or virtually equivalent. There were larger differences in mean scores by ethnicity. In every tested grade level, as a group, White students had the highest mean score, followed by Asian students, American Indian students, Hispanic students, and African American students. The mean score was higher among students who were Not Economically Disadvantaged than among students who were Economically Disadvantaged, in every grade. The mean raw score was also higher among students who were Not Disabled than among students who were Disabled. Taken as a whole, students who were Proficient in English had a higher mean score in every grade than students who were Not Proficient in English.

Science

The level of test difficulty in Science ranged from 0.62 in grade 10, to 0.74 in grade 8. Across all grades, alpha ranged from 0.85 to 0.88 and SEM ranged from 2.45 to 3.03. Science scores reflected small differences in mean scores by gender. In each grade, male students, as a group, had a higher mean score than female students. However, the differences in grades 4 and 8 were very small, less than one raw score point. There were larger differences in Science scores by ethnicity. As has been observed in other content areas, White students had the highest mean score, followed by Asian students, American Indian students, Hispanic students, and African American students. Science scores, like scores in other content areas, also differed by socioeconomic status. In every grade, the mean Science score was higher among students who were Not Economically Disadvantaged than among students who were Economically Disadvantaged. As has also been observed in other content areas, students who were Not Disabled scored higher, as a group, than students who were Disabled, in every grade. In every

Copyright © 2008 by the Wisconsin Department of Public Instruction

grade, the mean score in Science was higher among students who were Proficient in English than among students who were Not Proficient in English, as has also been observed in other content areas.

In summary, the raw scores and observed proportion of maximum raw scores as an indictor of the test difficulty indicate that the WKCE tests across grades and content areas were moderate in difficulty. The observed proportion of the maximum raw score ranged from a low of 0.57 in Mathematics grade 8 and 10 (relatively difficult tests) to a high of 0.76 in Social Studies in grade 4, a relatively easier test.

8.3 Summary Statistics for Scale Scores

As indicated previously, in addition to reporting raw score statistics, the WKCE program also reports scale scores based on the entire set of operational items in the test. The scale score of a student in a given content area represents the student's level of achievement in that content area. Note that higher scale scores indicate higher levels of achievement, and lower scale scores indicate lower levels of achievement.

As indicated in Part 6, the WKCE program used sample data for calibration purposes. As discussed there, the sample data were determined to be sufficiently representative of WI students based on both demographic and performance characteristics. The demographic comparisons between the sample data and the census data were provided in Part 6. The comparison based on scale scores is provided here, alongside the other testing results. In order to demonstrate the degree of similarity in the performance profiles of the sample data and the census data, Table 8-36 shows the means and standard deviations (SD) of scale scores from the 13 calibration district data, the census data, and the difference between the two, per grade and content area. As can be seen in the table, the difference in scale scores was small in all grades and content areas. Table 8-37 provides additional measures of the sample student data, such as N count, skewness, and kurtosis which can be compared to the full census table, which is Table 8-38.

Table 8-38 shows the mean scale score, standard deviation, skewness, kurtosis, minimum observed scale score (min), and maximum observed scale score (max) for the total population, plus lowest and highest obtainable scale scores (LOSS and HOSS) for all contents and grades. As reflected in the table, mean scale scores rise as grade level increases. This is the (intended) result of vertical scaling. The descriptive statistics of scale scores for the overall test and for the five NCLB groupings are described below. The LOSS and HOSS scores were discussed in Part 7, and they are included here to identify the lower and upper limits of the scale score range. The scale score results for the five NCLB groupings are drawn from tables 8-39 to 8-47.

Reading

Scale scores in Reading were moderately dispersed around the mean in every grade, as indicated by the standard deviation. The fact that the standard deviations generally increase by grade level here should be understood within the context of the fact that the range between the

Copyright © 2008 by the Wisconsin Department of Public Instruction

LOSS and HOSS also generally increases as grade level increases. The Minimum and Maximum columns indicate the minimum and maximum scores obtained. In every grade, at least one student scored at the LOSS and at least one student scored at the HOSS. As was observed in the context of raw scores, the scale scores in Reading also showed some consistent scoring patterns within the five NCLB groupings. For example, in every grade, female students had a slightly higher mean score than male students. In every grade, White students, as a group, had the highest mean score, followed by Asian students, American Indian students, Hispanic students and African American students. Taken as a group, those students who were Not Economically Disadvantaged scored higher than students who were Economically Disadvantaged. At every grade level, students who were Not Disabled had a higher mean Reading score than students who were Disabled. Taken as a whole, in every grade level, students who were Not Proficient in English had a lower mean Reading score, as a group, than students who were Proficient.

Mathematics

The standard deviations in Mathematics scores reflect a moderate degree of dispersion in every grade. At least one student scored at the LOSS and at the HOSS in each grade. The mean score was higher among male students than among female students in every grade, but these differences were very small. There were consistent differences in mean scores by ethnicity in Mathematics. White students, as a group, had the highest mean score, followed by Asian students, American Indian students, Hispanic students, and African American students. Socio-economic differences occurred here as well. In every grade level, the mean score among Economically Disadvantaged students was lower than the mean score for students who were Not Economically Disadvantaged. As was observed in other content areas, in every grade, those students who were Not Disabled, scored higher than students who were Disabled. As was observed in other content areas, in every grade level, students who were Not Proficient in English had a lower mean score than students who were Proficient in English.

Language Arts

In each grade level, the scale scores in Language Arts were moderately dispersed around the mean, as reflected in the standard deviation. At least one student scored at the LOSS and at least one student scored at the HOSS in each grade. In every grade, the mean score among female students was higher than that among male students. By ethnicity, White students had the highest mean score, followed by Asian students, American Indian Students, Hispanic students and African American students. In every grade level, those students who were Not Economically Disadvantaged, scored higher, as a group, than students who were Economically Disadvantaged. In Language Arts, as in other content areas, students who were Not Disabled had a higher mean score than students who were Disabled. Also as was observed in other content areas, students who were Proficient in English had a higher mean score, as a group, than those students who were Not Proficient in English, in every grade.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Social Studies

The distributions of scale scores for Social Studies reflect a moderate degree of dispersion around the mean in each grade level, as reflected in the standard deviation. In every grade, at least one student scored at the LOSS and at least one student scored at the HOSS. While readers will recall that some gender differences have been observed in the scale score results for other content areas, in Social Studies, the differences in mean scores by gender were only approximately a single scale score point, which is a very small difference, and this suggests performance levels that were virtually equivalent for each gender. However, larger differences in scale scores by ethnicity which were observed in other content areas occurred in Social Studies as well: White students had the highest mean score, followed by Asian students, American Indian Students, Hispanic students, and African American students. In Social Studies, as elsewhere, in every grade, students who were Not Economically Disadvantaged scored higher, as a group, than students who were Economically Disadvantaged. Those students who were Not Disabled had a higher mean score than students who were Disabled, in every grade. In every grade, students who were Not Proficient in English had a lower mean score than students who were Proficient.

Science

The standard deviation in scale scores for Science indicates a moderate degree of dispersion in every grade. In each grade level, at least one student scored at the LOSS and at least one student scored at the HOSS. In Science, male students had a higher mean score than female students in each grade, but that difference was very small in grade 4. The difference was also small in grade 8, though less so, and in grade 10, the difference grew to approximately seven scale score points. As was observed in other content areas, when students are grouped and compared by ethnicity, White students had the highest mean score, followed by Asian students, American Indian Students, Hispanic students, and African American students. As was also observed in other content areas, in every grade, those students who were Not Economically Disadvantaged scored higher, as a group, than students who were Economically Disadvantaged. In Science, as in other content areas, those students who were Not Disabled had a higher mean score than students who were Disabled, in every grade. As was also observed in other content areas, students who were Not Proficient in English scored lower, as a group, than students who were Proficient in English.

8.4 Cut Scores and Performance Level Classifications

When a test is used to make categorical decisions, such as pass/fail, the scores used to make that decision are called “cut scores.” Standard 4.19 of the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999) indicates that “[w]hen proposed score interpretation involves one or more cut scores, the rationale and procedures used for establishing cut scores should be clearly documented” (p. 59).

Copyright © 2008 by the Wisconsin Department of Public Instruction

In terms of the validity of the WKCE, it is essential to understand that the cut scores were established in a collaborative, participatory process. The two key activities in that process were *standard setting* and *descriptor writing*. Simply speaking, standard setting is a collaborative process of setting cut scores, and descriptor writing is a collaborative process of establishing a plain-language description of what students must know in order to fall into each of the performance levels established though cut scores.

Descriptors root cut scores and performance levels in the content that students are supposed to learn. Descriptors reflect expectations of what Wisconsin students should know and be able to do in each grade/content area. Descriptors and cut scores together define, in qualitative and quantitative terms, the difference between a student who is Proficient, and a student who is not. Descriptors are an important part of the validity evidence that the tests that make up the WKCE actually measure the content they purport to measure.

The *Wisconsin Model Academic Standards* guided the standard setting and descriptor writing process. These guided participatory processes served to ensure that the achievement levels in the WKCE reflect the achievement standards and abilities intended by the Wisconsin legislature, teachers, citizens, and DPI.

A special linking study that linked scores from the previous WKCE assessments (those which existed until Fall 2005) to the current WKCE (the assessments that began with the Fall 2005 administration) was also an important part of setting the cut scores. For details of the linking study, the standard setting activities, and the descriptor writing process, please refer to the Fall 2005 Technical Report (Part 11), the Fall 2006 Technical Report (Parts 8 and 12), or to the 2005 Standard Setting Technical Manual, all of which can be located at http://dpi.wi.gov/.

Student performance on the WKCE is reported in terms of four *proficiency categories* - *Advanced, Proficient, Basic,* and *Minimal Performance*. Table 8-48 shows the cut scores for each content and grade level. Tables 8-49 to 8-53 show the percentage of all students in each performance category, as well as subgroup comparisons by the five NCLB groupings. The results for each content area and grade are summarized below. For ease of reference, Tables 8-54 to 8-58 are also provided. There, readers can see the scale score ranges that define performance levels, together with the percentage of students in each performance level.

Reading

Table 8-49 shows the percentage of students in different proficiency levels in Reading. Results indicate that most students were classified in the “At or Above Proficient” category in Reading. The proportions of students at Proficient and Advanced categories were very similar across grades except grades 3 and 10, where slightly more students were in the Advanced category. As a group, more female students were classified into the Advanced category and fewer female student were classified into the in the Minimal or Basic proficiency categories than males. The results also indicated that more White students were classified in Advanced category, followed by Asian Students, American Indian students, Hispanic students, and African American students. Note that the American Indian population was smallest among the ethnic groups. Across all grades, those students Proficient in English were less likely to be classified in the

Copyright © 2008 by the Wisconsin Department of Public Instruction

Minimal Performance level than students who were Not Proficient in English. At the same time, and in every grade, a higher proportion of those students who were Proficient in English were classified in the Advanced level than students Not Proficient in English. As a group, a higher proportion of Disabled students were classified in the Minimal Performance category and lower proportion in Advanced category than Not Disabled students. Similarly, a higher proportion of Not Economically Disadvantaged students were classified in the Advanced level and a lower proportion at the Minimal Performance level. The opposite was true for Economically Disadvantaged students.

Mathematics

Table 8-50 provides the percentage of students in different proficiency levels in Mathematics. Results show that a high proportion of students were classified in the At or Above Proficient category. Looking across subgroups, there were some small but interesting differences in performance by gender. Slightly higher proportions of Female students were classified in Minimal category in grades 3, 4, and 10 compared to Male students although the difference was small. However, relatively higher proportions of Male students were classified in Advanced category across all grades compared to Female. Grouped and compared by ethnicity, the tendency was such that a higher proportion of White students were classified in the At or Above Proficient category, followed by Asian Students, American Indian students, Hispanic students, and African American students. As in Reading, those students Proficient in English were less likely to be classified at the Minimal Performance level and more likely to be classified in the At or Above Proficient category. Also, as we saw in Reading, as a group, Disabled students were more frequently classified at the Minimal category than Not Disabled students. Classification in the Advanced category in Mathematics was more common among Not Disabled students than among Disabled Students. Also, those students Not Economically Disadvantaged were more likely to classify at the Advanced level, and less likely to classify at Minimal Performance level in comparison to Economically Disadvantaged students.

Language Arts

Like Reading and Mathematics, a relatively high proportion of students in Language Arts were classified in the At or Above Proficient category. The proportion, however, was lower than the other two content areas. In terms of gender, a higher proportion of Female students were classified in the At or Above Proficient category compared to Males. Conversely, fewer female students were classified in the Minimal category in Language Arts. In terms of ethnicity, the prevailing tendency in performance was such that a higher proportion of White students were classified in the At or Above Proficient category, followed by Asian students, American Indian students, Hispanic students, and African American students. As in other content areas, a higher proportion of those students Proficient in English were classified at the Advanced category and a lower proportion at the Minimal category compared to students who were Not Proficient in English. Also as in other content areas, as a group, Disabled students were more often classified at the Minimal category and less often at the Advanced category as compared to the Not Disabled students. Also, those students Not Economically Disadvantaged were more often

Copyright © 2008 by the Wisconsin Department of Public Instruction

classified in the Advanced level and less often in the Minimal Performance level in comparison to their Economically Disadvantaged counterparts.

Social Studies

Relatively high proportions of students in Social Studies were classified in the Advanced category especially in grade 4 (see Table 8-52). In terms of gender, relatively high proportions of male students were classified in the Advanced category in grades 8 and 10. In grade 4, more females than males were classified in the Advanced category. The difference, however, was small. Like other content areas, the tendency was such that White students were classified in the At or Above Proficient category most often, followed by Asian Students, then American Indian students, Hispanic students, and African American students. Those students Proficient in English, Not Disabled, and Not Economically Disadvantaged were more likely to classify at the Advanced category and less likely to classify at the Minimal category as compared to their counterparts.

Science

Table 8-53 presents the percentage of students in different proficiency levels in Science. Once again, high proportions of WKCE students were classified into the At or Above Proficient category. Looking across gender, male students were classified in the Advanced category more often than female students. However, the passing rate (At or Above Proficient) was very similar in both gender categories. In terms of ethnicity, White students classified in the At or Above Proficient category most often, followed by Asian Students, American Indian students, Hispanic students, and African American students. Similarly, students Proficient in English, Not Disabled, and Not Economically Disadvantaged were more likely to be classified in the At or Above Proficient category and less likely to classify in Minimal category as compared to their corresponding counterparts.

8.5 Standard Performance Indicator (SPI) for Content Standard

In addition to raw scores and scale scores, teachers and educational decision-makers frequently need diagnostic information to inform instructional strategies and to help identify students' strengths and weaknesses. This kind of information can be derived from scores on subsets of test items which estimate how much a student knows in a clearly defined skill domain. These skill domains are called content standards (or simply standards, or objectives). The purpose of reporting SPI on the WKCE assessments is to show the relationship between the overall achievement being measured and represented by the test score and the skills within each of the content standards associated with the overall content area. Teachers may use the SPI for individual students as indicators of strengths and weaknesses, but they are best corroborated by other evidence, such as homework, class participation, diagnostic test scores, or observation. District and school administrators may compare the results by content area and grade level with the state mean percentage to better understand their strengths and weaknesses within a particular content area and grade level.

Copyright © 2008 by the Wisconsin Department of Public Instruction

An SPI score can be interpreted as an estimate of the number of items a student would be expected to answer correctly if there had been 100 similar items for a given reporting category. For example, an SPI of 77 for a given reporting category means that if the student were given 100 similar items, the student would be expected to answer 77 of them correctly. These are criterion-referenced scores, in that they estimate how much a student knows in a clearly defined skill domain (i.e., the criterion). Technical readers can refer to *TerraNova* 2nd Edition Technical Report (CTB/McGraw-Hill, 2000) for detailed SPI (referred as objective performance index, OPI) estimation procedures.

The SPI provides a more reliable estimate of student achievement on each content standard than is possible by simply reporting percent correct. However, *the SPI information should be used for low-stakes purposes because even the SPI can not be stable for any content standard with a small number of items.* Also, while the percentages belonging to each performance level based on SPI cut scores can be compared across years, the SPI value can not be compared across years. This approach, identifying student proficiency of each content standard, relates to the Wisconsin Model Academic Standards.

Note that the average difficulty of items belonging to each content standard within each grade was not considered when the test form was constructed. Also, note that the average difficulty of items for each content standard across grades was not controlled. There has not been an effort to make the average difficulty of items for all content standards similar within each grade, or more difficult as grade level increases. The difficulty of items is determined, in part, by the difficulty of the content being measured. The difficulty of concepts varies across content standards, within grades as well as across grades. The current test blueprints do not specify the average difficulty of items for each content standard within grades or across grades. The mean *p*-value and mean SPI scores cannot be compared across grades. The mean *p*-value or mean SPI scores simply show the relative difficulty among content standards within one grade and content.

Tables 8-59 to 8-63 identify the content standards, the number of MC and CR items within each standard, the total number of possible points per standard, the mean raw score, mean *p*-value, standard deviation of the raw scores, as well as the mean SPI score, and the standard deviation of SPI scores, for all content areas across grades. The results are summarized below.

Reading

Table 8-59 presents mean *p*-value and SPI scores for Reading, across content standards and grades. In general, the largest difference in the mean SPI scores between the content standards was observed in grade 3, followed by grades 5, 4, 10, 7, 8, and 6. The trend for most grades was that the standard 3, “Analyzes Text,” and standard 4, “Evaluate/Extends Text,” were relatively difficult, as shown by relatively lower mean SPI scores, in comparison to the other two Standards “Determines Meaning” and “Understands Text.” Note that the mean SPI corresponds to mean *p*-value for all grades and content standards. In grade 3, the mean SPI for standard 4 (48.49) was the lowest and standard 2 (74.17) was the highest, meaning that standard 4 was the most difficult and standard 2 was the easiest for students in grade 3. In grade 4, standard 1 was easiest followed by standards 2, 3, and 4. Grade 5 showed a similar picture, with standard 2

Copyright © 2008 by the Wisconsin Department of Public Instruction

being the easiest one. Grade 6 showed two basic levels of SPI scores. Content standards 2, 3, and 4 were relatively difficult in comparison to standard 1. The range of mean SPI scores, however, was small, with about 6 percentage points of difference. In grade 7, the SPI scores spanned a total range of the equivalent of approximately 12% percentage points. In grade 8, standards 1 and 2 were easier relative to standards 3 and 4, and the mean SPI scores spanned from 65.55 in standard 3 to 73.09 in standard 2, or the equivalent of approximately 8 percentage points. In grade 10, the mean SPI for standard 1 (73.87) was higher than the other standards.

Mathematics

Table 8-60 presents Mathematics *p*-values and SPI scores across grades and content standards. Results show that the mean *p*-values and SPI scores varied across standards in all grades, with largest difference in grade 3 (where SPI scores ranged from 39.25 to 80.08). The difference in difficulty, as measured by the mean SPI, decreased as grade level increased. Notably, one can observe that standard A, Mathematical Processes, was the most difficult standard across grades 3 through 8. The difficulty of other standards was in a similar in range. In grade 10, standards C, D, and F (Measurement, Algebraic Relationship, and Geometry) showed a lower mean SPI scores, indicating relatively difficult items in the standards.

Language Arts

The SPI data in Table 8-61 for Language Arts also shows variation in mean *p*-values and mean SPI scores across content standards, for each grade. The mean *p*-value indicated that standard B (writing) was easiest for students from grades 4 (72.44) and 8 (78.00) but difficult for students from grade 10 (58.96). Grade 10 students perceived standard D (Language) as the less difficult standard, where the mean SPI score was 71.93.

Social Studies

Social Studies mean *p*-value and SPI scores also varied across grades and content standards (see Table 8-62). It is interesting to note here that grade 4 students perceived standard C as most difficult (mean SPI of 66.60), whereas grade 8 students found standard E (Behavioral Science) most difficult, and grade 10 students found standard A (Geography) most difficult. However, the mean SPI scores for different standards were are relatively close to each other compared to the range among standards in Reading and Mathematics.

Science

As indicated in Table 8-63, mean *p*-value and SPI scores in Science varied in each grade. The mean SPI for all content standards in grades 4 and 8 were over 60%, and higher, with standard A (Science Connections) being most difficult in grade 4 and standard E (Earth and Space) being most difficult in grade 8. Grade 10 students perceived standard F (Life and

Copyright © 2008 by the Wisconsin Department of Public Instruction

Environment) as the most difficult standard. However, in all grades, the mean SPI percentages were very similar across standards.

Summary of Student Achievement Results

In the WKCE, the purpose of the Reading, Mathematics, Language Arts, Science, and Social Studies assessments is to demonstrate student achievement through test scores in these content areas. The results presented in Part 8, together with the validity evidence, indicate that the scale scores and performance levels reported in the WKCE program are valid and reliable evidence of student achievement in the tested content areas and grades. As such, these test scores can be used to classify students, school, districts, and the state with respect to how much achievement each shows in each content area. Classroom teachers may use these scores as evidence of student achievement in these content areas. District and school administrators may use this information for activities such as planning curriculum. At the state level, the overall results can be drawn upon for the accountability and reporting purposes associated with *No Child Left Behind* or school improvement initiatives.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 9: Reliability

Part 9 of the Technical Report builds upon existing analyses of the summary results by providing additional estimates of the reliability of those results. Reliability can be defined as the consistency of an assessment when the testing procedure is repeated with the same testing target group. A reliable assessment is one that would produce stable scores if the same group of students were to take the same test repeatedly, without any fatigue or memory of the test. As detailed below, the reliability of the Fall 2007 WKCE was estimated in four ways:

1. Internal consistency was assessed for all multiple-choice and constructed-response items using Chronbach's alpha
2. Standard error of measurement (SEM) was calculated for raw score and scale score
3. Classification consistency and accuracy were estimated for performance classification
4. Inter-rater reliability was estimated for all of the constructed-response items

As described below, the present chapter demonstrates adherence to AERA/APA/NCME standards: 2.1, 2.2, 2.10, 2.11, 2.14, and 2.15.

Standard 2.1 advises providing reliability estimates and the SEM for scores reported. Accordingly, Part 9 presents both a reliability analysis (using Chronbach's alpha) and SEM. Reliability and SEM are presented by both content area and sub-content area, as well as for both the total group of examinees and for each of the five NCLB groupings (gender, ethnicity, socioeconomic status, disability status, and English Language Proficiency), thereby speaking to standard 2.11, which advises that reliability and SEM should be assessed for all sub-population groups.

Standard 2.2 is also addressed. This standard advises reporting SEM. While SEM has already been discussed in Part 7, in the present chapter, the SEM is presented by content area, grade level, the five NCLB groupings, and on the sub-content level as well.

Standard 2.15 advises that when testing measures are used to make categorical decisions, the reliability of those decisions should be estimated. In the present context, Standard 2.15 applies specifically to performance level determinations, such as who is Proficient or Advanced. As described below, the Fall 2007 WKCE adhered to this standard by applying a detailed analysis of classification consistency and classification accuracy, two related measures used to evaluate the reliability of the performance level classifications used in the WKCE program. This analysis also addresses Standard 2.14, by providing a conditional SEM for the cut scores that separate the performance levels.

Standard 2.10 advises reporting measures of inter-rater consistency where subjective judgment is involved in scoring. As we saw in Part 5, CR items were scored by (human) raters; the process thus involved subjective judgment. As this section will show, a detailed assessment of inter-rater consistency was applied to the WKCE. The assessment conducted is termed inter-rater reliability; it measures the reliability of human raters as they score CR items.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Combined, Cronbach's alpha, SEM, classification consistency, classification accuracy, and inter-rater reliability provide several forms of evidence bearing on the reliability of the WKCE. Cronbach's alpha and SEM operate at the content level: they provide estimates of reliability for examinee scores in Reading, or Mathematics for example. Classification consistency and classification accuracy operate on the associated performance level classifications. These are of particular interest in the context of NCLB and the associated AYP requirements. Inter-rater reliability probes further, looking at individual items, and evaluating the reliability of the human raters as they assign scores, item by item.*

9.1 Measures of Internal Consistency and SEM

Cronbach's alpha is a frequently used measure of internal consistency for tests consisting of MC and CR items. Cronbach's alpha is computed as:

$$\hat{\alpha} = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_X^2}\right),$$

where k = number of items, σ_X^2 = the total score variance, and σ_i^2 = the variance of item i (Crocker & Algina, 1986). Standard error of measurement (SEM) is defined as follows:

$$\text{SEM} = SD\sqrt{1 - reliability},$$

where SD represents standard deviation of the raw score distribution, and reliability represents Cronbach's alpha.

Cronbach's alpha and the standard error of measurement (SEM) are shown in Tables 9-1 and 9-2 respectively. These tables include information for all students and for each NCLB grouping. As indicated in Table 9-1, reliability was highest in Reading and Mathematics. Looking at all examinees together, in the "Total" column, reliability ranges from 0.90 to 0.93 across grades for Reading, from 0.89 to 0.93 for Mathematics, from 0.82 to 0.85 for Language Arts, from 0.87 to 0.90 for Social Studies, and from 0.85 to 0.88 for Science. If 0.90 is considered a conservative criterion for an acceptable level of reliability, as measured by Cronbach's alpha, Language Arts, Social Studies and Science would not meet this criterion. These content areas have a relatively small maximum number of items (and score points), and, notably, fewer items (and score points) than either Reading or Mathematics. The number of items (and score points) has a close relationship with reliability, so these results, both for Reading and Mathematics, as well as Language Arts, Social Studies, and Science, can be understood within that context.

* Note that the field test items were not used in assessing the reliability and validity of the WKCE.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-1 shows that there were some differences in reliability for each NCLB grouping. Reliability for female students was often slightly lower than for male students. Reliability was also often higher for Asian students than for other students of ethnicities. In all grades and content areas, reliability was higher for the Proficient group than it was for the Not Proficient group. There were also some differences in reliability between Disabled and Not Disabled students, but those differences were very small, and not consistent: in some cases reliability was higher for the Disabled group, in other cases reliability was higher for the Not Disabled group. Differences were also mixed in the comparison by socio-economic status (SES), but these differences were consistently very small.

As indicated in Table 9-2, Reading and Mathematics both produced a larger SEM than the other content areas. Reading and Mathematics have more items than the other content areas. Readers should note that SEM tends to increase as the number of items (or score points) increases because, in general, the standard deviation is large for tests with large numbers of items or score points, and the SD is a part of the calculation for SEM. There were some observable differences in SEM looking across the five NCLB groupings. Differences between genders were small. Compared by ethnicity, SEM was somewhat smaller for White students than for students of every other ethnicity. The SEM was also smaller among students Proficient in English than among students Not Proficient in English, in all grades and contents. The SEM was also larger for students who were Disabled than those Not Disabled. In all grades and content areas, the SEM was larger for Economically Disadvantaged students than for those students who were Not Economically Disadvantaged.

Reliability, as measured by Cronbach's alpha, was also computed for each content standard within each content area. Table 9-3 shows this reliability by content standard. The last column presents the reliability for the total test per content area (with all content standards) for all examinees. It is clear that the reliability per content standard is lower than that for the total test per content area. As discussed above, the number of items (or score points) has a close relationship with reliability, and a small number of items (or score points) is often associated with lower reliability. A lower level of reliability statistics per content standard is therefore expected here, due to the smaller number of items for each standard than for the total test. The generally lower level of reliability per standard is one of the reasons why the information based on the content standards should be used for low stakes purposes only (this issue was previously discussed in the context of SPI).

By content standard then, reliability in Reading ranged from 0.45 to 0.85. The lowest reliability coefficients in Reading were for standard D in grades 3, 4, and 5. In Mathematics, reliability by content standard ranged from 0.48 to 0.76. Here, the lowest reliability statistics were found in standards A and C for grades 4 and 5. In Language Arts, reliability by content standard ranged from 0.31 in standard F of grade 8 to 0.80 in standard B of grade 8. Standard F showed the lowest level of reliability overall. In Social Studies, reliability by content standard ranged from 0.47 to 0.72. The lowest reliability statistics were found in standards C and D in grade 4, and in standard E in grade 8. At the content standard level, Science showed the lowest reliability statistics of all content areas, ranging from 0.23 to 0.59. The lowest level of reliability was found in grade 4, Standard D.

Copyright © 2008 by the Wisconsin Department of Public Instruction

The SEM associated with each content standard is presented in Table 9-4, by content area and grade level. Some differences in SEM by content standard can be observed. As indicated by the discussion above, these SEMs were smaller compared to those for the total test because of the small number of items within each content standard. In Reading, standard C showed a higher SEM than the other standards. In Mathematics, the SEM was generally lowest in standards B, C, and E. In Language Arts, the SEM for standard B was higher than for other standards. The SEM for Social Studies was higher in standard A for grade 10, in standard B for grades 8 and 10, and in standard C for grade 10. The SEM in Science was lower in grades 4 and 8 in standards A, B, C, D, G, and H than grade 10. Readers may wish to recall here that grade 10 Science has more items than either grade 4 or grade 8.

In summary, the reliability indices, as measured by Cronbach's alpha, at the test level, are in a reasonable range given number of items in each test. As indicated above, readers should also note that since the reliability is influenced by number of items one can expect lower reliability for the content standards with fewer items.

9.2 Classification Consistency and Accuracy

One of the cornerstones of the NCLB Act (2002) is the measurement of Adequate Yearly Progress (AYP) for states with respect to the percentage of students at or above the academic performance standards established by states. Because of a heavy emphasis on moving all students to or above the "Proficient" category by year 2014, the consistency and accuracy of the classification of students into these performance categories is of particular interest. The following section demonstrates how the consistency and accuracy of these classifications were assessed, and it provides evidence supporting the validity of these classifications.

Conceptually, classification consistency is defined as the extent to which two classifications of a single student agree, either based on two independent administrations of the same test, or one administration of two parallel test forms. However, it is difficult to obtain data from repeated administrations of the same form because of the cost, time, and student memory from prior administrations. It is also difficult to construct two psychometrically parallel forms. For these reasons, the common practice is to estimate classification consistency from a single administration.

A contingency table representing the probability of particular classification outcomes under specific scenarios is a convenient way to measure classification consistency. The table below is a contingency table of $(H+1) \times (H+1)$, where H is the number of cut scores. Three cut scores yield a 4×4 contingency table, as can be seen below (see Table a).

It is common to report two indices of classification consistency: the classification agreement "P" and the coefficient kappa. Hambleton and Novick (1973) proposed P as a measure of classification consistency, where P is defined as the sum of diagonal values of the contingency table:

$$P = P_{11} + P_{22} + P_{33} + P_{44.}$$

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table a
Contingency Table with 3 Cut Scores

	Level 1	Level 2	Level 3	Level 4	Sum
Level 1	P_{11}	P_{21}	P_{31}	P_{41}	$P_{.1}$
Level 2	P_{12}	P_{22}	P_{32}	P_{42}	$P_{.2}$
Level 3	P_{13}	P_{23}	P_{33}	P_{43}	$P_{.3}$
Level 4	P_{14}	P_{24}	P_{34}	P_{44}	$P_{.4}$
Sum	$P_{1.}$	$P_{2.}$	$P_{3.}$	$P_{4.}$	1.0

To reflect statistical chance agreement, Swaminathan, Hambleton, and Algina (1974) suggest using Cohen's kappa (1960):

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

where P_c is the chance probability of a consistent classification under two completely random assignments. This probability P_c is the sum of the probabilities obtained by multiplying the marginal probability of the first administration and the corresponding marginal probability of the second administration:

$$P_c = (P_{1.} \times P_{.1}) + (P_{2.} \times P_{.2}) + (P_{3.} \times P_{.3}) + (P_{4.} \times P_{.4}).$$

While classification *consistency* refers to the agreement between two observed scores, classification *accuracy* refers to the agreement between the observed score and the true score. Classification accuracy is defined as the extent to which the actual classifications of test takers agree with those that would be made on the basis of their true scores (Livingston & Lewis, 1995). It is common to estimate classification accuracy by assuming the psychometric model to find true scores corresponding to observed scores.

9.2.1 Kolen and Kim's Method for Pattern Scoring

As stated in Part 7, when item response theory (IRT) is applied to score examinees' responses, two types of scoring are available: number-correct scoring and item pattern scoring. WKCE is an example of a program that has applied item pattern scoring. Many methods of estimating the consistency and accuracy of classification based on number-correct scoring have been suggested in the psychometric literature. However, there have been relatively few studies dealing with item pattern scoring based on IRT. Kolen and Kim (2004) suggested a simple procedure for pattern scoring (KKM) based on IRT and simulated item responses. KKM requires a simulation of item responses as follows:

Step 1: Obtain item parameters (**I**) and the ability distribution weight ($\hat{g}(\theta)$) at each quadrature point.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Step 2: Compute two ability estimates at each quadrature point. At a given quadrature point θ_j, generate two sets of item responses using the item parameters from a test form, assuming that the same test form was administered twice to an examinee with the true ability θ_j.

$\theta_j \nearrow$ (1,1,0,0,…: Item response from the first administration, or Form 1) $\longrightarrow \hat{\theta}_{j1}$

$\theta_j \searrow$ (0,1,1,0,…: Item response from the second administration, or Form 2) $\longrightarrow \hat{\theta}_{j2}$

If two parallel (or alternative) forms, e.g., Form 1 and Form 2, are available, the two response patterns can be generated based on the item parameters from the two forms.

Step 3: Construct a classification matrix at each quadrature point. Determine the joint event for the cells in Table b using the two ability estimates obtained from Step 2.

Table b
Classification Table for One Cut Point (C_1)[1]

	First administration or Form 1: $\hat{\theta}_{ji} \geq C_1$	First administration or Form 1: $\hat{\theta}_{j1} < C_1$	
$\hat{\theta}_{j2} \geq C_1$			Second administration, or Form 2
$\hat{\theta}_{j2} < C_1$			

Step 4: Repeat Steps 2 and 3 *R* times and get average values over *R* replications. *R* should be a large number, e.g., 500, to obtain stable results.

Step 5: Multiply distribution weight ($\hat{g}(\theta)$) by the average values in Step 4 for each quadrature point, and sum across all quadrature points. From this, a final contingency table and classification consistency indices, such as kappa, can be computed.

Because examinees' abilities are estimated at each quadrature point, this quadrature point can be considered the true score. Therefore, classification accuracy is computed using both examinees' estimated abilities (observed scores) and quadrature point (true score).

As can be seen in Table 9-5, for Reading grade 3, there are two tables for each grade and content area. The first table is a contingency table with all three cut scores. This table was prepared based on the KKM procedure. The rows represent the first administration of an assessment, and the columns represent the second administration of the same assessment to the same students. As mentioned above, in the KKM procedure, the score distributions for the first

[1] This table is constructed for each quadrature point and replication. One, and only one, cell will have a value of 1 and zeros elsewhere.

Copyright © 2008 by the Wisconsin Department of Public Instruction

administration and the second administration are estimated using a simulation. So, the value in each cell represents the probability of belonging to a certain performance level in the first administration and the second administration. For example, in Reading grade 3, 0.04 represents the probability of belonging to "Minimal Performance" in the both first and second administrations. The 0.05 represents the probability of belonging to "Proficient" in the first administration and "Advanced" in the second administration. "Sum" is obtained simply by adding the four row values or the four column values.

The second table shows indices for classification consistency and classification accuracy. Because there are three cuts for the 2007 WKCE, four performance levels exist. The values in "All cuts" were obtained by applying all three cuts together. In Table 9-5 for Reading grade 3, when all three cuts were used for the computation, classification consistency (P) is 0.80, chance probability is 0.34, kappa is 0.70, and classification accuracy is 0.86. The values for cut 1 were obtained by applying only the first cut score. There are two levels whenever only one cut is applied. It is clear that the values for P, kappa, and classification accuracy with all three cuts are smaller than those with only one cut. The probability of assigning students to the incorrect performance level will increase with the number of cut scores.

Because the Proficient cut score is a criterion for the AYP report, the reliability values for this second cut need to be considered carefully. In Table 9-5, for example, the P for the second cut, which establishes the Proficient performance level, was 0.94, k was 0.81, and classification accuracy was 0.96. The interpretation of the table values outlined here is the same for Tables 9-6 to 9-27.

When only the Proficient cut score was applied, P was equal to or larger than 0.91, and kappa was equal to or larger than 0.75 for Reading and Mathematics. For Language Arts, the lowest P associated with the Proficient cut was 0.87 and the lowest kappa was 0.69. In Social Studies, the lowest P associated with the Proficient cut was 0.91 and the lowest kappa was 0.71. For Science, the lowest P was 0.89 and the lowest kappa was 0.73. If the criterion value of 0.90 is applied for P, each grade level in Language Arts was slightly lower than the criterion, as was Science (grades 8 and 10), but in all cases the value for P was larger than 0.80. Like inter-rater reliability, Landis and Koch (1977) suggest that values of Kappa greater than 0.75 indicate "excellent agreement," values between 0.40 and 0.74 represent "good agreement" beyond chance, and values below 0.40 denote "poor agreement." According to Landis and Koch's criteria for Kappa, all tests for Reading and Mathematics showed "excellent agreement" based on the cut for the Proficient performance level. For the other three content areas, Kappa represented "good agreement."

Figures 9-1 through 9-5 also show P, kappa, and classification accuracy when students were classified based on "All Cuts." These values are in Tables 9-5 to 9-27, but the plots are also provided for ease of understanding. As can be seen in the plots, all grades and content areas indicated classification consistency (P) based on all cuts over 0.70, except for Language Arts, where (P) was 0.69, 0.68, and 0.72 in grades 4, 8, and 10 respectively. The values of k were over 0.60 across all grades in Reading, Mathematics, and Social Studies. In Language Arts, the values of (k) were 0.54, 0.56, and 0.57 in grades 4, 8, and10 respectively. In Science, the values for (k) were 0.61, 0.58, and 0.59 respectively. Unlike other content areas, a Writing prompt of 9 points

Copyright © 2008 by the Wisconsin Department of Public Instruction

contributed to the Language Arts grade 10 scores, for which the total number of score points is 39. The impact of this Writing prompt may need to be examined in future administrations. In summary, based on Landis and Koch criteria, all test forms showed "good agreement."

Readers will recall that the SEM for the cut scores can be found in Part 8 in section 8.3.3. Alternatively, readers may note that the scoring tables (Tables 7-2 to 7-24) showed the SEM around all scores, including the cut scores. The SEM for each grade and content area was also plotted in Figures 7-1 to 7-5, and the location of the cut scores was provided in each plot so that the associated SEM could be located. These results showed that the SEM at the "Proficient" cut was low in all grades and content areas.

9.3 Inter-rater Reliability for CR Items and Writing Prompts

The reliability of handscoring may be measured in a variety of ways. Two of the most effective ways are 1) tabulations of exact and adjacent agreement, and 2) reliability coefficients. Reliability for constructed-response items is typically examined by calculating indices of inter-rater agreement: the degree of reliability with which different human raters assign scores to a given student response. Two indexes for inter-rater reliability are presented here: intraclass correlation, and weighted kappa. The formula for intraclass correlation and weighted kappa are from Rich Patz's 1998 unpublished paper, "Calculating Handscoring Reliability Coefficients."

Notation. To assess reliability, it is necessary to replicate the scoring process for a subset of papers. This is usually done with "blind double reads." Suppose that we have *N* responses, each of which is scored twice. We denote the two scores of response *n* by X_{n1} and X_{n2}, where *n*=1, 2, …, *N*. The resulting data may be presented in two ways, enumeration by response and cross-tabulation:

Data Structure 1: Enumeration by Response. Each row represents a single student response:

Response #	Score1	Score2	Mean Score
1	X_{11}	X_{12}	$\overline{X}_{1.}$
2	X_{21}	X_{12}	$\overline{X}_{2.}$
.	.	.	
.	.	.	
N	X_{N1}	X_{11}	$\overline{X}_{N.}$
Column Mean	$\overline{X}_{.1}$	$\overline{X}_{.2}$	$\overline{X}_{..}$

where:

$$\overline{X}_{.1} = (X_{11} + X_{12})/2$$

is the mean score for response 1 (similarly for responses 2,3, …*N*),

Copyright © 2008 by the Wisconsin Department of Public Instruction

$$\overline{X}_{.1} = \frac{1}{N}\sum_{n=1}^{N} X_{.1} = (X_{11} + X_{21} + ... + X_{N1})/N$$

is the mean of Score1 over all responses (similarly for Score2), and

$$\overline{X}_{..} = \frac{1}{N}\sum^{N} 1(X_{n1} + X_{n2})/2$$

is the overall mean score across both scores of all responses.

Data Structure 2: Cross-tabulation of Score1 and Score2. As an alternative, we may create a square table of counts for each Score1 by Score2 (i.e., $X_{n1} \times X_{n2}$) combination:

Data Structure 2: Cross-tabulation of Score1 and Score2

		Score2				Row Total
		0	1	…	K	
Score 1	0	n_{00}	n_{01}	…	n_{0K}	n_{0+}
	1	n_{10}	n_{11}	…	n_{1K}	n_{1+}
	.	.	.	…		
	.	.	.	…		
	K	n_{K0}	n_{K1}	…	n_{KK}	n_{K+}
Column Total		n_{+0}	n_{+1}	…	n_{+K}	n_{++}

where K is the maximum score (for a rubric including zero) obtainable for the item, n_{ij} is the number of responses for which Score1 = i and Score2 = j, n_{i+} is the number of responses for which Score1 = i, and n_{+j} is the number of responses for which Score2 = j.

Formulas for the two reliability coefficients of interest are now given:

1. **Intraclass correlation**, ρ_{IC}, describes the percent of overall score variance accounted for by the variance of mean response scores:

$$\rho_{IC} = \frac{Var_n(\overline{X}_{n.})}{Var_n(X_{n1}, X_{n2})} = \frac{\frac{1}{N-1}\sum_{n=1}^{N}(\overline{X}_{n.} - \overline{X}_{..})^2}{\frac{1}{2(N-1)}\sum_{n=1}^{N}[(X_{n1} - \overline{X}_{..})^2 + (X_{n2} - \overline{X}_{..})^2]}$$

Copyright © 2008 by the Wisconsin Department of Public Instruction

If agreement is perfect ρ_{IC} =1. Always, $0 \le \rho_{IC} \le 1$.

2. **Weighted Kappa**, *k*, is used in many contexts as a measure of association in square contingency tables:

$$k = \frac{\sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij}\frac{n_{ij}}{n_{++}} - \sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij}\frac{n_{i+}n_{+j}}{n^2_{++}}}{1-\sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij}\frac{n_{i+}n_{+j}}{n^2_{++}}}, \text{ where } w_{ij} = 1-\frac{(i-j)^2}{K^2}$$

If agreement is perfect, *k*=1. If agreement is what would be expected by chance, *k*=0. Always, $0 \le k \le 1$.

Ordinal rating scales (e.g., 0, 1, 2), used in scoring CR items contain a certain level of chance agreement that is expected. Although the intraclass correlation is reported in this report, it does not take into account the possibility of chance agreement between the two raters, but Cohen's Kappa does take this into consideration. In general, Kappa will have values equal to or smaller than the intraclass correlation. If agreement is perfect, then Kappa is +1. If agreement is at chance levels, Kappa is 0. Landis and Koch (1977) suggest that values of Kappa greater than 0.75 indicate "excellent agreement," values between 0.40 and 0.74 represent "good agreement" beyond chance, and values below 0.40 denote "poor agreement." Specific criteria for intraclass correlation or weighted kappa are not established.

Tables 9-28 through 9-30 present the rater agreement statistics for CR items and the Writing prompt. The evidence supporting inter-rater reliability is presented in terms of the percentage of agreement between raters, two indexes of inter-reliability, and the distributions of scores across score levels. In the table, "Perfect" agreement is defined as scores that are exactly the same. "Adjacent" agreement is defined as scores differing by one point. "Discrepant" cases are those cases where the scores of the two raters differed by more than one raw score point. The column for "codes" reflects the number of students who receive the condition codes, A, B, C, or D, which indicate illegible responses, responses that are off-topic, blank responses, or in another language. Mean reflects mean score, Number of Reads is exactly two times the number of papers submitted for the purposes of computing inter-rater reliability, as each paper submitted for that purpose was read twice. The Frequency column represents the scoring outcomes for the student responses, based on the raw scores given by each of the two raters. For example, in Table 9-28 for Reading grade 3, Form A, item 19, shows that the perfect agreement, adjacent agreement, discrepant agreement, and codes are 66%, 20%, 1%, and 14%, respectively. The percentage values for agreement (all forms of agreement) were computed after the students with condition codes were dropped. The 14% of the responses with condition codes means that a condition code was assigned by either one of the two readers, or both readers. The condition codes were transformed to the raw score of 0 when the three indexes for inter-reliability were estimated.

For Reading and Mathematics, all responses were read by a single rater, and a portion was submitted to a second rater for scoring. All Writing prompts were scored by two readers.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Reading

Inter-rater reliability results for Reading CR items are in Table 9-28. As indicated in the Table, the maximum score for any CR item in Reading is 3 score points. The percent of perfect agreement occurred in 54% to 77% of the cases. Adjacent agreement occurred in 15% to 40% of cases, but typically it occurred in less than 25% of cases. Discrepant agreement occurred far less often than any other mode of agreement. The percent of discrepant agreement ranges from 1% to 14%, but levels of discrepant agreement over 2% were not common. If we consider a difference of more than 1 score point between raters as critical, the rate of discrepant agreement indicates there were few critical differences. The percentage of responses with condition codes ranged from 2% to 14% across grades, but the percentage was not often more than 4%. Intraclass correlation ranges from 0.75 to 0.95, values were typically over 0.80. Weighted Kappa in Reading ranged from 0.49 to 0.90, though values less than 0.65 were not common.

Mathematics

Table 9-29 provides the inter-rater reliability indices for Mathematics. The CR items in Mathematics have a maximum score of either of 1 or 2 score points. In general, a lower maximum possible score produces better inter-rater agreement and reliability than a higher maximum score will. As may be expected then, compared to Reading, Mathematics produced, on average, a larger percent of perfect agreement and smaller percent of discrepant agreement. Perfect agreement occurred in 68% to 98% of reads. Over 90% of reads resulting in perfect agreement was common. Adjacent agreement ranged from less than 1% to 26% across grades, though typically the result was less than 10% adjacent agreement. Looking across all grades and items, from 1% to 16% of the students received condition codes, though more than 6% of reads resulting in condition codes was not common. Higher rates of codes were observed in grades 7 and 8. Intraclass correlation ranges from 0.84 to 1.00. Weighted Kappa ranged from 0.68 to 1.00. Both intraclass correlation and weighted kappa were, in general, higher for Mathematics than for Reading. As suggested above, better results for Mathematics do not mean that the raters for Mathematics did better than the raters for Reading. In many large scale assessments, it is often found that inter-reliability is higher for Mathematics than for Reading. Scoring Rubrics for Mathematics can be much clearer for raters than those for Reading.

Writing

Table 9-30 shows inter-rater reliability indices for the Writing Prompts. As indicated previously, the Writing prompts were scored on two rubrics, one for Composing (6 points) and one for Convention (3 points). Table 9-30 shows that the rate of perfect agreement was lower for the 6-point Composing element than for the 3-point Convention element. The difference here is due to the difference in score points. Perfect agreement is, as discussed above, less likely with a higher number of possible score points than with a lower number of possible score points. Adjacent and discrepant modes of agreement were, as may also be expected, more common where there were more possible score points. Perfect agreement here ranged from 56% to 62% for Composing, but from 88% to 96% for Convention. Adjacent agreement ranged from 33% to

Copyright © 2008 by the Wisconsin Department of Public Instruction

35% for Composing and from 2% to 6% for Convention. Notably, the sum of perfect agreement and adjacent agreement was similar for both Composing and Convention, and discrepant agreement was rare in both elements of the scoring rubric (not more than 3%). Codes were generated in 2 to 6% of cases. Intraclass correlation here ranges from 0.88 to 0.96, and weighted kappa ranges from 0.78 to 0.93.

The inter-rater reliability across CR items in terms of the weighted kappa and intraclass correlations is one way to measure the consistency of the hand score. The results of rater reliability measures, which assess the agreement rates within a given administration, demonstrate that the WKCE tests have relatively high reader reliability.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 10: Validity

The *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) defines validity as "the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing and evaluating tests" (p. 9). The purpose of test score validation is not to validate the test itself, but to validate interpretations of the test scores for particular purposes or uses. Test score validation is not a quantifiable property but an ongoing process, beginning at initial conceptualization and continuing throughout the entire assessment process. Every aspect of an assessment provides evidence in support of (or a challenge to) its validity, including design, content specifications, item development, psychometric quality, and inferences made from the results.

As the Technical Report has progressed, chapter by chapter, it has moved through the phases of the testing cycle. Each part of the Technical Report detailed the procedures and processes applied in the WKCE, as well as their results. Each part also highlighted the meaning and significance of the procedures, processes, and results, in terms of validity or a relationship to the *Standards.* Part 10 now addresses three final issues in validity: the issues of bias, construct validity, and test integrity. The analyses presented here add to the perspectives provided in Chapters 2 through 9. Below is a brief review:

Part 2 of the Technical Report described the involvement of Wisconsin educators, DPI, and CTB in the test development process. As indicated in Part 2, the test development process and the involvement of Wisconsin educators in that process formed an important part of the validity of the entire WKCE. The knowledge, expertise, and professional judgment offered by Wisconsin educators ultimately ensured that the content of the WKCE formed an adequate and representative sample of appropriate content, and that the content formed a legitimate basis upon which to valid derive conclusions about student achievement.

Part 3 of the Technical Report addressed the issue of Test Form development. Part 3 showed a general discussion of CTB's test book creation and editing process, the process of selecting operational test items, the content distribution of embedded field test item, and the process of obtaining DPI approvals. The test design process and the participation of Wisconsin educators in the process of test selection including item content and bias review provide a solid rationale for having confidence in the content and design of the WKCE as a tool from which to derive valid inferences about Wisconsin student performance. Parts 2 and 3 together addressed AERA/APA/NCME standards 1.2, 1.6, 3.1, 3.2, 3.3, 3.5, 3.6, 3.7, 3.9, 3.11, 3.16, 6.4, 6.15, 7.3, 7.4, 7.7, 13.3, and 13.5.

Part 4 of the Technical Report described the process, procedures, and policies that guided the administration of the WKCE, including accommodations, security, and the written procedures provided to test administrators and school personnel. The following AERA/APA/ NCME standards were addressed: 1.13, 3.3, 3.19, 3.20, 3.21, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 6.11, 6.15, 9.1, 10.1, and 10.2.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 5 of the Technical Report demonstrated adherence to AERA/APA/NCME standards 3.22, 3.23, 3.24, 5.8 and 5.9. It described how multiple-choice items and constructed-response items were scored, it described the hand scoring process, the training and selection of evaluators, the scoring rubrics used for scoring the constructed-response items, and the resulting score distributions.

Part 6 described the sample data used for calibration and scaling and compared demographic information of the sample against the WI population. It showed that the calibration sample data was sufficiently representative of WI student population providing a foundation for subsequent analyses. Part 6 thereby demonstrated adherence to AERA/APA/NCME standards 1.5, 1.13, 2.4, 4.7, and 6.1.

Part 7 of the Technical Report described the calibration and equating methods, as well as processes and procedures for deriving scale scores from response patterns. Some references to introductory and advanced discussions of IRT were provided. Several axes upon which to evaluate the calibration and equating procedures, such as the models and data used, software applied, the vertical relationship across grades, the successful estimation of parameters, fit, the standard error of measurement, and IRT scoring method were all discussed. Part 7 of this report thereby addressed AERA/APA/NCME standards 1.13, 4.1, 4.2, 4.3, 4.10, 4.11, 7.1, 7.2, and 7.10.

Part 8 presented classical item analysis, raw score results, scale score results, performance level information, and SPI scores. Scale score results provided a basic quantitative reference to student performance as derived through the IRT models applied. The performance level information reflected the performance level requirements of the NCLB policy environment, as well as interests of parents, students, and educators. The SPI scores then probed further, assessing specific skills and abilities. Combined, scale scores, performance levels, and SPI scores provided a comprehensive set of tools to assess WI student performance by content, grade level, and NCLB grouping (gender, ethnicity, socioeconomic status, disability status, and English Language Proficiency). Part 8 thus addressed AERA/APA/NCME standards 1.5, 3.18, 4.3, 4.5, 4.6, 4.7, 4.19, 7.1, 7.10, 13.15, and 13.19.

Part 9 demonstrated adherence to AERA/APA/NCME standards through several analyses of the reliability of the Fall 2007 WKCE. It presented a reliability analysis using Cronbach's alpha, SEM results, a detailed analysis of classification consistency and classification accuracy, and a full analysis of inter-rater reliability. The Fall 2007 WKCE thereby addressed AERA/APA/NCME standards 2.1, 2.2, 2.10, 2.11, 2.14, and 2.15.

In the subsequent pages, Part 10 will, as stated, present additional axes upon which to evaluate the validity of the WKCE program. As described below, the WKCE program formally assessed the issue of test bias through an analysis of differential item functioning (DIF). It is possible for items to function differently among different population groups, and it is also possible that results for an item do not reflect student ability, but instead reflect irrelevant information influenced by demographic factors. The DIF analysis provided below serves to determine if that possibility occurred, and to what degree, item by item for each of the five NCLB groupings. This analysis specifically addresses standards 7.1, 7.2, and 7.3.

Copyright © 2008 by the Wisconsin Department of Public Instruction

The present chapter will also provide estimations of construct validity. Two measures are provided: correlations among content area objectives, and an exploratory factor analysis. Both of these measures are provided to demonstrate the existence of a single underlying trait or ability for each content area, such as reading ability, or mathematics ability. The presence of a single underlying trait is a fundamental issue where scaling and analyzing results through IRT models. As such, these analyses are essential elements in assessing the validity of the WKCE.

10.1 Differential Item Functioning

The Fall 2007 WKCE tests were developed using procedures to minimize item and test bias. Expertise in this area is not, however, a substitute for statistical analyses of the items. Thus, an empirical differential item functioning (DIF) approach was used to examine potential item bias. The approach applied included systematic item analyses to determine if examinees with the same underlying level of ability had the same probability of correctly responding to the item. Items identified with DIF were examined to determine if item performance differences between identifiable subgroups of the population were due to extraneous or construct irrelevant information making the items unfairly difficult for one of the subgroups.

As indicated above, DIF analyses were conducted for the 5 NCLB groups: gender (male and female), ethnicity (White, African American, Hispanic, Asian, American Indian), socioeconomic status (Economically Disadvantaged, Not Economically Disadvantaged), disability status (Disabled, Not Disabled), and ELP (Proficient, and Not Proficient).

Three kinds of DIF statistics were used in this study, Linn-Harnisch, Mantel-Haenszel, and standardized mean difference. Linn-Harnisch is used for both MC and CR items. Mantel-Haenszel is applied to MC items and CR items. The standardized mean difference is applied to CR items. An item was flagged for DIF when any of these indexes crossed their respective thresholds.

(1) Linn-Harnisch (L-H)

Because the WKCE was built using item response theory (IRT), an appropriate procedure for examining item bias would need to reflect the IRT model. Several IRT-based procedures are available, such as a procedure that tests the equality of item parameters across groups (Lord, 1980), or any of the procedures that assess the differences in the area between the item characteristic curves (e.g., Linn, Levine, Hastings, & Wardrop, 1981). However, these procedures require a minimum of 800 to 1000 cases in each group to make reliable comparisons. A procedure that still relies on the predictions of the three-parameter model but does not require as many cases has been suggested by Linn and Harnisch (1981).

To take an example, in the case of gender DIF analyses, item parameters (e.g., discrimination, location, and guessing) and the scale score (θ) for each examinee were estimated using the three-parameter logistic model for MC items and the two-parameter partial credit model for CR items. The sample was then divided into male and female gender subgroups. The

Copyright © 2008 by the Wisconsin Department of Public Instruction

members in each group were sorted into ten equal score categories (deciles) based upon their location on the scale score (θ) scale. The expected proportion correct for each group based on the model prediction was compared to the observed (actual) proportion correct obtained by the group. The proportion of people in decile g who are expected to answer item i correctly is:

$$P_{ig} = \frac{1}{n_g} \sum_{j \varepsilon g} P_{ij} ,$$

where n_g is the number of examinees in decile g. To compute the proportion of people expected to answer item i correctly (over all deciles) for a specific subgroup (e.g., African American), the following statistic was computed:

$$P_i. = \frac{\sum_{g=1}^{10} n_g P_{ig}}{\sum_{g=1}^{10} n_g} .$$

The corresponding observed proportion correct for examinees in a decile (O_{ig}) is the number of examinees in decile g who answered item i correctly divided by the number of people in the decile (n_g). That is,

$$O_{ig} = \frac{\sum_{j \varepsilon g} u_{ij}}{n_g} ,$$

where u_{ij} is the dichotomous score for item i for examinee j. The corresponding formula to compute the observed proportion answering each item correctly (over all deciles) for a subgroup is given by:

$$O_i. = \frac{\sum_{g=1}^{10} n_g O_{ig}}{\sum_{g=1}^{10} n_g} .$$

After the values are calculated for these variables, the difference between the subgroup's observed proportion correct and expected proportion correct can be computed. The decile group difference (D_{ig}) for observed and expected proportion correctly answering item i in decile g is:

$$D_{ig} = O_{ig} - P_{ig} ,$$

and the overall group difference (D_i) between observed and expected proportion correct for item i in the complete group (over all deciles) is:

Copyright © 2008 by the Wisconsin Department of Public Instruction

$$D_{i\cdot} = O_{i\cdot} - P_{i\cdot}.$$

These indices are indicators of the degree to which subgroup members performed better or worse than expected on each item, based on the parameter estimates from all subgroups. Differences for decile groups provide an index for each of the ten regions on the scale score (θ) scale. The decile group difference (D_{ig}) can be either positive or negative. Use of the decile group differences as well as the overall group difference allows one to detect items that give a large positive difference in one range of θ and a large negative difference in another range of θ, yet have a small overall difference.

DIF is defined in terms of the decile group and total target subsample differences, the D_{i-} (sum of the negative group differences) and D_{i+} (sum of the positive group differences) values, and the corresponding standardized difference (Z_i) for the subsample (see Linn & Harnisch, 1981, p. 112). Items for which $|D_i| \geq 0.10$ and $|Z_i| \geq 2.58$ are flagged for DIF. If D_i is positive, the item is biased in favor of the target subsample. If D_i is negative, the item is biased against the target subsample.

(2) Mantel-Haenszel (M-H)

The Mantel-Haenszel statistic is computed as (Zwick, Donoghue, & Grima, 1993):

$$\text{Mantel } \chi^2 = \frac{\left(\sum_k F_k - \sum_k E(F_k)\right)^2}{\sum_k \text{Var}(F_k)},$$

where F_k is the sum of scores for the focal group at the kth level of the matching variable. Note that the Mantel-Haenszel statistic is sensitive to N such that larger sample sizes increase the value of chi square.

In addition to the Mantel-Haenszel chi-square statistic, the delta statistic (MH-D DIF) was computed for all items. Educational Testing Service (ETS) first developed the MH-D DIF statistic. To compute delta, alpha (the odds ratio) is first computed as:

$$\alpha_{MH} = \frac{\sum_{k=1}^{K} N_{r1k} N_{f0k} / N_k}{\sum_{k=1}^{K} N_{f1k} N_{r0k} / N_k},$$

where N_{r1k} is the number of correct responses in the reference group at ability level k, N_{f0k} is the number of incorrect responses in the focal group at ability level k, N_k is the total number of

Copyright © 2008 by the Wisconsin Department of Public Instruction

responses, N_{f1k} is the number of correct responses in the focal group at ability level k, and N_{r0k} is the number of incorrect responses in the reference group at ability level k. MH-D DIF is then computed as:

$$\text{MH-D DIF} = -2.35\ln(\alpha_{MH}) .$$

Positive values of MH-D DIF indicate items that favor the focal group, whereas negative values of MH-D DIF indicate items that favor the reference group. An item was flagged based on the delta statistic. An item is flagged when:

$$| MHD - DIF | \geq 1.5$$

Note that this procedure is applied to MC items only.

(3) Standardized Mean Difference (SMD)

A standardized mean difference statistic (SMD) was also computed for CR items. The SMD is an effect size index of DIF which is relatively easy to interpret (Zwick et al., 1993). The SMD compares the mean of the reference and focal group, adjusting for the distribution of reference and focal group members on the conditioning variable (Zwick et al., 1993). SMD is computed as (Zwick et al., 1993):

$$\text{ES } SMD = p_{Fk}(\sum_k m_{Fk} - \sum_k m_{Rk}) \quad ,$$

where p_{fk}= proportion of the focal group members at the kth level of the matching variable, m_{Fk}=1/N_{F1k} and m_{Rk}=1/N_{R1k}. A negative SMD value indicates an item on which the focal group has a lower mean than the reference group. A positive SMD value indicates an item on which the reference group has a lower mean than the focal group. An item is flagged when:

$$| ES - SMD | \geq 0.25 .$$

Results

Tables 10-1 to 10-8 show items flagged based on the criteria described above. As indicated, an item flagged for differential item functioning (DIF) is more difficult for a particular group of students than would be expected based on their total test scores. DIF analyses were conducted by gender, ethnicity, English Language Proficiency (ELP), disability status, and socio-economic status. The tables present Linn-Harnisch statistics (D+, D-, and *Z*) first, then the standardized mean difference is presented, and finally the Mantel-Haenszel statistic (Delta) is presented. After specifying these statistics for each item, two final columns provide a summary flag status. There is a column “LH Flag” to indicate where any of the Linn-Harnisch statistics produced a flag, and a “MH Flag” column to indicate where either Mantel-Haenszel or the

Copyright © 2008 by the Wisconsin Department of Public Instruction

standardized mean difference produced a flag.

For the DIF analysis by gender, the reference group is male, meaning that the results for female students are considered with reference to male student performance. For ethnicity, the reference group is white. This means that the performance of other ethnic groups is considered with reference to the performance of white students. No items are flagged for White students. The DIF analysis investigating item-functioning with reference to socioeconomic status defines Economically Disadvantaged students as the focal group, and Not Economically Disadvantaged students as the reference group. The DIF analysis for disability status uses the Not Disabled student population as a reference group to assess DIF within the Disabled student population. The DIF analysis for ELP compares item functioning among students Proficient and Not Proficient in English. Here, the Proficient group is the reference group, and the Not Proficient group is the focal group. MH is only computed for the focal group.

The summary flag information in the DIF tables is always expressed with reference to the focal group. That means that negative flags, (such as - B or - C) indicate that an item disadvantages the focal group, such as female, African American, or Economically Disadvantaged. A positive flag indicates that the item favors the focal group.

Note that a single item can be flagged for multiple NCLB groupings, such as for African American students, and for Hispanic students, and for Economically Disadvantaged students. Also note that the tables, and the analysis below, reflect DIF information for both operational and field test items.

Readers may wish to note that only items that were flagged with a “C” flag were included in the tables below. Readers can see B flagged items in the tables, but that occurs as a result of the fact that those items were also flagged with a “C” flag. The B flag represents a lower threshold for DIF.

Readers should also note that Linn-Harnisch DIF statistics can only be calculated for items with sufficient student N counts. In some cases (as is noted in the DIF table for American Indian students) the size of the tested population was too small to include valid Linn-Harnisch DIF statistics. This kind of result may be expected in some cases, especially for field-test items which are split across forms. Splitting items across forms reduces the number of students exposed to the items, and as a result, in some cases the number of students exposed to a particular item may be too small to produce valid Linn-Harnisch DIF statistics.

The DIF analysis produced flags on both MC and CR items for all NCLB groupings. For example, Table 10-1 shows that a total of eighteen (18) items were flagged for gender. One item was flagged in Reading grade 6, two Reading items were flagged in grades 7 and 10, and five items were flagged in Reading in grade 8. Similarly, in Mathematics, 5 items were flagged for DIF with reference to gender. One Mathematics item was flagged in grades 5, 7, and 8, and two items were flagged in grade 6. In Language Arts, one MC item was flagged in grade 10. Writing prompts were flagged for DIF in grades 4 and 8. Notice that most of these items favored female students and disfavored male students. No items were flagged in Social Studies or Science for gender.

Copyright © 2008 by the Wisconsin Department of Public Instruction

The same interpretation of the DIF flags applies to the other DIF tables. Each table references the grade and content area of the items flagged for DIF. The DIF flags for ethnicity are presented in Tables 10-2 through 10-5, English Proficiency in Table 10-6, Economically Disadvantaged in Table 10-7, and disability status in Table 10-8. Readers may note that some items are flagged by both Linn-Harnisch and Mantel-Haenszel methods, and some only by one of the methods. Also, a relatively large number of items were flagged for Asian students. Note that this student group consisted of a relatively small number of students across grades and content areas compared to other ethnic groups.

DIF studies include a systematic item analysis to determine if examinees with the same underlying level of ability have the same probability of getting the item correct. Items in the Fall 2007 WKCE that were flagged for DIF were notated as such in the item analyses and in the item pool so that content experts will be able to reevaluate these items in future item selection activities. These results have been shared with CTB's Content Development team, and have been made available for reference in future item selections.

10.2 Construct Validity

Construct validity is an indication of how well tests measure the skills or constructs they intend to measure, and it is the central concept underlying the 2007 WKCE assessment validation process. Evidence for construct validity is comprehensive and integrates evidence from both content and criterion-related validity. The WKCE test development was a multi-stage process which included specifications, item writing, review, field testing, and test construction.

In the context of IRT, for a test score to be meaningful for a given content area, the test should be essentially unidimensional. Threats to construct validity include unstandardized testing conditions, the internal factor structure, and unidimensionality of the tests. A test can be said to be unidimensional when all of the items in the test are measures of the same underlying ability or trait. Analyses of the internal structure of a test can indicate the extent to which the relationships among test items and components conform to the construct the test purports to measure. Educational assessments are usually designed to measure a single overall construct or domain (e.g., Reading achievement). Therefore, correlations between scores in content standards within a construct can be expected to be relatively high.

The correlation coefficients here reflect the degree of linear relationship and direction between any two given content standards. The correlation ranges from +1 to -1. A correlation of +1 indicates a perfect positive linear relationship and a correlation of -1 indicates a perfect negative linear relationship between two content standards. A correlation of 0 means there is no linear relationship. In general, the size of the correlation coefficient is influenced by the length of the test, the number of items, or score points. Readers are cautioned not to confuse correlation with causation - the presence of a high correlation between two content standards should not be taken as an indication that there is a causal relationship between them. Tables 10-9 to 10-13 show these correlations among content standards for each content area.

To assess the overall factor structure of the WKCE assessments, exploratory factor analyses were conducted for each content and grade. Factor analysis is a statistical technique

Copyright © 2008 by the Wisconsin Department of Public Instruction

commonly used to evaluate dimensionality. Factor analysis is used to detect simple patterns (dimensionality) within the patterns of relationships among items. The analysis is used to determine if the observed scores in a test can be explained largely or entirely in terms of a much smaller number of traits, called factors or dimensions. To take an example, if answering the mathematics items in a mathematics test involved a lot of reading ability, the mathematics test would not be only a measure of mathematics ability, it would be measuring reading ability as well. Such a test would be said to be multidimensional rather than essentially unidimensional. One way of evaluating the dimensions detected in factor analysis is by looking at the number of large Eigenvalues. The Eigenvalues can be considered as the projections of the data along the Eigenvectors. Previous research shows that the examination of the first two Eigenvalues can be useful in determining the existence of dominant factors. Specifically, where large ratios exist between the first and second Eigenvalues, a single dominant factor can be said to exist. Table 10-14 displays a summary of factor analysis results.

In terms of construct validity then, the presence of an appropriate level of correlation between the content standards and a single dominant factor can confirm that a test is unidimensional. These characteristics are indications of a test that measures what it purports to measure.

As may be observed in Tables 10-9 to 10-13, among all of the content areas, correlations within standards were generally highest in Reading. The correlations in Reading ranged from a low of 0.58 to a high of 0.84. Correlations for standard 4 were often lower than for other standards, but not always. This indicates that in some cases standard 4 may be slightly different than the other standards for Reading. Correlation coefficients in Mathematics tended to be lower than Reading, indicating that standards A-E for Mathematics were more unique than the set of standards in Reading. Correlations among standards for Mathematics ranged from 0.46 to 0.74. Grade 10 showed the highest reliability values. Language Arts had a small number of content standards and the correlations here ranged from 0.38 to 0.68. The D-B correlation in grade 8 and 10 was stronger than the other correlations. Correlations were often stronger in grade 10 Social Studies. The range in grade 4 was 0.48 to 0.65, in grade 8 it was 0.50 to 0.62, but in grade 10, the range jumped to 0.54 to 0.69. In Science, correlations with standard A and D tended to be lower than others in grade 4. Correlations ranged from 0.26 to 0.54 in grade 4. In grade 8 and grade 10, the range was slightly higher. In grade 8 the range was 0.38 to 0.51. The grade C correlation in grade 8 tended to be higher. In grade 10, the range was 0.32 to 0.56, though correlations for standards A, B, and C were often higher.

As may be seen in Table 10-14, the ratios of the first two Eigenvalues range from 2.86 to 14.97. That is, the variance of the first factor is approximately 3 to 15 times larger than the variance of the second largest factor. In Reading, ratios ranged from 8.86 to 14.97. Mathematics ratios were smaller. There, the range was from 2.86 to 12.00. Language Arts ratios ranged from 9.38 to13.93, Social Studies ranged from 7.96 to 9.13, and Science from 7.70 to 14.42.

Analyses of the internal structure of a test can indicate the extent to which the relationships among test items and components conform to the construct the test purports to measure. Educational assessments are usually designed to measure a single overall construct or domain (e.g., Reading achievement). WKCE test items are calibrated using unidimensional IRT

Copyright © 2008 by the Wisconsin Department of Public Instruction

models, which posits the presence of an essentially unidimensional construct underlying a group of test items and components. Within the context of the strength of IRT as a unidimensional model, in general, the ratios shown above can be understood as indicating that the content assessments in the WKCE were sufficiently unidimensional.

10.3 Erasure Analysis

The WKCE test results were subjected to a special program to analyze erasures in multiple-choice items where an incorrect answer choice was erased and replaced with the correct choice. A high rate of erasures can identify situations in which test integrity needs to be examined further. An erasure analysis was performed separately by grade and content area, per school. The list of flagged schools was released to DPI for evaluation.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Part 11: Summary and Recommendations

Results and key findings of the Fall 2007 WKCE tests are presented throughout the body of this report. Some issues of a technical nature that may require further attention in subsequent administrations are presented below.

1) During the item selection process for the 2008 assessments, CTB's Content Development noted that there were a relatively small number of quality items available in the item pool for some content standards. A review of the item pools, an identification of the content standards that may need additional quality items, and a proposal for a field-test plan in succeeding administrations may be recommended.

2) The percentage of students at the lowest obtainable scale score (LOSS) and the highest obtainable scale score (HOSS) was close to 1% for most grades and content areas. For grade 4 Social Studies, the percent at HOSS was close to 2%. The HOSS (possible ceiling effect) for this grade should be carefully tracked in the next administration.

3) Since WKCE Mathematics and Reading tests are in a vertical scale, longitudinal comparisons, such as the mean and standard deviation of scale scores and the percentage of students in each proficiency level, in 2005, 2006, and 2007, could provide a useful view of how different cohorts are performing across administrations. These longitudinal results would also allow an assessment of annual progress at the state level.

4) For handscored items (CR items), condition codes were assigned to student responses if their responses fell within the following categories: "A" denoted no response or no attempt, "B" represented illegible responses, "C" indicated another language, and "D" denoted a response that was off-topic. The percentage of students who received condition codes was generally small, but in some cases the percentage was larger. The items with a higher proportion of condition codes may require further investigation in terms of item position, speededness, or any other possible causes.

Copyright © 2008 by the Wisconsin Department of Public Instruction

References

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.

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

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

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

Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: some uses, misuses, and alternatives. *Educational and Psychological Measurement, 41*, 687-699.

Burket, G. R. (1991). *PARDUX* [Computer program]. Unpublished.

Burket, G. R. (2000). *ITEMWIN* [Computer program]. Unpublished.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46.

Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory.* Belmont, CA: Wadsworth Group/Thompson Learning.

Cronbach, L. J., Schönemann, P., & McKie, D. (1965). Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement, 25*, 291–312.

CTB/McGraw-Hill. (1997). TerraNova 1st Edition. Monterey, CA: Author.

CTB/McGraw-Hill. (2000). TerraNova 2nd Edition. Monterey, CA: Author.

CTB/McGraw-Hill. (1997–2001). WKCE. Monterey, CA: Author.

Fitzpatrick, A. R. (1991). *Status report on the results of preliminary analyses of dichotomous and multi-level items using the PARMATE program*. Monterey, CA: CTB/McGraw-Hill.

Fitzpatrick, A. R., & Julian, M. W. (1996*). Two studies comparing the parameter estimates produced by PARDUX and PARSCALE*. Unpublished manuscript.

Hambleton, R. K., and Novick. M. R. (1973). Toward an integration of theory and

Copyright © 2008 by the Wisconsin Department of Public Instruction

method for criterion-referenced tests. *Journal of Educational Measurement*, *10*, 159-96.

Kolen, M and Kim, D. (2004). Personal Correspondence

Kolen, M. J., & Brennan, R. L. (1995). *Test equating: Methods and Practices*. New York: Springer-Verlag.

Linn, R. L. (1989). *Educational measurement*. (3rd ed). New York: Macmillan.

Linn, R. L., & Harnisch, D. (1981). Interactions between item content and group membership in achievement test items. *Journal of Educational Measurement*, *18*, 109–118.

Linn, R. L., Levine, M. V., Hastings, C. N., & Wardrop, J. L. (1981). Item bias in a test of reading comprehension. *Applied Psychological Measurement*, *5*, 159–173.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, *32*, 179–197.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems* (71, 179–181). Hillsdale, NJ: Lawrence Erlbaum.

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

Patz, R. (1998). *Calculating Handscoring Reliability Coefficients.* Unpublished manuscript.

Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, *14*, 59–71.

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

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

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

Swaminathan, H., Hambleton, R. K., and Algina, J. (1974). Reliability of criterion referenced tests: A decision theoretic formulation. *Journal of Educational Measurement*, *11*, 263–268.

Thissen, D. (1982). Marginal maximum-likelihood estimation for the one-parameter

Copyright © 2008 by the Wisconsin Department of Public Instruction

logistic model. *Psychometrika*, *47*, 175–186.

Thissen, D. (1990). *MULTILOG: Multiple categorical item analysis and test scoring*. Version 6 [Computer program]. Mooresville, IN: Scientific Software.

van der Linden, W.J. & Hambleton, R. (1997). *Handbook of modern item response theory*. New York: Springer.

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

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

Yen, W. M. (1984). Obtaining maximum likelihood trait estimates from number-correct scores for the three-parameter logistic model. *Journal of Educational Measurement*, *21*, 93–111.

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

Yen, W. M. & Burket, G.R. (1997). Comparison of item response theory and Thurstone methods of vertical scaling. *Journal of Educational Measurement*, *34*, 293–313.

Yen, W.M. & Candell, G.L. (1991). Increasing score reliability with item-pattern scoring: An empirical study in five score metrics. *Applied Measurement in Education*, *4*, 209–228.

Zwick, R., Donoghue, J.R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, *30*, 233–251.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-1
Target Reading Test Blueprint: Grades 3–8, 10

	Category Title	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 10	
		MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
1	Determines meaning of words or phrases in context	12		11		11		11		11		11		7	
1.1	Uses context clues to determine meaning of words or phrases	8		8		7		7		7		7			
1.2	Uses knowledge of word structure to determine meaning of words	2		2		2		2		2		2			
1.3	Uses word reference materials to determine meaning of words and phrases	2		1		2		2		2		2			
2	Understands Text	17		17		15		14		14		14		7	
2.1	Demonstrates understanding of literal meaning by identifying stated information in literary text	7		7		6		6		6		6			
2.2	Demonstrates understanding of literal meaning by identifying stated information in informational text	7		8		6		6		6		6			
2.3	Demonstrates understanding of explicitly stated sequence of events in literary and informational text	3		2		3		2		2		2			
3	Analyzes Text	21	1	21	1	20	1	18	1	18	1	18	1	22	1
3.1	Analyzes literary text	10		10		9		8		8		8			
3.2	Analyzes informational text.	8		8		8		6		6		6			
33	Analyzes author's use of language in literary and informational text.	3		3		3		4		4		4			
4	Evaluates and Extends Text	4	1	5	1	11	1	11	1	11	1	11	1	14	1
4.1	Evaluates and extends literary text	2		2		3		3		3		3			
4.2	Evaluates and extends informational text	1		2		5		5		5		5			
4.3	Evaluates and extends author's use of language in literary and informational text	1		1		3		3		3		3			
	Number of Items	54	2	54	2	54	2	54	2	54	2	54	2	50	2
	Total Score Points for Test	60		60		60		60		60		60		56	

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-2
Target Mathematics Test Blueprint: Grades 3–8, 10

	Category Title	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 10	
		MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
A	Mathematical Processes	3	3	3	3	3	3	3	3	3	3	3	3	7	1
Aa	Reasoning														
Ab	Communication														
Ac	Connections														
Ad	Representation														
Ae	Problem Solving														
B	Number Operations and Relationships	11	1	11	0	11	0	12	0	12	0	7	0	7	0
Ba	Number Concepts	6		5		5		6		6		4		4	
Bb	Number Computation	5		6		6		6		6		3		3	
C	Geometry	9	1	8	1	9	1	9	1	10	2	8	1	8	1
Ca	Describing Figures	4		3		3		2		3		2		4	
Cb	Spatial Relationships and Transformations	4		4		4		4		4		4		2	
Cc	Coordinate System	1		1		2		3		3		2		2	
D	Measurement	8	0	8	1	9	1	9	1	9	0	11	1	9	1
Da	Measurable Attributes	3		3		4		2		3		2		1	
Db	Direct Measurement	4		4		3		3		3		3		2	
Dc	Indirect Measurement	1		1		2		4		3		6		6	
E	Statistics and Probability	7	1	7	1	9	1	8	1	8	1	8	1	9	0
Ea	Data Analysis and Statistics	5		4		6		5		5		5		4	
Eb	Probability	2		3		3		3		3		3		5	
F	Algebraic Relationships	8	1	9	1	10	1	10	1	9	1	14	1	10	1
Fa	Patterns, Relations, and Functions	4		5		5		5		2		7		5	
Fb	Expressions, Equations, and Inequalities	2		2		3		2		3		6		4	
Fc	Properties	2		2		2		3		4		1		1	
	Number of Items	46	4	46	4	51	4	51	4	51	4	51	4	50	4
	Total Score Points for Test	57		57		62		62		62		62		58	

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-3
Target Language Arts Test Blueprint: Grades 4, 8, 10

Content Standard		Grade 4		Grade 8		Grade 10	
		MC	Prompt	MC	Prompt	MC	Prompt
B	Writing	19	1	18	1	15	1
D	Language	5		6		9	
F	Research and Inquiry	6		6		6	
	Number of Items	30	1	30	1	30	1
	Total Number of Points	30	9	30	9	30	9

Table 2-4
Target Science Test Blueprint: Grades 4, 8, 10

Content Standard		Grade 4	Grade 8	Grade 10
A	Science Connections	4	3	5
B	Nature of Science	4	4	5
C	Science Inquiry	7	8	9
D	Physical Science	6	6	8
E	Earth and Space	6	6	6
F	Life and Environment	6	6	6
G	Science Applications	4	4	5
H	Personal/Social Perspectives	3	3	6
	Total Number of MC Items	40	40	50

*Note: Standard A, Science Connections, and Standard B, Nature of Science, are combined to form a reporting category; Standard G, Science Applications, and Standard H, Personal/Social Perspectives, are combined to form a reporting category.

Table 2-5
Target Social Studies Test Blueprint: Grades 4, 8, 10

Content Standard		Grade 4	Grade 8	Grade 10
A	Geography	9	10	10
B	History	8	13	12
C	Political Science	7	6	12
D	Economics	7	6	8
E	Behavioral Science	7	5	8
	Total Number of MC Items	38	40	50

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-6
Actual Reading Test Blueprint: Grades 3–8, 10

	Category Title	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade10	
		MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
1	Determines meaning of words or phrases in context	12		11		11		11		11		11		7	
1.1	Uses context clues to determine meaning of words or phrases	6		7		8		5		7		8		6	
1.2	Uses knowledge of word structure to determine meaning of words	4		2		1		4		2		3			
1.3	Uses word reference materials to determine meaning of words and phrases	2		2		2		2		2		0		1	
2	Understands Text	17		17		15		14		14		14		7	
2.1	Demonstrates understanding of literal meaning by identifying stated information in literary text	6		7		6		6		5		3		2	
2.2	Demonstrates understanding of literal meaning by identifying stated information in informational text	10		6		7		5		5		8		5	
2.3	Demonstrates understanding of explicitly stated sequence of events in literary and informational text	1		4		2		3		4		3			
3	Analyzes Text	21	1	21	1	20	1	18	1	18	1	18	1	22	1
3.1	Analyzes literary text	8		9		5		8		6		9		11	
3.2	Analyzes informational text.	8		7		11		6		8		6		3	
3.3	Analyzes author's use of language in literary and informational text.	5		5		4		4		4		3		8	
4	Evaluates and Extends Text	4	1	5	1	8	1	11	1	11	1	11	1	14	1
4.1	Evaluates and extends literary text	1		1		2		5		2		4		5	
4.2	Evaluates and extends informational text	1		2		4		1		5		4		6	
4.3	Evaluates and extends author's use of language in literary and informational text	2		2		2		5		4		3		3	

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-6 Cont'd
Actual Reading Test Blueprint: Grades 3–8, 10

	Category Title	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade10	
		MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
	Number of Items	54	2	54	2	54	2	54	2	54	2	54	2	50	2
	Total Score Points for Test	60		60		60		60		60		60		56	

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-7
Actual Mathematics Test Blueprint: Grades 3–8, 10

	Category Title	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 10	
		MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
A	Mathematical Processes	3	3	3	3	3	3	3	3	3	3	3	3	7	1
Aa	Reasoning														
Ab	Communication														
Ac	Connections														
Ad	Representation														
Ae	Problem Solving														
B	Number Operations and Relationships	11	1	11	0	11	0	12	0	12	0	7	0	7	0
Ba	Number Concepts	6		5		5		6		6		4		4	
Bb	Number Computation	5		6		6		6		6		3		3	
C	Geometry	9	1	8	1	9	1	9	1	10	2	8	1	8	1
Ca	Describing Figures	4		3		3		2		3		2		4	
Cb	Spatial Relationships and Transformations	4		4		4		4		4		4		2	
Cc	Coordinate System	1		1		2		3		3		2		2	
D	Measurement	8	0	8	1	9	1	9	1	9	0	11	1	9	1
Da	Measurable Attributes	3		3		4		2		3		2		1	
Db	Direct Measurement	4		4		3		3		3		3		2	
Dc	Indirect Measurement	1		1		2		4		3		6		6	
E	Statistics and Probability	7	1	7	1	9	1	8	1	8	1	8	1	9	0
Ea	Data Analysis and Statistics	5		4		6		5		5		5		4	
Eb	Probability	2		3		3		3		3		3		5	
F	Algebraic Relationships	8	1	9	1	10	1	10	1	9	1	14	1	10	1
Fa	Patterns, Relations, and Functions	4		5		5		5		2		7		5	
Fb	Expressions, Equations, and Inequalities	2		2		3		2		3		6		4	
Fc	Properties	2		2		2		3		4		1		1	
	Number of Items	46	4	46	4	51	4	51	4	51	4	51	4	50	4
	Total Score Points for Test	57		57		62		62		62		62		58	

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-8
Actual Language Arts Test Blueprint: Grades 4, 8, 10

Content Standard		Grade 4		Grade 8		Grade 10	
		MC	Prompt	MC	Prompt	MC	Prompt
B	Writing	19	1	18	1	15	1
D	Language	5		6		9	
F	Research and Inquiry	6		6		6	
	Total Number of Items	30	1	30	1	30	1
	Total Number of Points	30	9	30	9	30	9

Table 2-9
Actual Science Test Blueprint: Grades 4, 8, 10

Content Standard		Grade 4	Grade 8	Grade 10
A	Science Connections	4	3	5
B	Nature of Science	4	4	5
C	Science Inquiry	7	8	9
D	Physical Science	6	6	8
E	Earth and Space	6	6	6
F	Life and Environment	6	6	6
G	Science Applications	4	4	5
H	Personal/Social Perspectives	3	3	6
	Total Number of MC Items	40	40	50

*Note: Standard A, Science Connections, and Standard B, Nature of Science, are combined to form a reporting category; Standard G, Science Applications, and Standard H, Personal/Social Perspectives, are combined to form a reporting category.

Table 2-10
Actual Social Studies Test Blueprint: Grades 4, 8, 10

Content Standard		Grade 4	Grade 8	Grade 10
A	Geography	9	10	10
B	History	8	11	12
C	Political Science	7	6	12
D	Economics	7	8	8
E	Behavioral Science	7	5	8
	Total Number of MC Items	38	40	50

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-11
Reading: 2007 Item Development Plan

Grade	Reporting Category								Total	
	1		2		3		4			
	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
3	12		7		15	2	4	2	38	4
4	12		7		15	2	5	2	39	4
5	7		4		12	2	5	2	28	4
6	5		2		13	2	4	2	24	4
7	5		2		13	2	7	2	27	4
8	5		2		18	2	6	2	31	4
Total									187	24

Table 2-12
Mathematics: 2007 Item Development Plan

Grade	Reporting Category												Total	
	A		B		C		D		E		F			
	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
3	6	0	0	2	3	2	3		8	3		1	20	8
4	8		0		3	3	2	2	10	3		0	23	8
5	6	0	2		4	3	3	1	9	2	2	3	26	9
6	4	0	2		2	3	5	2	9	3	3	1	25	9
7	5	0	2		6	3	5		7	4	5	2	30	9
8	3	0	0		8	3	8	3	5	1	6	3	30	10
Total													154	53

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-13
Reading 2007 Item Development

Grade	Items Brought to Review Meeting	Items Written at Review Meeting	Total Items Reviewed
3	56	0	56
4	58	0	58
5	48	0	48
6	51	0	51
7	54	0	54
8	48	0	48
Reading Total	315	0	315

Table 2-14
Mathematics 2007 Item Development

Grade	Items Brought to Review Meeting	Items Written at Review Meeting	Total Items Reviewed
3	37	2	39
4	34	7	41
5	40	1	41
6	40	2	42
7	41	0	41
8	49	0	49
Mathematics Total	241	12	253

Table 2-15
Reading: 2007 Item Development by Reporting Category and Item Format

Grade	Reporting Category								Total	
	1		2		3		4			
	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
3	10		10		26	1	6	3	52	4
4	7		7		26	1	14	3	54	4
5	6		5		26	1	7	3	44	4
6	7		6	1	23	2	10	2	46	5
7	12		7		23	2	8	2	50	4
8	5		11		20		8	4	44	4
Total	47		46		144	7	53	17	139	25

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-16
Mathematics: 2007 Item Development by Reporting Category and Item Format

Grade	Reporting Category												Total	
	A		B		C		D		E		F			
	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
3	6	4	6	2	4	2	2		9	2	1	1	28	11
4	6	5	3		3	3	2	2	12	2	2	1	28	13
5	5	4	2		4	3	3	1	10	2	4	3	28	13
6	3	3	6		2	3	4	2	10	3	5	1	30	12
7	4	3	3		6	3	6		7	2	5	2	31	10
8	5	6	1		8	6	7	3	4	1	7	1	32	17
Total	29	25	21	2	27	20	24	8	52	12	24	9	177	76

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-17
Item Development Each Year and Total to Date

	MC items for 2004	CR items for 2004	MC items for 2005	CR items for 2005	MC items for 2006	CR items for 2006	MC items for 2007	CR items for 2007	MC items for 2008	CR items for 2008
Grade 3										
Reading	411	52	23	2	30	4	40	3	52	4
Math	317	36	33	14	18	2	30	4	28	11
Total	728	88	56	16	48	6	70	7	80	15
Grade 4										
Reading	380	56	32	3	34	3	25	4	54	4
Math	265	35	45	9	29	1	26	4	28	13
Language Arts	0	0	0	10	0	0	0	0	0	0
Science	0	0	0	0	123	34	0	0	0	0
Total	645	91	77	22	123	34	0	0	82	17
Grade 5										
Reading	433	59	36	6	29	5	29	7	44	4
Math	305	49	38	11	26	3	30	5	28	13
Total	738	108	74	17	55	8	59	12	72	17
Grade 6										
Reading	511	56	32	5	42	5	37	6	46	5
Math	310	41	53	16	7	2	28	4	30	12
Total	821	97	85	21	49	7	65	10	76	17
Grade 7										
Reading	359	44	35	4	38	4	25	5	50	4
Math	305	34	32	23	20	0	28	4	31	10
Total	664	78	67	27	58	4	53	9	81	14
Grade 8										
Reading	365	44	30	4	34	4	25	4	44	4
Math	289	51	47	25	20	2	28	4	32	17
Language Arts	0	0	0	10	0	0	0	0	0	0
Science	0	0	0	0	125	34	0	0	0	0
Total	654	95	77	39	125	34	0	0	76	21
Grade 10										
Reading	0	0	0	0	0	0	0	0	0	0
Math	0	0	0	0	0	0	0	0	0	0
Language Arts	0	0	0	0	0	0	0	0	0	0
Science	0	0	0	0	18	8	0	0	0	0
Total	0	0	0	0	18	8	0	0	0	0
TOTALS										
Reading	2,459	255	188	24	207	25	181	25	290	25
Mathematics	1,791	246	248	98	120	10	170	25	177	76
Language Arts	0	0	0	20	0	0	0	0	0	0
Science	0	0	0	0	266	76	0	0	0	0
Grand Total	4,250	557	436	142	476	101	247	38	177	76

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-18
Reading: 2007 Item Selection Review Results

Grade	Subskill	Subskill Indicator	Item Format	DoK	Accepted: No Revision	Accept w/ Revisions	Rejected	Grand Total
3	1	1.1	MC	2	5	4		9
		1.2	MC	1		1		1
	1 Total				5	5		10
	2	2.1	MC	1		3		3
				2	1			1
		2.2	MC	1		1		1
				2		1		1
		2.3	MC	1		2		2
				2	2			2
	2 Total				3	7		10
	3	3.1	BCR	3		1		1
			MC	2		11		11
				3	3	4		7
		3.2	MC	2		3		3
		3.3	MC	2		3		3
				3	1	1		2
	3 Total				4	23		27
	4	4.1	BCR	3		2		2
			MC	3		4		4
		4.2	BCR	3		1		1
			MC	3		1		1
		4.3	MC	3		1		1
	4 Total					9		9
3 Total					12	44		56
4	1	1.1	MC	2	5	2		7
	1 Total				5	2		7
	2	2.1	MC	1		1		1
		2.2	MC	1		3		3
				2		1		1
		2.3	MC	1		1		1
				2		1		1
	2 Total					7		7
	3	3.1	MC	2	2	9		11
				3	2			2
		3.2	MC	2		6		6
				3	2			2
		3.3	BCR	3		1		1
			MC	2		5		5
4	3 Total				6	21		27

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-18
Reading: 2007 Item Selection Review Results

Grade	Subskill	Subskill Indicator	Item Format	DoK	Accepted: No Revision	Accept w/ Revisions	Rejected	Grand Total
	4	4.1	BCR	3		1		1
			MC	2		1		1
				3		1		1
		4.2	BCR	3		2		2
			MC	2		4		4
				3		8		8
	4 Total					17		17
4 Total					11	47		58
5	1	1.1	MC	1	1	3		4
				2	1	1		2
	1 Total				2	4		6
	2	2.1	MC	2	1	2		3
		2.3	MC	2	1			1
				(blank)			1	1
	2 Total				2	2	1	5
	3	3.1	BCR	3		1		1
			MC	2	1	4		5
				3	2	4		6
		3.2	MC	2		9		9
				3		2		2
		3.3	MC	2	1	2		3
				3		1		1
	3 Total				4	23		27
	4	4.1	MC	3	1	1		2
		4.2	BCR	3	1			1
			MC	3		3		3
		4.3	BCR	3		2		2
			MC	3	1	1		2
	4 Total				3	7		10
5 Total					11	36	1	48
6	1	1.1	MC	1	5	2		7
	1 Total				5	2		7
	2	2.1	MC	1		1		1
		2.2	BCR	2		1		1
			MC	1		1		1
				2	1	1		2
		2.3	MC	1		1		1
				2		1		1
	2 Total				1	6		7

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-18
Reading: 2007 Item Selection Review Results

Grade	Subskill	Subskill Indicator	Item Format	DoK	Accepted: No Revision	Accept w/ Revisions	Rejected	Grand Total
	3	3.1	BCR	(blank)			1	1
6			MC	2	1	8		9
				3		3		3
		3.2	BCR	3		1		1
			MC	2	2	5		7
				3		1		1
				(blank)			1	1
		3.3	MC	2		1		1
				3		1		1
	3 Total				3	20	2	25
	4	4.1	BCR	3	1			1
			MC	3		2		2
		4.2	BCR	3		1		1
			MC	2		1		1
				3	1	2		3
		4.3	MC	2		3		3
				3		1		1
	4 Total				2	10		12
6 Total					11	38	2	51
7	1	1.1	MC	1	2	1		3
				2	1	1		2
		1.2	MC	1	1			1
				2		2		2
		1.3	MC	1	1	3		4
	1 Total				5	7		12
	2	2.1	MC	1	3	3		6
				2	1			1
	2 Total				4	3		7
	3	2.1	MC	1		1		1
		3.1	BCR	3		1		1
			MC	2		13		13
				3	1	2		3
		3.2	MC	2		2		2
				3	1			1
		3.3	BCR	3		1		1
			MC	1		1		1
				2		2		2
	3 Total				2	23		25
	4	4.1	BCR	3	2			2

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-18
Reading: 2007 Item Selection Review Results

Grade	Subskill	Subskill Indicator	Item Format	DoK	Accepted: No Revision	Accept w/ Revisions	Rejected	Grand Total
			MC	2		1	1	2
				3	2	3		5
		4.2	MC	3		1		1
7	4 Total				4	5	1	10
7 Total					15	38	1	54
8	1	1.1	MC	1	1			1
		1.2	MC	2		2		2
		1.3	MC	1		2		2
	1 Total				1	4		5
	2	2.1	MC	1		4		4
		2.2	MC	1	1	4		5
		2.3	MC	1	2			2
	2 Total				3	8		11
	3	3.1	MC	1		1		1
				2		1		1
				3	2	3	1	6
		3.2	MC	2		7		7
				3	2	3		5
	3 Total				4	15	1	20
	4	4.1	BCR	3		1		1
			MC	3	2			2
		4.2	BCR	3		3		3
			MC	2		2		2
				3		2		2
		4.3	MC	2		1		1
				3		1		1
	4 Total				2	10		12
8 Total					10	37	1	48
Grand Total					70	240	5	315

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-19
Mathematics: 2007 Item Selection Review Results

Grade	Reporting Category	Subskill	Item Format	DoK	Accepted: No Revision	Accepted: w/Revisions	Rejected	Grand Total
3	A	Aa	MC	1	1			1
				2	1	2		3
				3		1		1
			2pt-CR	3		1		1
			B-BCR	3		3		3
			MC	2		1		1
	A Total				2	8		10
	B	Ba	2pt-CR	3		1		1
			MC	1		1		1
				2		2		2
		Bb	A-BCR	2		1		1
			MC	1	2	1		3
	B Total				2	6		8
	C	Ca	MC	2	1	3		4
		Cb	2pt-CR	2		1		1
		Cc	2pt-CR	2		1		1
	C Total				1	5		6
	D	Dc	MC	2		2		2
	D Total					2		2
	E	Ea	A-BCR	3		1		1
			MC	2		2		2
				3	1	3		4
		Eb	A-BCR	2		1		1
			MC	3	3			3
	E Total				4	7		11
	F	Fb	2pt-CR	2		1		1
			MC	2	1			1
	F Total				1	1		2
3 Total					10	29		39
4	A	Aa	B-BCR	2	1	2		3
				3		2		2
			MC	2	1	1		2
				3		4		4
	A Total				2	9		11
	B	Ba	MC	1		2		2
		Bb	MC	2		1		1
	B Total					3		3
	C	Ca	MC	2	1			1
		Cb	2pt-CR	3		1		1

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-19
Mathematics: 2007 Item Selection Review Results

Grade	Reporting Category	Subskill	Item Format	DoK	Accepted: No Revision	Accepted: w/Revisions	Rejected	Grand Total
			A-BCR	2	1			1
			MC	2		1		1
		Cc	2pt-CR	3			1	1
			MC	1		1		1
	C Total				2	3	1	6
	D	Dc	A-BCR	2		2		2
			MC	2		2		2
	D Total					4		4
	E	Ea	2pt-CR	(blank)			1	1
			MC	1	1			1
				2		1		1
				3		5		5
		Eb	A-BCR	3		1		1
			MC	2		2		2
				3	1	2		3
	E Total				2	11	1	14
	F	Fa	A-BCR	1		1		1
			MC	2		1		1
		Fc	MC	2	1			1
	F Total				1	2		3
4 Total					7	32	2	41
5	A		B-BCR	2	1			1
				3	1	2		3
			MC	1		1		1
				2	2			2
				3		2		2
	A Total				4	5		9
	B	Bb	MC	2	1	1		2
	B Total				1	1		2
	C	Ca	2pt-CR	3		1		1
		Cb	2pt-CR	2		1		1
			A-BCR	2		1		1
		Cc	MC	1		1		1
				2	1	2		3
	C Total				1	6		7
	D	Dc	A-BCR	2		1		1
			MC	2	1	1	1	3
	D Total				1	2	1	4
	E	Ea	2pt-CR	3		1		1
			A-BCR	2	1			1
			MC	3		4		4

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-19
Mathematics: 2007 Item Selection Review Results

Grade	Reporting Category	Subskill	Item Format	DoK	Accepted: No Revision	Accepted: w/Revisions	Rejected	Grand Total
		Eb	MC	2	1	5		6
	E Total				2	10		12
	F	Fa	2pt-CR	3	1			1
			MC	2	1			1
				3	1			1
		Fb	2pt-CR	2		1		1
			MC	2	1	1		2
		Fc	A-BCR	2	1			1
	F Total				5	2		7
5 Total					14	26	1	41
	A		B-BCR	3		3		3
			MC	3		2	1	3
	A Total					5	1	6
	B	Ba	MC	2	1	2		3
				3	1			1
		Bb	MC	2		2		2
	B Total				2	4		6
	C	Ca	2pt-CR	2		1		1
			A-BCR	2		1		1
			MC	1	1			1
		Cc	2pt-CR	3		1		1
			MC	2		1		1
	C Total				1	4		5
	D	Da	2pt-CR	2		1		1
			A-BCR	2		1		1
6			MC	2	1			1
		Db	MC	3		1		1
		Dc	MC	2	1	1		2
	D Total				2	4		6
	E	Ea	MC	3		7		7
		Eb	2pt-CR	2		1		1
				3	1			1
			A-BCR	2		1		1
			MC	2	1	1		2
				3		1		1
	E Total				2	11		13
	F	Fa	2pt-CR	3		1		1
		Fb	MC	2	2			2
				3	1			1
		Fc	MC	2	1	1		2
	F Total				4	2		6

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-19
Mathematics: 2007 Item Selection Review Results

Grade	Reporting Category	Subskill	Item Format	DoK	Accepted: No Revision	Accepted: w/Revisions	Rejected	Grand Total
6 Total					11	30	1	42
7	A		MC	1		1		1
				2		3		3
			B-BCR	2		1		1
				3		1	1	2
	A Total					6	1	7
	B	Ba	MC	1	1	1		2
		Bb	MC	1	1			1
	B Total				2	1		3
	C	Ca	2pt-CR	1		1		1
			MC	1		2		2
		Cb	2pt-CR	1		1		1
			MC	2	1			1
		Cc	A-BCR	2		1		1
			MC	2		3		3
	C Total				1	8		9
	D	Da	MC	2	1			1
		Db	MC	1	1	1		2
				2	2			2
		Dc	MC	2		1		1
	D Total				4	2		6
	E	Ea	A-BCR	3			1	1
			MC	1		1		1
				3		2		2
		Eb	A-BCR	3		1		1
			MC	2		2		2
				3		2		2
	E Total					8	1	9
	F	Fa	MC	2	1			1
				3		1		1
		Fb	2pt-CR	2		1		1
			MC	1		1		1
				2		2		2
		Fc	2pt-CR	1	1			1
	F Total				2	5		7
7 Total					9	30	2	41
8	A		2pt-CR	3	2			2
			MC	1	1			1
				2		1		1
				3		2	1	3

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 2-19
Mathematics: 2007 Item Selection Review Results

Grade	Reporting Category	Subskill	Item Format	DoK	Accepted: No Revision	Accepted: w/Revisions	Rejected	Grand Total
			B-BCR	3	1	3		4
	A Total				4	6	1	11
	B	Ba	MC	2	1			1
	B Total				1			1
	C	Ca	2pt-CR	2		2		2
			A-BCR	1		2		2
			MC	1	1			1
		Cb	2pt-CR	2		2		2
			MC	1	1			1
				2	2	3		5
		Cc	MC	2	1			1
	C Total				5	9		14
	D	Da	A-BCR	2	1			1
			MC	1	1	1		2
		Db	2pt-CR	2	2			2
			MC	1	1			1
		Dc	MC	1		1		1
				2	2	1		3
	D Total				7	3		10
	E	Ea	MC	2	3			3
		Eb	A-BCR	1		1		1
			MC	3	1			1
	E Total				4	1		5
	F	Fb	A-BCR	2		1		1
			MC	2	2	1		3
		Fc	MC	1	1	1		2
				2		2		2
	F Total				3	5		8
8 Total					24	24	1	49
Grand Total					75	171	7	253

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 3-1
Fall 2007 Test Configuration for Operational (OP) Items

Content	Grade	No. of OP MC Items	No. of OP CR Items					Total Score Point	Total OP (MC + CR) Items
			1 point	2 point	3 point	4 point	6 point		
Reading	3	54			2			60	56
	4	54			2			60	56
	5	54			2			60	56
	6	54			2			60	56
	7	54			2			60	56
	8	54			2			60	56
	10	50			2			56	52
Math	3	46	3	4				57	53
	4	46	3	4				57	53
	5	51	3	4				62	58
	6	51	3	4				62	58
	7	51	3	4				62	58
	8	51	3	4				62	58
	10	50	0	4				58	54
Language Arts*	4	30						30	30
	8	30						30	30
	10	30			1		1	39	32
Social Studies	4	38						38	38
	8	40						40	40
	10	50						50	50
Science	4	40						40	40
	8	40						40	40
	10	50						50	50

*For Language Arts Grade 10, the two CR items are from the Writing Grade 10 prompts.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 3-2
Reading: 2007 Embedded Field Test Items

Grade	Objective	1	2	3		4		Total	
	Form	MC	MC	MC	CR	MC	CR	MC	CR
3	Form A	2	4	2		2	1	10	1
	Form B	1	3	3		3	1	10	1
	Form C	1	2	6		1	1	10	1
	Form D	4	3	3	1			10	1
3 Total		8	12	14	1	6	3	40	4
4	Form A		1	7		2	1	10	1
	Form B	1	1	6	1	2		10	1
	Form C		1	7	1	1		9	1
	Form D	3	1	3	1	1		8	1
4 Total		4	4	23	3	6	1	37	4
5	Form A	1	1	7	1	1		10	1
	Form B	3	1	4	1	2		10	1
	Form C			3	1	1	1	4	2
	Form D	1		3	2	1		5	2
5 Total		5	2	17	5	5	1	29	6
6	Form A	2		6		1	1	9	1
	Form B	3		6		1	1	10	1
	Form C		1	5	2			6	2
	Form D		1	8		1	1	10	1
6 Total		5	2	25	2	3	3	35	5
7	Form A		1	5		1	1	7	1
	Form B	1		7	1	1		9	1
	Form C		1	7		1	1	9	1
	Form D		3	4	1	2		9	1
7 Total		1	5	23	2	5	2	34	4
8	Form A	1	4	4	1			9	1
	Form B	1	2	4		3	1	10	1
	Form C		2	5	1	1		8	1
	Form D	1		4			2	5	2
8 Total		3	8	17	2	4	3	32	5
Reading Total		26	33	119	15	29	13	207	28

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 3-3
Mathematics: 2007 Embedded Field Test Items

Grade		A		B		C		D		E		F		Total	
	Form	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
3	Form A		1			3		1		3	1		1	7	3
	Form B	1	1		1	3		1		2			1	7	3
	Form C	2				3		1		4			1	10	1
	Form D	2	1	1		3	1	1		3				10	2
3 Total		5	3	1	1	12	1	4		12	1		3	34	9
4	Form A	1	1	1		1		1		3	1		1	7	3
	Form B	1	1	2		2		2	1	2		1		10	2
	Form C	2	1	3		2		1		2	1			10	2
	Form D			2		2		1		3		2	1	10	1
4 Total		4	3	8		7		5	1	10	2	3	2	37	8
5	Form A	1		2		1	1			4		2		10	1
	Form B	2	1	2		1	1			3		2		10	2
	Form C	1		1		1		1		4	1	2		10	1
	Form D	1	1	3		1		1	1	3		1		10	2
5 Total		5	2	8		4	2	2	1	14	1	7		40	6
6	Form A		1	2		2				3		1	1	8	2
	Form B	1	1	2		1				3		1	1	8	2
	Form C	1	1	2		1		1	1	2		1		8	2
	Form D			2		2	1			3		1		8	1
6 Total		2	3	8		6	1	1	1	11		4	2	32	7
7	Form A	1		2		1				3		1	1	8	1
	Form B	1	2	4		1				2				8	2
	Form C	1	1	2		1		1		3				8	1
	Form D	1		3		2				2			1	8	1
7 Total		4	3	11		5		1		10		1	2	32	5
8	Form A	1		1		1				3	1	2		8	1
	Form B	1		1						3	1	3		8	1
	Form C	2	1	1						3	1	2		8	2
	Form D	1	1	1		1		2		2	1	1		8	2
8 Total		5	2	4		2		2		11	4	8		32	6
Total		25	16	40	1	36	4	15	3	68	8	23	9	207	41

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 3-4
Unique Items Field Tested Each Year and Total to Date

Items Field-Tested by Item Type and Year										
	MC 2004	**CR 2004**	**MC 2005**	**CR 2005**	**MC 2006**	**CR 2006**	**MC 2007**	**CR 2007**	**Total MC to Date**	**Total CR to Date**
Grade 3										
Reading	242	12	24	2	27	2	40	4	333	20
Math	252	24	15	2	32	4	34	5	333	35
Total	494	36	39	4	59	6	74	9	666	55
Grade 4										
Reading	294	12	24	2	32	3	40	4	390	21
Math	231	29	15	2	32	4	34	4	312	39
Language Arts	0	0	0	6	0	0	0	0	0	6
Science	0	0	0	0	40	0	0	0	40	0
Social Studies	0	0	0	0	0	0	0	0	0	0
Total	525	41	39	10	104	7	74	8	742	66
Grade 5										
Reading	235	14	24	2	28	2	29	6	316	24
Math	257	34	15	2	32	4	40	4	344	44
Total	492	48	39	4	60	6	69	10	660	68
Grade 6										
Reading	259	14	24	1	33	3	35	5	351	23
Math	252	33	15	2	32	4	32	4	331	43
Total	511	47	39	3	65	7	67	9	682	66
Grade 7										
Reading	259	14	24	1	17	2	35	4	335	21
Math	243	33	15	2	32	4	32	3	322	42
Total	502	47	39	3	49	6	67	7	657	63
Grade 8										
Reading	274	14	24	1	33	4	32	5	363	24
Math	234	33	15	2	40	4	32	4	321	43
Language Arts	0	0	0	6	0	0	0	0	0	6
Science	0	0	0	0	40	0	0	0	40	0
Social Studies	0	0	0	0	0	0	0	0	0	0
Total	508	47	39	9	113	8	64	9	724	73
Grade 10										
Reading	0	0	0	0	0	0	0	0	0	0
Math	0	0	0	0	0	0	0	0	0	0
Language Arts	0	0	0	0	0	0	0	0	0	0
Science	0	0	0	0	10	0	0	0	10	0
Social Studies	0	0	0	0	0	0	0	0	0	0
Total	0	0	0	0	10	0	0	0	10	0
TOTALS										
Grand Totals	3,032	266	234	33	460	40	415	52	4,141	391

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Used a Scribe	3	59886	Reading	780	1.30
		59886	Math	1065	1.78
	4	59988	Reading	780	1.30
		59988	Math	946	1.58
		59988	Science	420	0.70
		59988	Language Arts	370	0.62
		59988	Writing	430	0.72
		59988	Social Studies	422	0.70
	5	59526	Reading	634	1.07
		59526	Math	694	1.17
	6	60810	Reading	384	0.63
		60810	Math	390	0.64
	7	61951	Reading	306	0.49
		61951	Math	340	0.55
	8	64044	Reading	245	0.38
		64044	Math	263	0.41
		64044	Science	184	0.29
		64044	Language Arts	169	0.26
		64044	Writing	230	0.36
		64044	Social Studies	183	0.29
	10	70112	Reading	84	0.12
		70112	Math	96	0.14
		70112	Science	63	0.09
		70112	Language Arts	62	0.09
		70112	Writing	106	0.15
		70112	Social Studies	64	0.09

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Provided Extra Time	3	59886	Reading	5873	9.81
		59886	Math	6248	10.43
	4	59988	Reading	6410	10.69
		59988	Math	6700	11.17
		59988	Science	6402	10.67
		59988	Language Arts	6183	10.31
		59988	Writing	6142	10.24
		59988	Social Studies	6368	10.62
	5	59526	Reading	6693	11.24
		59526	Math	6875	11.55
	6	60810	Reading	6709	11.03
		60810	Math	6912	11.37
	7	61951	Reading	6642	10.72
		61951	Math	6827	11.02
	8	64044	Reading	6810	10.63
		64044	Math	7034	10.98
		64044	Science	6743	10.53
		64044	Language Arts	6605	10.31
		64044	Writing	6525	10.19
		64044	Social Studies	6727	10.50
	10	70112	Reading	5361	7.65
		70112	Math	5374	7.66
		70112	Science	5198	7.41
		70112	Language Arts	5161	7.36
		70112	Writing	5188	7.40
		70112	Social Studies	5163	7.36

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Read Test Questions and Content to Student	3	59886	Reading	0	0.00
		59886	Math	5546	9.26
	4	59988	Reading	0	0.00
		59988	Math	5909	9.85
		59988	Science	5767	9.61
		59988	Language Arts	4788	7.98
		59988	Writing	5050	8.42
		59988	Social Studies	5700	9.50
	5	59526	Reading	0	0.00
		59526	Math	5848	9.82
	6	60810	Reading	0	0.00
		60810	Math	5359	8.81
	7	61951	Reading	0	0.00
		61951	Math	5133	8.29
	8	64044	Reading	0	0.00
		64044	Math	5210	8.14
		64044	Science	5168	8.07
		64044	Language Arts	4684	7.31
		64044	Writing	4779	7.46
		64044	Social Studies	5093	7.95
	10	70112	Reading	0	0.00
		70112	Math	3358	4.79
		70112	Science	3355	4.79
		70112	Language Arts	3142	4.48
		70112	Writing	3305	4.71
		70112	Social Studies	3344	4.77

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Used DPI-Provided Test Translation	3	59886	Reading	0	0.00
		59886	Math	921	1.54
	4	59988	Reading	0	0.00
		59988	Math	640	1.07
		59988	Science	497	0.83
		59988	Language Arts	0	0.00
		59988	Writing	334	0.56
		59988	Social Studies	489	0.82
	5	59526	Reading	0	0.00
		59526	Math	514	0.86
	6	60810	Reading	0	0.00
		60810	Math	428	0.70
	7	61951	Reading	0	0.00
		61951	Math	467	0.75
	8	64044	Reading	0	0.00
		64044	Math	410	0.64
		64044	Science	327	0.51
		64044	Language Arts	0	0.00
		64044	Writing	217	0.34
		64044	Social Studies	319	0.50
	10	70112	Reading	0	0.00
		70112	Math	207	0.30
		70112	Science	167	0.24
		70112	Language Arts	0	0.00
		70112	Writing	149	0.21
		70112	Social Studies	167	0.24

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Used Locally Provided Test Translation	3	59886	Reading	0	0.00
		59886	Math	0	0.00
	4	59988	Reading	0	0.00
		59988	Math	201	0.34
		59988	Science	173	0.29
		59988	Language Arts	0	0.00
		59988	Writing	128	0.21
		59988	Social Studies	178	0.30
	5	59526	Reading	0	0.00
		59526	Math	0	0.00
	6	60810	Reading	0	0.00
		60810	Math	0	0.00
	7	61951	Reading	0	0.00
		61951	Math	0	0.00
	8	64044	Reading	0	0.00
		64044	Math	74	0.12
		64044	Science	62	0.10
		64044	Language Arts	0	0.00
		64044	Writing	46	0.07
		64044	Social Studies	64	0.10
	10	70112	Reading	0	0.00
		70112	Math	80	0.11
		70112	Science	61	0.09
		70112	Language Arts	0	0.00
		70112	Writing	53	0.08
		70112	Social Studies	58	0.08

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Used DPI-Provided Glossary of Terms	3	59886	Reading	0	0.00
		59886	Math	0	0.00
	4	59988	Reading	0	0.00
		59988	Math	254	0.42
		59988	Science	245	0.41
		59988	Language Arts	0	0.00
		59988	Writing	0	0.00
		59988	Social Studies	246	0.41
	5	59526	Reading	0	0.00
		59526	Math	0	0.00
	6	60810	Reading	0	0.00
		60810	Math	0	0.00
	7	61951	Reading	0	0.00
		61951	Math	0	0.00
	8	64044	Reading	0	0.00
		64044	Math	224	0.35
		64044	Science	213	0.33
		64044	Language Arts	0	0.00
		64044	Writing	0	0.00
		64044	Social Studies	211	0.33
	10	70112	Reading	0	0.00
		70112	Math	124	0.18
		70112	Science	127	0.18
		70112	Language Arts	0	0.00
		70112	Writing	0	0.00
		70112	Social Studies	90	0.13

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Used Text Talker	3	59886	Reading	0	0.00
		59886	Math	32	0.05
	4	59988	Reading	0	0.00
		59988	Math	4	0.01
		59988	Science	3	0.01
		59988	Language Arts	4	0.01
		59988	Writing	6	0.01
		59988	Social Studies	5	0.01
	5	59526	Reading	0	0.00
		59526	Math	7	0.01
	6	60810	Reading	0	0.00
		60810	Math	12	0.02
	7	61951	Reading	0	0.00
		61951	Math	10	0.02
	8	64044	Reading	0	0.00
		64044	Math	3	0.00
		64044	Science	3	0.00
		64044	Language Arts	3	0.00
		64044	Writing	11	0.02
		64044	Social Studies	4	0.01
	10	70112	Reading	0	0.00
		70112	Math	12	0.02
		70112	Science	9	0.01
		70112	Language Arts	7	0.01
		70112	Writing	6	0.01
		70112	Social Studies	7	0.01

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Signed Test Questions and Content to Student	3	59886	Reading	0	0.00
		59886	Math	26	0.04
	4	59988	Reading	0	0.00
		59988	Math	14	0.02
		59988	Science	14	0.02
		59988	Language Arts	10	0.02
		59988	Writing	14	0.02
		59988	Social Studies	14	0.02
	5	59526	Reading	0	0.00
		59526	Math	15	0.03
	6	60810	Reading	0	0.00
		60810	Math	12	0.02
	7	61951	Reading	0	0.00
		61951	Math	27	0.04
	8	64044	Reading	0	0.00
		64044	Math	18	0.03
		64044	Science	16	0.03
		64044	Language Arts	17	0.03
		64044	Writing	16	0.03
		64044	Social Studies	17	0.03
	10	70112	Reading	0	0.00
		70112	Math	8	0.01
		70112	Science	8	0.01
		70112	Language Arts	8	0.01
		70112	Writing	7	0.01
		70112	Social Studies	8	0.01

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Used Another DPI-Approved Accommodation	3	59886	Reading	1348	2.25
		59886	Math	1521	2.54
	4	59988	Reading	1601	2.67
		59988	Math	1745	2.91
		59988	Science	1677	2.80
		59988	Language Arts	1526	2.54
		59988	Writing	1537	2.56
		59988	Social Studies	1659	2.77
	5	59526	Reading	1675	2.81
		59526	Math	1786	3.00
	6	60810	Reading	1735	2.85
		60810	Math	1708	2.81
	7	61951	Reading	1995	3.22
		61951	Math	2082	3.36
	8	64044	Reading	1730	2.70
		64044	Math	1752	2.74
		64044	Science	1699	2.65
		64044	Language Arts	1630	2.55
		64044	Writing	1624	2.54
		64044	Social Studies	1695	2.65
	10	70112	Reading	630	0.90
		70112	Math	663	0.95
		70112	Science	695	0.99
		70112	Language Arts	637	0.91
		70112	Writing	689	0.98
		70112	Social Studies	685	0.98

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Used DPI-Provided Braille Test	3	59886	Reading	21	0.04
		59886	Math	21	0.04
	4	59988	Reading	12	0.02
		59988	Math	12	0.02
		59988	Science	12	0.02
		59988	Language Arts	12	0.02
		59988	Writing	12	0.02
		59988	Social Studies	12	0.02
	5	59526	Reading	5	0.01
		59526	Math	5	0.01
	6	60810	Reading	5	0.01
		60810	Math	5	0.01
	7	61951	Reading	4	0.01
		61951	Math	4	0.01
	8	64044	Reading	10	0.02
		64044	Math	10	0.02
		64044	Science	10	0.02
		64044	Language Arts	10	0.02
		64044	Writing	10	0.02
		64044	Social Studies	10	0.02
	10	70112	Reading	8	0.01
		70112	Math	8	0.01
		70112	Science	8	0.01
		70112	Language Arts	8	0.01
		70112	Writing	8	0.01
		70112	Social Studies	8	0.01

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 4-1
Test Accommodations Cont'd

Accommodation	Grade	N Count	Content Area	Number of Students	Percent of Total
Used a Non-Allowed Accommodation	3	59886	Reading	23	0.04
		59886	Math	23	0.04
	4	59988	Reading	28	0.05
		59988	Math	22	0.04
		59988	Science	13	0.02
		59988	Language Arts	5	0.01
		59988	Writing	9	0.02
		59988	Social Studies	5	0.01
	5	59526	Reading	9	0.02
		59526	Math	3	0.01
	6	60810	Reading	10	0.02
		60810	Math	11	0.02
	7	61951	Reading	14	0.02
		61951	Math	8	0.01
	8	64044	Reading	15	0.02
		64044	Math	6	0.01
		64044	Science	6	0.01
		64044	Language Arts	14	0.02
		64044	Writing	14	0.02
		64044	Social Studies	6	0.01
	10	70112	Reading	19	0.03
		70112	Math	9	0.01
		70112	Science	10	0.01
		70112	Language Arts	10	0.01
		70112	Writing	11	0.02
		70112	Social Studies	10	0.01

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-1
Reading Rubric, Grades 3–8 and 10

Reading items at all grade levels were scored using item-specific scoring guides that are based on a generic, 0–3 holistic rubric.

3 points

- The response demonstrates *thorough understanding* of the reading concept embodied in the task.
- The response is *accurate, complete, insightful, and fulfills all the requirements* of the task.
- Necessary support and/or examples are included.
- Information is clearly *text-based*.

2 points

- The response demonstrates *partial understanding* of the reading concept embodied in the task.
- The response is *accurate* and *fulfills most of the requirements* of the task.
- Necessary support and/or examples may not be complete or clearly text-based.

1 point

- The response demonstrates *an incomplete understanding* of the reading concept embodied in the task.
- The response provides *some information that is text-based*, but does not fulfill the requirements of the task.
- Information provided is *too general* or *too simplistic*.
- Necessary support and/or examples may be incomplete or omitted.

0 points

- The response demonstrates *no understanding* of the reading concept embodied in the task.
- The response is *inaccurate, confused*, or *irrelevant.*
- The student has *failed to respond to the task.*

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-2
Mathematics Rubric, Grades 3–8 and 10

Mathematics constructed response operational items each have two parts. Part A is scored as correct/incorrect. Part B is scored using a 2-point holistic rubric.

2 points

- The student demonstrates a thorough understanding of the mathematical concepts and/or procedures represented in the problem. The student responds correctly to the problem, uses mathematical procedures and/or concepts, and provides clear and complete explanations and interpretations containing words, diagrams, or calculations unless otherwise specified. The response may contain minor flaws that do not detract from the demonstration of a thorough understanding of the problem.

1 point

- The student provides a response that is only partially correct. The student provides a correct solution, but may demonstrate a misunderstanding of the underlying mathematical concepts and/or procedures. The student provides a correct solution, but in place of showing his/her work writes, "I used my calculator." The student provides a thorough demonstration of understanding the problem, but states an incorrect solution or conclusion.

0 points

- The student provides a completely incorrect solution, a response that cannot be interpreted, or no response at all.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-3
Writing Rubric, Conventions of Written English, Grade 4

3 points **Advanced Control**

The response demonstrates advanced control of a wide range of conventions identified in the 4th grade Wisconsin Model Academic Standards in English Language Arts:

- Uses parts of speech effectively, including nouns, pronouns, and adjectives
- Uses adverbials effectively, including words and phrases
- Employs principles of agreement related to number, gender, and case
- Capitalizes proper nouns, titles, and initial words of sentences
- Uses punctuation marks and conjunctions, as appropriate, to separate sentences and connect independent clauses
- Uses commas correctly to punctuate appositives and lists
- Spells correctly in general and usually on more difficult words
- Uses word order and punctuation marks to distinguish statements, questions, exclamations, and commands
- Makes errors that are infrequent and minor

2 points **Proficient Control**

The response demonstrates proficient control of the essential conventions identified in the 4th grade Wisconsin Model Academic Standards in English Language Arts:

- Generally controls grammar and usage (principles of agreement, noun and verb forms, superlative and comparative forms)
- Capitalizes proper nouns, titles, and initial words of sentences
- Uses end-stop punctuation correctly most of the time; internal punctuation (commas, apostrophes) is sometimes missing or wrong.
- Generally uses correct spelling with common words but more difficult words are problematic
- Makes errors typical of those commonly found in a rough draft; errors do not significantly distract the reader

1 point **Minimal Control**

The response demonstrates minimal control of the essential conventions identified in the 4th grade Wisconsin Model Academic Standards in English Language Arts:

- Contains numerous serious end-stop punctuation errors, resulting in fragments, comma splices, run-ons
- Shows poor control of subject/verb agreement, possessive forms, capitalization, superlatives and comparatives
- Spelling errors are frequent, even on common words
- Makes errors that are frequent, varied, and distracting

0 points **Off Topic, No Response, Illegible, Another Language**

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-4
Writing Rubric, Composing, Grade 4

Wisconsin Writing Grade 4 Rubric 6-Point Scoring Guide					
Elements of Rubric	**Purpose & Focus**	**Organization & Coherence**	**Development of Content**	**Sentence Fluency**	**Word Choice**
Element Description	Consistently focuses on the topic and maintains a unified purpose Demonstrates understanding of the requirements of the assigned task	Uses a logical plan of development with an effective beginning, middle, and end Keeps relationships among ideas clear Paragraphs logically and uses appropriate transitional devices	Expands and supports main ideas with specific details, examples, and/or reasons that are 1) clearly related to the topic and purpose, and 2) effective for audience	Uses varied sentence structures, creating a fluent, effective, and readable style	Controls word choice with respect to both denotation and connotation Demonstrates attention to context (audience, purpose, situation, tone) Evidences some control over figurative language for rhetorical effect (e.g. metaphors, similes)
Positive Descriptors	Focused, unified, controlled, relevant	Well organized, integrated, smooth, controlled, coherent	Thorough, specific, well-developed, well-supported, well-illustrated, insightful, convincing	Fluid, varied, controlled, effective	Vivid, precise, concrete, concise
Negative Descriptors	Rambling, loosely related, redundant, irrelevant, lacks purpose	Disorganized, hard to follow, mechanical, illogical shifts, incoherent	Vague, general, simplistic, superficial, incomplete, illogical, inadequately supported, lacks illustration	Choppy, simple, repetitive, garbled, ineffective, awkward	Awkward, imprecise, vague wordy, repetitive

Rubric Holistic Scoring Scale						
Scores	**6**	**5**	**4**	**3**	**2**	**1**
Description	Exemplary control of the domain	Advanced control of the domain	Proficient control of the domain	Adequate control of the domain	Basic control of the domain	Minimal control of the domain

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-5
Writing Rubric, Conventions of Written English, Grade 8

3 points **Advanced Control**

The response demonstrates advanced control of a wide range of conventions identified in the 8th grade Wisconsin Model Academic Standards in English Language Arts:

- Uses words, phrases, and clauses effectively, including coordinate and subordinate conjunctions, relative pronouns, and comparative adjectives
- Uses correct tenses to indicate the relative order of events
- Employs principles of agreement, including subject-verb, pronoun-noun, and preposition-pronoun
- Punctuates compound, complex, and compound-complex sentences correctly
- Employs the conventions of capitalization
- Spells frequently used words correctly and uses effective strategies for spelling unfamiliar words
- Makes errors that are infrequent and minor

2 points **Proficient Control**

The response demonstrates proficient control of the conventions identified in the 8th grade Wisconsin Model Academic Standards in English Language Arts:

- Generally controls grammar and usage (principles of agreement, noun and verb forms, pronoun reference, superlative and comparative forms)
- Generally uses phrases, dependent and independent clauses clearly and correctly
- Capitalizes most words correctly; control over more sophisticated capitalization skills may be spotty
- Uses end-stop punctuation correctly most of the time; internal punctuation (commas, apostrophes, semicolons) is sometimes missing or wrong
- Generally uses correct spelling with grade-level words and reasonable phonetic approaches to more difficult words
- Makes errors typical of those commonly found in a rough draft; errors do not seriously distract the reader

1 point **Minimal Control**

The response demonstrates minimal control of the conventions identified in the 8th grade Wisconsin Model Academic Standards in English Language Arts:

- Contains numerous serious end-stop or internal punctuation errors, resulting in fragments, comma splices, run-ons
- Shows poor control of grammar and usage (principles of agreement; verb and/or noun forms including possessives; pronoun reference; superlative and comparative forms; appropriate use of phrases/independent, dependent clauses, capitalization)
- Frequently misspells words, even those on grade-level
- Makes errors that are frequent, varied, and distracting

0 points **Off topic, No response, Illegible, Another language**

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-6
Writing Rubric, Composing, Grade 8

Wisconsin Writing Rubric Grade 8 6-Point Scoring Guide					
Elements of Rubric	**Purpose & Focus**	**Organization & Coherence**	**Development of Content**	**Sentence Fluency**	**Word Choice**
Element Description	Clearly presents and maintains a unified purpose, focus, and/or thesis Demonstrates understanding of the requirements of the assigned task	Frames the discussion with an effective introduction and conclusion Creates a logical structure of development for the topic, thesis, and purpose Uses transitional strategies (from idea to idea, paragraph to paragraph, and sentence to sentence)	Demonstrates ***quality*** of invented content (e.g. of explanations, arguments, rationale, ideas, details, examples, illustrations) Demonstrates ***thoroughness*** in the elaboration of content	Demonstrates use of varied syntactic structures including simple, compound, complex, and compound/complex sentences Evidences some control over stylistic effects (e.g. variety, readability**)**	Controls word choice with respect to both denotation and connotation Demonstrates attention to context (audience, purpose, situation, tone) Evidences some control over figurative language for rhetorical effect (e.g. similes, metaphors, personification)
Positive Descriptors	Focused, unified, controlled, relevant	Well organized, integrated, smooth, controlled, coherent	Quality: clear, convincing, accurate, effective, well-reasoned, insightful Thoroughness: specific, well-developed, well-supported, well-illustrated	Fluid, varied, controlled, effective	Apt, discriminating, vivid, precise, concrete, concise
Negative Descriptors	Rambling, loosely related, redundant, irrelevant, lacks purpose	Disorganized, hard to follow, mechanical, illogical shifts, incoherent	Quality: vague, imprecise, inaccurate, simplistic, poorly reasoned, superficial Thoroughness: incomplete, general, inadequately, developed, inadequately supported, lacks illustration	Choppy, monotonous, garbled, ineffective, awkward	Inappropriate, cliched, awkward, imprecise, vague wordy

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-6 Cont'd
Writing Rubric, Composing, Grade 8

Rubric Holistic Scoring Scale						
Scores	**6**	**5**	**4**	**3**	**2**	**1**
Description	Exemplary control of the domain	Advanced control of the domain	Proficient control of the domain	Adequate control of the domain	Basic control of the domain	Minimal control of the domain

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-7
Writing Rubric, Conventions of Written English, Grade 10

3 points **Advanced Control**

The response demonstrates advanced control of a wide range of conventions identified in the 12th grade Wisconsin Model Academic Standards in English Language Arts:

- Uses words, phrases, and clauses effectively, including interrelated clauses in complex sentences
- Uses correct tenses, including conditionals, to indicate the relative order and relationship of events
- Employs principles of agreement, including subject-verb, pronoun-noun, and preposition-pronoun
- Punctuates compound, complex, and compound-complex sentences correctly, including appropriate use of colons, hyphens, dashes, ellipses, and italics; punctuates dialogue correctly; follows citation conventions
- Employs the conventions of capitalization
- Spells frequently used words correctly and uses effective strategies for spelling unfamiliar words
- Makes errors that are infrequent and minor

2 points **Proficient Control**

The response demonstrates proficient control of essential conventions identified in the 12th grade Wisconsin Model Academic Standards in English Language Arts:

- Generally controls grammar and usage (principles of agreement, noun and verb forms, pronoun references, superlative and comparative forms)
- Generally uses phrases, dependent and independent clauses clearly and correctly
- Uses end-stop punctuation correctly most of the time; internal punctuation (commas, apostrophes, semicolons, colons) is sometimes missing or wrong; sometimes fails to punctuate dialogue correctly or to accurately follow citation conventions
- Employs the conventions of capitalization
- Generally uses correct spelling with grade-level words and reasonable phonetic approaches to more difficult words
- Makes errors typical of those commonly found in a rough draft; errors do not seriously distract the reader

1 point **Minimal Control**

The response demonstrates minimal control of essential conventions identified in the 12th grade Wisconsin Model Academic Standards in English Language Arts

- Contains numerous serious end-stop or internal punctuation errors, resulting in fragments, comma splices, run-ons
- Shows poor control of grammar and usage (principles of agreement, verb and/or noun forms; pronoun reference; superlative and comparative forms)
- Shows poor control of spelling, even on grade-level words
- Makes errors that are frequent, varied, and distracting

0 points **Off Topic, No Response, Another Language, Illegible**

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-8
Writing Rubric, Composing, Grade 10

Wisconsin Writing Grade 10 Rubric 6-Point Scoring Guide					
Elements of Rubric	**Purpose & Focus**	**Organization & Coherence**	**Development of Content**	**Sentence Fluency**	**Word Choice**
Element Description	Explicitly states, or strongly implies, a thesis or unifying purpose which firmly guides the paper Demonstrates understanding of the requirements of the assigned task	Frames the discussion with an effective introduction and conclusion Creates a logical structure of development for the topic, thesis, and purpose Uses effective and varied transitional strategies (from idea to idea, paragraph to paragraph, and sentence to sentence)	Demonstrates ***quality*** of invented content (e.g. of explanations, arguments, rationale, ideas, details, examples, illustrations) Demonstrates ***thoroughness*** in the elaboration of content	Demonstrates syntactic control of simple, compound, complex, and compound/complex sentences Evidences some control over stylistic effects (e.g. flow, cadence, parallelism, variety, readability, judicious use of active and passive voice, effective repetition**)**	Controls word choice with respect to both denotation and connotation Demonstrates attention to context (audience, purpose, situation, tone) Evidences some control over figurative language for rhetorical effect (e.g. metaphors, similes, hyperbole, analogies)
Positive Descriptors	Focused, unified, controlled, relevant	Well organized, integrated, smooth, controlled, coherent	Quality: clear, precise, accurate, effective, well-reasoned, insightful Thoroughness: complete, specific, well-developed, well-supported, well-illustrated	Fluid, varied, controlled, effective, skilled	Apt, discriminating, vivid, precise, concrete, concise
Negative Descriptors	Rambling, loosely related, redundant, irrelevant, lacks purpose	Disorganized, hard to follow, mechanical, illogical shifts, incoherent	Quality: vague, imprecise, Inaccurate, simplistic, poorly reasoned, superficial Thoroughness: incomplete, general, inadequately, developed, inadequately supported, lacks illustration	Choppy, monotonous, garbled, ineffective, awkward	Inappropriate, cliched, awkward, imprecise, vague wordy

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-8 Cont'd
Writing Rubric, Composing, Grade 10

Rubric Holistic Scoring Scale						
Scores	**6**	**5**	**4**	**3**	**2**	**1**
Description	Exemplary control of the domain	Advanced control of the domain	Proficient control of the domain	Adequate control of the domain	Basic control of the domain	Minimal control of the domain

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-9
Score Distribution for Reading CR Items*

Grade	Form	Test Book Item No.	N	Scores				Condition Codes**			
				0	1	2	3	A	B	C	D
3		19	13170	4448	4355	2662	71	1572	12	18	32
		31	13170	1608	6001	4686	131	690	8	21	25
	A	63	2767	1048	1086	371	29	203	0	11	19
	B	67	2728	953	681	847	31	209	0	0	7
	C	64	2723	861	962	602	125	166	0	0	7
	D	67	2728	886	737	783	160	158	0	2	2
4		13	13177	583	3465	6756	1893	451	8	6	15
		47	13177	1403	4313	5880	978	545	25	8	25
	A	67	2711	868	772	701	136	207	0	1	26
	B	64	2726	470	835	965	331	120	0	0	5
	C	67***									
	D	62	2709	838	683	958	130	97	0	0	3
5		13	13089	835	4218	5980	1549	476	9	5	17
		47	13089	1299	3803	6059	1445	460	7	3	13
	A	67	2757	689	1150	653	110	147	0	2	6
	B	67***									
	C	57	2761	842	1434	334	60	90	0	0	1
	C	63	2708	768	297	1405	135	101	0	0	2
	D	57***									
	D	63	2747	968	781	685	181	129	0	1	2
6		18	19415	1238	9309	7312	969	574	6	1	6
		24	19415	4108	8510	5467	731	586	8	2	3
	A	67	2870	857	1380	448	33	147	0	2	3
	B	67	2894	815	897	864	171	146	0	0	1
	C	57	2879	674	1585	508	20	89	0	0	3
	C	65	2864	621	1073	875	159	129	0	0	7
	D	67	2902	646	1300	731	46	172	1	0	6
7		31	20243	1412	7644	8482	1957	696	9	3	40
		50	20243	5840	6071	6182	1454	674	7	3	12
	A	67	2911	855	1223	392	275	164	0	1	1
	B	63	2966	603	541	847	861	111	0	0	3
	C	67	2901	365	1854	405	132	144	0	0	1
	D	67	2902	525	1368	753	110	146	0	0	0
8		19	20273	1150	7383	8479	2621	629	7	0	4
		56	20273	2285	4510	8997	3707	752	7	0	15
	A	65	2890	527	1065	818	322	157	0	0	1
	B	67	2929	884	704	1003	203	132	0	0	3
	C	63	2921	759	1440	507	116	98	0	0	1
	D	60	2890	439	1086	907	367	90	0	0	1
	D	63	2900	343	903	1325	233	95	0	0	1

*This is the score distribution of the first read.
**A: No response or no attempt, B: Illegible, C: Another Language, D: Off-topic.
***This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-9 Cont'd
Score Distribution for Reading CR Items*

Grade	Form	Test Book Item No.	N	Scores				Condition Codes**			
				0	1	2	3	A	B	C	D
10		7	13045	2814	3708	3630	1558	1306	5	0	24
		52	13045	1035	5072	4473	1546	900	3	0	16

*This is the score distribution of the first read.
**A: No response or no attempt, B: Illegible, C: Another Language, D: Off-topic.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-10
Score Distribution for Mathematics CR Items*

Grade	Form	Test Book Item No.	Part	N	Scores			Condition Codes**			
					0	1	2	A	B	C	D
3		5		13170	6038	1644	5147	338	1	1	1
		20	A	13170	5880	6170	0	1117	0	3	0
		20	B	13170	7915	2181	2377	685	1	8	3
		27	A	13170	4597	7963	0	601	1	7	1
		27	B	13170	7134	2230	2863	924	0	9	10
		39	A	13170	8176	4135	0	859	0	0	0
		39	B	13170	7678	2446	2456	573	2	10	5
	A	51		1696	587	258	785	65	0	0	1
	A	57	A	2755	436	2241	0	78	0	0	0
	A	57	B	2755	1250	467	895	138	0	5	0
	B	51		1668	566	246	777	79	0	0	0
	B	57***	A								
	B	57***	B								
	C	60		2685	114	796	1643	132	0	0	0
	D	58	A	2704	388	2250	0	66	0	0	0
	D	58	B	2704	2108	475	28	91	0	2	0
4		11	A	13177	9975	2884	0	304	3	11	0
		11	B	13177	6421	4517	1585	609	4	18	23
		22		13177	770	3527	8518	346	0	16	0
		25	A	13177	6975	5917	0	283	1	1	0
		25	B	13177	6027	4688	1879	555	0	19	9
		39	A	13177	5731	7169	0	275	1	1	0
		39	B	13177	7534	1325	3816	472	1	23	6
	A	51		2752	1024	227	1397	104	0	0	0
	A	58***	A								
	A	58***	B								
	B	59	A	2742	242	2427	0	73	0	0	0
	B	59	B	2742	562	1344	736	100	0	0	0
	C	58***	A								
	C	58***	B								
	D	59***									
5		12		13089	3901	1316	7499	364	4	5	0
		20	A	13089	1879	10911	0	291	6	2	0
		20	B	13089	8700	2205	1736	430	5	13	0
		27	A	13089	5258	7493	0	332	3	2	1
		27	B	13089	2765	1830	8053	428	6	7	0
		46	A	13089	8243	4557	0	281	6	2	0
		46	B	13089	7858	770	4110	340	7	4	0
	A	63		2763	1287	838	523	109	1	4	1

*This is the score distribution of the first read.
**A: No response or no attempt, B: Illegible, C: Another Language, D: Off-topic.
***This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-10 Cont'd
Score Distribution for Mathematics CR Items*

Grade	Form	Test Book Item No.	Part	N	Scores			Condition Codes**			
					0	1	2	A	B	C	D
5	B	64	A	2737	906	1642	0	189	0	0	0
	B	64	B	2737	1235	1386	44	70	0	1	1
	C	64		2768	699	272	1714	82	0	0	1
	D	64	A	2731	1440	1229	0	62	0	0	0
	D	64	B	2731	283	1246	1122	80	0	0	0
6		7	A	19415	8679	9519	0	1208	8	1	0
		7	B	19415	10405	1394	6282	1319	4	9	2
		22	A	19415	2152	16823	0	433	4	1	2
		22	B	19415	9589	8072	1163	574	2	12	3
		39	A	19415	10496	8351	0	561	5	1	1
		39	B	19415	8671	5254	4795	678	4	11	2
		51		19415	7151	7531	4299	428	3	3	0
	A	63	A	2905	1080	1734	0	91	0	0	0
	A	63	B	2904	551	543	1697	108	0	5	0
	B	63	A	2908	1656	1120	0	132	0	0	0
	B	63	B	2906	1587	236	916	167	0	0	0
	C	62***	A								
	C	62***	B								
	D	62***									
7		7	A	20243	5827	13877	0	528	8	3	0
		7	B	20243	5961	2473	11033	747	13	15	1
		19	A	20243	8755	10815	0	655	7	10	1
		19	B	20243	9908	7014	2327	965	12	15	2
		38	A	20243	7229	11143	0	1867	1	0	3
		38	B	20243	7068	9966	1921	1270	9	5	4
		52		20243	8884	6793	3970	593	2	1	0
	A	61		1673	848	318	416	87	0	3	1
	B	61	A	2937	1708	993	0	236	0	0	0
	B	61	B	2937	1378	1073	287	199	0	0	0
	C	62		2909	1920	380	418	190	0	1	0
	D	61		1619	687	400	444	88	0	0	0

*This is the score distribution of the first read.
**A: No response or no attempt, B: Illegible, C: Another Language, D: Off-topic.
***This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-10 Cont'd
Score Distribution for Mathematics CR Items*

Grade	Form	Test Book Item No.	Part	N	Scores			Condition Codes**			
					0	1	2	A	B	C	D
8		10		20273	13034	3403	2838	988	6	1	3
		19	A	20273	5820	13769	0	680	3	0	1
		19	B	20273	5959	3709	9321	1271	3	6	4
		33	A	20273	10439	8253	0	1581	0	0	0
		33	B	20273	7765	7759	3193	1546	3	5	2
		49	A	20273	14080	5433	0	757	2	0	1
		49	B	20273	13507	2423	3334	999	4	5	1
	A	62		2913	2543	154	92	122	0	2	0
	B	62		2887	1100	1438	189	160	0	0	0
	C	62	A	2890	974	1777	0	139	0	0	0
	C	62	B	2889	1191	880	602	216	0	0	0
	D	61	A	2897	1103	1719	0	74	0	1	0
	D	61	B	2897	1522	203	1055	116	0	1	0
10		23		13045	4186	4415	3245	1194	0	4	1
		33		13045	6271	2496	2080	2198	0	0	0
		36		13045	4749	2693	3989	1613	0	0	1
		46		13045	4736	4593	2530	1176	0	5	5

*This is the score distribution of the first read.
**A: No response or no attempt, B: Illegible, C: Another Language, D: Off-topic.
***This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-11
Score Distribution for Grades 4, 8, and 10 Writing: Composing Rubric

Grade	Form	Total N	Scores						Condition Codes**			
			1	2	3	4	5	6	A	B	C	D
4	Rater1	59977	1658	12380	25510	16422	1791	142	1305	45	28	696
	Rater2	59977	1663	12308	25460	16560	1792	121	1305	52	26	690
	Diff*	0	-5	72	50	-138	-1	21	0	-7	2	6
8	Rater1	64047	572	8525	23622	24980	3546	303	1744	1	12	742
	Rater2	64047	595	8481	23534	25040	3628	293	1733	2	13	728
	Diff*	64070	-23	44	88	-60	-82	10	11	-1	-1	14
10	Rater1	70031	2683	16482	26441	17359	2314	180	3659	3	9	901
	Rater2	70031	2718	16453	26686	17104	2375	148	3640	6	10	891
	Diff*	0	-35	29	-245	255	-61	32	19	-3	-1	10

*Diff = N of Rater1 – N of Rater 2.
**A: No response or no attempt, B: Illegible, C: Another Language, D: Off-topic.

Table 5-12
Percentage Distribution of Scores, Grades 4, 8, and 10 Writing: Composing Rubric

Grade	Form	Total N	Scores						Condition Codes**			
			1	2	3	4	5	6	A	B	C	D
4	Rater1	59977	2.76	20.64	42.53	27.38	2.99	0.24	2.18	0.08	0.05	1.16
	Rater2	59977	2.77	20.52	42.45	27.61	2.99	0.20	2.18	0.09	0.04	1.15
8	Rater1	64047	0.89	13.31	36.88	39.00	5.54	0.47	2.72	0.00	0.02	1.16
	Rater2	64047	0.93	13.24	36.74	39.10	5.66	0.46	2.71	0.00	0.02	1.14
10	Rater1	70031	3.83	23.54	37.76	24.79	3.30	0.26	5.22	0.00	0.01	1.29
	Rater2	70031	3.88	23.49	38.11	24.42	3.39	0.21	5.20	0.01	0.01	1.27

**A: No response or no attempt, B: Illegible, C: Another Language, D: Off-topic.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-13
Score Distribution, Grades 4, 8, and 10 Writing: Convention Rubric

Grade		Total N	Scores			Condition Codes**		
			1	2	3	A	B	C
4	Rater1	59977	2434	55578	587	1305	45	28
	Rater2	59977	2421	55600	573	1305	52	26
	Diff*	0	13	-22	14	0	-7	2
8	Rater1	64047	670	61437	183	1744	1	12
	Rater2	64047	662	61465	172	1733	2	13
	Diff	0	8	-28	11	11	-1	-1
10	Rater1	70031	1596	62744	2020	3659	3	9
	Rater2	70031	1657	62734	1984	3640	6	10
	Diff	0	-61	10	36	19	-3	-1

*Diff = N of Rater 1 – N of Rater 2.
**A: No response or no attempt, B: Illegible, C: Another Language.

Table 5-14
Percentage Distribution of Scores for Grades 4, 8, and 10 Writing: Convention Rubric

Grade		Total N	Scores			Condition Codes**		
			1	2	3	A	B	C
4	Rater1	59977	4.06	92.67	0.98	2.18	0.08	0.05
	Rater2	59977	4.04	92.70	0.96	2.18	0.09	0.04
8	Rater1	64047	1.05	95.92	0.29	2.72	0.00	0.02
	Rater2	64047	1.03	95.97	0.27	2.71	0.00	0.02
10	Rater1	70031	2.28	89.59	2.88	5.22	0.00	0.01
	Rater2	70031	2.37	89.58	2.83	5.20	0.01	0.01

**A: No response or no attempt, B: Illegible, C: Another Language.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 5-15
Score Distribution for Grades 4, 8, and 10 Writing: Total Score, Composing and Convention Combined

Grade		Total N	Scores									
			0	1	2	3	4	5	6	7	8	9
4	Rater1	59977	1378	73	1689	1549	11723	25197	16329	1570	366	103
	Rater2	59977	1383	74	1675	1547	11661	25165	16462	1544	383	83
	Diff*	0	-5	-1	14	2	62	32	-133	26	-17	20
8	Rater1	64047	1757	36	1146	305	8369	23596	24965	3511	275	87
	Rater2	64047	1748	37	1114	351	8312	23511	25029	3599	277	69
	Diff	0	9	-1	32	-46	57	85	-64	-88	-2	18
10	Rater1	70031	3671	213	1785	1846	16226	26288	16808	2014	1031	149
	Rater2	70031	3656	220	1803	1865	16182	26563	16565	1954	1101	122
	Diff	0	15	-7	-18	-19	44	-275	243	60	-70	27

*Diff = N of Rater 1 – N of Rater 2.

Table 5-16
Percent for Grades 4, 8, and 10 Writing: Total Score, Composing and Convention Combined

Grade		Total N	Scores									
			0	1	2	3	4	5	6	7	8	9
4	Rater1	59977	2.30	0.12	2.82	2.58	19.55	42.01	27.23	2.62	0.61	0.17
	Rater2	59977	2.31	0.12	2.79	2.58	19.44	41.96	27.45	2.57	0.64	0.14
8	Rater1	64047	2.74	0.06	1.79	0.48	13.07	36.84	38.98	5.48	0.43	0.14
	Rater2	64047	2.73	0.06	1.74	0.55	12.98	36.71	39.08	5.62	0.43	0.11
10	Rater1	70031	5.24	0.30	2.55	2.64	23.17	37.54	24.00	2.88	1.47	0.21
	Rater 2	70031	5.22	0.31	2.57	2.66	23.11	37.93	23.65	2.79	1.57	0.17

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 6-1
The Current 13 Calibration Districts

	District Name
1	BUTTERNUT
2	KENOSHA
3	LA CROSSE
4	MADISON
5	PLATTEVILLE
6	RICHLAND
7	SHEBOYGAN
8	SHOREWOOD
9	VERONA
10	WABENO
11	WATERTOWN
12	WAUSAU
13	WAUWATOSA

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 6-2
Total Number of Students for Census and CD 13

Grade	Census	CD 13
3	59886	6635
4	59988	6488
5	59526	6414
6	60810	6516
7	61951	6673
8	64044	6778
10	70112	7453

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 6-3
Number and Percent of Census and 13 CD Students, by Gender*

Grade	Census		13 CD	
	M	F	M	F
3	30820	29066	3457	3178
4	30761	29227	3348	3140
5	30463	29060	3293	3121
6	31292	29515	3282	3234
7	31774	30161	3468	3203
8	32730	31314	3506	3272
10	35920	34192	3806	3647

* Students of unspecified gender or ethnicity are not counted

Grade	Census		13 CD	
	M	F	M	F
3	51	49	52	48
4	51	49	52	48
5	51	49	51	49
6	51	49	50	50
7	51	49	52	48
8	51	49	52	48
10	51	49	51	49

* Students of unspecified gender or ethnicity are not counted

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 6-4
Number and Percent of Census and 13 CD Students, by Ethnicity*

Grade	Census					13 CD				
	W	B	H	A	AI	W	B	H	A	AI
3	44954	6492	5381	2175	884	4258	905	892	516	64
4	45391	6566	5059	2107	865	4100	985	804	544	55
5	45478	6346	4718	2171	809	4172	881	719	598	44
6	46779	6415	4522	2259	832	4295	853	738	593	37
7	47662	6624	4512	2277	869	4364	909	747	590	61
8	49665	6907	4250	2270	949	4406	985	694	615	76
10	56162	6519	3996	2384	1042	5195	901	635	652	70

* Students of unspecified ethnicity are not counted

Grade	Census**					13 CD				
	W	B	H	A	AI	W	B	H	A	AI
3	75	11	9	4	1	64	14	13	8	1
4	76	11	8	4	1	63	15	12	8	1
5	76	11	8	4	1	65	14	11	9	1
6	77	11	7	4	1	66	13	11	9	1
7	77	11	7	4	1	65	14	11	9	1
8	78	11	7	4	1	65	15	10	9	1
10	80	9	6	3	1	70	12	9	9	1

* Students of unspecified ethnicity are not counted
**Percentages may not total to exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 6-5
Number and Percent of Census and 13 CD Students, by Socioeconomic Status (SES)

Grade	Census		13 CD	
	Yes	No	Yes	No
3	21903	37983	2607	4028
4	21698	38290	2592	3896
5	20668	38858	2449	3965
6	20400	40410	2400	4116
7	20339	41612	2404	4269
8	20425	43619	2464	4314
10	18564	51548	2229	5224

Grade	Census**		13 CD	
	Yes	No	Yes	No
3	37	63	39	61
4	36	64	40	60
5	35	65	38	62
6	34	66	37	63
7	33	67	36	64
8	32	68	36	64
10	26	74	30	70

**Percentages may not total to exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 6-6
Number and Percent of Census and 13 CD Students, by Disability

Grade	Census		13 CD	
	Yes	No	Yes	No
3	8306	51580	894	5741
4	8522	51466	970	5518
5	8419	51107	978	5436
6	8560	52250	919	5597
7	8541	53410	993	5680
8	9213	54831	1068	5710
10	9463	60649	1126	6327

Grade	Census**		13 CD	
	Yes	No	Yes	No
3	14	86	13	87
4	14	86	15	85
5	14	86	15	85
6	14	86	14	86
7	14	86	15	85
8	14	86	16	84
10	14	87	15	85

**Percentages may not total to exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 6-7
Number and Percent of Census and 13 CD Students, by ELP

Grade	Census		13 CD	
	Yes	No	Yes	No
3	55463	4423	5655	980
4	56015	3973	5596	892
5	55781	3745	5569	845
6	57352	3458	5663	853
7	58637	3314	5824	849
8	60964	3080	5953	825
10	67611	2501	6742	711

Grade	Census**		13 CD	
	Yes	No	Yes	No
3	93	7	85	15
4	93	7	86	14
5	94	6	87	13
6	94	6	87	13
7	95	5	87	13
8	95	5	88	12
10	96	4	90	10

**Percentages may not total to exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-1
Item Flagged Based on Yen's Q_1

Content	Grade	Form	Test Book ID	CR Part	Status	Type	N	Z	Critical Z
RD	3	3	64		FT	CR	2555	16.87	6.8
	3	4	62		FT	MC	14462	62.11	38.6
	4	1	67		FT	CR	2503	24.15	6.67
	4	3	61		FT	MC	14649	52.32	39.06
	5	1	65		FT	MC	14364	40.36	38.30
	6	All	44		OP	MC	6219	16.84	16.58
	6	All	46		OP	MC	6217	19.92	16.31
	6	2	58		FT	MC	14805	46.94	39.48
	6	4	58		FT	MC	14824	62.35	39.53
	7	All	11		OP	MC	6249	43.24	16.66
	7	All	30		OP	MC	6373	40.03	16.99
	7	All	51		OP	MC	6371	32.17	16.99
	7	1	67		FT	CR	2740	10.45	7.31
	8	All	4		OP	MC	6476	17.31	17.27
	8	All	46		OP	MC	6472	21.38	17.26
	10	All	20		OP	MC	7199	75.9	19.20
MA	4	1	59		FT	MC	14699	42.27	39.20
	4	3	54		FT	MC	14677	52.49	39.14
	6	All	39	B	OP	CR	6104	22.95	16.28
	6	1	56		FT	MC	14934	268.49	39.82
	8	All	51		OP	MC	6481	17.74	17.28
	8	3	64		FT	MC	15527	52.77	41.41
	10	All	46		OP	CR	6618	25.21	17.65
LA	8	(NA)	28		OP	MC	6266	37.43	16.71
	10	(NA)	1	B	OP	CR	7140	50.06	19.04
SS	4	(NA)	5		OP	MC	6080	17.82	16.21

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-2
Scoring Table for Reading Grade 3

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	270	120	31	436	7
1	270	120	32	438	7
2	270	120	33	440	7
3	270	120	34	443	7
4	270	120	35	445	7
5	270	120	36	447	7
6	270	120	37	449	7
7	270	120	38	451	7
8	270	120	39	454	7
9	270	120	40	456	7
10	270	120	41	459	7
11	270	120	42	461	7
12	343	47	43	464	7
13	363	28	**44**	**466**	**8**
14	375	21	45	469	8
15	383	18	46	472	8
16	390	15	47	475	8
17	**395**	**14**	48	478	8
18	400	12	49	482	9
19	404	11	50	485	9
20	407	11	51	490	10
21	411	10	52	495	11
22	414	9	53	500	12
23	417	9	54	507	13
24	419	8	55	516	16
25	422	8	56	529	20
26	424	8	57	548	28
27	427	8	58	582	40
28	429	7	59	640	56
29	**431**	**7**	60	640	56
30	434	7			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-3
Scoring Table for Reading Grade 4

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	280	91	31	449	9
1	280	91	32	452	9
2	280	91	33	455	9
3	280	91	34	458	9
4	280	91	35	461	9
5	280	91	36	464	9
6	280	91	37	467	9
7	280	91	38	470	9
8	280	91	39	473	9
9	280	91	40	476	9
10	280	91	41	479	10
11	295	79	42	482	10
12	335	52	43	486	10
13	356	38	**44**	**490**	**11**
14	370	30	45	494	11
15	380	24	46	498	11
16	388	21	47	502	12
17	395	18	48	507	12
18	**401**	**16**	49	512	13
19	406	15	50	517	13
20	411	13	51	523	14
21	415	13	52	529	15
22	419	12	53	536	16
23	423	11	54	544	17
24	427	11	55	554	19
25	430	11	56	565	22
26	434	10	57	581	26
27	437	10	58	602	33
28	**440**	**10**	59	639	49
29	443	10	60	650	55
30	446	10			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-4
Scoring Table for Reading Grade 5

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	290	100	31	456	10
1	290	100	32	459	10
2	290	100	33	462	10
3	290	100	34	465	10
4	290	100	35	468	10
5	290	100	36	472	9
6	290	100	37	475	9
7	290	100	38	478	9
8	290	100	39	481	9
9	290	100	40	484	10
10	290	100	41	487	10
11	290	100	42	491	10
12	290	100	43	494	10
13	347	52	**44**	**498**	**10**
14	368	35	45	502	11
15	381	27	46	506	11
16	391	22	47	510	11
17	399	18	48	514	12
18	**405**	**16**	49	519	12
19	411	15	50	525	13
20	416	14	51	530	13
21	421	13	52	537	14
22	425	12	53	544	15
23	429	12	54	552	17
24	433	12	55	562	19
25	436	11	56	574	22
26	440	11	57	591	27
27	443	11	58	615	36
28	**447**	**10**	59	661	59
29	450	10	60	690	79
30	453	10			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-5
Scoring Table for Reading Grade 6

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	300	93	31	472	12
1	300	93	32	476	12
2	300	93	33	480	12
3	300	93	34	484	12
4	300	93	35	488	12
5	300	93	36	491	12
6	300	93	37	495	12
7	300	93	38	499	12
8	300	93	39	503	12
9	300	93	40	507	12
10	300	93	41	511	12
11	300	93	**42**	**516**	**13**
12	300	93	43	520	13
13	338	66	44	525	13
14	364	47	45	529	13
15	381	35	46	534	14
16	393	28	47	539	14
17	402	23	48	545	14
18	410	21	49	551	15
19	417	19	50	557	16
20	**424**	**17**	51	564	16
21	429	16	52	571	17
22	435	15	53	580	18
23	440	14	54	589	20
24	444	14	55	601	22
25	449	13	56	615	25
26	453	13	57	633	30
27	**457**	**13**	58	659	39
28	461	12	59	705	61
29	465	12	60	730	77
30	469	12			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-6
Scoring Table for Reading Grade 7

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	310	108	31	482	12
1	310	108	32	486	12
2	310	108	33	490	12
3	310	108	34	494	12
4	310	108	35	498	12
5	310	108	36	502	12
6	310	108	37	506	12
7	310	108	38	509	12
8	310	108	39	513	12
9	310	108	40	517	12
10	310	108	41	522	12
11	310	108	**42**	**526**	**12**
12	314	104	43	530	13
13	359	59	44	534	13
14	381	40	45	539	13
15	395	30	46	544	13
16	406	25	47	549	14
17	415	21	48	554	14
18	422	19	49	560	15
19	428	17	50	566	15
20	**434**	**16**	51	573	16
21	440	15	52	580	17
22	445	14	53	588	19
23	449	14	54	598	21
24	454	13	55	609	23
25	458	13	56	623	26
26	462	13	57	641	31
27	466	13	58	666	39
28	**470**	**12**	59	709	57
29	474	12	60	780	107
30	478	12			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-7
Scoring Table for Reading Grade 8

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	330	68	31	485	12
1	330	68	32	489	12
2	330	68	33	493	12
3	330	68	34	497	12
4	330	68	35	501	12
5	330	68	36	504	12
6	330	68	37	508	12
7	330	68	38	512	12
8	330	68	39	516	12
9	330	68	40	520	12
10	330	68	41	524	12
11	330	68	42	528	12
12	330	68	43	532	12
13	330	68	44	536	12
14	330	68	**45**	**541**	**13**
15	357	53	46	546	13
16	380	42	47	550	13
17	396	36	48	556	14
18	409	31	49	561	14
19	419	27	50	567	14
20	428	24	51	573	15
21	436	22	52	579	15
22	443	20	53	587	16
23	**449**	**18**	54	595	18
24	454	17	55	606	20
25	460	15	56	618	23
26	464	15	57	636	29
27	469	14	58	663	41
28	473	14	59	714	69
29	477	13	60	790	127
30	**481**	**13**			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-8
Scoring Table for Reading Grade 10

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	350	71	29	501	14
1	350	71	**30**	**506**	**14**
2	350	71	31	510	14
3	350	71	32	515	14
4	350	71	33	519	14
5	350	71	34	524	14
6	350	71	35	528	14
7	350	71	36	533	14
8	350	71	37	538	14
9	350	71	38	542	14
10	350	71	39	547	14
11	350	71	40	552	14
12	372	52	**41**	**556**	**14**
13	393	39	42	561	14
14	408	32	43	566	14
15	419	27	44	572	14
16	429	23	45	577	14
17	437	21	46	583	14
18	444	19	47	589	15
19	451	17	48	595	15
20	**457**	**17**	49	602	16
21	462	16	50	611	18
22	468	15	51	620	20
23	473	15	52	632	22
24	478	15	53	647	27
25	483	14	54	669	34
26	488	14	55	708	53
27	492	14	56	820	156
28	497	14			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-9
Scoring Table for Mathematics Grade 3

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	220	101	29	388	12
1	220	101	**30**	**392**	**11**
2	220	101	31	395	11
3	220	101	32	399	11
4	220	101	33	403	11
5	220	101	34	406	11
6	220	101	**35**	**410**	**11**
7	220	101	36	414	11
8	220	101	37	417	11
9	220	101	38	421	11
10	224	97	39	425	11
11	266	55	40	429	11
12	286	37	41	433	12
13	300	29	42	437	12
14	311	24	43	441	12
15	319	21	44	445	12
16	327	19	45	450	12
17	334	18	**46**	**454**	**12**
18	340	16	47	459	13
19	345	16	48	464	13
20	350	15	49	469	14
21	355	14	50	475	14
22	360	14	51	482	15
23	364	13	52	489	16
24	368	13	53	498	18
25	372	13	54	508	21
26	376	12	55	523	26
27	380	12	56	548	37
28	384	12	57	630	117

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-10
Scoring Table for Mathematics Grade 4

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	240	68	29	417	13
1	240	68	**30**	**421**	**13**
2	240	68	31	425	13
3	240	68	32	429	12
4	240	68	33	434	12
5	240	68	**34**	**438**	**12**
6	240	68	35	442	12
7	240	68	36	446	12
8	240	68	37	450	12
9	240	68	38	454	12
10	240	68	39	458	12
11	256	58	40	462	12
12	284	44	41	466	12
13	303	37	42	471	12
14	318	32	43	475	12
15	330	28	44	480	13
16	341	25	**45**	**484**	**13**
17	349	23	46	490	13
18	357	20	47	495	14
19	365	19	48	501	14
20	371	18	49	507	15
21	377	17	50	514	16
22	383	16	51	521	17
23	388	15	52	530	18
24	393	15	53	541	21
25	398	14	54	555	24
26	403	14	55	574	31
27	408	14	56	608	47
28	412	13	57	650	75

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-11
Scoring Table for Mathematics Grade 5

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	270	92	32	454	13
1	270	92	33	458	13
2	270	92	34	462	12
3	270	92	**35**	**466**	**12**
4	270	92	36	469	12
5	270	92	37	473	12
6	270	92	38	477	12
7	270	92	39	480	11
8	270	92	40	484	11
9	270	92	41	487	11
10	270	92	42	491	11
11	270	92	43	495	11
12	308	54	44	499	11
13	329	38	45	502	11
14	344	31	**46**	**506**	**12**
15	356	26	47	510	12
16	366	23	48	514	12
17	375	21	49	519	12
18	382	20	50	523	13
19	390	19	51	528	13
20	396	18	52	533	13
21	402	18	53	539	14
22	408	17	54	545	15
23	414	16	55	552	16
24	419	16	56	559	17
25	424	15	57	568	19
26	429	15	58	579	22
27	433	15	59	592	25
28	438	14	60	612	32
29	442	14	61	647	49
30	**446**	**13**	62	680	70
31	450	13			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-12
Scoring Table for Mathematics Grade 6

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	310	67	**32**	**486**	**11**
1	310	67	33	490	11
2	310	67	34	493	11
3	310	67	35	497	11
4	310	67	36	500	11
5	310	67	37	503	11
6	310	67	38	507	11
7	310	67	39	510	11
8	310	67	40	513	11
9	310	67	41	516	11
10	310	67	42	520	11
11	310	67	43	523	11
12	330	55	44	527	11
13	358	42	45	530	11
14	377	35	**46**	**534**	**11**
15	392	30	47	537	11
16	403	26	48	541	11
17	413	23	49	545	11
18	421	21	50	549	12
19	428	19	51	553	12
20	434	18	52	558	12
21	440	16	53	563	13
22	446	15	54	568	13
23	451	15	55	574	14
24	455	14	56	581	15
25	460	13	57	589	17
26	**464**	**13**	58	598	19
27	468	13	59	611	23
28	472	12	60	628	28
29	476	12	61	658	42
30	479	12	62	700	70
31	483	12			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-13
Scoring Table for Mathematics Grade 7

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	330	91	32	507	10
1	330	91	33	510	10
2	330	91	34	513	10
3	330	91	35	516	10
4	330	91	36	519	10
5	330	91	37	523	10
6	330	91	38	526	10
7	330	91	39	529	10
8	330	91	40	532	10
9	330	91	41	535	10
10	330	91	42	538	10
11	337	84	43	541	10
12	376	45	44	545	10
13	396	32	45	548	10
14	410	26	46	552	10
15	421	23	**47**	**555**	**10**
16	430	21	48	559	11
17	438	19	49	563	11
18	446	18	50	567	11
19	452	17	51	571	11
20	458	16	52	576	12
21	464	15	53	580	12
22	469	14	54	586	13
23	473	13	55	592	13
24	478	13	56	598	14
25	**482**	**12**	57	606	16
26	486	12	58	615	17
27	490	11	59	627	20
28	493	11	60	645	27
29	497	11	61	679	44
30	500	11	62	710	65
31	**504**	**10**			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-14
Scoring Table for Mathematics Grade 8

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	350	105	32	535	11
1	350	105	33	538	11
2	350	105	34	541	11
3	350	105	35	545	11
4	350	105	36	548	11
5	350	105	37	551	11
6	350	105	38	554	11
7	350	105	39	558	10
8	350	105	40	561	10
9	350	105	41	564	10
10	350	105	42	567	10
11	350	105	43	570	10
12	388	67	**44**	**574**	**10**
13	418	40	45	577	10
14	435	31	46	580	10
15	448	26	47	584	11
16	458	22	48	588	11
17	466	20	49	591	11
18	473	18	50	595	11
19	479	17	51	599	11
20	**485**	**16**	52	604	12
21	491	15	53	608	12
22	495	14	54	613	13
23	500	14	55	619	14
24	505	13	56	625	15
25	509	13	57	633	16
26	**513**	**13**	58	642	18
27	517	12	59	654	22
28	521	12	60	671	28
29	524	12	61	703	44
30	528	12	62	730	63
31	531	11			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-15
Scoring Table for Mathematics Grade 10

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	410	82	30	557	9
1	410	82	31	560	9
2	410	82	32	563	9
3	410	82	33	565	9
4	410	82	34	568	9
5	410	82	35	571	9
6	410	82	36	573	8
7	410	82	37	576	8
8	410	82	38	578	8
9	410	82	39	581	8
10	410	82	40	583	8
11	419	73	41	586	8
12	453	39	42	589	8
13	471	28	43	591	8
14	483	23	44	594	8
15	492	20	**45**	**597**	**9**
16	500	18	46	600	9
17	507	16	47	603	9
18	512	15	48	607	9
19	**518**	**14**	49	610	10
20	522	13	50	614	10
21	527	13	51	619	11
22	531	12	52	624	12
23	535	12	53	630	13
24	538	11	54	637	14
25	**542**	**11**	55	646	17
26	545	10	56	659	21
27	548	10	57	682	32
28	551	10	58	750	96
29	554	10			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-16
Scoring Table for Language Arts Grade 4

Raw Score	Scale Score	SEM
0	140	110
1	140	110
2	140	110
3	140	110
4	140	110
5	140	110
6	140	110
7	210	40
8	229	22
9	239	17
10	247	14
11	**254**	**12**
12	260	11
13	265	11
14	270	10
15	274	10
16	**278**	**10**
17	283	9
18	287	9
19	291	9
20	295	9
21	300	9
22	304	9
23	**309**	**10**
24	315	10
25	321	11
26	328	12
27	337	14
28	350	19
29	374	31
30	420	67

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-17
Scoring Table for Language Arts Grade 8

Raw Score	Scale Score	SEM
0	250	90
1	250	90
2	250	90
3	250	90
4	250	90
5	250	90
6	250	90
7	250	90
8	287	53
9	309	31
10	322	22
11	331	17
12	339	15
13	345	13
14	351	13
15	356	12
16	**362**	**12**
17	367	12
18	372	12
19	377	12
20	382	11
21	**387**	**11**
22	393	11
23	399	12
24	405	12
25	412	13
26	**420**	**14**
27	431	17
28	448	23
29	473	28
30	520	61

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-18
Scoring Table for Language Arts Grade 10

Raw Score	Scale Score	SEM
0	290	75
1	290	75
2	290	75
3	290	75
4	290	75
5	290	75
6	296	69
7	331	34
8	347	22
9	358	17
10	366	15
11	373	15
12	379	14
13	385	14
14	390	14
15	**396**	**14**
16	401	13
17	406	13
18	411	13
19	416	13
20	421	13
21	426	12
22	**431**	**13**
23	437	13
24	442	13
25	447	13
26	453	13
27	459	14
28	466	14
29	473	15
30	481	16
31	**489**	**17**
32	498	18
33	508	18
34	520	19
35	533	21
36	548	23
37	567	27
38	598	39
39	630	58

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-19
Scoring Table for Social Studies Grade 4

Raw Score	Scale Score	SEM
0	170	71
1	170	71
2	170	71
3	170	71
4	170	71
5	170	71
6	170	71
7	170	71
8	198	43
9	218	23
10	227	16
11	234	13
12	240	11
13	**244**	**10**
14	248	9
15	252	9
16	256	8
17	259	8
18	262	8
19	**265**	**8**
20	268	7
21	271	7
22	274	7
23	276	7
24	279	7
25	282	7
26	284	7
27	287	7
28	**290**	**7**
29	293	7
30	296	7
31	299	7
32	302	7
33	306	8
34	310	8
35	316	10
36	324	12
37	338	19
38	400	81

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-20
Scoring Table for Social Studies Grade 8

Raw Score	Scale Score	SEM
0	230	99
1	230	99
2	230	99
3	230	99
4	230	99
5	230	99
6	230	99
7	230	99
8	230	99
9	230	99
10	271	58
11	299	30
12	312	21
13	321	16
14	329	14
15	**335**	**13**
16	341	13
17	346	12
18	351	12
19	356	12
20	360	11
21	**365**	**11**
22	369	11
23	373	10
24	377	10
25	381	10
26	385	10
27	389	10
28	393	10
29	397	10
30	401	10
31	**405**	**10**
32	410	11
33	415	11
34	420	11
35	426	12
36	433	13
37	442	15
38	455	19
39	477	29
40	530	74

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-21
Scoring Table for Social Studies Grade 10

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	240	130	26	425	11
1	240	130	27	429	11
2	240	130	28	433	11
3	240	130	29	437	11
4	240	130	30	440	11
5	240	130	31	444	10
6	240	130	32	448	10
7	240	130	33	451	10
8	240	130	**34**	**455**	**10**
9	240	130	35	458	10
10	240	130	36	462	10
11	277	93	37	466	10
12	327	43	38	470	10
13	346	27	39	474	11
14	358	22	40	479	11
15	368	19	41	483	11
16	375	17	42	489	12
17	382	16	43	495	13
18	388	15	44	501	14
19	394	14	45	509	15
20	399	14	46	519	17
21	404	13	47	532	21
22	**409**	**13**	48	549	25
23	413	12	49	575	33
24	417	12	50	620	62
25	**421**	**12**			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-22
Scoring Table for Science Grade 4

Raw Score	Scale Score	SEM
0	170	57
1	170	57
2	170	57
3	170	57
4	170	57
5	170	57
6	170	57
7	170	57
8	170	57
9	170	57
10	196	31
11	211	22
12	221	18
13	229	16
14	236	14
15	241	13
16	246	12
17	**251**	**11**
18	255	10
19	259	9
20	263	9
21	266	9
22	270	9
23	274	9
24	277	9
25	**281**	**9**
26	284	9
27	288	9
28	291	9
29	295	9
30	299	9
31	303	9
32	307	9
33	312	9
34	316	9
35	**322**	**10**
36	329	12
37	338	14
38	353	21
39	385	40
40	440	85

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-23
Scoring Table for Science Grade 8

Raw Score	Scale Score	SEM
0	230	80
1	230	80
2	230	80
3	230	80
4	230	80
5	230	80
6	230	80
7	230	80
8	230	80
9	230	80
10	265	45
11	283	27
12	295	21
13	305	18
14	313	16
15	320	16
16	326	15
17	333	15
18	338	14
19	344	14
20	**349**	**14**
21	354	13
22	359	13
23	364	13
24	369	12
25	373	12
26	**378**	**12**
27	382	12
28	387	11
29	392	11
30	396	11
31	401	12
32	406	12
33	412	12
34	418	13
35	**425**	**14**
36	434	16
37	444	18
38	459	22
39	484	33
40	560	109

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-24
Scoring Table for Science Grade 10

Raw Score	Scale Score	SEM	Raw Score	Scale Score	SEM
0	240	141	26	434	12
1	240	141	27	438	12
2	240	141	28	442	12
3	240	141	29	446	12
4	240	141	30	450	12
5	240	141	31	454	12
6	240	141	32	458	12
7	240	141	33	462	12
8	240	141	**34**	**466**	**12**
9	240	141	35	470	12
10	240	141	36	474	12
11	306	75	37	478	12
12	337	44	38	483	12
13	354	31	39	488	12
14	366	25	40	493	13
15	375	21	41	498	13
16	383	19	42	503	13
17	390	17	43	510	14
18	396	16	44	516	15
19	402	15	45	524	16
20	407	15	46	533	17
21	**412**	**14**	47	544	19
22	417	13	48	559	23
23	421	13	49	584	34
24	425	13	50	610	51
25	**430**	**12**			

* **Bold** represents SEM around cut score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 7-25
The Number of Students and Percents at LOSS and HOSS

Content	Grade	LOSS	N	Percent	HOSS	N	Percent
RD	3	270	628	1.07	640	11	0.02
	4	280	660	1.12	650	17	0.03
	5	290	834	1.43	690	10	0.02
	6	300	567	0.95	730	3	0.01
	7	310	524	0.86	780	2	0.00
	8	330	560	0.89	790	19	0.03
	10	350	865	1.27	820	84	0.12
MA	3	220	102	0.17	630	173	0.29
	4	240	82	0.14	650	102	0.17
	5	270	119	0.20	680	82	0.14
	6	310	152	0.25	700	32	0.05
	7	330	158	0.26	710	48	0.08
	8	350	540	0.86	730	35	0.06
	10	410	963	1.41	750	145	0.21
LA	4	140	218	0.37	420	223	0.38
	8	250	700	1.12	520	941	1.50
	10	290	289	0.43	630	5	0.01
SS	4	170	168	0.28	400	1241	2.10
	8	230	310	0.49	530	491	0.78
	10	240	370	0.54	620	129	0.19
SC	4	170	132	0.22	440	211	0.36
	8	230	148	0.24	560	789	1.26
	10	240	467	0.69	610	61	0.09

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-1
Item Analysis Grade 3 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.60	0.22	0.08%				
OP		2	MC	0.67	0.41	0.13%				
OP		3	MC	0.84	0.52	0.22%				
OP		4	MC	0.64	0.54	0.45%				
OP		5	MC	0.60	0.42	0.58%				
OP		6	MC	0.54	0.40	0.75%				
OP		7	MC	0.73	0.54	1.22%				
OP		8	MC	0.84	0.54	0.38%				
OP		9	MC	0.73	0.51	1.21%				
OP		10	MC	0.77	0.39	0.74%				
OP		11	MC	0.67	0.55	1.24%				
OP		12	MC	0.75	0.51	1.69%				
OP		13	MC	0.80	0.51	1.94%				
OP		14	MC	0.66	0.49	2.60%				
OP		15	MC	0.75	0.46	3.27%				
OP		16	MC	0.69	0.49	5.28%			+	
OP		17	MC	0.41	0.44	2.91%				
OP		18	MC	0.62	0.48	3.37%				
OP		19	CR	0.25	0.41	0.00%				+
OP		20	MC	0.57	0.44	0.17%				
OP		21	MC	0.82	0.44	0.28%				
OP		22	MC	0.75	0.52	0.88%				
OP		23	MC	0.55	0.37	0.42%				
OP		24	MC	0.68	0.43	0.49%				
OP		25	MC	0.61	0.53	1.30%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-1 Cont'd
Item Analysis Grade 3 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-Value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.72	0.53	0.38%				
OP		27	MC	0.42	0.24	0.75%		+		
OP		28	MC	0.77	0.50	1.14%				
OP		29	MC	0.70	0.55	0.61%				
OP		30	MC	0.65	0.36	1.35%				
OP		31	CR	0.40	0.48	0.00%				
OP		32	MC	0.95	0.30	1.50%				
OP		33	MC	0.78	0.46	1.86%				
OP		34	MC	0.82	0.51	2.35%				
OP		35	MC	0.85	0.45	2.46%				
OP		36	MC	0.68	0.39	2.22%				
OP		37	MC	0.69	0.39	2.38%				
OP		38	MC	0.45	0.41	2.65%				
OP		39	MC	0.74	0.47	0.17%				
OP		40	MC	0.64	0.49	0.20%				
OP		41	MC	0.48	0.34	0.36%				
OP		42	MC	0.74	0.58	0.75%				
OP		43	MC	0.79	0.51	2.05%				
OP		44	MC	0.56	0.48	0.28%				
OP		45	MC	0.79	0.45	0.33%				
OP		46	MC	0.85	0.56	0.31%				
OP		47	MC	0.53	0.25	1.30%				
OP		48	MC	0.81	0.43	0.36%				
OP		49	MC	0.62	0.36	0.53%				
OP		50	MC	0.81	0.50	0.74%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-1 Cont'd
Item Analysis Grade 3 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		51	MC	0.72	0.49	0.56%				
OP		52	MC	0.75	0.56	0.91%				
OP		53	MC	0.47	0.28	2.00%				
OP		54	MC	0.62	0.49	0.74%				
OP		55	MC	0.71	0.58	1.61%				
OP		56	MC	0.51	0.42	0.85%				
FT	A	57	MC	0.73	0.44	0.32%				
FT	A	58	MC	0.55	0.41	0.40%				
FT	A	59	MC	0.83	0.55	0.57%				
FT	A	60	MC	0.71	0.44	0.43%				
FT	A	61	MC	0.72	0.58	0.62%				
FT	A	62	MC	0.65	0.48	0.95%				
FT	A	63	CR	0.25	0.50	0.75%				+
FT	A	64	MC	0.61	0.43	1.60%				
FT	A	65	MC	0.73	0.45	1.78%				
FT	A	66	MC	0.43	0.24	2.28%				
FT	A	67	MC	0.47	0.29	3.00%				
FT	B	57	MC	0.71	0.46	1.48%				
FT	B	58	MC	0.70	0.52	0.33%				
FT	B	59	MC	0.70	0.50	0.58%				
FT	B	60	MC	0.70	0.36	1.24%				
FT	B	61	MC	0.52	0.41	0.52%				
FT	B	62	MC	0.77	0.52	0.73%				
FT	B	63	MC	0.52	0.38	1.52%				
FT	B	64	MC	0.55	0.36	1.10%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-1 Cont'd
Item Analysis Grade 3 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	65	MC	0.70	0.52	1.25%				
FT	B	66	MC	0.75	0.51	1.42%				
FT	B	67	CR	0.33	0.55	1.01%				
FT	C	57	MC	0.30	0.03	0.36%	+	+		
FT	C	58	MC	0.62	0.51	0.38%				
FT	C	59	MC	0.67	0.46	0.66%				
FT	C	60	MC	0.61	0.43	0.51%				
FT	C	61	MC	0.66	0.50	0.79%				
FT	C	62	MC	0.76	0.52	0.83%				
FT	C	63	MC	0.57	0.46	1.35%				
FT	C	64	CR	0.33	0.53	0.78%				
FT	C	65	MC	0.58	0.22	2.86%		+		
FT	C	66	MC	0.70	0.44	2.99%				
FT	C	67	MC	0.28	0.26	3.26%				+
FT	D	57	MC	0.69	0.34	0.38%				
FT	D	58	MC	0.73	0.59	0.60%				
FT	D	59	MC	0.70	0.38	0.43%				
FT	D	60	MC	0.52	0.29	0.81%				
FT	D	61	MC	0.69	0.50	0.59%				
FT	D	62	MC	0.54	-0.04	0.86%	+	+		
FT	D	63	MC	0.43	0.31	1.05%				
FT	D	64	MC	0.63	0.36	0.93%				
FT	D	65	MC	0.57	0.42	1.06%				
FT	D	66	MC	0.58	0.41	1.84%				
FT	D	67	CR	0.36	0.49	0.75%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-2
Item Analysis Grade 4 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.79	0.52	0.00%				
OP		2	MC	0.59	0.52	0.11%				
OP		3	MC	0.60	0.28	0.40%				
OP		4	MC	0.83	0.51	0.79%				
OP		5	MC	0.72	0.46	0.18%				
OP		6	MC	0.64	0.49	0.50%				
OP		7	MC	0.79	0.49	0.26%				
OP		8	MC	0.85	0.43	1.38%				
OP		9	MC	0.71	0.57	0.21%				
OP		10	MC	0.74	0.57	0.80%				
OP		11	MC	0.88	0.51	0.48%				
OP		12	MC	0.57	0.30	1.35%		+		
OP		13	CR	0.59	0.46	0.00%				
OP		14	MC	0.75	0.40	1.14%				
OP		15	MC	0.68	0.19	1.59%				
OP		16	MC	0.69	0.33	1.41%				
OP		17	MC	0.83	0.31	1.86%				
OP		18	MC	0.87	0.43	1.51%				
OP		19	MC	0.39	0.37	0.08%				
OP		20	MC	0.78	0.51	0.18%				
OP		21	MC	0.57	0.25	0.21%				
OP		22	MC	0.66	0.23	0.67%				
OP		23	MC	0.84	0.38	1.32%				
OP		24	MC	0.49	0.23	0.21%				
OP		25	MC	0.58	0.44	0.30%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-2 Cont'd
Item Analysis Grade 4 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.88	0.50	0.71%				
OP		27	MC	0.73	0.41	0.32%				
OP		28	MC	0.39	0.24	0.56%		+		
OP		29	MC	0.44	0.34	0.66%				
OP		30	MC	0.45	0.27	0.55%				
OP		31	MC	0.34	0.27	1.01%				
OP		32	MC	0.57	0.45	0.75%				
OP		33	MC	0.69	0.23	0.91%				
OP		34	MC	0.79	0.54	1.07%				
OP		35	MC	0.52	0.43	1.09%				
OP		36	MC	0.75	0.54	1.88%				
OP		37	MC	0.85	0.45	0.11%				
OP		38	MC	0.59	0.32	0.24%				
OP		39	MC	0.56	0.29	0.45%				
OP		40	MC	0.62	0.57	0.47%				
OP		41	MC	0.89	0.40	0.16%				
OP		42	MC	0.87	0.50	0.34%				
OP		43	MC	0.45	0.32	0.42%				
OP		44	MC	0.78	0.44	0.51%				
OP		45	MC	0.65	0.49	1.19%				
OP		46	MC	0.36	0.25	1.96%				
OP		47	CR	0.49	0.48	0.00%				
OP		48	MC	0.57	0.42	0.77%				
OP		49	MC	0.72	0.59	1.06%				
OP		50	MC	0.68	0.36	0.96%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-2 Cont'd
Item Analysis Grade 4 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		51	MC	0.65	0.46	0.82%				
OP		52	MC	0.76	0.51	1.17%				
OP		53	MC	0.56	0.41	1.07%				
OP		54	MC	0.69	0.58	1.07%				
OP		55	MC	0.61	0.43	1.24%				
OP		56	MC	0.65	0.43	1.57%				
FT	A	57	MC	0.63	0.35	0.22%				
FT	A	58	MC	0.31	0.31	0.27%		+		
FT	A	59	MC	0.53	0.40	0.55%				
FT	A	60	MC	0.38	0.27	0.80%		+		
FT	A	61	MC	0.42	0.31	0.37%				
FT	A	62	MC	0.69	0.56	0.65%				
FT	A	63	MC	0.68	0.36	0.70%				
FT	A	64	MC	0.68	0.20	0.74%				
FT	A	65	MC	0.48	0.41	0.95%				
FT	A	66	MC	0.65	0.44	1.20%				
FT	A	67	CR	0.35	0.54	0.80%				
FT	B	57	MC	0.63	0.51	0.32%				
FT	B	58	MC	0.82	0.46	0.26%				
FT	B	59	MC	0.87	0.43	0.46%				
FT	B	60	MC	0.63	0.37	0.37%				
FT	B	61	MC	0.55	0.52	0.48%				
FT	B	62	MC	0.46	0.35	0.64%				
FT	B	63	MC	0.59	0.43	0.83%				
FT	B	64	CR	0.48	0.45	0.34%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-2 Cont'd
Item Analysis Grade 4 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	65	MC	0.77	0.47	0.85%				
FT	B	66	MC	0.87	0.45	0.92%				
FT	B	67	MC	0.47	0.19	1.02%				
FT	C	57	MC	0.74	0.48	0.27%				
FT	C	58	MC	0.61	0.37	0.42%				
FT	C	59	MC	0.92	0.46	0.46%				
FT	C	60	MC	0.87	0.55	0.29%				
FT	C	61	MC	0.21	-0.02	0.42%	+	+		+
FT	C	62	MC	0.55	0.25	0.84%		+		
FT	C	63	MC	0.60	0.42	0.99%				
FT	C	64	MC	0.53	0.45	1.25%				
FT	C	65	MC	0.74	0.56	1.50%				
FT	C	66	MC	0.31	0.33	1.79%		+		
FT	C	67*	CR							
FT	D	57	MC	0.72	0.53	0.38%				
FT	D	58	MC	0.82	0.46	0.38%				
FT	D	59	MC	0.46	0.33	0.47%				
FT	D	60	MC	0.59	0.43	0.56%				
FT	D	61	MC	0.39	0.32	0.93%		+		
FT	D	62	CR	0.38	0.29	0.31%				
FT	D	63	MC	0.79	0.42	0.88%				
FT	D	64	MC	0.72	0.46	1.05%				
FT	D	65	MC	0.78	0.52	1.17%				
FT	D	66	MC	0.62	0.51	1.30%				
FT	D	67	MC	0.83	0.57	1.53%				

*This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-3
Item Analysis Grade 5 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.87	0.46	0.06%				
OP		2	MC	0.88	0.42	0.11%				
OP		3	MC	0.79	0.50	0.08%				
OP		4	MC	0.66	0.45	0.32%				
OP		5	MC	0.30	0.36	0.40%				
OP		6	MC	0.70	0.38	0.53%				
OP		7	MC	0.79	0.52	0.66%				
OP		8	MC	0.80	0.48	1.27%				
OP		9	MC	0.35	0.18	0.16%		+		
OP		10	MC	0.66	0.47	0.31%				
OP		11	MC	0.57	0.39	0.44%				
OP		12	MC	0.84	0.54	0.32%				
OP		13	CR	0.56	0.49	0.00%				
OP		14	MC	0.85	0.43	1.71%				
OP		15	MC	0.77	0.41	1.79%				
OP		16	MC	0.51	0.35	2.13%				
OP		17	MC	0.69	0.40	1.98%				
OP		18	MC	0.83	0.48	2.18%				
OP		19	MC	0.48	0.16	2.34%		+		
OP		20	MC	0.68	0.35	0.26%				
OP		21	MC	0.39	0.38	0.31%				
OP		22	MC	0.61	0.32	0.44%				
OP		23	MC	0.68	0.48	0.77%				
OP		24	MC	0.42	0.21	1.35%		+		
OP		25	MC	0.72	0.38	0.37%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-3 Cont'd
Item Analysis Grade 5 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.52	0.24	0.32%				
OP		27	MC	0.68	0.42	0.73%				
OP		28	MC	0.52	0.42	0.21%				
OP		29	MC	0.57	0.51	0.55%				
OP		30	MC	0.73	0.45	0.50%				
OP		31	MC	0.70	0.50	0.29%				
OP		32	MC	0.71	0.35	0.58%				
OP		33	MC	0.70	0.50	0.60%				
OP		34	MC	0.65	0.49	1.32%				
OP		35	MC	0.61	0.48	0.50%				
OP		36	MC	0.44	0.23	0.64%				
OP		37	MC	0.48	0.14	0.81%	+			
OP		38	MC	0.60	0.32	1.92%				
OP		39	MC	0.94	0.40	0.13%				
OP		40	MC	0.83	0.48	0.92%				
OP		41	MC	0.91	0.51	0.24%				
OP		42	MC	0.86	0.48	0.15%				
OP		43	MC	0.86	0.45	0.47%				
OP		44	MC	0.81	0.52	0.73%				
OP		45	MC	0.81	0.41	1.35%				
OP		46	MC	0.50	0.16	2.84%		+		
OP		47	CR	0.55	0.56	0.00%				
OP		48	MC	0.87	0.45	0.40%				
OP		49	MC	0.60	0.50	0.64%				
OP		50	MC	0.54	0.27	0.81%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-3 Cont'd
Item Analysis Grade 5 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		51	MC	0.56	0.38	0.64%				
OP		52	MC	0.74	0.55	0.48%				
OP		53	MC	0.67	0.53	0.60%				
OP		54	MC	0.73	0.49	3.61%				
OP		55	MC	0.67	0.60	0.82%				
OP		56	MC	0.64	0.40	1.16%				
FT	A	57	MC	0.69	0.54	0.26%				
FT	A	58	MC	0.44	0.23	0.34%				
FT	A	59	MC	0.75	0.58	0.43%				
FT	A	60	MC	0.77	0.51	0.65%				
FT	A	61	MC	0.64	0.33	0.36%				
FT	A	62	MC	0.71	0.45	0.48%				
FT	A	63	MC	0.63	0.54	0.48%				
FT	A	64	MC	0.71	0.54	0.38%				
FT	A	65	MC	0.70	0.28	0.60%				
FT	A	66	MC	0.68	0.52	0.69%				
FT	A	67	CR	0.36	0.41	0.52%				
FT	B	57	MC	0.93	0.38	0.18%				
FT	B	58	MC	0.39	0.29	0.24%				
FT	B	59	MC	0.42	0.25	0.56%				
FT	B	60	MC	0.85	0.52	0.46%				
FT	B	61	MC	0.41	0.15	0.61%		+		
FT	B	62	MC	0.42	0.23	0.75%				
FT	B	63	MC	0.69	0.44	0.53%				
FT	B	64	MC	0.43	0.08	0.46%	+	+		

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-3 Cont'd
Item Analysis Grade 5 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	65	MC	0.52	0.29	0.31%		+		
FT	B	66	MC	0.57	0.44	0.62%				
FT	B	67*	CR							
FT	C	57	CR	0.30	0.36	0.32%				
FT	C	58	MC	0.68	0.37	0.26%				
FT	C	59	MC	0.48	0.26	0.32%				
FT	C	60	MC	0.83	0.24	0.33%				
FT	C	61	MC	0.59	0.32	0.31%				
FT	C	62	MC	0.49	0.29	0.25%				
FT	C	63	CR	0.45	0.43	0.30%				
FT	D	57*	CR							
FT	D	58	MC	0.82	0.49	0.24%				
FT	D	59	MC	0.88	0.48	0.39%				
FT	D	60	MC	0.82	0.32	0.42%				
FT	D	61	MC	0.77	0.50	0.46%				
FT	D	62	MC	0.82	0.51	0.65%				
FT	D	63	CR	0.35	0.41	0.53%				

*This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-4
Item Analysis Grade 6 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.53	0.27	0.19%				
OP		2	MC	0.91	0.39	0.11%				
OP		3	MC	0.62	0.46	0.11%				
OP		4	MC	0.78	0.53	0.19%				
OP		5	MC	0.68	0.23	0.75%				
OP		6	MC	0.68	0.46	0.11%				
OP		7	MC	0.47	0.28	0.08%				
OP		8	MC	0.83	0.44	2.36%				
OP		9	MC	0.55	0.17	3.65%				
OP		10	MC	0.36	0.23	0.19%				
OP		11	MC	0.81	0.36	0.18%				
OP		12	MC	0.66	0.32	0.38%				
OP		13	MC	0.47	0.38	0.48%				
OP		14	MC	0.81	0.37	0.16%				
OP		15	MC	0.82	0.34	0.46%				
OP		16	MC	0.72	0.45	0.18%				
OP		17	MC	0.77	0.42	0.27%				
OP		18	CR	0.50	0.43	0.00%				
OP		19	MC	0.36	0.17	0.19%		+		
OP		20	MC	0.85	0.15	0.18%				
OP		21	MC	0.69	0.48	0.51%				
OP		22	MC	0.35	0.20	0.40%				
OP		23	MC	0.55	0.36	0.72%				
OP		24	CR	0.40	0.58	0.00%				
OP		25	MC	0.55	0.34	0.26%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-4 Cont'd
Item Analysis Grade 6 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.42	0.17	0.77%		+		
OP		27	MC	0.54	0.49	0.40%				
OP		28	MC	0.76	0.29	2.29%				
OP		29	MC	0.45	0.30	0.51%				
OP		30	MC	0.60	0.35	0.98%				
OP		31	MC	0.69	0.42	1.41%				
OP		32	MC	0.45	0.29	2.85%				
OP		33	MC	0.46	0.36	3.43%				
OP		34	MC	0.58	0.38	1.88%				
OP		35	MC	0.84	0.45	2.69%				
OP		36	MC	0.77	0.45	2.23%				
OP		37	MC	0.81	0.52	1.96%				
OP		38	MC	0.84	0.54	0.21%				
OP		39	MC	0.64	0.44	0.26%				
OP		40	MC	0.72	0.31	0.38%				
OP		41	MC	0.73	0.41	0.37%				
OP		42	MC	0.84	0.51	0.61%				
OP		43	MC	0.66	0.36	2.08%				
OP		44	MC	0.70	0.17	0.32%				
OP		45	MC	0.78	0.41	0.35%				
OP		46	MC	0.87	0.40	1.97%				
OP		47	MC	0.77	0.41	2.90%				
OP		48	MC	0.63	0.34	0.37%				
OP		49	MC	0.77	0.50	0.37%				
OP		50	MC	0.41	0.37	0.85%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-4 Cont'd
Item Analysis Grade 6 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		51	MC	0.68	0.37	1.23%				
OP		52	MC	0.56	0.39	0.51%				
OP		53	MC	0.65	0.44	0.71%				
OP		54	MC	0.94	0.39	0.38%				
OP		55	MC	0.85	0.45	1.22%				
OP		56	MC	0.50	0.38	0.38%				
FT	A	57	MC	0.77	0.49	0.29%				
FT	A	58	MC	0.50	0.31	0.34%				
FT	A	59	MC	0.84	0.44	0.60%				
FT	A	60	MC	0.54	0.18	0.47%				
FT	A	61	MC	0.70	0.50	0.57%				
FT	A	62	MC	0.90	0.36	0.40%				
FT	A	63*	MC							
FT	A	64	MC	0.79	0.49	0.52%				
FT	A	65	MC	0.59	0.22	0.61%				
FT	A	66	MC	0.47	0.32	1.14%				
FT	A	67	CR	0.29	0.40	0.47%				+
FT	B	57	MC	0.76	0.43	0.62%				
FT	B	58	MC	0.69	0.21	0.67%				
FT	B	59	MC	0.45	0.08	0.75%	+			
FT	B	60	MC	0.80	0.45	0.42%				
FT	B	61	MC	0.81	0.53	0.52%				
FT	B	62	MC	0.61	0.47	0.52%				
FT	B	63	MC	0.72	0.40	0.53%				
FT	B	64	MC	0.65	0.45	0.48%				

*This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-4 Cont'd
Item Analysis Grade 6 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	65	MC	0.63	0.48	0.69%				
FT	B	66	MC	0.76	0.45	0.70%				
FT	B	67	CR	0.38	0.50	0.59%				
FT	C	57	CR	0.33	0.47	0.41%				
FT	C	58	MC	0.57	0.43	0.25%				
FT	C	59	MC	0.49	0.26	0.43%				
FT	C	60	MC	0.66	0.28	0.33%				
FT	C	61	MC	0.57	0.28	1.30%		+		
FT	C	62	MC	0.56	0.49	0.53%				
FT	C	63	MC	0.44	0.37	0.59%				
FT	C	64	MC	0.72	0.41	0.60%				
FT	C	65	CR	0.41	0.58	0.39%				
FT	D	57	MC	0.65	0.24	0.34%				
FT	D	58	MC	0.53	0.07	0.36%	+	+		
FT	D	59	MC	0.62	0.22	0.44%		+		
FT	D	60	MC	0.40	0.26	0.54%				
FT	D	61	MC	0.54	0.44	0.44%				
FT	D	62	MC	0.46	0.29	0.59%		+		
FT	D	63	MC	0.38	0.33	0.87%				
FT	D	64	MC	0.61	0.42	0.74%				
FT	D	65	MC	0.73	0.45	0.81%				
FT	D	66	MC	0.63	0.43	0.97%				
FT	D	67	CR	0.35	0.48	0.92%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-5
Item Analysis Grade 7 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.71	0.56	0.22%				
OP		2	MC	0.72	0.33	0.19%				
OP		3	MC	0.70	0.35	0.19%				
OP		4	MC	0.56	0.38	0.09%				
OP		5	MC	0.67	0.32	0.06%				
OP		6	MC	0.89	0.43	0.19%				
OP		7	MC	0.33	0.34	0.25%				
OP		8	MC	0.86	0.51	0.33%				
OP		9	MC	0.62	0.26	1.17%				
OP		10	MC	0.69	0.35	1.20%				
OP		11	MC	0.72	0.32	2.27%				
OP		12	MC	0.77	0.38	0.33%				
OP		13	MC	0.33	0.20	0.75%		+		
OP		14	MC	0.71	0.41	1.17%				
OP		15	MC	0.60	0.36	1.33%				
OP		16	MC	0.63	0.33	1.30%				
OP		17	MC	0.48	0.34	1.33%				
OP		18	MC	0.79	0.42	0.36%				
OP		19	MC	0.61	0.42	0.16%				
OP		20	MC	0.45	0.43	0.47%				
OP		21	MC	0.45	0.31	0.42%		+		
OP		22	MC	0.79	0.47	0.59%				
OP		23	MC	0.45	0.43	0.67%				
OP		24	MC	0.94	0.34	0.22%				
OP		25	MC	0.72	0.50	0.45%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-5 Cont'd
Item Analysis Grade 7 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.62	0.50	0.94%				
OP		27	MC	0.47	0.34	1.92%				
OP		28	MC	0.75	0.33	2.44%				
OP		29	MC	0.66	0.40	0.22%				
OP		30	MC	0.73	0.36	0.33%				
OP		31	CR	0.51	0.43	0.00%				
OP		32	MC	0.92	0.37	0.41%				
OP		33	MC	0.81	0.44	0.53%				
OP		34	MC	0.41	0.25	0.59%				
OP		35	MC	0.76	0.29	0.63%				
OP		36	MC	0.84	0.46	0.63%				
OP		37	MC	0.68	0.42	0.27%				
OP		38	MC	0.64	0.26	0.36%				
OP		39	MC	0.81	0.53	0.33%				
OP		40	MC	0.88	0.44	1.19%				
OP		41	MC	0.76	0.49	1.49%				
OP		42	MC	0.58	0.40	0.30%				
OP		43	MC	0.65	0.34	0.47%				
OP		44	MC	0.49	0.30	0.38%				
OP		45	MC	0.43	0.09	1.44%	+	+		
OP		46	MC	0.82	0.47	0.28%				
OP		47	MC	0.60	0.38	0.48%				
OP		48	MC	0.68	0.34	0.83%				
OP		49	MC	0.40	0.27	1.45%				
OP		50	CR	0.39	0.45	0.00%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-5 Cont'd
Item Analysis Grade 7 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		51	MC	0.35	0.29	0.36%				
OP		52	MC	0.71	0.48	0.38%				
OP		53	MC	0.70	0.36	0.52%				
OP		54	MC	0.79	0.48	0.84%				
OP		55	MC	0.92	0.44	0.38%				
OP		56	MC	0.83	0.49	0.39%				
FT	A	57	MC	0.93	0.39	0.40%				
FT	A	58	MC	0.83	0.49	0.53%				
FT	A	59	MC	0.87	0.40	0.53%				
FT	A	60	MC	0.37	0.23	0.41%				
FT	A	61	MC	0.57	0.21	0.84%				
FT	A	62	MC	0.75	0.47	0.42%				
FT	A	63	MC	0.21	0.09	0.48%	+	+		+
FT	A	64	MC	0.31	0.08	0.54%	+	+		
FT	A	65	MC	0.92	0.40	0.56%				
FT	A	66	MC	0.40	0.27	0.69%				
FT	A	67	CR	0.34	0.59	0.60%				
FT	B	57	MC	0.78	0.41	1.49%				
FT	B	58	MC	0.58	0.40	0.49%				
FT	B	59	MC	0.51	0.26	0.47%				
FT	B	60	MC	0.82	0.53	0.41%				
FT	B	61	MC	0.25	0.23	0.63%				+
FT	B	62	MC	0.65	0.46	0.81%				
FT	B	63	CR	0.56	0.65	0.41%				
FT	B	64	MC	0.83	0.49	0.49%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-5 Cont'd
Item Analysis Grade 7 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	65	MC	0.90	0.44	0.55%				
FT	B	66	MC	0.66	0.40	0.59%				
FT	B	67	MC	0.34	0.17	0.63%				
FT	C	57	MC	0.73	0.48	0.36%				
FT	C	58	MC	0.74	0.55	0.36%				
FT	C	59	MC	0.47	0.35	0.41%				
FT	C	60	MC	0.80	0.56	0.43%				
FT	C	61	MC	0.26	0.30	0.45%		+		+
FT	C	62	MC	0.20	-0.01	0.62%	+	+		+
FT	C	63	MC	0.82	0.34	0.39%				
FT	C	64	MC	0.43	0.36	0.77%				
FT	C	65	MC	0.85	0.52	0.53%				
FT	C	66	MC	0.83	0.46	0.58%				
FT	C	67	CR	0.37	0.44	0.65%				
FT	D	57	MC	0.69	0.27	0.44%				
FT	D	58	MC	0.90	0.40	0.31%				
FT	D	59	MC	0.70	0.38	0.35%				
FT	D	60	MC	0.84	0.31	0.35%				
FT	D	61	MC	0.80	0.25	0.29%				
FT	D	62	MC	0.49	0.43	0.38%				
FT	D	63	MC	0.59	0.39	0.39%				
FT	D	64	MC	0.73	0.36	0.38%				
FT	D	65	MC	0.38	0.17	0.55%		+		
FT	D	66	MC	0.59	0.38	0.88%				
FT	D	67	CR	0.39	0.46	0.63%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-6
Item Analysis Grade 8 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.96	0.22	0.18%				
OP		2	MC	0.82	0.41	0.15%				
OP		3	MC	0.78	0.34	0.03%				
OP		4	MC	0.86	0.28	0.23%				
OP		5	MC	0.79	0.41	0.11%				
OP		6	MC	0.79	0.38	0.17%				
OP		7	MC	0.52	0.14	1.00%	+			
OP		8	MC	0.79	0.43	1.14%				
OP		9	MC	0.65	0.34	0.08%				
OP		10	MC	0.67	0.37	1.37%				
OP		11	MC	0.71	0.49	0.26%				
OP		12	MC	0.90	0.42	0.15%				
OP		13	MC	0.76	0.49	0.18%				
OP		14	MC	0.91	0.33	0.17%				
OP		15	MC	0.49	0.31	0.75%				
OP		16	MC	0.62	0.36	0.32%				
OP		17	MC	0.76	0.43	0.25%				
OP		18	MC	0.70	0.29	0.40%				
OP		19	CR	0.56	0.52	0.00%				
OP		20	MC	0.68	0.31	0.15%				
OP		21	MC	0.82	0.49	0.11%				
OP		22	MC	0.55	0.37	0.17%				
OP		23	MC	0.84	0.48	0.28%				
OP		24	MC	0.74	0.51	0.49%				
OP		25	MC	0.68	0.57	0.65%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-6 Cont'd
Item Analysis Grade 8 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.42	0.32	1.11%				
OP		27	MC	0.49	0.40	0.35%				
OP		28	MC	0.58	0.34	0.77%				
OP		29	MC	0.60	0.20	1.40%				
OP		30	MC	0.63	0.49	0.45%				
OP		31	MC	0.76	0.40	0.55%				
OP		32	MC	0.82	0.54	0.22%				
OP		33	MC	0.45	0.23	0.45%		+		
OP		34	MC	0.62	0.26	0.60%		+		
OP		35	MC	0.63	0.38	0.35%				
OP		36	MC	0.82	0.50	0.52%				
OP		37	MC	0.74	0.35	0.43%				
OP		38	MC	0.78	0.49	0.40%				
OP		39	MC	0.73	0.39	0.42%				
OP		40	MC	0.69	0.40	0.32%				
OP		41	MC	0.65	0.44	0.20%				
OP		42	MC	0.52	0.37	1.34%				
OP		43	MC	0.77	0.52	0.20%				
OP		44	MC	0.43	0.16	0.35%		+		
OP		45	MC	0.76	0.34	0.52%				
OP		46	MC	0.66	0.13	0.29%	+	+		
OP		47	MC	0.63	0.34	0.31%				
OP		48	MC	0.79	0.53	0.51%				
OP		49	MC	0.45	0.31	0.60%				
OP		50	MC	0.89	0.46	0.88%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-6 Cont'd
Item Analysis Grade 8 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		51	MC	0.84	0.45	0.28%				
OP		52	MC	0.76	0.30	0.35%				
OP		53	MC	0.57	0.48	0.77%				
OP		54	MC	0.47	0.36	0.46%				
OP		55	MC	0.57	0.28	0.52%		+		
OP		56	CR	0.58	0.50	0.00%				
FT	A	57	MC	0.68	0.41	0.55%				
FT	A	58	MC	0.74	0.47	0.54%				
FT	A	59	MC	0.84	0.44	0.52%				
FT	A	60	MC	0.75	0.35	0.67%				
FT	A	61	MC	0.36	0.32	0.55%		+		
FT	A	62	MC	0.81	0.46	0.54%				
FT	A	63	MC	0.73	0.41	0.72%				
FT	A	64	MC	0.72	0.47	0.74%				
FT	A	65	CR	0.45	0.59	0.53%				
FT	A	66	MC	0.58	0.17	4.53%		+		
FT	A	67	MC	0.56	0.21	4.57%				
FT	B	57	MC	0.54	0.45	0.30%				
FT	B	58	MC	0.95	0.38	0.26%				
FT	B	59	MC	0.73	0.49	0.28%				
FT	B	60	MC	0.63	0.49	0.32%				
FT	B	61	MC	0.85	0.44	0.30%				
FT	B	62	MC	0.75	0.28	0.29%				
FT	B	63	MC	0.59	0.40	0.33%				
FT	B	64	MC	0.72	0.37	0.36%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-6 Cont'd
Item Analysis Grade 8 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	65	MC	0.68	0.28	0.58%				
FT	B	66*	MC							
FT	B	67	CR	0.40	0.55	0.41%				
FT	C	57	MC	0.75	0.34	0.37%				
FT	C	58	MC	0.45	0.27	0.51%				
FT	C	59	MC	0.88	0.43	0.62%				
FT	C	60	MC	0.69	0.46	0.46%				
FT	C	61	MC	0.53	0.29	0.47%				
FT	C	62	MC	0.53	0.21	0.50%				
FT	C	63	CR	0.33	0.49	0.39%				
FT	C	64	MC	0.67	0.31	1.86%				
FT	C	65	MC	0.60	0.19	1.91%		+		
FT	C	66	MC	0.62	0.31	1.90%				
FT	C	67	MC	0.55	0.34	1.96%				
FT	D	57	MC	0.71	0.35	0.31%				
FT	D	58	MC	0.41	0.07	0.30%	+	+		
FT	D	59	MC	0.64	0.36	0.42%				
FT	D	60	CR	0.48	0.56	0.32%				
FT	D	61	MC	0.42	0.28	0.33%		+		
FT	D	62	MC	0.40	0.24	0.37%				
FT	D	63	CR	0.51	0.50	0.22%				

*This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-7
Item Analysis Grade 10 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.87	0.39	0.11%				
OP		2	MC	0.90	0.39	0.07%				
OP		3	MC	0.62	0.31	0.23%				
OP		4	MC	0.82	0.42	0.18%				
OP		5	MC	0.45	0.40	0.15%				
OP		6	MC	0.85	0.39	0.75%				
OP		7	CR	0.42	0.53	0.00%				
OP		8	MC	0.48	0.35	0.17%				
OP		9	MC	0.74	0.51	0.28%				
OP		10	MC	0.55	0.28	0.25%				
OP		11	MC	0.69	0.41	0.48%				
OP		12	MC	0.75	0.48	0.40%				
OP		13	MC	0.74	0.45	0.46%				
OP		14	MC	0.80	0.49	0.65%				
OP		15	MC	0.84	0.52	0.68%				
OP		16	MC	0.39	0.47	0.72%				
OP		17	MC	0.36	0.33	0.41%		+		
OP		18	MC	0.48	0.38	0.50%				
OP		19	MC	0.90	0.36	0.43%				
OP		20	MC	0.69	0.40	0.43%				
OP		21	MC	0.73	0.54	0.84%				
OP		22	MC	0.51	0.50	0.90%				
OP		23	MC	0.89	0.50	0.43%				
OP		24	MC	0.79	0.48	0.55%				
OP		25	MC	0.79	0.46	0.57%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-7 Cont'd
Item Analysis Grade 10 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.80	0.52	0.47%				
OP		27	MC	0.71	0.46	0.58%				
OP		28	MC	0.66	0.53	1.92%				
OP		29	MC	0.79	0.39	2.07%				
OP		30	MC	0.51	0.45	1.44%				
OP		31	MC	0.41	0.47	2.29%				
OP		32	MC	0.80	0.54	0.63%				
OP		33	MC	0.82	0.50	0.70%				
OP		34	MC	0.78	0.39	1.39%				
OP		35	MC	0.64	0.39	1.46%				
OP		36	MC	0.66	0.39	0.84%				
OP		37	MC	0.61	0.43	0.75%				
OP		38	MC	0.56	0.43	0.87%				
OP		39	MC	0.42	0.34	0.87%				
OP		40	MC	0.71	0.46	0.95%				
OP		41	MC	0.58	0.34	0.73%				
OP		42	MC	0.81	0.52	1.93%				
OP		43	MC	0.69	0.39	0.73%				
OP		44	MC	0.91	0.47	0.72%				
OP		45	MC	0.56	0.43	0.80%				
OP		46	MC	0.48	0.39	0.87%				
OP		47	MC	0.72	0.57	0.81%				
OP		48	MC	0.58	0.37	1.15%				
OP		49	MC	0.51	0.28	0.94%		+		
OP		50	MC	0.65	0.51	0.90%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-7 Cont'd
Item Analysis Grade 10 Reading

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		51	MC	0.49	0.29	1.01%				
OP		52	CR	0.49	0.60	0.00%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-8
Item Analysis Grade 3 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.94	0.31	0.26%				
OP		2	MC	0.93	0.40	0.23%				
OP		3	MC	0.82	0.47	0.22%				
OP		4	MC	0.72	0.34	0.22%				
OP		5	CR	0.49	0.51	0.00%				
OP		6	MC	0.83	0.42	0.42%				
OP		7	MC	0.64	0.45	0.47%				
OP		8	MC	0.87	0.36	0.44%				
OP		9	MC	0.88	0.33	1.39%				
OP		10	MC	0.82	0.28	1.18%				
OP		11	MC	0.91	0.31	0.81%				
OP		12	MC	0.62	0.35	0.84%				
OP		13	MC	0.81	0.43	1.40%				
OP		14	MC	0.95	0.26	1.81%				
OP		15	MC	0.89	0.27	0.42%				
OP		16	MC	0.73	0.40	0.42%				
OP		17	MC	0.60	0.27	3.21%				
OP		18	MC	0.89	0.36	0.37%				
OP		19	MC	0.71	0.36	0.68%				
OP		20A	CR	0.46	0.42	0.00%				
OP		20B	CR	0.28	0.45	0.00%				+
OP		21	MC	0.83	0.45	0.65%				
OP		22	MC	0.78	0.30	0.59%				
OP		23	MC	0.66	0.40	0.65%				
OP		24	MC	0.80	0.25	1.34%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-8 Cont'd
Item Analysis Grade 3 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		25	MC	0.63	0.45	1.15%				
OP		26	MC	0.83	0.41	1.85%				
OP		27A	CR	0.61	0.51	0.00%				
OP		27B	CR	0.29	0.47	0.00%				+
OP		28	MC	0.73	0.40	3.10%				
OP		29	MC	0.51	0.34	3.73%				
OP		30	MC	0.72	0.42	3.64%				
OP		31	MC	0.84	0.42	0.33%				
OP		32	MC	0.71	0.48	1.21%				
OP		33	MC	0.88	0.38	0.78%				
OP		34	MC	0.59	0.32	0.56%				
OP		35	MC	0.87	0.30	0.47%				
OP		36	MC	0.29	0.35	2.79%				+
OP		37	MC	0.88	0.20	0.59%				
OP		38	MC	0.91	0.25	0.73%				
OP		39A	CR	0.31	0.49	0.00%				
OP		39B	CR	0.29	0.55	0.00%				+
OP		40	MC	0.89	0.33	0.75%				
OP		41	MC	0.93	0.33	0.81%				
OP		42	MC	0.59	0.35	0.89%				
OP		43	MC	0.68	0.43	0.70%				
OP		44	MC	0.92	0.28	0.79%				
OP		45	MC	0.77	0.43	1.00%				
OP		46	MC	0.69	0.30	0.87%				
OP		47	MC	0.68	0.43	1.15%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-8 Cont'd
Item Analysis Grade 3 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		48	MC	0.73	0.40	1.37%				
OP		49	MC	0.72	0.45	1.62%				
OP		50	MC	0.93	0.30	1.35%				
FT	A	51	CR	0.51	0.43	0.40%				
FT	A	52	MC	0.37	0.30	0.55%				
FT	A	53	MC	0.79	0.34	0.82%				
FT	A	54	MC	0.91	0.38	0.67%				
FT	A	55	MC	0.80	0.28	0.77%				
FT	A	56	MC	0.89	0.42	0.72%				
FT	A	57A	CR	0.86	0.33	0.43%				
FT	A	57B	CR	0.41	0.42	0.41%				
FT	A	58	MC	0.61	0.25	0.99%				
FT	A	59	MC	0.69	0.27	1.18%				
FT	B	51	CR	0.57	0.50	7.76%			+	
FT	B	52	MC	0.15	0.04	0.45%	+	+		+
FT	B	53	MC	0.96	0.21	0.65%				
FT	B	54	MC	0.42	0.37	0.77%				
FT	B	55	MC	0.95	0.32	0.74%				
FT	B	56	MC	0.60	0.45	0.70%				
FT	B	57A*	CR							
FT	B	57B*	CR							
FT	B	58	MC	0.48	0.33	1.12%				
FT	B	59	MC	0.22	0.22	1.48%		+		+
FT	C	51	MC	0.62	0.55	0.56%				
FT	C	52	MC	0.74	0.44	0.72%				

*This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-8 Cont'd
Item Analysis Grade 3 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	C	53	MC	0.73	0.28	1.84%				
FT	C	54	MC	0.69	0.32	0.60%				
FT	C	55	MC	0.95	0.30	0.58%				
FT	C	56	MC	0.93	0.26	0.65%				
FT	C	57	MC	0.94	0.32	1.61%				
FT	C	58	MC	0.82	0.39	2.50%				
FT	C	59	MC	0.46	0.31	2.98%				
FT	C	60	CR	0.80	0.42	0.70%				
FT	C	61	MC	0.35	0.18	1.10%		+		
FT	D	51	MC	0.84	0.30	0.47%				
FT	D	52	MC	0.83	0.25	0.61%				
FT	D	53	MC	0.80	0.37	0.74%				
FT	D	54	MC	0.58	0.50	0.59%				
FT	D	55	MC	0.75	0.39	0.58%				
FT	D	56	MC	0.84	0.37	0.77%				
FT	D	57	MC	0.50	0.48	0.67%				
FT	D	58A	CR	0.85	0.28	0.23%				
FT	D	58B	CR	0.10	0.31	0.40%				+
FT	D	59	MC	0.63	0.32	1.40%				
FT	D	60	MC	0.89	0.34	1.45%				
FT	D	61	MC	0.91	0.36	1.76%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-9
Item Analysis Grade 4 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.97	0.14	0.08%	+			
OP		2	MC	0.90	0.21	0.14%				
OP		3	MC	0.89	0.27	0.19%				
OP		4	MC	0.90	0.42	0.33%				
OP		5	MC	0.86	0.44	0.21%				
OP		6	MC	0.77	0.41	0.56%				
OP		7	MC	0.74	0.36	0.69%				
OP		8	MC	0.90	0.39	0.24%				
OP		9	MC	0.62	0.44	0.46%				
OP		10	MC	0.96	0.33	0.51%				
OP		11A	CR	0.22	0.33	0.00%				+
OP		11B	CR	0.29	0.37	0.00%				+
OP		12	MC	0.64	0.56	1.12%				
OP		13	MC	0.75	0.54	0.88%				
OP		14	MC	0.82	0.44	1.05%				
OP		15	MC	0.93	0.33	0.14%				
OP		16	MC	0.70	0.43	0.41%				
OP		17	MC	0.67	0.35	0.26%				
OP		18	MC	0.86	0.34	0.53%				
OP		19	MC	0.70	0.29	0.33%				
OP		20	MC	0.98	0.15	0.30%				
OP		21	MC	0.72	0.42	0.77%				
OP		22	CR	0.79	0.24	0.00%				
OP		23	MC	0.72	0.32	0.49%				
OP		24	MC	0.82	0.32	0.57%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-9 Cont'd
Item Analysis Grade 4 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		25A	CR	0.47	0.53	0.00%				
OP		25B	CR	0.33	0.47	0.00%				
OP		26	MC	0.65	0.34	0.49%				
OP		27	MC	0.90	0.30	0.57%				
OP		28	MC	0.82	0.40	0.70%				
OP		29	MC	0.77	0.40	0.85%				
OP		30	MC	0.43	0.38	0.85%				
OP		31	MC	0.91	0.31	0.24%				
OP		32	MC	0.93	0.32	0.89%				
OP		33	MC	0.94	0.26	0.30%				
OP		34	MC	0.76	0.37	0.46%				
OP		35	MC	0.87	0.26	0.38%				
OP		36	MC	0.97	0.15	1.13%				
OP		37	MC	0.54	0.36	0.53%				
OP		38	MC	0.70	0.49	5.96%			+	
OP		39A	CR	0.56	0.43	0.00%				
OP		39B	CR	0.36	0.48	0.00%				
OP		40	MC	0.56	0.40	0.26%				
OP		41	MC	0.90	0.20	0.54%				
OP		42	MC	0.71	0.26	0.45%				
OP		43	MC	0.80	0.36	0.54%				
OP		44	MC	0.45	0.25	0.72%				
OP		45	MC	0.56	0.45	0.75%				
OP		46	MC	0.64	0.42	1.16%				
OP		47	MC	0.59	0.47	0.80%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-9 Cont'd
Item Analysis Grade 4 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		48	MC	0.77	0.27	0.86%				
OP		49	MC	0.89	0.41	1.13%				
OP		50	MC	0.86	0.32	1.24%				
FT	A	51	CR	0.57	0.55	2.76%				
FT	A	52	MC	0.66	0.24	0.30%				
FT	A	53	MC	0.64	0.48	0.31%				
FT	A	54	MC	0.71	0.26	0.59%				
FT	A	55	MC	0.76	0.39	0.40%				
FT	A	56*	MC							
FT	A	57	MC	0.57	0.36	0.64%				
FT	A	58A*	CR							
FT	A	58B*	CR							
FT	A	59	MC	0.70	0.40	0.62%				
FT	B	51	MC	0.82	0.28	0.31%				
FT	B	52	MC	0.59	0.42	0.36%				
FT	B	53	MC	0.89	0.34	0.47%				
FT	B	54	MC	0.93	0.26	0.30%				
FT	B	55	MC	0.58	0.45	0.35%				
FT	B	56	MC	0.73	0.58	0.51%				
FT	B	57	MC	0.27	0.23	0.47%		+		+
FT	B	58	MC	0.38	0.11	0.70%	+			
FT	B	59A	CR	0.91	0.40	0.14%				
FT	B	59B	CR	0.53	0.43	0.32%				
FT	B	60	MC	0.77	0.55	0.66%				
FT	B	61	MC	0.79	0.45	1.00%				

*This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-9 Cont'd
Item Analysis Grade 4 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	C	51	MC	0.85	0.38	0.36%				
FT	C	52	MC	0.79	0.35	0.40%				
FT	C	53	MC	0.96	0.29	0.42%				
FT	C	54	MC	0.10	-0.07	0.79%	+	+		+
FT	C	55	MC	0.31	0.22	0.63%				
FT	C	56	MC	0.60	0.35	0.96%				
FT	C	57	MC	0.80	0.31	0.68%				
FT	C	58A*	CR							
FT	C	58B*	CR							
FT	C	59	MC	0.67	0.37	0.94%				
FT	C	60	MC	0.62	0.32	1.37%				
FT	C	61	MC	0.58	0.43	1.70%				
FT	D	51	MC	0.77	0.38	0.29%				
FT	D	52	MC	0.83	0.40	0.41%				
FT	D	53	MC	0.64	0.51	0.51%				
FT	D	54	MC	0.93	0.25	0.29%				
FT	D	55	MC	0.59	0.43	0.32%				
FT	D	56	MC	0.73	0.57	0.44%				
FT	D	57	MC	0.83	0.45	0.37%				
FT	D	58	MC	0.80	0.35	0.58%				
FT	D	59*	CR							
FT	D	60	MC	0.20	0.16	0.78%		+		+
FT	D	61	MC	0.61	0.29	1.27%		+		

*This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-10
Item Analysis Grade 5 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.89	0.32	0.14%				
OP		2	MC	0.84	0.39	0.16%				
OP		3	MC	0.68	0.34	0.11%				
OP		4	MC	0.59	0.48	0.74%				
OP		5	MC	0.94	0.27	0.16%				
OP		6	MC	0.75	0.53	0.22%				
OP		7	MC	0.43	0.58	0.51%				
OP		8	MC	0.71	0.43	0.61%				
OP		9	MC	0.50	0.25	0.69%				
OP		10	MC	0.67	0.54	0.80%				
OP		11	MC	0.62	0.40	0.43%				
OP		12	CR	0.65	0.59	0.00%				
OP		13	MC	0.51	0.42	0.82%				
OP		14	MC	0.67	0.45	1.22%				
OP		15	MC	0.88	0.32	0.24%				
OP		16	MC	0.63	0.30	0.37%				
OP		17	MC	0.76	0.34	0.22%				
OP		18	MC	0.73	0.38	0.66%				
OP		19	MC	0.59	0.18	0.34%				
OP		20A	CR	0.85	0.24	0.00%				
OP		20B	CR	0.25	0.43	0.00%				+
OP		21	MC	0.38	0.38	0.24%				
OP		22	MC	0.87	0.32	0.18%				
OP		23	MC	0.67	0.53	0.34%				
OP		24	MC	0.95	0.24	0.37%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-10 Cont'd
Item Analysis Grade 5 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		25	MC	0.70	0.45	0.30%				
OP		26	MC	0.51	0.39	0.71%				
OP		27A	CR	0.60	0.45	0.00%				
OP		27B	CR	0.71	0.47	0.00%				
OP		28	MC	0.81	0.37	0.48%				
OP		29	MC	0.88	0.31	0.27%				
OP		30	MC	0.50	0.45	0.74%				
OP		31	MC	0.65	0.33	0.53%				
OP		32	MC	0.92	0.28	0.56%				
OP		33	MC	0.95	0.29	0.71%				
OP		34	MC	0.61	0.47	0.50%				
OP		35	MC	0.92	0.38	0.71%				
OP		36	MC	0.68	0.36	0.19%				
OP		37	MC	0.56	0.34	0.37%				
OP		38	MC	0.38	0.25	0.30%		+		
OP		39	MC	0.91	0.30	0.30%				
OP		40	MC	0.89	0.27	0.32%				
OP		41	MC	0.36	0.16	0.56%		+		
OP		42	MC	0.91	0.20	0.30%				
OP		43	MC	0.79	0.29	0.43%				
OP		44	MC	0.75	0.18	0.40%				
OP		45	MC	0.84	0.42	0.90%				
OP		46A	CR	0.35	0.45	0.00%				
OP		46B	CR	0.36	0.43	0.00%				
OP		47	MC	0.84	0.38	0.39%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-10 Cont'd
Item Analysis Grade 5 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		48	MC	0.81	0.50	0.74%				
OP		49	MC	0.57	0.31	0.77%				
OP		50	MC	0.90	0.37	0.56%				
OP		51	MC	0.49	0.44	1.54%				
OP		52	MC	0.84	0.26	0.43%				
OP		53	MC	0.75	0.38	0.26%				
OP		54	MC	0.98	0.24	0.26%				
OP		55	MC	0.81	0.47	0.48%				
FT	A	56	MC	0.80	0.40	0.26%				
FT	A	57	MC	0.90	0.34	0.29%				
FT	A	58	MC	0.79	0.47	0.31%				
FT	A	59	MC	0.87	0.32	0.34%				
FT	A	60	MC	0.76	0.48	0.28%				
FT	A	61	MC	0.60	0.35	0.38%				
FT	A	62	MC	0.79	0.29	0.49%				
FT	A	63	CR	0.36	0.50	0.43%				
FT	A	64	MC	0.61	0.29	0.49%				
FT	A	65	MC	0.55	0.45	1.41%				
FT	A	66	MC	0.20	0.11	0.73%	+	+		+
FT	B	56	MC	0.58	0.22	0.32%				
FT	B	57	MC	0.67	0.42	0.50%				
FT	B	58	MC	0.52	0.39	0.46%				
FT	B	59	MC	0.56	0.50	0.41%				
FT	B	60	MC	0.44	0.33	0.52%				
FT	B	61	MC	0.86	0.37	0.50%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-10 Cont'd
Item Analysis Grade 5 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	62	MC	0.04	0.14	0.41%	+			+
FT	B	63	MC	0.84	0.47	0.64%				
FT	B	64A	CR	0.64	0.42	1.06%				
FT	B	64B	CR	0.28	0.43	0.26%				+
FT	B	65	MC	0.64	0.47	0.82%				
FT	B	66	MC	0.75	0.48	1.01%				
FT	C	56	MC	0.87	0.34	0.38%				
FT	C	57	MC	0.14	-0.05	0.31%	+	+		+
FT	C	58	MC	0.42	0.37	0.44%				
FT	C	59	MC	0.75	0.46	0.44%				
FT	C	60	MC	0.64	0.26	0.62%				
FT	C	61	MC	0.82	0.41	0.69%				
FT	C	62	MC	0.45	0.39	0.51%				
FT	C	63	MC	0.68	0.33	1.22%				
FT	C	64	CR	0.69	0.43	0.29%				
FT	C	65	MC	0.47	0.26	0.84%				
FT	C	66	MC	0.82	0.46	1.07%				
FT	D	56	MC	0.54	0.51	0.29%				
FT	D	57	MC	0.90	0.27	0.32%				
FT	D	58	MC	0.88	0.40	0.37%				
FT	D	59	MC	0.91	0.31	0.48%				
FT	D	60	MC	0.56	0.55	0.45%				
FT	D	61	MC	0.49	0.23	0.44%				
FT	D	62	MC	0.76	0.50	0.51%				
FT	D	63	MC	0.71	0.52	0.98%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-10 Cont'd
Item Analysis Grade 5 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	D	64A	CR	0.46	0.49	0.16%				
FT	D	64B	CR	0.66	0.50	0.28%				
FT	D	65	MC	0.65	0.53	0.86%				
FT	D	66	MC	0.81	0.42	1.30%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-11
Item Analysis Grade 6 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.79	0.41	0.30%				
OP		2	MC	0.85	0.43	0.08%				
OP		3	MC	0.64	0.40	0.61%				
OP		4	MC	0.47	0.25	0.27%		+		
OP		5	MC	0.88	0.35	0.54%				
OP		6	MC	0.59	0.45	0.27%				
OP		7A	CR	0.49	0.36	0.00%				
OP		7B	CR	0.37	0.44	0.00%				
OP		8	MC	0.68	0.39	0.81%				
OP		9	MC	0.87	0.42	0.59%				
OP		10	MC	0.67	0.38	1.44%				
OP		11	MC	0.86	0.36	1.18%				
OP		12	MC	0.53	0.34	1.53%				
OP		13	MC	0.53	0.38	2.04%				
OP		14	MC	0.85	0.47	1.50%				
OP		15	MC	0.82	0.47	1.60%				
OP		16	MC	0.52	0.45	0.73%				
OP		17	MC	0.74	0.37	0.24%				
OP		18	MC	0.69	0.40	0.35%				
OP		19	MC	0.91	0.34	0.24%				
OP		20	MC	0.88	0.35	0.57%				
OP		21	MC	0.48	0.38	0.85%				
OP		22A	CR	0.88	0.31	0.00%				
OP		22B	CR	0.30	0.50	0.00%				+
OP		23	MC	0.67	0.31	0.78%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-11 Cont'd
Item Analysis Grade 6 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		24	MC	0.61	0.35	0.51%				
OP		25	MC	0.39	0.35	0.53%				
OP		26	MC	0.81	0.37	0.62%				
OP		27	MC	0.69	0.42	0.54%				
OP		28	MC	0.60	0.52	0.57%				
OP		29	MC	0.80	0.33	0.94%				
OP		30	MC	0.46	0.38	0.35%				
OP		31	MC	0.79	0.40	0.30%				
OP		32	MC	0.58	0.46	0.37%				
OP		33	MC	0.95	0.18	0.30%				
OP		34	MC	0.79	0.28	0.38%				
OP		35	MC	0.97	0.17	0.24%				
OP		36	MC	0.70	0.45	0.46%				
OP		37	MC	0.77	0.14	0.89%	+			
OP		38	MC	0.90	0.30	0.48%				
OP		39A	CR	0.41	0.53	0.00%				
OP		39B	CR	0.36	0.51	0.00%				
OP		40	MC	0.54	0.46	0.96%				
OP		41	MC	0.66	0.41	0.49%				
OP		42	MC	0.98	0.13	0.34%	+			
OP		43	MC	0.74	0.29	0.29%				
OP		44	MC	0.68	0.54	0.37%				
OP		45	MC	0.78	0.43	0.40%				
OP		46	MC	0.72	0.52	0.56%				
OP		47	MC	0.80	0.45	0.38%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-11 Cont'd
Item Analysis Grade 6 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		48	MC	0.78	0.44	0.49%				
OP		49	MC	0.57	0.42	0.62%				
OP		50	MC	0.57	0.46	1.10%				
OP		51	CR	0.41	0.52	0.00%				
OP		52	MC	0.45	0.26	0.49%				
OP		53	MC	0.81	0.25	0.62%				
OP		54	MC	0.56	0.34	0.59%				
OP		55	MC	0.58	0.50	0.30%				
FT	A	56	MC	0.23	0.45	0.47%		+		+
FT	A	57	MC	0.72	0.38	0.67%				
FT	A	58	MC	0.95	0.15	0.67%				
FT	A	59	MC	0.35	0.29	0.47%				
FT	A	60	MC	0.36	0.28	0.43%				
FT	A	61	MC	0.53	0.34	0.39%				
FT	A	62	MC	0.80	0.21	0.52%				
FT	A	63A	CR	0.62	0.58	0.23%				
FT	A	63B	CR	0.71	0.61	0.35%				
FT	A	64	MC	0.31	0.26	1.49%				
FT	B	56	MC	0.39	0.51	0.31%				
FT	B	57	MC	0.94	0.24	0.42%				
FT	B	58	MC	0.80	0.43	0.52%				
FT	B	59	MC	0.69	0.41	0.58%				
FT	B	60	MC	0.19	-0.05	0.68%	+	+		+
FT	B	61	MC	0.52	0.41	0.42%				
FT	B	62	MC	0.51	0.35	0.66%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-11 Cont'd
Item Analysis Grade 6 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	63A	CR	0.40	0.48	0.58%				
FT	B	63B	CR	0.38	0.51	0.82%				
FT	B	64	MC	0.52	0.10	1.14%	+	+		
FT	C	56	MC	0.84	0.34	0.39%				
FT	C	57	MC	0.42	0.57	0.41%				
FT	C	58	MC	0.86	0.22	0.47%				
FT	C	59	MC	0.60	0.48	0.47%				
FT	C	60	MC	0.49	0.32	0.53%				
FT	C	61	MC	0.70	0.38	0.44%				
FT	C	62A*	CR							
FT	C	62B*	CR							
FT	C	63	MC	0.89	0.37	0.81%				
FT	C	64	MC	0.36	0.24	0.78%				
FT	D	56	MC	0.27	0.49	0.31%				+
FT	D	57	MC	0.88	0.41	0.50%				
FT	D	58	MC	0.87	0.43	0.45%				
FT	D	59	MC	0.53	0.43	0.37%				
FT	D	60	MC	0.36	0.08	0.50%	+	+		
FT	D	61	MC	0.71	0.40	0.60%				
FT	D	62*	CR							
FT	D	63	MC	0.77	0.24	0.88%				
FT	D	64	MC	0.57	0.46	0.84%				

*This item was dropped. See Part 7 for more information.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-12
Item Analysis Grade 7 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.71	0.47	0.25%				
OP		2	MC	0.90	0.39	0.05%				
OP		3	MC	0.62	0.51	0.16%				
OP		4	MC	0.71	0.18	0.12%		+		
OP		5	MC	0.78	0.39	0.12%				
OP		6	MC	0.75	0.39	0.11%				
OP		7A	CR	0.68	0.52	0.00%				
OP		7B	CR	0.60	0.57	0.00%				
OP		8	MC	0.73	0.31	0.48%				
OP		9	MC	0.57	0.47	0.31%				
OP		10	MC	0.64	0.55	0.64%				
OP		11	MC	0.58	0.59	0.83%				
OP		12	MC	0.55	0.18	0.53%		+		
OP		13	MC	0.55	0.59	0.61%				
OP		14	MC	0.89	0.31	0.48%				
OP		15	MC	0.80	0.40	0.64%				
OP		16	MC	0.61	0.23	0.33%				
OP		17	MC	0.59	0.29	0.56%				
OP		18	MC	0.85	0.50	0.25%				
OP		19A	CR	0.54	0.40	0.00%				
OP		19B	CR	0.29	0.52	0.00%				+
OP		20	MC	0.97	0.20	0.47%				
OP		21	MC	0.53	0.28	0.41%				
OP		22	MC	0.70	0.34	0.31%				
OP		23	MC	0.55	0.48	0.30%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-12 Cont'd
Item Analysis Grade 7 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		24	MC	0.86	0.48	0.16%				
OP		25	MC	0.41	0.44	0.28%				
OP		26	MC	0.92	0.18	0.42%				
OP		27	MC	0.68	0.44	0.41%				
OP		28	MC	0.82	0.51	0.42%				
OP		29	MC	0.83	0.28	0.42%				
OP		30	MC	0.83	0.34	0.25%				
OP		31	MC	0.48	0.39	0.67%				
OP		32	MC	0.64	0.49	0.55%				
OP		33	MC	0.54	0.34	0.58%				
OP		34	MC	0.40	0.32	0.50%				
OP		35	MC	0.68	0.55	0.48%				
OP		36	MC	0.65	0.35	0.31%				
OP		37	MC	0.95	0.24	0.28%				
OP		38A	CR	0.56	0.38	0.00%				
OP		38B	CR	0.34	0.38	0.00%				
OP		39	MC	0.65	0.46	0.33%				
OP		40	MC	0.89	0.33	0.42%				
OP		41	MC	0.75	0.43	0.33%				
OP		42	MC	0.90	0.34	0.48%				
OP		43	MC	0.61	0.47	0.27%				
OP		44	MC	0.98	0.22	0.25%				
OP		45	MC	0.38	0.44	0.30%				
OP		46	MC	0.84	0.43	0.34%				
OP		47	MC	0.83	0.29	0.62%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-12 Cont'd
Item Analysis Grade 7 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		48	MC	0.86	0.34	0.17%				
OP		49	MC	0.48	0.51	0.34%				
OP		50	MC	0.67	0.60	0.83%				
OP		51	MC	0.51	0.36	0.75%				
OP		52	CR	0.36	0.60	0.00%				
OP		53	MC	0.83	0.30	0.30%				
OP		54	MC	0.55	0.42	0.41%		+		
OP		55	MC	0.41	0.10	0.61%	+	+		
FT	A	56	MC	0.74	0.33	0.51%				
FT	A	57	MC	0.23	0.38	0.47%		+		+
FT	A	58	MC	0.81	0.27	0.52%				
FT	A	59	MC	0.55	0.48	0.56%				
FT	A	60	MC	0.45	0.20	0.54%				
FT	A	61	CR	0.36	0.43	0.40%				
FT	A	62	MC	0.47	0.21	1.18%				
FT	A	63	MC	0.11	0.06	1.18%	+	+		+
FT	A	64	MC	0.54	0.37	1.37%				
FT	B	56	MC	0.37	0.55	0.41%				
FT	B	57	MC	0.56	0.29	0.52%				
FT	B	58	MC	0.49	0.61	0.41%				
FT	B	59	MC	0.89	0.34	0.51%				
FT	B	60	MC	0.46	0.39	0.49%				
FT	B	61A	CR	0.37	0.50	1.27%				
FT	B	61B	CR	0.30	0.43	1.02%				
FT	B	62	MC	0.28	0.21	0.64%				+

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-12 Cont'd
Item Analysis Grade 7 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	63	MC	0.21	0.15	0.90%	+			+
FT	B	64	MC	0.41	0.02	1.04%	+	+		
FT	C	56	MC	0.70	0.36	0.44%				
FT	C	57	MC	0.66	0.26	0.41%				
FT	C	58	MC	0.57	0.22	0.43%				
FT	C	59	MC	0.59	0.48	0.50%				
FT	C	60	MC	0.42	0.26	0.56%				
FT	C	61	MC	0.33	0.23	0.49%		+		
FT	C	62	CR	0.22	0.39	0.93%				+
FT	C	63	MC	0.67	0.22	1.20%				
FT	C	64	MC	0.53	0.30	1.07%				
FT	D	56	MC	0.52	0.42	0.53%				
FT	D	57	MC	0.71	0.38	0.41%				
FT	D	58	MC	0.56	0.32	0.47%				
FT	D	59	MC	0.71	0.39	0.38%				
FT	D	60	MC	0.34	0.27	0.87%		+		
FT	D	61	CR	0.42	0.45	0.38%				
FT	D	62	MC	0.63	0.12	0.65%	+			
FT	D	63	MC	0.41	0.04	0.79%	+	+		
FT	D	64	MC	0.71	0.25	0.75%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-13
Item Analysis Grade 8 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.40	0.19	0.18%				
OP		2	MC	0.53	0.48	0.20%				
OP		3	MC	0.69	0.49	0.21%				
OP		4	MC	0.56	0.53	0.48%				
OP		5	MC	0.46	0.31	0.41%		+		
OP		6	MC	0.87	0.42	0.21%				
OP		7	MC	0.63	0.52	0.26%				
OP		8	MC	0.52	0.35	0.51%				
OP		9	MC	0.53	0.43	0.23%				
OP		10	CR	0.21	0.51	0.00%				+
OP		11	MC	0.85	0.40	0.26%				
OP		12	MC	0.40	0.37	0.80%				
OP		13	MC	0.67	0.49	0.69%				
OP		14	MC	0.88	0.35	0.51%				
OP		15	MC	0.58	0.29	0.61%				
OP		16	MC	0.71	0.43	0.35%				
OP		17	MC	0.41	0.34	0.57%				
OP		18	MC	0.70	0.26	0.32%				
OP		19A	CR	0.69	0.41	0.00%				
OP		19B	CR	0.57	0.50	0.00%				
OP		20	MC	0.73	0.38	0.43%				
OP		21	MC	0.34	0.37	0.35%				
OP		22	MC	0.38	0.39	0.49%				
OP		23	MC	0.81	0.40	0.40%				
OP		24	MC	0.94	0.26	0.26%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-13 Cont'd
Item Analysis Grade 8 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		25	MC	0.52	0.43	0.54%				
OP		26	MC	0.65	0.37	0.54%				
OP		27	MC	0.58	0.40	1.12%				
OP		28	MC	0.72	0.40	0.81%				
OP		29	MC	0.41	0.38	0.40%				
OP		30	MC	0.47	0.34	0.35%				
OP		31	MC	0.54	0.34	0.43%				
OP		32	MC	0.49	0.36	0.61%				
OP		33A	CR	0.41	0.51	0.00%				
OP		33B	CR	0.38	0.59	0.00%				
OP		34	MC	0.52	0.39	0.58%				
OP		35	MC	0.66	0.31	0.80%				
OP		36	MC	0.81	0.45	0.77%				
OP		37	MC	0.41	0.48	0.35%				
OP		38	MC	0.64	0.34	0.64%				
OP		39	MC	0.72	0.33	0.88%				
OP		40	MC	0.64	0.16	0.80%		+		
OP		41	MC	0.87	0.36	0.51%				
OP		42	MC	0.66	0.35	0.61%				
OP		43	MC	0.73	0.34	0.81%				
OP		44	MC	0.78	0.38	0.43%				
OP		45	MC	0.66	0.52	0.48%				
OP		46	MC	0.33	0.25	0.69%				
OP		47	MC	0.81	0.29	0.43%				
OP		48	MC	0.43	0.46	0.48%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-13 Cont'd
Item Analysis Grade 8 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		49A	CR	0.25	0.44	0.00%				+
OP		49B	CR	0.22	0.44	0.00%				+
OP		50	MC	0.73	0.37	0.54%				
OP		51	MC	0.24	0.29	0.46%				+
OP		52	MC	0.59	0.55	0.58%				
OP		53	MC	0.50	0.47	0.74%				
OP		54	MC	0.84	0.47	0.46%				
OP		55	MC	0.66	0.25	0.63%				
FT	A	56	MC	0.51	0.41	0.50%				
FT	A	57	MC	0.58	0.56	0.47%				
FT	A	58	MC	0.51	0.50	0.52%				
FT	A	59	MC	0.40	0.29	0.51%				
FT	A	60	MC	0.47	0.37	0.61%				
FT	A	61	MC	0.09	0.25	0.70%		+		+
FT	A	62	CR	0.06	0.40	0.40%				+
FT	A	63	MC	0.40	0.61	0.68%				
FT	A	64	MC	0.37	0.29	0.88%				
FT	B	56	MC	0.51	0.27	0.46%				
FT	B	57	MC	0.89	0.18	0.49%				
FT	B	58	MC	0.19	0.09	0.61%	+	+		+
FT	B	59	MC	0.73	0.38	0.54%				
FT	B	60	MC	0.31	0.12	0.56%	+	+		
FT	B	61	MC	0.38	0.34	0.53%				
FT	B	62	CR	0.33	0.46	0.57%				
FT	B	63	MC	0.15	0.21	1.70%		+		+

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-13 Cont'd
Item Analysis Grade 8 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
FT	B	64	MC	0.76	0.40	1.81%				
FT	C	56	MC	0.25	0.47	0.47%				+
FT	C	57	MC	0.50	0.41	0.54%				
FT	C	58	MC	0.54	0.31	0.52%				
FT	C	59	MC	0.88	0.29	0.53%				
FT	C	60	MC	0.72	0.41	0.61%				
FT	C	61	MC	0.17	0.30	0.93%				+
FT	C	62A	CR	0.65	0.50	0.63%				
FT	C	62B	CR	0.39	0.57	1.12%				
FT	C	63	MC	0.45	0.40	1.49%				
FT	C	64	MC	0.14	0.12	1.37%	+	+		+
FT	D	56	MC	0.44	0.25	0.50%				
FT	D	57	MC	0.79	0.40	0.43%				
FT	D	58	MC	0.44	0.26	0.55%				
FT	D	59	MC	0.35	0.31	0.72%		+		
FT	D	60	MC	0.79	0.43	0.50%				
FT	D	61A	CR	0.61	0.52	0.16%				
FT	D	61B	CR	0.42	0.52	0.43%				
FT	D	62	MC	0.82	0.38	1.18%				
FT	D	63	MC	0.24	0.13	1.32%	+	+		+
FT	D	64	MC	0.83	0.38	1.37%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-14
Item Analysis Grade 10 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.83	0.43	0.18%				
OP		2	MC	0.67	0.40	0.19%				
OP		3	MC	0.46	0.30	0.52%				
OP		4	MC	0.57	0.50	0.41%				
OP		5	MC	0.48	0.40	0.11%				
OP		6	MC	0.57	0.44	0.46%				
OP		7	MC	0.40	0.42	0.46%				
OP		8	MC	0.65	0.48	0.48%				
OP		9	MC	0.42	0.36	0.73%		+		
OP		10	MC	0.45	0.34	1.23%				
OP		11	MC	0.48	0.47	1.23%				
OP		12	MC	0.71	0.44	0.47%				
OP		13	MC	0.80	0.35	0.36%				
OP		14	MC	0.50	0.41	0.46%				
OP		15	MC	0.61	0.41	0.44%				
OP		16	MC	0.79	0.51	0.52%				
OP		17	MC	0.67	0.63	0.51%				
OP		18	MC	0.71	0.48	0.30%				
OP		19	MC	0.49	0.42	0.80%				
OP		20	MC	0.55	0.50	0.55%				
OP		21	MC	0.72	0.46	0.62%				
OP		22	MC	0.44	0.55	0.66%				
OP		23	CR	0.42	0.63	0.00%				
OP		24	MC	0.70	0.40	0.80%				
OP		25	MC	0.71	0.44	1.04%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-14 Cont'd
Item Analysis Grade 10 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.57	0.59	1.20%				
OP		27	MC	0.50	0.42	1.62%				
OP		28	MC	0.49	0.49	0.46%				
OP		29	MC	0.76	0.48	0.37%				
OP		30	MC	0.58	0.41	0.65%				
OP		31	MC	0.65	0.24	0.54%				
OP		32	MC	0.47	0.42	0.54%				
OP		33	CR	0.26	0.59	0.00%				+
OP		34	MC	0.72	0.41	0.62%				
OP		35	MC	0.93	0.28	0.55%				
OP		36	CR	0.42	0.64	0.00%				
OP		37	MC	0.60	0.32	0.68%				
OP		38	MC	0.56	0.40	0.79%				
OP		39	MC	0.66	0.50	0.66%				
OP		40	MC	0.44	0.47	0.59%				
OP		41	MC	0.54	0.65	0.61%				
OP		42	MC	0.60	0.49	0.52%				
OP		43	MC	0.49	0.56	0.80%				
OP		44	MC	0.75	0.47	0.76%				
OP		45	MC	0.52	0.45	0.81%		+		
OP		46	CR	0.39	0.54	0.00%				
OP		47	MC	0.55	0.48	0.65%				
OP		48	MC	0.84	0.41	0.51%				
OP		49	MC	0.69	0.41	0.68%				
OP		50	MC	0.76	0.47	0.90%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-14 Cont'd
Item Analysis Grade 10 Mathematics

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		51	MC	0.36	0.43	0.91%				
OP		52	MC	0.44	0.52	0.91%				
OP		53	MC	0.53	0.39	1.08%				
OP		54	MC	0.67	0.56	1.05%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-15
Item Analysis Grade 4 Language Arts

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.40	0.33	0.34%				
OP		2	MC	0.92	0.29	0.32%				
OP		3	MC	0.82	0.29	0.66%				
OP		4	MC	0.94	0.25	0.47%				
OP		5	MC	0.40	0.16	0.61%				
OP		6	MC	0.83	0.36	0.13%				
OP		7	MC	0.59	0.37	0.21%				
OP		8	MC	0.84	0.28	0.14%				
OP		9	MC	0.83	0.40	0.21%				
OP		10	MC	0.46	0.33	0.83%				
OP		11	MC	0.90	0.40	0.47%				
OP		12	MC	0.94	0.31	0.58%				
OP		13	MC	0.80	0.41	0.35%				
OP		14	MC	0.69	0.41	0.55%				
OP		15	MC	0.52	0.21	0.40%				
OP		16	MC	0.65	0.47	2.92%				
OP		17	MC	0.68	0.46	0.61%				
OP		18	MC	0.70	0.45	2.55%				
OP		19	MC	0.45	0.31	1.30%				
OP		20	MC	0.69	0.41	0.67%				
OP		21	MC	0.73	0.44	0.67%				
OP		22	MC	0.66	0.40	1.86%				
OP		23	MC	0.52	0.32	1.59%				
OP		24	MC	0.58	0.43	1.01%				
OP		25	MC	0.45	0.23	1.24%		+		

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-15 Cont'd
Item Analysis Grade 4 Language Arts

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.54	0.38	1.65%				
OP		27	MC	0.63	0.40	1.70%				
OP		28	MC	0.33	0.28	3.03%				
OP		29	MC	0.76	0.44	3.27%				
OP		30	MC	0.45	0.13	3.59%	+	+		
*OP WR	A	1A	CR	0.51	0.40	0.26%				
*OP WR	A	1B	CR	0.65	0.29	0.26%				

* Operational Writing prompt items are included here. Writing raw score is based on these items, but they do not contribute to the operational (scale) score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-16
Item Analysis Grade 8 Language Arts

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.65	0.17	0.16%				
OP		2	MC	0.92	0.13	0.19%	+			
OP		3	MC	0.68	0.32	0.99%				
OP		4	MC	0.85	0.38	0.03%				
OP		5	MC	0.74	0.41	0.09%				
OP		6	MC	0.53	0.31	0.37%				
OP		7	MC	0.88	0.39	0.47%				
OP		8	MC	0.66	0.48	0.23%				
OP		9	MC	0.81	0.42	0.36%				
OP		10	MC	0.64	0.34	0.19%				
OP		11	MC	0.83	0.37	0.09%				
OP		12	MC	0.69	0.43	0.81%				
OP		13	MC	0.49	0.36	0.14%				
OP		14	MC	0.81	0.49	0.09%				
OP		15	MC	0.87	0.41	0.17%				
OP		16	MC	0.87	0.37	0.28%				
OP		17	MC	0.66	0.46	0.36%				
OP		18	MC	0.84	0.49	0.84%				
OP		19	MC	0.81	0.42	0.29%				
OP		20	MC	0.55	0.56	4.32%				
OP		21	MC	0.90	0.46	0.50%				
OP		22	MC	0.61	0.44	1.05%				
OP		23	MC	0.68	0.45	1.01%				
OP		24	MC	0.78	0.41	0.53%				
OP		25	MC	0.70	0.41	2.57%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-16 Cont'd
Item Analysis Grade 8 Language Arts

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.46	0.36	0.76%				
OP		27	MC	0.73	0.33	1.55%				
OP		28	MC	0.36	-0.03	1.38%	+	+		
OP		29	MC	0.66	0.40	1.64%				
OP		30	MC	0.64	0.43	2.02%				
*OP WR	A	1A	CR	0.55	0.47	0.56%				
*OP WR	A	1B	CR	0.66	0.20	0.56%				

* Operational Writing prompt items are included here. Writing raw score is based on these items, but they do not contribute to the operational (scale) score.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-17
Item Analysis Grade 10 Language Arts

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.73	0.40	0.07%				
OP		2	MC	0.83	0.44	0.10%				
OP		3	MC	0.67	0.32	0.13%				
OP		4	MC	0.67	0.28	1.25%				
OP		5	MC	0.77	0.34	0.28%				
OP		6	MC	0.66	0.36	0.14%				
OP		7	MC	0.43	0.21	0.29%				
OP		8	MC	0.81	0.38	0.24%				
OP		9	MC	0.59	0.33	0.36%				
OP		10	MC	0.69	0.42	0.27%				
OP		11	MC	0.75	0.44	1.16%				
OP		12	MC	0.64	0.27	0.35%				
OP		13	MC	0.49	0.18	0.31%		+		
OP		14	MC	0.79	0.50	0.17%				
OP		15	MC	0.60	0.40	0.22%				
OP		16	MC	0.69	0.33	0.29%				
OP		17	MC	0.87	0.41	0.98%				
OP		18	MC	0.69	0.27	0.64%				
OP		19	MC	0.66	0.48	0.32%				
OP		20	MC	0.52	0.34	0.50%				
OP		21	MC	0.42	0.21	0.57%		+		
OP		22	MC	0.65	0.45	0.73%				
OP		23	MC	0.44	0.31	0.57%				
OP		24	MC	0.33	0.23	0.70%				
OP		25	MC	0.75	0.55	0.95%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-17 Cont'd
Item Analysis Grade 10 Language Arts

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.69	0.42	0.66%				
OP		27	MC	0.24	0.15	0.88%		+		+
OP		28	MC	0.85	0.42	0.87%				
OP		29	MC	0.66	0.40	1.11%				
OP		30	MC	0.71	0.50	1.19%				
*OP WR		1A	CR	0.49	0.58	0.00%				
*OP WR		1B	CR	0.66	0.42	0.00%				

*The Writing Prompt in Grade 10 Language Arts is part of the scale score for Language Arts Grade 10.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-18
Item Analysis Grade 4 Social Studies

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.50	0.42	0.08%				
OP		2	MC	0.90	0.34	0.10%				
OP		3	MC	0.94	0.30	0.32%				
OP		4	MC	0.50	0.43	0.11%				
OP		5	MC	0.82	0.21	0.16%				
OP		6	MC	0.58	0.21	0.29%		+		
OP		7	MC	0.78	0.41	0.62%				
OP		8	MC	0.89	0.35	0.13%				
OP		9	MC	0.81	0.39	0.70%				
OP		10	MC	0.76	0.29	0.27%				
OP		11	MC	0.74	0.41	0.46%				
OP		12	MC	0.59	0.47	0.51%				
OP		13	MC	0.93	0.30	0.88%				
OP		14	MC	0.94	0.27	2.06%				
OP		15	MC	0.83	0.39	0.94%				
OP		16	MC	0.51	0.30	2.30%				
OP		17	MC	0.78	0.32	1.01%				
OP		18	MC	0.90	0.34	0.11%				
OP		19	MC	0.96	0.31	0.83%				
OP		20	MC	0.64	0.46	1.05%				
OP		21	MC	0.55	0.44	0.48%				
OP		22	MC	0.72	0.48	0.35%				
OP		23	MC	0.86	0.42	0.42%				
OP		24	MC	0.93	0.29	0.42%				
OP		25	MC	0.90	0.41	0.85%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-18 Cont'd
Item Analysis Grade 4 Social Studies

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.70	0.48	2.70%				
OP		27	MC	0.94	0.39	0.38%				
OP		28	MC	0.79	0.42	0.66%				
OP		29	MC	0.65	0.41	0.54%				
OP		30	MC	0.66	0.42	0.88%				
OP		31	MC	0.64	0.43	0.93%				
OP		32	MC	0.76	0.37	0.69%				
OP		33	MC	0.59	0.35	0.99%				
OP		34	MC	0.69	0.42	2.22%				
OP		35	MC	0.89	0.42	0.59%				
OP		36	MC	0.63	0.47	1.25%				
OP		37	MC	0.90	0.43	0.83%				
OP		38	MC	0.73	0.46	0.99%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-19
Item Analysis Grade 8 Social Studies

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.95	0.30	0.02%				
OP		2	MC	0.92	0.32	0.34%				
OP		3	MC	0.65	0.38	0.26%				
OP		4	MC	0.38	0.18	0.26%		+		
OP		5	MC	0.79	0.32	0.28%				
OP		6	MC	0.96	0.34	0.37%				
OP		7	MC	0.79	0.28	0.49%				
OP		8	MC	0.91	0.36	0.09%				
OP		9	MC	0.57	0.42	0.22%				
OP		10	MC	0.81	0.47	0.83%				
OP		11	MC	0.69	0.41	1.48%				
OP		12	MC	0.75	0.40	0.05%				
OP		13	MC	0.59	0.34	0.31%				
OP		14	MC	0.73	0.46	0.46%				
OP		15	MC	0.84	0.39	0.60%				
OP		16	MC	0.76	0.55	0.62%				
OP		17	MC	0.66	0.49	1.65%				
OP		18	MC	0.78	0.35	0.62%				
OP		19	MC	0.50	0.38	0.88%				
OP		20	MC	0.47	0.32	0.22%				
OP		21	MC	0.57	0.41	1.11%				
OP		22	MC	0.94	0.36	0.17%				
OP		23	MC	0.71	0.38	0.19%				
OP		24	MC	0.59	0.40	0.28%				
OP		25	MC	0.83	0.39	0.54%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-19 Cont'd
Item Analysis Grade 8 Social Studies

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.60	0.39	4.10%				
OP		27	MC	0.63	0.42	0.29%				
OP		28	MC	0.66	0.53	0.35%				
OP		29	MC	0.58	0.40	1.02%				
OP		30	MC	0.46	0.27	0.43%				
OP		31	MC	0.64	0.33	1.27%				
OP		32	MC	0.80	0.48	0.37%				
OP		33	MC	0.75	0.46	0.86%				
OP		34	MC	0.81	0.44	0.91%				
OP		35	MC	0.67	0.41	0.60%				
OP		36	MC	0.47	0.31	0.43%				
OP		37	MC	0.57	0.24	0.80%				
OP		38	MC	0.58	0.35	0.68%				
OP		39	MC	0.46	0.39	0.77%				
OP		40	MC	0.64	0.39	0.60%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-20
Item Analysis Grade 10 Social Studies

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.79	0.47	0.11%				
OP		2	MC	0.95	0.32	0.91%				
OP		3	MC	0.63	0.34	0.06%				
OP		4	MC	0.52	0.35	0.18%				
OP		5	MC	0.67	0.48	0.34%				
OP		6	MC	0.42	0.41	0.29%				
OP		7	MC	0.57	0.38	0.35%				
OP		8	MC	0.60	0.27	0.49%				
OP		9	MC	0.58	0.40	0.63%				
OP		10	MC	0.80	0.33	0.62%				
OP		11	MC	0.59	0.42	0.60%				
OP		12	MC	0.47	0.26	0.31%				
OP		13	MC	0.52	0.24	0.29%				
OP		14	MC	0.76	0.38	0.32%				
OP		15	MC	0.36	0.32	0.17%				
OP		16	MC	0.66	0.46	0.79%				
OP		17	MC	0.55	0.40	0.24%				
OP		18	MC	0.45	0.24	0.39%				
OP		19	MC	0.88	0.40	0.22%				
OP		20	MC	0.88	0.38	0.41%				
OP		21	MC	0.67	0.36	0.93%				
OP		22	MC	0.77	0.47	0.31%				
OP		23	MC	0.50	0.45	0.84%				
OP		24	MC	0.68	0.51	0.62%				
OP		25	MC	0.96	0.24	0.31%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-20 Cont'd
Item Analysis Grade 10 Social Studies

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.68	0.42	0.49%				
OP		27	MC	0.76	0.48	1.42%				
OP		28	MC	0.65	0.40	0.51%				
OP		29	MC	0.46	0.20	0.48%				
OP		30	MC	0.64	0.43	0.46%				
OP		31	MC	0.77	0.44	0.39%				
OP		32	MC	0.60	0.44	1.56%				
OP		33	MC	0.63	0.42	0.76%				
OP		34	MC	0.75	0.42	0.53%				
OP		35	MC	0.62	0.39	1.05%				
OP		36	MC	0.78	0.46	0.94%				
OP		37	MC	0.63	0.42	0.74%				
OP		38	MC	0.74	0.51	0.76%				
OP		39	MC	0.81	0.41	0.79%				
OP		40	MC	0.17	0.09	0.51%	+	+		+
OP		41	MC	0.28	0.20	1.03%		+		+
OP		42	MC	0.63	0.37	0.53%				
OP		43	MC	0.40	0.42	1.22%		+		
OP		44	MC	0.67	0.49	0.72%				
OP		45	MC	0.55	0.52	0.49%				
OP		46	MC	0.78	0.50	0.51%				
OP		47	MC	0.84	0.45	0.76%				
OP		48	MC	0.66	0.48	1.07%				
OP		49	MC	0.48	0.42	0.62%				
OP		50	MC	0.86	0.48	1.31%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-21
Item Analysis Grade 4 Science

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.99	0.18	0.08%				
OP		2	MC	0.83	0.29	0.10%				
OP		3	MC	0.88	0.44	0.50%				
OP		4	MC	0.81	0.34	0.43%				
OP		5	MC	0.91	0.18	0.05%				
OP		6	MC	0.79	0.28	0.80%				
OP		7	MC	0.93	0.33	0.22%				
OP		8	MC	0.70	0.42	0.32%				
OP		9	MC	0.79	0.35	0.40%				
OP		10	MC	0.93	0.23	0.64%				
OP		11	MC	0.75	0.37	1.36%				
OP		12	MC	0.76	0.32	0.89%				
OP		13	MC	0.69	0.32	0.18%				
OP		14	MC	0.67	0.12	0.88%	+			
OP		15	MC	0.85	0.49	3.32%				
OP		16	MC	0.82	0.38	3.77%				
OP		17	MC	0.80	0.45	0.34%				
OP		18	MC	0.53	0.38	1.34%				
OP		19	MC	0.48	0.45	0.58%				
OP		20	MC	0.62	0.38	0.58%				
OP		21	MC	0.66	0.40	0.86%				
OP		22	MC	0.81	0.51	1.02%				
OP		23	MC	0.86	0.35	0.46%				
OP		24	MC	0.41	0.13	0.89%	+	+		
OP		25	MC	0.85	0.32	0.30%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-21 Cont'd
Item Analysis Grade 4 Science

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.55	0.36	0.18%				
OP		27	MC	0.69	0.37	0.38%				
OP		28	MC	0.90	0.40	0.62%				
OP		29	MC	0.67	0.28	1.01%				
OP		30	MC	0.79	0.46	0.62%				
OP		31	MC	0.46	0.37	2.28%				
OP		32	MC	0.33	0.10	0.48%	+	+		
OP		33	MC	0.67	0.40	1.12%				
OP		34	MC	0.95	0.27	1.47%				
OP		35	MC	0.77	0.50	1.66%				
OP		36	MC	0.67	0.36	2.62%				
OP		37	MC	0.45	0.39	0.83%				
OP		38	MC	0.47	0.29	1.12%				
OP		39	MC	0.48	0.45	1.25%				
OP		40	MC	0.48	0.36	1.71%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-22
Item Analysis Grade 8 Science

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.87	0.32	0.11%				
OP		2	MC	0.97	0.27	0.06%				
OP		3	MC	0.96	0.26	0.22%				
OP		4	MC	0.86	0.28	0.97%				
OP		5	MC	0.92	0.36	0.26%				
OP		6	MC	0.58	0.30	0.29%				
OP		7	MC	0.92	0.33	0.26%				
OP		8	MC	0.77	0.35	0.43%				
OP		9	MC	0.53	0.36	0.79%				
OP		10	MC	0.86	0.38	0.38%				
OP		11	MC	0.94	0.30	1.05%				
OP		12	MC	0.95	0.30	1.09%				
OP		13	MC	0.76	0.36	0.26%				
OP		14	MC	0.66	0.39	0.31%				
OP		15	MC	0.70	0.38	0.17%				
OP		16	MC	0.76	0.45	0.42%				
OP		17	MC	0.73	0.33	0.26%				
OP		18	MC	0.89	0.29	0.18%				
OP		19	MC	0.69	0.45	0.43%				
OP		20	MC	0.73	0.41	0.48%				
OP		21	MC	0.70	0.35	0.46%				
OP		22	MC	0.72	0.40	0.32%				
OP		23	MC	0.86	0.38	0.57%				
OP		24	MC	0.63	0.49	0.80%				
OP		25	MC	0.80	0.48	0.45%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-22 Cont'd
Item Analysis Grade 8 Science

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.58	0.35	0.35%				
OP		27	MC	0.58	0.45	5.16%			+	
OP		28	MC	0.72	0.48	0.42%				
OP		29	MC	0.78	0.40	0.37%				
OP		30	MC	0.74	0.31	1.06%				
OP		31	MC	0.42	0.30	1.82%				
OP		32	MC	0.59	0.42	0.42%				
OP		33	MC	0.79	0.35	0.66%				
OP		34	MC	0.67	0.44	0.42%				
OP		35	MC	0.53	0.30	0.40%				
OP		36	MC	0.40	0.24	6.65%		+	+	
OP		37	MC	0.62	0.47	0.72%				
OP		38	MC	0.63	0.35	1.25%				
OP		39	MC	0.56	0.16	0.68%		+		
OP		40	MC	0.57	0.27	0.54%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-23
Item Analysis Grade 10 Science

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		1	MC	0.73	0.38	0.13%				
OP		2	MC	0.87	0.36	0.08%				
OP		3	MC	0.44	0.10	0.08%	+	+		
OP		4	MC	0.40	0.16	2.94%				
OP		5	MC	0.66	0.37	0.17%				
OP		6	MC	0.84	0.41	0.08%				
OP		7	MC	0.68	0.43	0.31%				
OP		8	MC	0.82	0.41	0.25%				
OP		9	MC	0.88	0.37	0.28%				
OP		10	MC	0.68	0.33	0.42%				
OP		11	MC	0.69	0.44	0.52%				
OP		12	MC	0.38	0.26	0.34%				
OP		13	MC	0.86	0.42	0.34%				
OP		14	MC	0.71	0.29	0.21%				
OP		15	MC	0.48	0.33	0.55%				
OP		16	MC	0.52	0.29	0.21%				
OP		17	MC	0.63	0.36	0.28%				
OP		18	MC	0.57	0.28	0.34%				
OP		19	MC	0.75	0.40	0.27%				
OP		20	MC	0.68	0.41	0.28%				
OP		21	MC	0.86	0.28	0.24%				
OP		22	MC	0.44	0.42	0.55%				
OP		23	MC	0.46	0.43	0.45%				
OP		24	MC	0.41	0.27	0.81%				
OP		25	MC	0.54	0.43	0.32%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-23 Cont'd
Item Analysis Grade 10 Science

Item ID Field			Item Statistic Fields				Flag			
OP/FT	**Form**	**Test Book Item**	**Item Type**	***p*-value**	**Corr**	**Omit Rate**	**Flag Corr**	**Flag Distracter**	**Flag Omit**	**Flag *p*-value**
OP		26	MC	0.41	0.32	0.35%				
OP		27	MC	0.70	0.45	0.31%				
OP		28	MC	0.67	0.51	0.38%				
OP		29	MC	0.32	0.23	0.35%				
OP		30	MC	0.66	0.46	0.48%				
OP		31	MC	0.40	0.19	0.69%				
OP		32	MC	0.69	0.18	0.32%				
OP		33	MC	0.33	0.31	0.45%				
OP		34	MC	0.62	0.35	0.39%				
OP		35	MC	0.69	0.48	0.63%				
OP		36	MC	0.70	0.40	0.76%				
OP		37	MC	0.35	0.31	0.96%		+		
OP		38	MC	0.66	0.35	0.46%				
OP		39	MC	0.67	0.50	1.24%				
OP		40	MC	0.67	0.36	0.53%				
OP		41	MC	0.45	0.40	0.70%				
OP		42	MC	0.57	0.40	0.45%				
OP		43	MC	0.45	0.50	0.58%				
OP		44	MC	0.71	0.41	0.59%				
OP		45	MC	0.38	0.30	1.91%		+		
OP		46	MC	0.54	0.45	0.56%				
OP		47	MC	0.59	0.38	0.65%				
OP		48	MC	0.46	0.32	0.58%				
OP		49	MC	0.88	0.27	0.91%				
OP		50	MC	0.69	0.42	0.96%				

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-24
The Number of Items Flagged

Content	Grade	OP Items				FT Items			
		Flag Corr	Flag Distracter	Flag Omit	Flag *p*-value	Flag Corr	Flag Distracter	Flag Omit	Flag *p*-value
RD	3		1	1	1	2	3		2
	4		2			1	6		1
	5	1	4			1	3		
	6		2			2	4		1
	7	1	3			3	5		4
	8	2	5			1	5		
	10		2						
MA	3				4	1	3	1	3
	4	1		1	2	2	4		3
	5		2		1	3	2		4
	6	2	1		1	3	4		3
	7	1	4		1	5	6		5
	8		2		4	4	7		8
	10		2		1				
LA	4	1	2						
	8	2	1						
	10		3		1				
SS	4		1						
	8		1						
	10	1	3		2				
SC	4	3	2						
	8		2	2					
	10	1	3						

*Note that the number of FT items flagged reflects the total number flagged across all forms. FT items flagged for a grade and content area are not necessarily from a single form; the table entries reflect the total from across all forms.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-25
Raw Score Descriptive Statistics based on Census Data

Content	Grade	N Count	Mean # of Items Correct	Test Difficulty	SD	Skewness	Kurtosis	Min Obtained	Max Obtained	Max Possible	Alpha	SEM
Reading	3	58543	39.48	0.66	11.84	-0.62	-0.54	0	59	60	0.93	3.07
	4	58726	39.19	0.65	11.14	-0.63	-0.39	0	60	60	0.92	3.17
	5	58399	39.40	0.66	10.76	-0.65	-0.31	0	60	60	0.91	3.16
	6	59683	38.50	0.64	9.96	-0.60	-0.23	0	60	60	0.90	3.19
	7	60777	38.54	0.64	10.14	-0.61	-0.20	0	60	60	0.90	3.22
	8	62782	41.05	0.68	10.17	-0.73	-0.03	0	60	60	0.90	3.17
	10	68161	36.63	0.65	10.75	-0.56	-0.42	0	56	56	0.92	3.04
Mathematics	3	59019	39.72	0.70	9.56	-0.60	-0.07	0	57	57	0.90	2.99
	4	59115	40.27	0.71	8.93	-0.65	0.07	0	57	57	0.89	2.94
	5	58678	41.86	0.68	10.56	-0.47	-0.37	0	62	62	0.90	3.26
	6	59907	39.95	0.64	11.21	-0.34	-0.64	0	62	62	0.92	3.27
	7	60967	39.99	0.64	11.27	-0.35	-0.62	0	62	62	0.92	3.24
	8	62914	35.22	0.57	11.79	-0.08	-0.74	0	62	62	0.92	3.39
	10	68330	33.13	0.57	12.77	-0.01	-1.04	0	58	58	0.93	3.26
Language Arts	4	58755	19.66	0.66	5.29	-0.37	-0.55	0	30	30	0.82	2.25
	8	62581	21.64	0.72	5.49	-0.72	-0.14	1	30	30	0.85	2.15
	10	67803	24.37	0.62	6.33	-0.71	0.07	0	39	39	0.84	2.52
Social Studies	4	59022	28.97	0.76	6.40	-0.93	0.45	1	38	38	0.87	2.28
	8	62701	27.88	0.70	7.15	-0.50	-0.47	1	40	40	0.87	2.54
	10	67916	32.25	0.64	9.20	-0.38	-0.62	0	50	50	0.90	2.91
Science	4	59057	28.76	0.72	6.43	-0.68	-0.02	0	40	40	0.85	2.46
	8	62721	29.57	0.74	6.58	-0.72	0.03	0	40	40	0.86	2.45
	10	67953	30.80	0.62	8.78	-0.29	-0.56	0	50	50	0.88	3.03

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-26
Raw Score Descriptive Statistics based on 13 Districts

Content	Grade	N Count	Mean # of Items Correct	Test Difficulty	SD	Skewness	Kurtosis	Min Obtained	Max Obtained	Max Possible	Alpha	SEM
Reading	3	6507	38.82	0.65	12.39	-0.53	-0.72	3	59	60	0.94	3.08
	4	6353	38.97	0.65	11.56	-0.58	-0.52	1	60	60	0.92	3.18
	5	6309	39.50	0.66	11.45	-0.66	-0.42	1	59	60	0.92	3.16
	6	6391	38.45	0.64	10.51	-0.52	-0.44	0	60	60	0.91	3.21
	7	6537	38.58	0.64	10.75	-0.57	-0.38	4	59	60	0.91	3.23
	8	6633	40.82	0.68	10.79	-0.60	-0.35	1	60	60	0.91	3.20
	10	7250	36.20	0.65	11.57	-0.48	-0.62	0	56	56	0.93	3.09
Mathematics	3	6548	39.60	0.69	9.94	-0.52	-0.24	4	57	57	0.91	3.02
	4	6385	40.40	0.71	9.14	-0.63	0.01	1	57	57	0.90	2.93
	5	6334	42.39	0.68	10.75	-0.51	-0.36	7	62	62	0.91	3.25
	6	6416	40.63	0.66	11.40	-0.34	-0.67	0	62	62	0.92	3.25
	7	6560	40.38	0.65	11.58	-0.34	-0.69	4	62	62	0.92	3.22
	8	6650	35.37	0.57	12.08	-0.05	-0.81	1	62	62	0.92	3.38
	10	7248	33.01	0.57	13.35	0.05	-1.12	0	58	58	0.94	3.26
Language Arts	4	6348	19.65	0.65	5.47	-0.39	-0.53	1	30	30	0.83	2.24
	8	6598	21.35	0.71	5.68	-0.64	-0.33	2	30	30	0.86	2.16
	10	7141	24.15	0.62	6.79	-0.69	0.01	2	39	39	0.86	2.57
Social Studies	4	6365	28.74	0.76	6.64	-0.88	0.29	1	38	38	0.88	2.29
	8	6617	27.54	0.69	7.48	-0.41	-0.64	3	40	40	0.88	2.55
	10	7124	32.11	0.64	9.66	-0.35	-0.70	1	50	50	0.91	2.91
Science	4	6375	28.38	0.71	6.67	-0.59	-0.26	5	40	40	0.86	2.48
	8	6626	29.02	0.73	6.87	-0.62	-0.19	4	40	40	0.87	2.47
	10	7123	30.24	0.60	9.28	-0.19	-0.70	2	50	50	0.89	3.03

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-27
Raw Score Descriptive Statistics by Gender

Content	Grade	Male					Female				
		N Count	Mean	Test Difficulty	SD	Alpha	N Count	Mean	Test Difficulty	SD	Alpha
Reading	3	29976	38.39	0.64	12.20	0.93	28567	40.62	0.68	11.34	0.93
	4	29962	38.36	0.64	11.44	0.92	28764	40.06	0.67	10.75	0.91
	5	29764	38.61	0.64	11.04	0.92	28635	40.22	0.67	10.40	0.91
	6	30590	37.78	0.63	10.24	0.90	29093	39.26	0.65	9.59	0.89
	7	31022	37.97	0.63	10.50	0.91	29747	39.14	0.65	9.70	0.89
	8	31942	39.88	0.66	10.46	0.91	30840	42.27	0.70	9.71	0.90
	10	34790	35.67	0.64	11.04	0.92	33371	37.64	0.67	10.34	0.92
Mathematics	3	30264	39.77	0.70	9.61	0.90	28755	39.67	0.70	9.51	0.90
	4	30198	40.48	0.71	8.95	0.89	28917	40.06	0.70	8.90	0.89
	5	29938	42.12	0.68	10.73	0.91	28740	41.59	0.67	10.38	0.90
	6	30725	40.02	0.65	11.42	0.92	29182	39.87	0.64	10.99	0.91
	7	31146	40.16	0.65	11.50	0.92	29813	39.82	0.64	11.02	0.91
	8	32018	35.73	0.58	12.00	0.92	30896	34.68	0.56	11.54	0.91
	10	34884	33.65	0.58	13.00	0.94	33446	32.60	0.56	12.50	0.93
Language Arts	4	29989	19.06	0.64	5.39	0.82	28766	20.28	0.68	5.11	0.81
	8	31829	20.74	0.69	5.72	0.85	30752	22.58	0.75	5.07	0.83
	10	34557	23.34	0.60	6.56	0.85	33162	25.46	0.65	5.87	0.83
Social Studies	4	30161	28.79	0.76	6.51	0.88	28861	29.17	0.77	6.26	0.87
	8	31913	28.03	0.70	7.33	0.88	30788	27.72	0.69	6.96	0.87
	10	34654	32.26	0.65	9.66	0.91	33262	32.23	0.64	8.70	0.89
Science	4	30178	28.89	0.72	6.55	0.86	28879	28.62	0.72	6.29	0.85
	8	31931	29.80	0.74	6.84	0.88	30790	29.34	0.73	6.28	0.85
	10	34665	31.59	0.63	9.20	0.90	33288	29.97	0.60	8.23	0.86

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-28
Raw Score Descriptive Statistics for Reading by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	Test Difficulty	SD	Alpha
Reading	W	3	44294	41.57	0.69	11.02	0.93
		4	44722	41.16	0.69	10.28	0.91
		5	44833	41.32	0.69	9.84	0.90
		6	46125	40.33	0.67	9.06	0.88
		7	46991	40.38	0.67	9.24	0.88
		8	48976	42.86	0.71	9.21	0.89
		10	55153	38.43	0.69	9.86	0.91
	AA	3	6313	31.03	0.52	11.86	0.92
		4	6346	30.72	0.51	11.36	0.91
		5	6178	30.78	0.51	11.07	0.90
		6	6214	30.14	0.50	10.01	0.88
		7	6416	30.28	0.50	10.35	0.89
		8	6637	32.57	0.54	10.67	0.89
		10	5989	26.65	0.48	10.64	0.91
	H	3	4970	33.01	0.55	11.41	0.91
		4	4778	33.59	0.56	10.98	0.91
		5	4495	33.73	0.56	10.77	0.90
		6	4335	33.30	0.55	9.88	0.88
		7	4298	32.90	0.55	10.24	0.89
		8	4036	35.34	0.59	10.49	0.90
		10	3725	29.73	0.53	10.94	0.91
	A	3	2108	37.52	0.63	11.83	0.93
		4	2032	37.17	0.62	11.43	0.92
		5	2097	37.29	0.62	11.41	0.92
		6	2195	35.54	0.59	10.94	0.91
		7	2218	36.06	0.60	10.66	0.90
		8	2200	38.43	0.64	10.51	0.90
		10	2292	32.78	0.59	10.93	0.91
	AI	3	858	35.74	0.60	11.67	0.92
		4	848	35.40	0.59	10.96	0.91
		5	795	35.99	0.60	10.21	0.90
		6	814	34.38	0.57	9.85	0.89
		7	854	34.51	0.58	9.93	0.89
		8	931	37.37	0.62	10.21	0.89
		10	994	32.16	0.57	10.43	0.91

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-29
Raw Score Descriptive Statistics for Mathematics by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	Test Difficulty	SD	Alpha
Mathematics	W	3	44377	41.42	0.73	8.74	0.89
		4	44823	41.84	0.73	8.08	0.87
		5	44906	43.70	0.70	9.76	0.89
		6	46186	41.92	0.68	10.39	0.90
		7	47015	42.14	0.68	10.36	0.91
		8	48985	37.37	0.60	11.05	0.91
		10	55172	35.33	0.61	12.12	0.93
	AA	3	6335	31.67	0.56	10.25	0.91
		4	6380	32.76	0.57	9.77	0.90
		5	6188	32.73	0.53	10.56	0.89
		6	6225	29.65	0.48	10.58	0.90
		7	6415	29.36	0.47	10.28	0.89
		8	6608	24.14	0.39	10.08	0.89
		10	5990	20.32	0.35	9.31	0.88
	H	3	5296	35.38	0.62	9.17	0.88
		4	4986	36.16	0.63	8.81	0.88
		5	4648	36.60	0.59	10.09	0.89
		6	4457	34.42	0.56	10.46	0.89
		7	4432	33.51	0.54	10.44	0.89
		8	4152	28.77	0.46	10.33	0.89
		10	3835	24.41	0.42	10.42	0.90
	A	3	2147	40.30	0.71	9.12	0.89
		4	2074	40.69	0.71	8.94	0.89
		5	2138	42.32	0.68	10.40	0.90
		6	2223	40.69	0.66	11.25	0.92
		7	2255	39.77	0.64	11.45	0.92
		8	2235	34.65	0.56	11.78	0.92
		10	2329	31.10	0.54	12.66	0.93
	AI	3	864	36.55	0.64	9.03	0.88
		4	852	37.14	0.65	8.52	0.87
		5	797	38.73	0.62	9.64	0.88
		6	816	35.16	0.57	10.44	0.89
		7	850	35.68	0.58	10.65	0.90
		8	932	30.68	0.49	10.80	0.90
		10	996	26.81	0.46	11.47	0.92

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-30
Raw Score Descriptive Statistics for Language Arts by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	Test Difficulty	SD	Alpha
Language Arts	W	4	44757	20.44	0.68	5.02	0.81
		8	48921	22.41	0.75	5.18	0.84
		10	54990	25.40	0.65	5.78	0.82
	AA	4	6339	16.12	0.54	5.30	0.80
		8	6530	18.03	0.60	5.70	0.83
		10	5795	18.26	0.47	6.53	0.82
	H	4	4779	17.58	0.59	5.07	0.78
		8	4000	19.14	0.64	5.53	0.82
		10	3666	20.53	0.53	6.38	0.83
	A	4	2034	19.06	0.64	5.44	0.83
		8	2197	20.74	0.69	5.40	0.83
		10	2250	22.95	0.59	6.12	0.82
	AI	4	846	17.74	0.59	5.20	0.80
		8	931	19.37	0.65	5.56	0.83
		10	978	21.23	0.54	6.44	0.83

Table 8-31
Raw Score Descriptive Statistics for Social Studies by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	Test Difficulty	SD	Alpha
Social Studies	W	4	44890	30.11	0.79	5.72	0.85
		8	48970	29.15	0.73	6.58	0.86
		10	55107	33.69	0.67	8.62	0.89
	AA	4	6361	23.85	0.63	7.32	0.88
		8	6522	21.65	0.54	7.02	0.84
		10	5834	23.82	0.48	8.64	0.87
	H	4	4877	26.04	0.69	6.48	0.85
		8	4072	24.06	0.60	6.94	0.84
		10	3719	26.60	0.53	8.77	0.88
	A	4	2042	27.86	0.73	6.56	0.87
		8	2202	26.14	0.65	7.23	0.87
		10	2267	29.88	0.60	8.86	0.88
	AI	4	852	26.70	0.70	6.50	0.86
		8	933	25.31	0.63	7.03	0.85
		10	981	27.96	0.56	8.89	0.88

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-32
Raw Score Descriptive Statistics for Science by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	Test Difficulty	SD	Alpha
Science	W	4	44890	29.98	0.75	5.79	0.83
		8	48979	30.83	0.77	5.89	0.84
		10	55100	32.34	0.65	8.12	0.86
	AA	4	6379	23.25	0.58	6.89	0.85
		8	6533	23.27	0.58	6.80	0.84
		10	5868	21.58	0.43	7.69	0.83
	H	4	4889	25.61	0.64	6.35	0.83
		8	4072	25.92	0.65	6.55	0.83
		10	3716	24.95	0.50	8.05	0.84
	A	4	2044	27.55	0.69	6.41	0.84
		8	2207	27.95	0.70	6.52	0.84
		10	2273	28.49	0.57	8.43	0.86
	AI	4	855	26.54	0.66	6.35	0.83
		8	928	27.37	0.68	6.56	0.84
		10	988	26.87	0.54	8.46	0.86

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-33
Raw Score Descriptive Statistics by Socioeconomic Status

Content	Grade	Economically Disadvantaged					Not Economically Disadvantaged				
		N Count	Mean	Test Difficulty	SD	Alpha	N Count	Mean	Test Difficulty	SD	Alpha
Reading	3	21027	34.02	0.57	11.94	0.92	37516	42.54	0.71	10.61	0.92
	4	20939	33.87	0.56	11.37	0.91	37787	42.14	0.70	9.84	0.90
	5	20011	34.06	0.57	11.01	0.91	38388	42.18	0.70	9.51	0.90
	6	19766	33.07	0.55	10.10	0.89	39917	41.20	0.69	8.70	0.87
	7	19684	33.10	0.55	10.34	0.89	41093	41.15	0.69	8.93	0.88
	8	19713	35.64	0.59	10.59	0.90	43069	43.53	0.73	8.94	0.88
	10	17552	30.41	0.54	10.94	0.91	50609	38.79	0.69	9.79	0.91
Mathematics	3	21432	35.50	0.62	9.73	0.90	37587	42.13	0.74	8.58	0.88
	4	21232	36.23	0.64	9.24	0.89	37883	42.54	0.75	7.89	0.87
	5	20216	36.74	0.59	10.55	0.90	38462	44.55	0.72	9.53	0.89
	6	19943	34.05	0.55	10.99	0.90	39964	42.89	0.69	10.11	0.90
	7	19826	33.79	0.55	10.85	0.90	41141	42.98	0.69	10.20	0.90
	8	19818	28.82	0.46	10.99	0.90	43096	38.16	0.62	10.95	0.91
	10	17670	25.69	0.44	11.36	0.91	50660	35.73	0.62	12.19	0.93
Language Arts	4	20937	17.46	0.58	5.22	0.80	37818	20.87	0.70	4.92	0.80
	8	19569	19.12	0.64	5.65	0.83	43012	22.79	0.76	5.00	0.83
	10	16214	20.89	0.54	6.51	0.83	51589	25.46	0.65	5.86	0.82
Social Studies	4	21141	26.07	0.69	6.80	0.87	37881	30.59	0.81	5.53	0.85
	8	19659	24.11	0.60	7.18	0.86	43042	29.60	0.74	6.44	0.86
	10	17375	27.07	0.54	8.99	0.88	50541	34.03	0.68	8.58	0.89
Science	4	21159	25.80	0.64	6.65	0.84	37898	30.41	0.76	5.66	0.83
	8	19660	26.08	0.65	6.90	0.85	43061	31.17	0.78	5.75	0.83
	10	17398	25.84	0.52	8.62	0.87	50555	32.51	0.65	8.16	0.87

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-34
Raw Score Descriptive Statistics by Disability

Content	Grade	Disabled					Not Disabled				
		N Count	Mean	Test Difficulty	SD	Alpha	N Count	Mean	Test Difficulty	SD	Alpha
Reading	3	7362	29.76	0.50	12.67	0.93	51181	40.88	0.68	11.03	0.92
	4	7565	28.91	0.48	12.11	0.92	51161	40.71	0.68	10.14	0.91
	5	7544	28.70	0.48	11.50	0.91	50855	40.99	0.68	9.69	0.90
	6	7678	27.72	0.46	10.21	0.89	52005	40.09	0.67	8.87	0.87
	7	7665	27.61	0.46	10.42	0.89	53112	40.12	0.67	9.06	0.88
	8	8271	29.20	0.49	10.19	0.88	54511	42.85	0.71	8.88	0.88
	10	8367	24.51	0.44	10.01	0.89	59794	38.33	0.68	9.71	0.90
Mathematics	3	7541	34.49	0.61	10.34	0.91	51478	40.49	0.71	9.20	0.90
	4	7744	33.93	0.60	10.11	0.90	51371	41.23	0.72	8.33	0.88
	5	7647	33.12	0.53	11.21	0.91	51031	43.17	0.70	9.81	0.89
	6	7757	29.38	0.47	10.99	0.90	52150	41.52	0.67	10.36	0.90
	7	7699	28.70	0.46	10.56	0.90	53268	41.62	0.67	10.40	0.91
	8	8272	23.77	0.38	9.92	0.88	54642	36.95	0.60	11.06	0.91
	10	8378	20.08	0.35	9.23	0.87	59952	34.96	0.60	12.11	0.93
Language Arts	4	7639	16.17	0.54	5.21	0.79	51116	20.18	0.67	5.10	0.81
	8	8220	16.08	0.54	5.57	0.81	54361	22.48	0.75	4.96	0.82
	10	7909	17.46	0.45	6.08	0.80	59894	25.28	0.65	5.77	0.82
Social Studies	4	7866	25.15	0.66	6.93	0.87	51156	29.56	0.78	6.10	0.86
	8	8280	21.35	0.53	7.13	0.84	54421	28.87	0.72	6.61	0.86
	10	8314	23.17	0.46	8.48	0.86	59602	33.51	0.67	8.57	0.89
Science	4	7865	25.43	0.64	6.99	0.86	51192	29.27	0.73	6.18	0.84
	8	8295	23.65	0.59	7.25	0.86	54426	30.47	0.76	5.97	0.84
	10	8323	22.62	0.45	8.22	0.85	59630	31.94	0.64	8.23	0.87

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-35
Raw Score Descriptive Statistics by English Language Proficiency

Content	Grade	Proficient					Not Proficient				
		N Count	Mean	Test Difficulty	SD	Alpha	N Count	Mean	Test Difficulty	SD	Alpha
Reading	3	54543	40.08	0.67	11.69	0.93	4000	31.31	0.52	10.83	0.90
	4	55053	39.75	0.66	10.98	0.92	3673	30.91	0.52	10.13	0.88
	5	54903	39.94	0.67	10.57	0.91	3496	30.89	0.51	10.11	0.88
	6	56446	39.02	0.65	9.76	0.89	3237	29.56	0.49	9.02	0.85
	7	57675	39.03	0.65	9.94	0.90	3102	29.45	0.49	9.45	0.86
	8	59899	41.48	0.69	10.00	0.90	2883	32.24	0.54	9.68	0.87
	10	65899	37.04	0.66	10.56	0.92	2262	24.76	0.44	9.18	0.87
Mathematics	3	54658	40.06	0.70	9.53	0.90	4361	35.52	0.62	9.00	0.88
	4	55193	40.60	0.71	8.86	0.89	3922	35.65	0.63	8.67	0.87
	5	54986	42.27	0.68	10.47	0.90	3692	35.76	0.58	10.00	0.88
	6	56504	40.34	0.65	11.15	0.91	3403	33.36	0.54	10.15	0.88
	7	57698	40.45	0.65	11.16	0.92	3269	31.81	0.51	10.00	0.88
	8	59882	35.60	0.57	11.75	0.92	3032	27.55	0.44	9.78	0.87
	10	65908	33.53	0.58	12.71	0.93	2422	22.25	0.38	9.05	0.86
Language Arts	4	55079	19.86	0.66	5.26	0.82	3676	16.58	0.55	4.78	0.75
	8	59724	21.82	0.73	5.44	0.85	2857	17.85	0.59	5.12	0.78
	10	66339	24.48	0.63	6.29	0.84	1464	19.02	0.49	5.52	0.76
Social Studies	4	55257	29.26	0.77	6.31	0.87	3765	24.81	0.65	6.22	0.84
	8	59771	28.16	0.70	7.07	0.87	2930	22.23	0.56	6.28	0.80
	10	65662	32.55	0.65	9.11	0.90	2254	23.43	0.47	7.39	0.82
Science	4	55281	29.05	0.73	6.35	0.85	3776	24.44	0.61	5.98	0.79
	8	59785	29.82	0.75	6.50	0.86	2936	24.48	0.61	6.01	0.79
	10	65699	31.08	0.62	8.70	0.88	2254	22.62	0.45	6.96	0.78

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-36
Scale Score Mean and Standard Deviation for Census and 13 CD Results

Content	Grade	13 CD Mean	Census Mean	Diff= 13 CD-Census	13 CD Standard deviation	Census Standard deviation	Diff= 13 CD-Census
Reading	3	455.80	457.50	-1.69	43.61	40.32	3.30
	4	475.51	476.18	-0.67	50.41	47.23	3.19
	5	484.56	484.48	0.08	51.40	46.47	4.93
	6	503.37	503.15	0.22	52.67	48.95	3.73
	7	514.01	513.83	0.18	52.92	48.42	4.51
	8	526.81	527.56	-0.74	56.73	51.98	4.75
	10	536.39	538.56	-2.17	67.03	60.03	7.00
Mathematics	3	431.61	431.74	-0.14	47.36	44.42	2.95
	4	467.88	466.33	1.55	47.35	45.46	1.89
	5	495.93	493.08	2.86	50.47	48.65	1.82
	6	517.17	513.85	3.33	48.31	46.52	1.79
	7	535.58	533.44	2.14	46.21	44.62	1.59
	8	542.50	541.46	1.04	51.19	49.89	1.30
	10	560.68	561.27	-0.58	52.50	48.67	3.84
Language Arts	4	295.54	295.46	0.08	32.05	29.75	2.30
	8	394.89	396.00	-1.11	43.78	41.86	1.92
	10	446.63	447.27	-0.64	43.22	38.59	4.63
Social Studies	4	296.53	296.92	-0.38	28.35	27.26	1.09
	8	393.66	394.93	-1.27	40.35	37.84	2.51
	10	448.19	448.79	-0.60	48.85	43.99	4.86
Science	4	296.16	296.92	-0.76	31.45	30.22	1.23
	8	397.11	399.47	-2.36	44.32	41.94	2.38
	10	449.29	452.02	-2.72	49.19	44.99	4.20

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-37
Scale Score Descriptive Statistics based on 13 Districts

Content	Grade	N Count	Mean	SD	Skewness	Kurtosis	Min	Max	LOSS	HOSS
Reading	3	6507	455.80	43.61	-0.91	3.63	270	623	270	640
	4	6353	475.51	50.41	-0.90	2.54	280	650	280	650
	5	6309	484.56	51.40	-1.15	2.92	290	651	290	690
	6	6391	503.37	52.67	-0.70	1.98	300	730	300	730
	7	6537	514.01	52.92	-0.78	1.85	310	688	310	780
	8	6633	526.81	56.73	-0.50	1.66	330	790	330	790
	10	7250	536.39	67.03	-0.33	0.93	350	820	350	820
Mathematics	3	6548	431.61	47.36	-0.04	1.90	220	630	220	630
	4	6385	467.88	47.35	-0.27	1.77	240	650	240	650
	5	6334	495.93	50.47	-0.22	1.40	270	680	270	680
	6	6416	517.17	48.31	-0.30	1.21	310	700	310	700
	7	6560	535.58	46.21	-0.21	0.96	330	710	330	710
	8	6650	542.50	51.19	-0.59	1.35	350	730	350	730
	10	7248	560.68	52.50	-0.32	0.68	410	750	410	750
Language Arts	4	6348	295.54	32.05	-0.48	3.98	140	420	140	420
	8	6598	394.89	43.78	-0.17	1.38	250	520	250	520
	10	7138	446.63	43.22	-0.43	1.28	290	630	290	630
Social Studies	4	6365	296.53	28.35	0.59	3.70	170	400	170	400
	8	6617	393.66	40.35	-0.16	1.96	230	530	230	530
	10	7124	448.19	48.85	-0.66	2.68	240	620	240	620
Science	4	6375	296.16	31.45	0.28	2.30	170	440	170	440
	8	6626	397.11	44.32	0.34	2.43	230	560	230	560
	10	7123	449.29	49.19	-0.84	2.75	240	610	240	610

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-38
Scale Score Descriptive Statistics based on Census Data

Content	Grade	N Count	Mean	SD	Skewness	Kurtosis	Min	Max	LOSS	HOSS
Reading	3	58543	457.50	40.32	-1.05	4.28	270	640	270	640
	4	58726	476.18	47.23	-0.93	2.75	280	650	280	650
	5	58399	484.48	46.47	-1.09	3.37	290	690	290	690
	6	59683	503.15	48.95	-0.90	2.45	300	730	300	730
	7	60777	513.83	48.42	-0.82	2.21	310	780	310	780
	8	62782	527.56	51.98	-0.75	2.11	330	790	330	790
	10	68161	538.56	60.03	-0.44	1.20	350	820	350	820
Mathematics	3	59019	431.74	44.42	-0.13	1.99	220	630	220	630
	4	59115	466.33	45.46	-0.29	1.88	240	650	240	650
	5	58678	493.08	48.65	-0.30	1.56	270	680	270	680
	6	59907	513.85	46.52	-0.46	1.39	310	700	310	700
	7	60967	533.44	44.62	-0.34	1.39	330	710	330	710
	8	62914	541.46	49.89	-0.69	1.56	350	730	350	730
	10	68330	561.27	48.67	-0.44	0.95	410	750	410	750
Language Arts	4	58755	295.46	29.75	-0.34	3.88	140	420	140	420
	8	62581	396.00	41.86	-0.17	1.73	250	520	250	520
	10	67837	447.27	38.59	-0.43	1.11	290	630	290	630
Social Studies	4	59022	296.92	27.26	0.54	4.27	170	400	170	400
	8	62701	394.93	37.84	-0.19	2.25	230	530	230	530
	10	67916	448.79	43.99	-0.64	2.89	240	620	240	620
Science	4	59057	296.92	30.22	0.20	2.67	170	440	170	440
	8	62721	399.47	41.94	0.31	2.49	230	560	230	560
	10	67953	452.02	44.99	-0.91	3.18	240	610	240	610

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-39
Scale Score Descriptive Statistics by Gender

Content	Grade	Male					Female				
		N Count	Mean	SD	Min	Max	N Count	Mean	SD	Min	Max
Reading	3	29976	453.79	41.82	270	640	28567	461.39	38.29	270	640
	4	29962	472.62	49.26	280	650	28764	479.90	44.71	280	650
	5	29764	481.16	48.44	290	690	28635	487.93	44.06	290	690
	6	30590	499.59	51.15	300	730	29093	506.88	46.22	300	730
	7	31022	511.04	51.18	310	720	29747	516.76	45.16	310	780
	8	31942	522.19	53.73	330	790	30840	533.11	49.49	330	790
	10	34790	533.22	61.98	350	820	33371	544.12	57.41	350	820
Mathematics	3	30264	432.11	44.99	220	630	28755	431.35	43.81	220	630
	4	30198	467.88	46.21	240	650	28917	464.72	44.62	240	650
	5	29938	494.75	50.47	270	680	28740	491.33	46.62	270	680
	6	30725	514.19	48.01	310	700	29182	513.48	44.90	310	700
	7	31146	534.42	46.37	330	710	29813	532.44	42.69	330	710
	8	32018	543.13	51.42	350	730	30896	539.73	48.19	350	730
	10	34884	562.83	50.31	410	750	33446	559.63	46.84	410	750
Language Arts	4	29989	292.04	30.54	140	420	28766	299.02	28.48	140	420
	8	31829	389.18	42.75	250	520	30752	403.06	39.70	250	520
	10	34625	441.03	39.89	290	630	33212	453.78	36.06	290	630
Social Studies	4	30161	296.11	27.68	170	400	28861	297.77	26.80	170	400
	8	31913	395.58	39.72	230	530	30788	394.26	35.78	230	530
	10	34654	448.66	47.38	240	620	33262	448.92	40.15	240	620
Science	4	30178	297.21	31.50	170	440	28879	296.62	28.82	170	440
	8	31931	401.37	45.05	230	560	30790	397.50	38.36	230	560
	10	34665	455.33	48.59	240	610	33288	448.56	40.62	240	610

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-40
Scale Score Descriptive Statistics for Reading by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	SD	Min	Max
Reading	W	3	44294	464.16	37.29	270	640
		4	44722	484.14	43.26	280	650
		5	44833	492.33	41.96	290	690
		6	46125	511.94	43.93	300	730
		7	46991	522.46	43.64	310	780
		8	48976	536.62	47.13	330	790
		10	55153	548.33	55.03	350	820
	AA	3	6313	429.77	44.46	270	582
		4	6346	441.06	52.43	280	608
		5	6178	448.57	52.52	290	639
		6	6214	462.67	53.68	300	613
		7	6416	474.95	52.61	310	640
		8	6637	484.70	56.01	330	659
		10	5989	483.87	61.43	350	693
	H	3	4970	437.44	39.50	270	582
		4	4778	454.07	45.84	280	643
		5	4495	461.51	46.95	290	606
		6	4335	478.13	48.87	300	672
		7	4298	487.03	49.84	310	688
		8	4036	499.23	52.90	330	693
		10	3725	500.63	61.30	350	721
	A	3	2108	452.31	38.95	270	582
		4	2032	468.87	48.43	280	650
		5	2097	476.51	50.64	290	690
		6	2195	489.42	54.76	300	652
		7	2218	502.32	51.51	310	780
		8	2200	514.67	53.62	330	737
		10	2292	518.28	61.48	350	820
	AI	3	858	446.38	36.51	270	557
		4	848	461.66	45.82	280	602
		5	795	471.41	42.67	290	582
		6	814	483.91	48.78	300	618
		7	854	496.09	46.42	310	627
		8	931	509.37	49.87	330	646
		10	994	514.96	58.12	350	820

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-41
Scale Score Descriptive Statistics for Mathematics by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	SD	Min	Max
Mathematics	**W**	3	44377	439.43	41.39	220	630
		4	44823	473.97	41.86	240	650
		5	44906	501.22	45.02	270	680
		6	46186	521.91	42.41	310	700
		7	47015	541.75	40.66	330	710
		8	48985	550.88	44.16	350	730
		10	55172	569.82	44.21	410	750
	AA	3	6335	395.64	46.64	220	630
		4	6380	429.61	49.68	240	650
		5	6188	451.96	50.90	270	658
		6	6225	471.05	48.64	310	615
		7	6415	491.63	44.65	330	662
		8	6608	491.41	54.74	350	669
		10	5990	508.79	47.75	410	662
	H	3	5296	411.86	39.26	220	630
		4	4986	446.26	41.42	240	603
		5	4648	469.88	44.14	270	680
		6	4457	491.49	43.38	310	654
		7	4432	508.95	40.34	330	710
		8	4152	514.92	47.32	350	655
		10	3835	529.42	44.78	410	678
	A	3	2147	434.33	43.39	220	630
		4	2074	468.74	46.55	240	650
		5	2138	496.39	50.10	270	680
		6	2223	517.54	48.29	310	700
		7	2255	533.34	46.91	330	710
		8	2235	539.95	50.64	350	730
		10	2329	555.58	49.10	410	750
	AI	3	864	417.07	38.38	220	529
		4	852	450.97	40.73	292	603
		5	797	479.84	40.94	270	647
		6	816	495.68	40.75	347	628
		7	850	517.29	40.34	330	635
		8	932	523.51	47.20	350	640
		10	996	539.14	45.54	410	674

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-42
Scale Score Descriptive Statistics for Language Arts by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	SD	Min	Max
Language Arts	W	4	44757	299.71	28.19	140	420
		8	48921	401.57	40.41	250	520
		10	55067	453.17	35.94	290	630
	AA	4	6339	276.06	31.35	140	420
		8	6530	369.83	40.95	250	520
		10	5855	411.78	39.05	290	578
	H	4	4779	284.17	28.01	140	420
		8	4000	377.57	39.12	250	520
		10	3676	425.35	36.95	290	568
	A	4	2034	292.88	31.10	140	420
		8	2197	389.82	40.93	250	520
		10	2255	439.31	36.32	290	588
	AI	4	846	285.97	26.91	140	420
		8	931	380.58	38.25	250	520
		10	977	428.53	38.37	290	568

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-43
Scale Score Descriptive Statistics for Social Studies by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	SD	Min	Max
Social Studies	W	4	44890	301.24	26.09	170	400
		8	48970	401.40	35.08	230	530
		10	55107	455.51	40.35	240	620
	AA	4	6361	277.48	27.23	170	400
		8	6522	362.86	38.19	230	530
		10	5834	408.10	49.10	240	589
	H	4	4877	285.60	23.57	170	400
		8	4072	375.80	35.28	230	530
		10	3719	423.96	42.61	240	586
	A	4	2042	293.23	26.48	170	400
		8	2202	386.84	38.07	230	530
		10	2267	438.87	42.81	240	620
	AI	4	852	287.86	24.40	170	400
		8	933	382.27	35.96	230	530
		10	981	430.34	42.10	240	589

Table 8-44
Scale Score Descriptive Statistics for Science by Ethnicity

Content	Ethnicity	Grade	N Count	Mean	SD	Min	Max
Science	W	4	44890	302.23	28.41	170	440
		8	48979	406.69	39.55	230	560
		10	55100	459.76	40.04	240	610
	AA	4	6379	273.12	29.93	170	440
		8	6533	363.38	38.70	230	560
		10	5868	404.35	51.31	240	610
	H	4	4889	282.72	26.84	170	440
		8	4072	378.11	36.79	230	560
		10	3716	423.57	45.84	240	587
	A	4	2044	292.52	30.04	170	440
		8	2207	390.81	40.93	230	560
		10	2273	442.19	40.77	240	610
	AI	4	855	287.05	27.37	170	440
		8	928	386.64	40.25	230	560
		10	988	432.97	44.28	240	535

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-45
Scale Score Descriptive Statistics by Socioeconomic Status

Content	Grade	Economically Disadvantaged					Not Economically Disadvantaged				
		N Count	Mean	SD	Min	Max	N Count	Mean	SD	Min	Max
Reading	3	21027	439.89	41.54	270	623	37516	467.36	36.03	270	640
	4	20939	454.46	49.31	280	643	37787	488.22	41.38	280	650
	5	20011	462.45	49.12	290	666	38388	495.96	40.53	290	690
	6	19766	477.16	50.93	300	657	39917	516.02	42.40	300	730
	7	19684	488.58	50.36	310	688	41093	525.93	42.43	310	780
	8	19713	500.63	54.11	330	713	43069	539.88	45.99	330	790
	10	17552	504.58	60.96	350	820	50609	550.34	55.01	350	820
Mathematics	3	21432	412.43	42.81	220	630	37587	442.75	41.46	220	630
	4	21232	446.32	45.15	240	650	37883	477.55	41.63	240	650
	5	20216	470.08	47.82	270	680	38462	505.16	44.56	270	680
	6	19943	489.73	46.73	310	700	39964	525.88	41.47	310	700
	7	19826	509.42	43.16	330	688	41141	545.02	40.51	330	710
	8	19818	514.18	51.83	350	730	43096	554.01	43.56	350	730
	10	17670	532.89	48.42	410	750	50660	571.16	44.70	410	750
Language Arts	4	20937	283.51	29.17	140	420	37818	302.07	27.95	140	420
	8	19569	377.60	40.55	250	520	43012	404.37	39.71	250	520
	10	17323	425.83	38.29	290	578	50514	454.63	35.86	290	630
Social Studies	4	21141	285.48	25.24	170	400	37881	303.30	26.25	170	400
	8	19659	375.74	37.14	230	530	43042	403.70	34.80	230	530
	10	17375	425.02	44.75	240	620	50541	456.95	40.63	240	620
Science	4	21159	283.66	28.63	170	440	37898	304.32	28.51	170	440
	8	19660	378.91	39.78	230	560	43061	408.86	39.48	230	560
	10	17398	427.33	48.10	240	610	50555	460.51	40.53	240	610

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-46
Scale Score Descriptive Statistics by Disability

Content	Grade	Disabled					Not Disabled				
		N Count	Mean	SD	Min	Max	N Count	Mean	SD	Min	Max
Reading	3	7362	424.84	49.39	270	586	51181	462.19	36.51	270	640
	4	7565	432.39	58.98	280	650	51161	482.66	41.48	280	650
	5	7544	437.76	59.60	290	666	50855	491.41	39.76	290	690
	6	7678	449.59	58.93	300	672	52005	511.05	41.84	300	730
	7	7665	462.36	55.87	310	720	53112	521.26	42.36	310	780
	8	8271	468.27	58.30	330	705	54511	536.55	44.52	330	790
	10	8367	471.89	60.39	350	702	59794	547.89	53.75	350	820
Mathematics	3	7541	407.98	46.77	220	630	51478	435.22	42.97	220	630
	4	7744	434.61	51.66	240	650	51371	471.12	42.45	240	650
	5	7647	452.81	54.11	270	680	51031	499.11	44.76	270	680
	6	7757	469.31	51.25	310	654	52150	520.47	41.91	310	700
	7	7699	488.58	46.48	330	710	53268	539.92	40.42	330	710
	8	8272	489.79	53.82	350	730	54642	549.28	44.29	350	730
	10	8378	507.05	47.69	410	750	59952	568.84	43.75	410	750
Language Arts	4	7639	276.12	32.09	140	420	51116	298.35	28.27	140	420
	8	8220	356.47	42.31	250	520	54361	401.98	38.40	250	520
	10	8287	406.58	36.34	290	592	59550	452.94	35.36	290	630
Social Studies	4	7866	282.72	25.80	170	400	51156	299.10	26.82	170	400
	8	8280	361.47	39.84	230	530	54421	400.02	34.82	230	530
	10	8314	406.37	46.82	240	620	59602	454.70	40.17	240	620
Science	4	7865	282.74	30.69	170	440	51192	299.09	29.55	170	440
	8	8295	364.70	43.41	230	560	54426	404.77	39.09	230	560
	10	8323	409.43	51.26	240	610	59630	457.96	40.64	240	610

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-47
Scale Score Descriptive Statistics by English Language Proficiency

Content	Grade	Proficient					Not Proficient				
		N Count	Mean	SD	Min	Max	N Count	Mean	SD	Min	Max
Reading	3	54543	459.32	39.89	270	640	4000	432.59	37.75	270	552
	4	55053	478.36	46.65	280	650	3673	443.61	43.66	280	594
	5	54903	486.67	45.58	290	690	3496	450.00	46.63	290	566
	6	56446	505.57	47.97	300	730	3237	460.94	46.49	300	607
	7	57675	516.13	47.32	310	780	3102	471.07	48.58	310	688
	8	59899	529.64	51.12	330	790	2883	484.21	50.74	330	790
	10	65899	540.78	58.96	350	820	2262	473.85	54.91	350	646
Mathematics	3	54658	433.31	44.53	220	630	4361	412.14	37.91	220	630
	4	55193	467.96	45.32	240	650	3922	443.42	41.10	240	617
	5	54986	494.89	48.39	270	680	3692	466.05	44.30	270	638
	6	56504	515.48	46.21	310	700	3403	486.68	43.09	310	623
	7	57698	535.20	44.24	330	710	3269	502.28	39.61	330	670
	8	59882	543.06	49.53	350	730	3032	509.93	46.37	350	730
	10	65908	562.77	48.22	410	750	2422	520.37	42.61	410	682
Language Arts	4	55079	296.58	29.55	140	420	3676	278.67	27.60	140	420
	8	59724	397.30	41.70	250	520	2857	368.76	35.40	250	520
	10	65639	448.42	38.25	290	630	2198	413.00	32.48	290	527
Social Studies	4	55257	297.98	27.32	170	400	3765	281.36	21.03	170	400
	8	59771	396.30	37.56	230	530	2930	367.07	32.33	230	530
	10	65662	450.13	43.48	240	620	2254	409.49	40.42	240	544
Science	4	55281	298.22	30.10	170	440	3776	277.90	25.14	170	381
	8	59785	400.93	41.80	230	560	2936	369.79	32.91	230	560
	10	65699	453.38	44.48	240	610	2254	412.17	41.28	240	558

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-48
Performance Level Cut Scores for all Contents

Content	3			4			5			6			7			8			10		
	B	**P**	**A**	**B**	**P**	**A**	**B**	**P**	**A**	**B**	**P**	**A**	**B**	**P**	**A**	**B**	**P**	**A**	**B**	**P**	**A**
Reading	394	430	466	396	440	489	401	444	497	418	457	514	434	467	523	445	480	539	456	503	555
Mathematics	392	407	452	421	438	484	445	463	505	464	485	532	480	504	555	483	513	573	516	541	595
Language Arts				252	277	308										358	385	418	393	428	484
Social Studies				242	263	288										334	364	403	408	420	455
Science				249	279	320										349	375	419	411	429	466

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-49
Percentage of Students in Each Performance Level by Sub-Group (Reading)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
3	M	2860	4.89	3.47	6.23	3.18	13.69	8.67	4.70	6.88	4.55	9.50	18.76	2.89	9.24	2.45
	B	8622	14.73	12.75	16.62	11.00	28.91	26.94	19.26	20.98	13.57	30.58	31.15	12.37	24.07	9.49
	P	20887	35.68	35.64	35.72	33.78	39.97	43.74	41.08	42.08	34.92	46.03	32.48	36.14	41.33	32.51
	A	26174	44.71	48.15	41.43	52.04	17.44	20.64	34.96	30.07	46.97	13.90	17.62	48.61	25.36	55.55
Total		58543	100	28567	29976	44294	6313	4970	2108	858	54543	4000	7362	51181	21027	37516
4	M	2636	4.49	3.20	5.72	2.79	13.54	7.97	4.87	6.01	4.12	10.02	20.37	2.14	8.98	2.00
	B	7806	13.29	12.25	14.29	9.57	29.17	23.59	18.55	20.40	12.21	29.49	30.83	10.70	22.69	8.08
	P	23612	40.21	40.38	40.04	38.88	41.87	47.80	43.65	46.93	39.58	49.61	33.89	41.14	45.14	37.48
	A	24672	42.01	44.17	39.94	48.77	15.43	20.64	32.92	26.65	44.09	10.89	14.91	46.02	23.19	52.44
Total		58726	100	28764	29962	44722	6346	4778	2032	848	55053	3673	7565	51161	20939	37787
5	M	2171	3.72	2.74	4.65	2.13	12.38	7.07	4.91	3.65	3.34	9.58	17.90	1.61	7.74	1.62
	B	6534	11.19	9.95	12.38	7.88	26.32	20.85	15.07	15.35	10.24	26.03	31.00	8.25	19.73	6.74
	P	25096	42.97	43.09	42.86	41.11	47.04	51.86	47.11	55.35	42.28	53.86	38.14	43.69	50.03	39.29
	A	24598	42.12	44.21	40.11	48.88	14.26	20.22	32.90	25.66	44.13	10.53	12.96	46.45	22.50	52.35
Total		58399	100	28635	29764	44833	6178	4495	2097	795	54903	3496	7544	50855	20011	38388

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-49 Cont'd
Percentage of Students in Each Performance Level by Sub-Group (Reading)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
6	M	2706	4.53	3.22	5.78	2.53	14.93	8.84	7.61	7.25	4.06	12.85	23.09	1.79	9.78	1.93
	B	5728	9.60	8.79	10.37	6.38	23.69	19.01	16.26	15.97	8.58	27.40	27.08	7.02	18.60	5.14
	P	24857	41.65	41.90	41.41	39.85	47.33	50.08	43.51	50.00	41.15	50.42	38.96	42.05	49.11	37.95
	A	26392	44.22	46.09	42.44	51.23	14.05	22.08	32.62	26.78	46.22	9.33	10.88	49.14	22.50	54.97
Total		59683	100	29093	30590	46125	6214	4335	2195	814	56446	3237	7678	52005	19766	39917
7	M	3125	5.14	3.67	6.55	2.79	17.10	11.21	7.71	7.61	4.53	16.44	24.55	2.34	11.20	2.24
	B	5549	9.13	8.11	10.10	6.45	21.17	17.66	12.76	14.05	8.39	22.86	26.64	6.60	16.72	5.49
	P	24193	39.81	41.11	38.55	37.69	45.70	49.02	45.85	49.88	39.22	50.64	36.45	40.29	47.85	35.96
	A	27910	45.92	47.10	44.80	53.08	16.04	22.10	33.68	28.45	47.85	10.06	12.35	50.77	24.24	56.31
Total		60777	100	29747	31022	46991	6416	4298	2218	854	57675	3102	7665	53112	19684	41093
8	M	3601	5.74	4.01	7.40	3.29	18.83	12.29	7.59	8.16	5.22	16.55	27.94	2.37	12.33	2.72
	B	5560	8.86	7.75	9.93	6.15	21.26	16.82	14.23	15.79	8.20	22.44	25.60	6.32	16.43	5.39
	P	25682	40.91	40.47	41.33	39.18	45.44	49.55	46.68	48.23	40.41	51.23	37.42	41.44	47.77	37.77
	A	27939	44.50	47.77	41.34	51.38	14.46	21.33	31.50	27.82	46.17	9.78	9.04	49.88	23.48	54.12
Total		62782	100	30840	31942	48976	6637	4036	2200	931	59899	2883	8271	54511	19713	43069

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-49 Cont'd
Percentage of Students in Each Performance Level by Sub-Group (Reading)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
10	M	5876	8.62	6.53	10.63	5.26	29.40	20.56	13.39	13.98	7.80	32.45	36.88	4.67	19.16	4.97
	B	10401	15.26	13.78	16.68	12.32	30.42	27.68	22.86	22.74	14.61	34.17	32.63	12.83	25.58	11.68
	P	23128	33.93	34.98	32.93	34.22	28.80	33.05	39.31	39.74	34.10	29.00	22.35	35.55	34.86	33.61
	A	28756	42.19	44.71	39.77	48.20	11.37	18.71	24.43	23.54	43.49	4.38	8.14	46.95	20.40	49.74
Total		68161	100	33371	34790	55153	5989	3725	2292	994	65899	2262	8367	59794	17552	50609

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-50
Percentage of Students in Each Performance Level by Sub-Group (Mathematics)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
3	M	9456	16.02	16.28	15.78	10.76	42.65	28.17	13.65	22.45	15.13	27.20	33.96	13.39	27.95	9.22
	B	5807	9.84	9.74	9.94	8.43	14.03	15.07	11.41	15.51	9.39	15.46	13.57	9.29	13.59	7.70
	P	24933	42.25	42.71	41.81	43.12	34.21	43.41	43.50	45.83	42.08	44.32	36.15	43.14	42.37	42.18
	A	18823	31.89	31.27	32.48	37.69	9.11	13.35	31.44	16.20	33.40	13.02	16.32	34.17	16.09	40.90
Total		59019	100	28755	30264	44377	6335	5296	2147	864	54658	4361	7541	51478	21432	37587
4	M	7929	13.41	14.11	12.75	8.50	38.73	24.15	12.87	20.54	12.56	25.45	36.07	10.00	24.68	7.10
	B	5725	9.69	10.03	9.35	8.17	15.03	15.24	10.61	14.79	9.22	16.17	13.48	9.11	14.03	7.25
	P	25017	42.32	42.72	41.94	43.28	34.56	43.30	41.47	46.36	42.19	44.14	35.14	43.40	42.72	42.09
	A	20444	34.58	33.14	35.96	40.05	11.68	17.31	35.05	18.31	36.03	14.25	15.32	37.49	18.58	43.56
Total		59115	100	28917	30198	44823	6380	4986	2074	852	55193	3922	7744	51371	21232	37883
5	M	8238	14.04	13.83	14.24	9.17	40.56	25.90	12.39	17.69	13.07	28.41	40.96	10.01	26.07	7.71
	B	5616	9.57	10.36	8.81	8.12	15.56	15.02	9.59	13.17	9.13	16.17	14.29	8.86	14.32	7.08
	P	20537	35.00	37.03	33.05	35.04	31.01	38.15	36.30	41.53	34.84	37.35	29.28	35.86	37.01	33.95
	A	24287	41.39	38.78	43.89	47.67	12.86	20.93	41.72	27.60	42.96	18.07	15.47	45.27	22.61	51.26
Total		58678	100	28740	29938	44906	6188	4648	2138	797	54986	3692	7647	51031	20216	38462

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-50 Cont'd
Percentage of Students in Each Performance Level by Sub-Group (Mathematics)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
6	M	7626	12.73	12.28	13.16	7.89	40.64	23.09	11.20	21.08	11.94	25.89	43.33	8.18	25.61	6.30
	B	6655	11.11	11.36	10.87	9.10	19.37	18.87	11.70	17.77	10.59	19.69	17.97	10.09	17.11	8.11
	P	24105	40.24	41.53	39.01	41.37	30.96	41.28	39.90	42.03	40.19	41.02	28.68	41.96	39.93	40.39
	A	21521	35.92	34.83	36.96	41.63	9.03	16.76	37.20	19.12	37.28	13.40	10.02	39.78	17.34	45.20
Total		59907	100	29182	30725	46186	6225	4457	2223	816	56504	3403	7757	52150	19943	39964
7	M	6403	10.50	10.16	10.81	5.95	35.96	21.10	10.42	14.94	9.63	25.82	39.71	6.28	21.95	4.99
	B	7476	12.26	12.99	11.57	9.53	23.41	22.74	14.01	19.88	11.60	23.92	22.52	10.78	19.98	8.54
	P	27829	45.65	46.97	44.38	47.20	34.68	44.68	45.41	48.12	45.84	42.31	31.52	47.69	44.66	46.12
	A	19259	31.59	29.88	33.23	37.32	5.94	11.48	30.16	17.06	32.93	7.95	6.25	35.25	13.41	40.35
Total		60967	100	29813	31146	47015	6415	4432	2255	850	57698	3269	7699	53268	19826	41141
8	M	6938	11.03	10.97	11.08	6.30	38.91	21.53	10.20	16.95	10.38	23.81	40.63	6.55	23.44	5.32
	B	8237	13.09	13.74	12.47	10.55	24.17	22.59	15.84	19.42	12.52	24.41	24.96	11.30	20.81	9.54
	P	31123	49.47	51.05	47.94	51.97	32.08	47.04	50.07	50.64	49.66	45.75	30.10	52.40	45.24	51.42
	A	16616	26.41	24.24	28.51	31.18	4.84	8.84	23.89	12.98	27.44	6.04	4.30	29.76	10.51	33.72
Total		62914	100	30896	32018	48985	6608	4152	2235	932	59882	3032	8272	54642	19818	43096

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-50 Cont'd
Percentage of Students in Each Performance Level by Sub-Group (Mathematics)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
10	M	10829	15.85	15.96	15.74	10.24	52.70	34.29	18.25	28.21	14.93	40.71	56.05	10.23	32.63	10.00
	B	9296	13.61	14.16	13.07	11.72	21.37	23.75	18.46	20.88	13.12	26.92	20.20	12.68	20.31	11.27
	P	31719	46.42	47.90	45.00	49.82	23.46	36.17	44.10	40.96	47.04	29.69	21.17	49.95	38.82	49.07
	A	16486	24.13	21.98	26.19	28.22	2.47	5.79	19.19	9.94	24.92	2.68	2.58	27.14	8.25	29.67
Total		68330	100	33446	34884	55172	5990	3835	2329	996	65908	2422	8378	59952	17670	50660

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-51
Percentage of Students in Each Performance Level by Sub-Group (Language Arts)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
4	M	3019	5.14	3.68	6.54	3.25	14.72	8.98	6.49	8.27	4.71	11.51	15.29	3.62	9.61	2.67
	B	10576	18.00	15.44	20.45	14.67	33.22	26.34	20.80	26.12	17.19	30.20	32.20	15.88	27.32	12.84
	P	26044	44.33	44.22	44.43	44.59	39.94	47.14	44.84	46.34	44.06	48.31	40.23	44.94	45.59	43.63
	A	19116	32.54	36.66	28.58	37.49	12.12	17.54	27.88	19.27	34.04	9.98	12.28	35.56	17.49	40.87
Total		58755	100	28766	29989	44757	6339	4779	2034	846	55079	3676	7639	51116	20937	37818
8	M	9015	14.41	9.72	18.93	10.94	32.21	24.75	15.57	24.38	13.66	29.89	46.28	9.59	25.57	9.33
	B	13928	22.26	20.34	24.11	19.87	30.75	32.13	28.90	29.97	21.52	37.63	30.86	20.95	30.21	18.64
	P	22778	36.40	37.76	35.08	37.92	28.59	31.62	36.19	32.01	36.80	28.00	18.19	39.15	31.83	38.48
	A	16860	26.94	32.18	21.88	31.27	8.45	11.50	19.34	13.64	28.02	4.48	4.67	30.31	12.40	33.56
Total		62581	100	30752	31829	48921	6530	4000	2197	931	59724	2857	8220	54361	19569	43012
10	M	5579	8.22	5.19	11.14	5.31	28.63	17.27	8.12	16.58	7.72	23.20	32.88	4.79	17.99	4.87
	B	13254	19.54	16.38	22.57	16.29	36.31	32.75	28.96	30.30	18.76	42.68	40.32	16.65	31.46	15.45
	P	38050	56.09	58.91	53.38	59.64	32.57	45.40	52.11	46.57	56.85	33.30	25.37	60.37	45.23	59.81
	A	10954	16.15	19.52	12.91	18.76	2.49	4.57	10.82	6.55	16.66	0.82	1.44	18.19	5.32	19.86
Total		67837	100	33212	34625	55067	5855	3676	2255	977	65639	2198	8287	59550	17323	50514

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-52
Percentage of Students in Each Performance Level by Sub-Group (Social Studies)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
4	M	1086	1.84	1.48	2.18	0.88	7.77	2.93	1.71	2.23	1.75	3.16	4.44	1.44	3.85	0.72
	B	3284	5.56	5.16	5.95	3.57	15.53	9.88	6.61	9.16	5.17	11.37	11.99	4.58	10.16	3.00
	P	15136	25.65	25.48	25.80	21.38	40.50	40.17	33.40	37.91	24.20	46.91	41.37	23.23	37.61	18.96
	A	39516	66.95	67.88	66.06	74.18	36.20	47.02	58.28	50.70	68.89	38.57	42.21	70.76	48.38	77.31
Total		59022	100	28861	30161	44890	6361	4877	2042	852	55257	3765	7866	51156	21141	37881
8	M	2703	4.31	3.73	4.87	2.29	16.16	8.60	5.45	6.11	3.96	11.40	18.06	2.22	9.45	1.96
	B	8615	13.74	14.01	13.48	9.85	32.81	24.83	19.62	22.29	12.92	30.51	32.44	10.89	24.66	8.75
	P	25015	39.90	41.64	38.22	39.16	38.82	46.54	44.78	45.66	39.51	47.85	37.67	40.23	44.17	37.94
	A	26368	42.05	40.63	43.43	48.70	12.20	20.04	30.15	25.94	43.61	10.24	11.84	46.65	21.72	51.34
Total		62701	100	30788	31913	48970	6522	4072	2202	933	59771	2930	8280	54421	19659	43042
10	M	10098	14.87	12.91	16.75	10.39	44.17	30.55	18.04	25.48	13.95	41.61	48.65	10.16	29.88	9.71
	B	4574	6.74	6.83	6.64	5.68	11.54	11.67	9.84	11.32	6.47	14.42	12.55	5.92	10.45	5.46
	P	21428	31.55	34.62	28.60	30.94	30.67	35.57	40.05	35.98	31.40	35.89	26.92	32.20	35.60	30.16
	A	31816	46.85	45.63	48.01	52.99	13.63	22.21	32.07	27.22	48.18	8.07	11.88	51.72	24.07	54.68
Total		67916	100	33262	34654	55107	5834	3719	2267	981	65662	2254	8314	59602	17375	50541

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-53
Percentage of Students in Each Performance Level by Sub-Group (Science)

Grade	Proficiency Level	Examinees		Gender		Ethnicity					ELP		Disability		SES	
		N	%	Female	Male	White	African-American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
4	M	2973	5.03	4.59	5.46	2.62	18.47	9.43	4.60	7.37	4.67	10.41	10.93	4.13	9.88	2.33
	B	11391	19.29	19.48	19.10	14.94	36.73	32.05	25.64	29.36	18.11	36.60	31.09	17.48	30.55	13.00
	P	33424	56.60	57.80	55.44	59.57	40.34	51.97	54.01	54.15	57.06	49.87	49.15	57.74	51.48	59.45
	A	11269	19.08	18.13	19.99	22.86	4.47	6.55	15.75	9.12	20.17	3.13	8.82	20.66	8.09	25.22
Total		59057	100	28879	30178	44890	6379	4889	2044	855	55281	3776	7865	51192	21159	37898
8	M	5865	9.35	8.71	9.97	5.55	30.75	18.15	12.14	14.01	8.69	22.72	32.51	5.82	19.37	4.78
	B	9639	15.37	16.11	14.65	11.92	30.92	26.50	21.93	23.28	14.58	31.51	26.67	13.65	24.82	11.05
	P	29043	46.31	49.27	43.45	48.34	32.60	44.20	45.54	46.34	46.56	41.04	32.26	48.45	42.73	47.94
	A	18174	28.98	25.91	31.93	34.19	5.72	11.15	20.39	16.38	30.17	4.73	8.56	32.09	13.09	36.23
Total		62721	100	30790	31931	48979	6533	4072	2207	928	59785	2936	8295	54426	19660	43061
10	M	10104	14.87	14.67	15.06	9.44	50.61	32.97	19.71	26.22	13.89	43.52	47.62	10.30	31.18	9.26
	B	7534	11.09	12.38	9.84	9.40	18.68	19.24	16.45	17.11	10.71	22.05	17.93	10.13	16.83	9.11
	P	23215	34.16	38.46	30.04	35.29	23.16	32.94	37.62	33.50	34.36	28.39	24.03	35.58	32.68	34.68
	A	27100	39.88	34.48	45.06	45.88	7.55	14.85	26.22	23.18	41.04	6.03	10.43	43.99	19.32	46.96
Total		67953	100	33288	34665	55100	5868	3716	2273	988	65699	2254	8323	59630	17398	50555

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-54
Cut scores and Associated Impact Data for WKCE-CRT Reading

	Score Range				Impact Data				
Grade	Minimal	Basic	Proficient	Advanced	Minimal	Basic	Proficient	Advanced	Proficient +Advanced
3	270-393	394-429	430-465	466-640	4.89%	14.73%	35.68%	44.71%	80.39%
4	280-395	396-439	440-488	489-650	4.49%	13.29%	40.21%	42.01%	82.22%
5	290-400	401-443	444-496	497-690	3.72%	11.19%	42.97%	42.12%	85.09%
6	300-417	418-456	457-513	514-730	4.53%	9.60%	41.65%	44.22%	85.87%
7	310-433	434-466	467-522	523-780	5.14%	9.13%	39.81%	45.92%	85.73%
8	330-444	445-479	480-538	539-790	5.74%	8.86%	40.91%	44.50%	85.41%
10	350-455	456-502	503-554	555-820	8.62%	15.26%	33.93%	42.19%	76.12%

Table 8-55
Cut scores and Associated Impact Data for WKCE-CRT Mathematics

	Score Range				Impact Data				
Grade	Minimal	Basic	Proficient	Advanced	Minimal	Basic	Proficient	Advanced	Proficient +Advanced
3	220-391	392-406	407-451	452-630	16.02%	9.84%	42.25%	31.89%	74.14%
4	240-420	421-437	438-483	484-650	13.41%	9.69%	42.32%	34.58%	76.90%
5	270-444	445-462	463-504	505-680	14.04%	9.57%	35.00%	41.39%	76.39%
6	310-463	464-484	485-531	532-700	12.73%	11.11%	40.24%	35.92%	76.16%
7	330-479	480-503	504-554	555-710	10.50%	12.26%	45.65%	31.59%	77.24%
8	350-482	483-512	513-572	573-730	11.03%	13.09%	49.47%	26.41%	75.88%
10	410-515	516-540	541-594	595-750	15.85%	13.61%	46.42%	24.13%	70.55%

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-56
Cut scores and Associated Impact Data for WKCE-CRT Language Arts

	Score Range				Impact Data				
Grade	Minimal	Basic	Proficient	Advanced	Minimal	Basic	Proficient	Advanced	Proficient +Advanced
4	140-251	252-276	277-307	308-420	5.14%	18.00%	44.33%	32.54%	76.86%
8	250-357	358-384	385-417	418-520	14.41%	22.26%	36.40%	26.94%	63.34%
10	290-392	393-427	428-483	484-630	8.22%	19.54%	56.09%	16.15%	72.24%

Table 8-57
Cut scores and Associated Impact Data for WKCE-CRT Social Studies

	Score Range				Impact Data				
Grade	Minimal	Basic	Proficient	Advanced	Minimal	Basic	Proficient	Advanced	Proficient +Advanced
4	170-241	242-262	263-287	288-400	1.8%	5.6%	25.6%	67.0%	92.6%
8	230-333	334-363	364-402	403-530	4.3%	13.7%	39.9%	42.1%	82.0%
10	240-407	408-419	420-454	455-620	14.9%	6.7%	31.6%	46.8%	78.4%

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-58
Cut scores and Associated Impact Data for WKCE-CRT Science

	Score Range				Impact Data				
Grade	Minimal	Basic	Proficient	Advanced	Minimal	Basic	Proficient	Advanced	Proficient +Advanced
4	170-248	249-278	279-319	320-440	5.03%	19.29%	56.60%	19.08%	75.68%
8	230-348	349-374	375-418	419-560	9.35%	15.37%	46.31%	28.98%	75.28%
10	240-410	411-428	429-465	466-610	14.87%	11.09%	34.16%	39.88%	74.04%

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-59
Summary Statistics for Reading Content Standards Raw and SPI Scores

Grade	N	Content Standard	Standard	No. of Items		Total Score Points	Mean	Mean *p*-value	SD	SPI	
				MC	CR					Mean	SD
3	58543	1	Determines Meaning	12	0	12	7.91	0.66	2.74	65.39	20.39
	58543	2	Understands Text	17	0	17	12.63	0.74	3.80	74.17	21.24
	58543	3	Analyzes Text	21	1	24	15.62	0.65	4.94	65.45	19.73
	58543	4	Evaluates/Extends Text	4	1	7	3.32	0.47	1.62	48.49	17.42
4	58726	1	Determines Meaning	11	0	11	7.93	0.72	2.47	71.66	20.05
	58726	2	Understands Text	17	0	17	11.81	0.69	3.62	69.30	19.90
	58726	3	Analyzes Text	21	1	24	14.93	0.62	4.74	62.60	18.77
	58726	4	Evaluates/Extends Text	5	1	8	4.52	0.57	1.63	56.45	15.55
5	58399	1	Determines Meaning	11	0	11	7.57	0.69	2.64	68.28	21.61
	58399	2	Understands Text	15	0	15	11.72	0.78	2.97	77.80	18.61
	58399	3	Analyzes Text	20	1	23	13.91	0.60	4.23	61.05	17.02
	58399	4	Evaluates/Extends Text	8	1	11	6.20	0.56	2.24	57.11	16.83
6	59683	1	Determines Meaning	11	0	11	7.60	0.69	2.31	68.35	18.04
	59683	2	Understands Text	14	0	14	8.98	0.64	2.82	63.97	17.58
	59683	3	Analyzes Text	18	1	21	12.96	0.62	3.67	62.14	16.18
	59683	4	Evaluates/Extends Text	11	1	14	8.97	0.64	2.57	64.66	15.90
7	60777	1	Determines Meaning	11	0	11	7.87	0.72	2.17	71.45	16.86
	60777	2	Understands Text	14	0	14	9.28	0.66	2.76	66.30	17.22
	60777	3	Analyzes Text	18	1	21	12.37	0.59	3.89	59.47	17.01
	60777	4	Evaluates/Extends Text	11	1	14	9.02	0.64	2.74	64.69	16.69
8	62782	1	Determines Meaning	11	0	11	8.04	0.73	2.29	72.02	18.11
	62782	2	Understands Text	14	0	14	10.23	0.73	2.77	73.09	17.80
	62782	3	Analyzes Text	18	1	21	13.65	0.65	3.60	65.55	15.65
	62782	4	Evaluates/Extends Text	11	1	14	9.13	0.65	2.89	66.22	18.26

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-59 Cont'd
Summary Statistics for Reading Content Standards Raw and SPI Scores

Grade	N	Content Standard	Standard	No. of Items		Total Score Points	Mean	Mean *p*-value	SD	SPI	
				MC	CR					Mean	SD
10	68161	1	Determines Meaning	7	0	7	5.21	0.74	1.51	73.87	17.47
	68161	2	Understands Text	7	0	7	4.42	0.63	1.84	63.21	21.54
	68161	3	Analyzes Text	22	1	25	14.81	0.59	5.07	59.88	19.24
	68161	4	Evaluates/Extends Text	14	1	17	12.19	0.72	3.52	71.85	19.33

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-60
Summary Statistics for Mathematics Content Standards Raw and SPI Scores

Grade	N	Content Standard	Standard	No. of Items		Total Score Points	Mean	Mean *p*-value	SD	SPI	
				MC	CR					Mean	SD
3	59019	A	Mathematical Processes	3	3	9	3.40	0.38	2.24	39.25	22.14
	59019	B	Number Operations	11	1	12	9.01	0.75	2.49	75.00	18.53
	59019	C	Geometry	9	1	10	8.04	0.80	1.71	80.08	13.94
	59019	D	Measurement	8	0	8	5.99	0.75	1.62	74.03	15.44
	59019	E	Statistics/Probability	7	1	8	5.65	0.71	1.86	69.97	19.62
	59019	F	Algebraic Relationships	8	1	10	7.63	0.76	1.93	77.27	16.01
4	59115	A	Mathematical Processes	3	3	9	4.18	0.46	2.04	46.84	19.15
	59115	B	Number Operations	11	0	11	9.10	0.83	1.88	82.14	14.52
	59115	C	Geometry	8	1	10	7.53	0.75	1.82	75.44	13.44
	59115	D	Measurement	8	1	9	6.77	0.75	1.75	74.94	16.37
	59115	E	Statistics/Probability	7	1	8	5.07	0.63	1.70	62.19	16.69
	59115	F	Algebraic Relationships	9	1	10	7.64	0.76	2.07	76.90	18.53
5	58678	A	Mathematical Processes	3	3	9	5.04	0.56	2.03	57.22	18.29
	58678	B	Number Operations	11	0	11	8.59	0.78	1.96	77.46	15.14
	58678	C	Geometry	9	1	10	7.75	0.78	1.62	76.99	11.89
	58678	D	Measurement	9	1	11	7.05	0.64	2.50	64.52	19.42
	58678	E	Statistics/Probability	9	1	10	6.12	0.61	2.23	60.52	18.61
	58678	F	Algebraic Relationships	10	1	11	7.30	0.66	2.65	66.85	21.78

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-60 Cont'd
Summary Statistics for Mathematics Content Standards Raw and SPI Scores

Grade	N	Content Standard	Standard	No. of Items		Total Score Points	Mean	Mean *p*-value	SD	SPI	
				MC	CR					Mean	SD
6	59907	A	Mathematical Processes	3	3	9	4.13	0.46	2.14	46.55	20.16
	59907	B	Number Operations	12	0	12	7.71	0.64	2.60	63.77	18.87
	59907	C	Geometry	9	1	10	7.64	0.76	1.96	76.10	15.71
	59907	D	Measurement	9	1	11	6.70	0.61	2.41	61.61	19.42
	59907	E	Statistics/Probability	8	1	9	5.90	0.66	2.19	64.79	20.11
	59907	F	Algebraic Relationships	10	1	11	7.86	0.71	2.28	72.29	17.94
7	60967	A	Mathematical Processes	3	3	9	4.26	0.47	2.18	47.86	20.50
	60967	B	Number Operations	12	0	12	8.18	0.68	2.64	67.86	19.56
	60967	C	Geometry	10	2	12	8.01	0.67	2.52	67.26	17.64
	60967	D	Measurement	9	0	9	6.57	0.73	1.77	72.44	16.32
	60967	E	Statistics/Probability	8	1	10	6.50	0.65	2.05	64.43	17.87
	60967	F	Algebraic Relationships	9	1	10	6.48	0.65	2.28	65.36	18.93
8	62914	A	Mathematical Processes	3	3	9	4.15	0.46	2.31	47.74	22.12
	62914	B	Number Operations	7	0	7	3.74	0.53	1.91	53.95	22.24
	62914	C	Geometry	8	1	9	6.00	0.67	1.96	65.71	17.43
	62914	D	Measurement	11	1	12	5.94	0.50	2.65	50.12	19.20
	62914	E	Statistics/Probability	8	1	9	5.41	0.60	2.04	58.53	18.85
	62914	F	Algebraic Relationships	14	1	16	9.97	0.62	3.30	62.15	18.74
10	68330	A	Mathematical Processes	7	1	9	5.33	0.59	2.37	59.01	23.43
	68330	B	Number Operations	7	0	7	4.61	0.66	1.88	64.91	22.77
	68330	C	Geometry	8	1	10	5.60	0.56	2.54	55.96	22.06
	68330	D	Measurement	9	1	11	5.65	0.51	2.99	52.14	24.56
	68330	E	Statistics/Probability	9	0	9	5.56	0.62	2.07	61.66	19.13
	68330	F	Algebraic Relationships	10	1	12	6.38	0.53	2.91	53.95	21.67

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-61
Summary Statistics for Language Arts Content Standards Raw and SPI Scores

Grade	N	Content Standard	Standard	No. of Items		Total Score Points	Mean	Mean *p*-value	SD	SPI	
				MC	CR					Mean	SD
4	58755	B	Writing	19	0	19	13.74	0.72	3.51	72.44	17.83
	58755	D	Language	5	0	5	2.78	0.56	1.32	54.84	18.12
	58755	F	Research and Inquiry	6	0	6	3.14	0.52	1.49	53.40	18.29
8	62581	B	Writing	18	0	18	14.07	0.78	3.45	78.00	18.60
	62581	D	Language	6	0	6	3.56	0.59	1.68	60.14	23.43
	62581	F	Research and Inquiry	6	0	6	4.02	0.67	1.27	66.24	14.04
10	67837	B	Writing	15	2	24	9.26	0.39	2.83	58.96	14.62
	67837	D	Language	9	0	9	6.46	0.72	2.11	71.93	20.18
	67837	F	Research and Inquiry	6	0	6	3.79	0.63	1.45	62.74	16.83

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-62
Summary Statistics for Social Studies Content Standards Raw and SPI Scores

Grade	N	Content Standard	Standard	No. of Items		Total Score Points	Mean	Mean *p*-value	SD	SPI	
				MC	CR					Mean	SD
4	59022	A	Geography	9	0	9	7.16	0.80	1.85	79.15	18.06
	59022	B	History	8	0	8	6.64	0.83	1.48	82.66	16.01
	59022	C	Political Science	7	0	7	4.64	0.66	1.59	66.60	17.37
	59022	D	Economics	7	0	7	5.33	0.76	1.32	76.35	15.05
	59022	E	Behavioral Science	7	0	7	5.20	0.74	1.67	74.89	19.93
8	62701	A	Geography	10	0	10	8.28	0.83	1.85	82.24	16.49
	62701	B	History	11	0	11	7.18	0.65	2.33	66.10	18.13
	62701	C	Political Science	6	0	6	3.92	0.65	1.53	65.55	19.94
	62701	D	Economics	8	0	8	5.47	0.68	1.81	68.18	18.71
	62701	E	Behavioral Science	5	0	5	3.03	0.61	1.34	61.17	19.80
10	67916	A	Geography	10	0	10	6.00	0.60	1.97	59.05	15.23
	67916	B	History	12	0	12	7.67	0.64	2.65	64.74	19.17
	67916	C	Political Science	12	0	12	7.32	0.61	2.78	61.64	20.53
	67916	D	Economics	8	0	8	5.12	0.64	1.93	64.03	20.12
	67916	E	Behavioral Science	8	0	8	6.13	0.77	1.64	76.25	17.07

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 8-63
Summary Statistics for Science Content Standards Raw and SPI Scores

Grade	N	Content Standard	Standard	No. of Items		Total Score Points	Mean	Mean *p*-value	SD	SPI	
				MC	CR					Mean	SD
4	59057	A	Science Connections	4	0	4	2.64	0.66	0.99	65.96	15.41
	59057	B	Nature of Science	4	0	4	2.95	0.74	1.06	74.40	20.05
	59057	C	Science Inquiry	7	0	7	4.86	0.69	1.56	70.03	18.17
	59057	D	Physical Science	6	0	6	4.49	0.75	0.98	75.76	10.23
	59057	E	Earth and Space	6	0	6	4.20	0.70	1.29	68.64	14.44
	59057	F	Life and Environment	6	0	6	4.38	0.73	1.39	71.41	18.25
	59057	G	Science Applications	4	0	4	3.15	0.79	1.06	79.30	20.83
	59057	H	Personal/Social Perspectives	3	0	3	2.07	0.69	0.86	69.32	19.98
8	62721	A	Science Connections	3	0	3	2.25	0.75	0.88	74.52	20.84
	62721	B	Nature of Science	4	0	4	2.70	0.67	1.10	67.40	19.38
	62721	C	Science Inquiry	8	0	8	6.39	0.80	1.52	79.97	15.63
	62721	D	Physical Science	6	0	6	4.50	0.75	1.17	74.93	14.19
	62721	E	Earth and Space	6	0	6	3.88	0.65	1.45	65.14	18.26
	62721	F	Life and Environment	6	0	6	4.16	0.69	1.40	69.42	17.14
	62721	G	Science Applications	4	0	4	3.46	0.86	0.82	85.83	15.70
	62721	H	Personal/Social Perspectives	3	0	3	2.25	0.75	0.84	74.85	19.59
10	67953	A	Science Connections	5	0	5	3.07	0.61	1.34	60.97	19.54
	67953	B	Nature of Science	5	0	5	3.41	0.68	1.25	68.71	19.64
	67953	C	Science Inquiry	9	0	9	5.40	0.60	1.95	60.39	17.17
	67953	D	Physical Science	8	0	8	5.00	0.63	1.86	61.86	18.36
	67953	E	Earth and Space	6	0	6	3.54	0.59	1.30	58.12	13.85
	67953	F	Life and Environment	6	0	6	3.34	0.56	1.29	57.76	15.26
	67953	G	Science Applications	5	0	5	2.96	0.59	1.43	58.94	21.98
	67953	H	Personal/Social Perspectives	6	0	6	4.07	0.68	1.54	68.12	20.74

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-1
Reliability for Total Group and Subgroups Using Cronbach's Alpha

Content	Grade	Total	Gender		Ethnicity					ELP		Disability		SES	
			Female	Male	White	African American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
Reading	3	0.93	0.93	0.93	0.93	0.92	0.91	0.93	0.92	0.93	0.90	0.93	0.92	0.92	0.92
	4	0.92	0.91	0.92	0.91	0.91	0.91	0.92	0.91	0.92	0.88	0.92	0.91	0.91	0.90
	5	0.91	0.91	0.92	0.90	0.90	0.90	0.92	0.90	0.91	0.88	0.91	0.90	0.91	0.90
	6	0.90	0.89	0.90	0.88	0.88	0.88	0.91	0.89	0.89	0.85	0.89	0.87	0.89	0.87
	7	0.90	0.89	0.91	0.88	0.89	0.89	0.90	0.89	0.90	0.86	0.89	0.88	0.89	0.88
	8	0.90	0.90	0.91	0.89	0.89	0.90	0.90	0.89	0.90	0.87	0.88	0.88	0.90	0.88
	10	0.92	0.92	0.92	0.91	0.91	0.91	0.91	0.91	0.92	0.87	0.89	0.90	0.91	0.91
Mathematics	3	0.90	0.90	0.90	0.89	0.91	0.88	0.89	0.88	0.90	0.88	0.91	0.90	0.90	0.88
	4	0.89	0.89	0.89	0.87	0.90	0.88	0.89	0.87	0.89	0.87	0.90	0.88	0.89	0.87
	5	0.90	0.90	0.91	0.89	0.89	0.89	0.90	0.88	0.90	0.88	0.91	0.89	0.90	0.89
	6	0.92	0.91	0.92	0.90	0.90	0.89	0.92	0.89	0.91	0.88	0.90	0.90	0.90	0.90
	7	0.92	0.91	0.92	0.91	0.89	0.89	0.92	0.90	0.92	0.88	0.90	0.91	0.90	0.90
	8	0.92	0.91	0.92	0.91	0.89	0.89	0.92	0.90	0.92	0.87	0.88	0.91	0.90	0.91
	10	0.93	0.93	0.94	0.93	0.88	0.90	0.93	0.92	0.93	0.86	0.87	0.93	0.91	0.93
Language Arts	4	0.82	0.81	0.82	0.81	0.80	0.78	0.83	0.80	0.82	0.75	0.79	0.81	0.80	0.80
	8	0.85	0.83	0.85	0.84	0.83	0.82	0.83	0.83	0.85	0.78	0.81	0.82	0.83	0.83
	10	0.82	0.81	0.83	0.80	0.80	0.81	0.80	0.80	0.82	0.74	0.77	0.79	0.82	0.80
Social Studies	4	0.87	0.87	0.88	0.85	0.88	0.85	0.87	0.86	0.87	0.84	0.87	0.86	0.87	0.85
	8	0.87	0.87	0.88	0.86	0.84	0.84	0.87	0.85	0.87	0.80	0.84	0.86	0.86	0.86
	10	0.90	0.89	0.91	0.89	0.87	0.88	0.88	0.88	0.90	0.82	0.86	0.89	0.88	0.89
Science	4	0.85	0.85	0.86	0.83	0.85	0.83	0.84	0.83	0.85	0.79	0.86	0.84	0.84	0.83
	8	0.86	0.85	0.88	0.84	0.84	0.83	0.84	0.84	0.86	0.79	0.86	0.84	0.85	0.83
	10	0.88	0.86	0.90	0.86	0.83	0.84	0.86	0.86	0.88	0.78	0.85	0.87	0.87	0.87

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-2
Standard Error of Measurement for Total Group and Subgroups

Content	Grade	Total	Gender		Ethnicity					ELP		Disability		SES	
			Female	Male	White	African American	Hispanic	Asian	American Indian	Proficient	Not Proficient	Disabled	Not Disabled	Economically Disadvantaged	Not Economically Disadvantaged
Reading	3	3.13	3.00	3.23	2.92	3.35	3.42	3.13	3.30	3.09	3.42	3.35	3.12	3.38	3.00
	4	3.15	3.23	3.24	3.08	3.41	3.29	3.23	3.29	3.11	3.51	3.43	3.04	3.41	3.11
	5	3.23	3.12	3.12	3.11	3.50	3.41	3.23	3.23	3.17	3.50	3.45	3.06	3.30	3.01
	6	3.15	3.18	3.24	3.14	3.47	3.42	3.28	3.27	3.24	3.49	3.39	3.20	3.35	3.14
	7	3.21	3.22	3.15	3.20	3.43	3.40	3.37	3.29	3.14	3.54	3.46	3.14	3.43	3.09
	8	3.22	3.07	3.14	3.05	3.54	3.32	3.32	3.39	3.16	3.49	3.53	3.08	3.35	3.10
	10	3.04	2.92	3.12	2.96	3.19	3.28	3.28	3.13	2.99	3.31	3.32	3.07	3.28	2.94
Mathematics	3	3.02	3.01	3.04	2.90	3.08	3.18	3.02	3.13	3.01	3.12	3.10	2.91	3.08	2.97
	4	2.96	2.95	2.97	2.91	3.09	3.05	2.97	3.07	2.94	3.13	3.20	2.89	3.06	2.84
	5	3.34	3.28	3.22	3.24	3.50	3.35	3.29	3.34	3.31	3.46	3.36	3.25	3.34	3.16
	6	3.17	3.30	3.23	3.29	3.35	3.47	3.18	3.46	3.35	3.52	3.48	3.28	3.48	3.20
	7	3.19	3.31	3.25	3.11	3.41	3.46	3.24	3.37	3.16	3.46	3.34	3.12	3.43	3.23
	8	3.33	3.46	3.39	3.32	3.34	3.43	3.33	3.42	3.32	3.53	3.44	3.32	3.48	3.29
	10	3.38	3.31	3.18	3.21	3.23	3.30	3.35	3.24	3.36	3.39	3.33	3.20	3.41	3.23
Language Arts	4	2.24	2.23	2.29	2.19	2.37	2.38	2.24	2.33	2.23	2.39	2.39	2.22	2.33	2.20
	8	2.13	2.09	2.22	2.07	2.35	2.35	2.23	2.29	2.11	2.40	2.43	2.10	2.33	2.06
	10	2.34	2.24	2.35	2.26	2.49	2.41	2.39	2.47	2.30	2.50	2.53	2.30	2.41	2.25
Social Studies	4	2.31	2.26	2.26	2.22	2.54	2.51	2.37	2.43	2.28	2.49	2.50	2.28	2.45	2.14
	8	2.58	2.51	2.54	2.46	2.81	2.78	2.61	2.72	2.55	2.81	2.85	2.47	2.69	2.41
	10	2.91	2.89	2.90	2.86	3.12	3.04	3.07	3.08	2.88	3.14	3.17	2.84	3.11	2.85
Science	4	2.49	2.44	2.45	2.39	2.67	2.62	2.56	2.62	2.46	2.74	2.62	2.47	2.66	2.33
	8	2.46	2.43	2.37	2.36	2.72	2.70	2.61	2.62	2.43	2.75	2.71	2.39	2.67	2.37
	10	3.04	3.08	2.91	3.04	3.17	3.22	3.15	3.17	3.01	3.26	3.18	2.97	3.11	2.94

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-3
Cronbach Alpha Reliability Coefficients for Content Standards

Content Area	Grade	Content Standard								
		A/1 Alpha	B/2 Alpha	C/3 Alpha	D/4 Alpha	E Alpha	F Alpha	G Alpha	H Alpha	Total Alpha
Reading	3	0.73	0.84	0.85	0.52					0.93
	4	0.73	0.79	0.81	0.45					0.92
	5	0.74	0.78	0.77	0.58					0.91
	6	0.66	0.69	0.73	0.66					0.90
	7	0.65	0.69	0.74	0.67					0.90
	8	0.67	0.74	0.72	0.69					0.90
	10	0.56	0.65	0.82	0.81					0.92
Mathematics	3	0.62	0.73	0.58	0.56	0.65	0.58			0.90
	4	0.54	0.67	0.52	0.59	0.58	0.69			0.89
	5	0.51	0.63	0.48	0.66	0.64	0.75			0.90
	6	0.56	0.68	0.63	0.67	0.70	0.69			0.92
	7	0.60	0.71	0.69	0.62	0.64	0.66			0.92
	8	0.64	0.63	0.61	0.67	0.62	0.76			0.92
	10	0.72	0.68	0.72	0.76	0.62	0.74			0.93
Language Arts	4		0.77		0.44		0.46			0.82
	8		0.80		0.61		0.31			0.85
	10		0.66		0.68		0.45			0.84
Social Studies	4	0.67	0.61	0.48	0.50	0.64				0.87
	8	0.68	0.64	0.54	0.60	0.47				0.87
	10	0.56	0.70	0.72	0.63	0.60				0.90
Science	4	0.32	0.48	0.55	0.23	0.38	0.52	0.56	0.37	0.85
	8	0.40	0.41	0.57	0.40	0.46	0.47	0.44	0.38	0.86
	10	0.47	0.48	0.55	0.57	0.33	0.36	0.55	0.59	0.88

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-4
Standard Error of Measurement per Content Standard

Content Area	Grade	SEM Per Content Standard								
		A/1	B/2	C/3	D/4	E	F	G	H	Total
Reading	3	1.42	1.52	1.91	1.12					3.07
	4	1.28	1.66	2.07	1.21					3.17
	5	1.35	1.39	2.03	1.45					3.16
	6	1.35	1.57	1.91	1.50					3.19
	7	1.28	1.54	1.98	1.57					3.22
	8	1.32	1.41	1.90	1.61					3.17
	10	1.00	1.09	2.15	1.53					3.04
Mathematics	3	1.38	1.29	1.11	1.07	1.10	1.25			2.99
	4	1.38	1.08	1.26	1.12	1.10	1.15			2.94
	5	1.42	1.19	1.17	1.46	1.34	1.33			3.26
	6	1.42	1.47	1.19	1.38	1.20	1.27			3.27
	7	1.38	1.42	1.40	1.09	1.23	1.33			3.24
	8	1.39	1.16	1.22	1.52	1.26	1.62			3.39
	10	1.25	1.06	1.34	1.46	1.28	1.48			3.26
Language Arts	4		1.68		0.99		1.09			2.25
	8		1.54		1.05		1.05			2.15
	10		1.65		1.19		1.08			2.52
Social Studies	4	1.06	0.92	1.15	0.93	1.00				2.28
	8	1.05	1.40	1.04	1.14	0.98				2.54
	10	1.31	1.45	1.47	1.17	1.04				2.91
Science	4	0.82	0.76	1.05	0.86	1.02	0.96	0.70	0.68	2.46
	8	0.68	0.84	1.00	0.91	1.07	1.02	0.61	0.66	2.45
	10	0.98	0.90	1.31	1.22	1.06	1.03	0.96	0.99	3.03

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-5
Classification Consistency and Classification Accuracy for Reading Grade 3

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.04	0.02	0.00	0.00	0.06
Basic Proficient	0.02	0.10	0.04	0.00	0.16
Proficient	0.00	0.03	0.26	0.05	0.34
Advanced	0.00	0.00	0.05	0.40	0.44
Sum	0.06	0.15	0.34	0.45	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.97	0.94	0.90	0.80
Probability of Chance	0.89	0.67	0.51	0.34
Kappa (k)	0.70	0.81	0.80	0.70
Classification Accuracy	0.98	0.96	0.93	0.86

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-6
Classification Consistency and Classification Accuracy for Reading Grade 4

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.04	0.02	0.00	0.00	0.05
Basic Proficient	0.01	0.09	0.03	0.00	0.13
Proficient	0.00	0.03	0.30	0.05	0.38
Advanced	0.00	0.00	0.05	0.38	0.44
Sum	0.05	0.13	0.39	0.43	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.97	0.94	0.90	0.81
Probability of Chance	0.90	0.70	0.51	0.35
Kappa (k)	0.69	0.81	0.79	0.70
Classification Accuracy	0.98	0.96	0.92	0.86

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-7
Classification Consistency and Classification Accuracy for Reading Grade 5

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.03	0.01	0.00	0.00	0.05
Basic Proficient	0.01	0.07	0.03	0.00	0.11
Proficient	0.00	0.03	0.30	0.06	0.38
Advanced	0.00	0.00	0.05	0.41	0.46
Sum	0.05	0.11	0.37	0.47	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.97	0.94	0.89	0.81
Probability of Chance	0.91	0.73	0.50	0.37
Kappa (k)	0.71	0.79	0.78	0.70
Classification Accuracy	0.98	0.96	0.92	0.87

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-8
Classification Consistency and Classification Accuracy for Reading Grade 6

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.04	0.02	0.00	0.00	0.06
Basic Proficient	0.01	0.06	0.03	0.00	0.10
Proficient	0.00	0.03	0.29	0.07	0.38
Advanced	0.00	0.00	0.06	0.40	0.46
Sum	0.05	0.10	0.38	0.47	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.97	0.94	0.88	0.78
Probability of Chance	0.90	0.74	0.50	0.37
Kappa (k)	0.69	0.77	0.75	0.66
Classification Accuracy	0.98	0.96	0.91	0.85

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-9
Classification Consistency and Classification Accuracy for Reading Grade 7

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.05	0.01	0.00	0.00	0.06
Basic Proficient	0.01	0.05	0.03	0.00	0.09
Proficient	0.00	0.03	0.28	0.06	0.37
Advanced	0.00	0.00	0.06	0.41	0.47
Sum	0.06	0.10	0.37	0.47	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.97	0.94	0.88	0.79
Probability of Chance	0.88	0.74	0.50	0.37
Kappa (k)	0.74	0.77	0.76	0.67
Classification Accuracy	0.98	0.96	0.92	0.85

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-10
Classification Consistency and Classification Accuracy for Reading Grade 8

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.05	0.02	0.00	0.00	0.07
Basic Proficient	0.02	0.05	0.03	0.00	0.10
Proficient	0.00	0.03	0.30	0.06	0.39
Advanced	0.00	0.00	0.05	0.39	0.44
Sum	0.07	0.10	0.39	0.45	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.96	0.94	0.89	0.79
Probability of Chance	0.87	0.72	0.51	0.36
Kappa (k)	0.72	0.78	0.77	0.68
Classification Accuracy	0.97	0.95	0.92	0.85

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-11
Classification Consistency and Classification Accuracy for Reading Grade 10

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.08	0.02	0.00	0.00	0.11
Basic Proficient	0.02	0.10	0.04	0.00	0.16
Proficient	0.00	0.03	0.21	0.05	0.30
Advanced	0.00	0.00	0.05	0.38	0.43
Sum	0.11	0.16	0.31	0.43	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.96	0.93	0.90	0.78
Probability of Chance	0.81	0.61	0.51	0.31
Kappa (k)	0.77	0.81	0.79	0.68
Classification Accuracy	0.97	0.95	0.92	0.84

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-12
Classification Consistency and Classification Accuracy for Mathematics Grade 3

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.13	0.02	0.01	0.00	0.17
Basic Proficient	0.03	0.04	0.04	0.00	0.10
Proficient	0.01	0.04	0.28	0.06	0.39
Advanced	0.00	0.00	0.06	0.29	0.34
Sum	0.17	0.10	0.39	0.35	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.93	0.91	0.88	0.74
Probability of Chance	0.72	0.61	0.55	0.31
Kappa (k)	0.75	0.76	0.74	0.62
Classification Accuracy	0.95	0.93	0.92	0.81

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-13
Classification Consistency and Classification Accuracy for Mathematics Grade 4

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.10	0.02	0.01	0.00	0.14
Basic Proficient	0.02	0.04	0.03	0.00	0.09
Proficient	0.01	0.04	0.28	0.07	0.39
Advanced	0.00	0.00	0.06	0.32	0.39
Sum	0.13	0.10	0.38	0.39	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.94	0.91	0.87	0.74
Probability of Chance	0.77	0.65	0.52	0.33
Kappa (k)	0.73	0.76	0.72	0.61
Classification Accuracy	0.95	0.93	0.91	0.80

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-14
Classification Consistency and Classification Accuracy for Mathematics Grade 5

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.11	0.02	0.01	0.00	0.14
Basic Proficient	0.02	0.03	0.04	0.00	0.09
Proficient	0.01	0.04	0.22	0.05	0.31
Advanced	0.00	0.00	0.06	0.40	0.46
Sum	0.14	0.09	0.32	0.46	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.94	0.91	0.89	0.76
Probability of Chance	0.77	0.65	0.50	0.34
Kappa (k)	0.74	0.75	0.78	0.63
Classification Accuracy	0.96	0.94	0.92	0.83

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-15
Classification Consistency and Classification Accuracy for Mathematics Grade 6

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.09	0.03	0.01	0.00	0.12
Basic Proficient	0.02	0.04	0.03	0.00	0.10
Proficient	0.01	0.03	0.28	0.05	0.37
Advanced	0.00	0.00	0.05	0.35	0.40
Sum	0.12	0.11	0.38	0.40	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.94	0.92	0.89	0.76
Probability of Chance	0.79	0.65	0.52	0.33
Kappa (k)	0.71	0.77	0.78	0.65
Classification Accuracy	0.96	0.94	0.92	0.83

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-16
Classification Consistency and Classification Accuracy for Mathematics Grade 7

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.08	0.02	0.00	0.00	0.11
Basic Proficient	0.02	0.06	0.03	0.00	0.12
Proficient	0.00	0.03	0.33	0.05	0.42
Advanced	0.00	0.00	0.05	0.30	0.35
Sum	0.11	0.12	0.42	0.35	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.95	0.93	0.90	0.78
Probability of Chance	0.81	0.65	0.54	0.33
Kappa (k)	0.74	0.79	0.78	0.67
Classification Accuracy	0.96	0.95	0.93	0.84

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-17
Classification Consistency and Classification Accuracy for Mathematics Grade 8

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.08	0.03	0.00	0.00	0.11
Basic Proficient	0.03	0.07	0.04	0.00	0.13
Proficient	0.00	0.04	0.38	0.04	0.47
Advanced	0.00	0.00	0.04	0.24	0.29
Sum	0.11	0.13	0.47	0.29	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.94	0.92	0.91	0.78
Probability of Chance	0.80	0.63	0.59	0.33
Kappa (k)	0.71	0.78	0.78	0.67
Classification Accuracy	0.96	0.94	0.94	0.84

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-18
Classification Consistency and Classification Accuracy for Mathematics Grade 10

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.14	0.03	0.00	0.00	0.17
Basic Proficient	0.03	0.07	0.04	0.00	0.14
Proficient	0.00	0.04	0.34	0.04	0.42
Advanced	0.00	0.00	0.03	0.24	0.27
Sum	0.18	0.14	0.42	0.27	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.93	0.92	0.93	0.78
Probability of Chance	0.71	0.57	0.60	0.30
Kappa (k)	0.75	0.81	0.82	0.69
Classification Accuracy	0.95	0.94	0.95	0.84

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-19
Classification Consistency and Classification Accuracy for Language Arts Grade 4

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.04	0.02	0.00	0.00	0.07
Basic Proficient	0.02	0.09	0.05	0.00	0.16
Proficient	0.00	0.05	0.27	0.08	0.41
Advanced	0.00	0.00	0.07	0.29	0.37
Sum	0.07	0.16	0.40	0.37	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.95	0.89	0.85	0.69
Probability of Chance	0.87	0.65	0.53	0.33
Kappa (k)	0.60	0.69	0.67	0.54
Classification Accuracy	0.96	0.92	0.89	0.77

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-20
Classification Consistency and Classification Accuracy for Language Arts Grade 8

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.13	0.03	0.00	0.00	0.17
Basic Proficient	0.03	0.12	0.05	0.00	0.21
Proficient	0.00	0.06	0.19	0.07	0.32
Advanced	0.00	0.00	0.06	0.24	0.30
Sum	0.17	0.22	0.31	0.31	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.92	0.87	0.87	0.68
Probability of Chance	0.72	0.53	0.58	0.27
Kappa (k)	0.73	0.73	0.70	0.56
Classification Accuracy	0.94	0.91	0.91	0.76

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-21
Classification Consistency and Classification Accuracy for Language Arts Grade 10

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.08	0.02	0.00	0.00	0.10
Basic Proficient	0.02	0.11	0.06	0.00	0.20
Proficient	0.00	0.05	0.38	0.06	0.49
Advanced	0.00	0.00	0.06	0.15	0.21
Sum	0.10	0.19	0.50	0.21	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.95	0.89	0.87	0.72
Probability of Chance	0.82	0.58	0.67	0.34
Kappa (k)	0.73	0.74	0.62	0.57
Classification Accuracy	0.96	0.92	0.91	0.79

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-22
Classification Consistency and Classification Accuracy for Social Studies Grade 4

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.01	0.01	0.00	0.00	0.02
Basic Proficient	0.01	0.03	0.02	0.00	0.06
Proficient	0.00	0.02	0.17	0.07	0.26
Advanced	0.00	0.00	0.06	0.60	0.66
Sum	0.02	0.06	0.25	0.67	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.98	0.96	0.88	0.82
Probability of Chance	0.95	0.85	0.55	0.51
Kappa (k)	0.63	0.71	0.72	0.63
Classification Accuracy	0.99	0.97	0.91	0.87

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-23
Classification Consistency and Classification Accuracy for Social Studies Grade 8

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.04	0.02	0.00	0.00	0.06
Basic Proficient	0.02	0.08	0.04	0.00	0.14
Proficient	0.00	0.04	0.27	0.06	0.38
Advanced	0.00	0.00	0.07	0.35	0.42
Sum	0.06	0.15	0.38	0.41	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.96	0.91	0.87	0.75
Probability of Chance	0.89	0.67	0.51	0.34
Kappa (k)	0.64	0.73	0.74	0.62
Classification Accuracy	0.97	0.94	0.91	0.82

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-24
Classification Consistency and Classification Accuracy for Social Studies Grade 10

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.13	0.02	0.01	0.00	0.17
Basic Proficient	0.02	0.02	0.03	0.00	0.07
Proficient	0.01	0.03	0.19	0.05	0.29
Advanced	0.00	0.00	0.05	0.42	0.48
Sum	0.17	0.07	0.29	0.48	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.93	0.91	0.89	0.76
Probability of Chance	0.72	0.64	0.50	0.34
Kappa (k)	0.75	0.76	0.79	0.64
Classification Accuracy	0.95	0.94	0.92	0.83

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-25
Classification Consistency and Classification Accuracy for Science Grade 4

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.04	0.02	0.00	0.00	0.06
Basic Proficient	0.02	0.12	0.05	0.00	0.19
Proficient	0.00	0.05	0.42	0.05	0.52
Advanced	0.00	0.00	0.06	0.17	0.23
Sum	0.06	0.19	0.53	0.22	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.96	0.90	0.89	0.75
Probability of Chance	0.89	0.62	0.65	0.37
Kappa (k)	0.64	0.73	0.68	0.61
Classification Accuracy	0.97	0.93	0.92	0.82

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-26
Classification Consistency and Classification Accuracy for Science Grade 8

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.08	0.03	0.00	0.00	0.12
Basic Proficient	0.03	0.08	0.05	0.00	0.16
Proficient	0.00	0.05	0.31	0.06	0.43
Advanced	0.00	0.00	0.06	0.23	0.30
Sum	0.12	0.16	0.43	0.30	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.93	0.89	0.87	0.71
Probability of Chance	0.80	0.60	0.58	0.31
Kappa (k)	0.68	0.73	0.70	0.58
Classification Accuracy	0.95	0.92	0.91	0.78

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-27
Classification Consistency and Classification Accuracy for Science Grade 10

Contingency Table with All Cut Scores

	Minimal Performance	Basic Proficient	Proficient	Advanced	Sum
Minimal Performance	0.14	0.03	0.01	0.00	0.18
Basic Proficient	0.03	0.04	0.04	0.00	0.11
Proficient	0.01	0.04	0.21	0.06	0.32
Advanced	0.00	0.00	0.06	0.32	0.39
Sum	0.18	0.11	0.32	0.38	

Indexes for Classification Consistency and Classification Accuracy

	Cut 1	Cut 2	Cut 3	All cuts
Classification Consistency (P)	0.91	0.89	0.88	0.71
Probability of Chance	0.70	0.59	0.53	0.30
Kappa (k)	0.71	0.74	0.74	0.59
Classification Accuracy	0.94	0.92	0.91	0.79

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-28
Inter-Rater Reliability Reading*

				Percentage Absolute Difference								Frequency			
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra. Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2	3
3	A	19	3	66	20	1	14	0.92	0.84	0.72	1,632	801	495	333	3
3	B	19	3	66	23	2	9	0.90	0.80	0.75	1,510	686	518	300	6
3	C	19	3	68	19	2	10	0.91	0.83	0.77	1,490	698	446	343	3
3	D	19	3	66	21	3	10	0.90	0.80	0.72	1,454	690	486	272	6
3	A	31	3	72	20	2	7	0.89	0.79	1.16	1,632	312	750	561	9
3	B	31	3	71	22	2	5	0.88	0.75	1.17	1,510	267	739	491	13
3	C	31	3	72	23	2	3	0.86	0.72	1.24	1,490	233	683	557	17
3	D	31	3	73	22	2	4	0.87	0.75	1.22	1,454	232	676	539	7
3	A	63	3	76	15	2	7	0.91	0.81	0.71	1,132	507	462	147	16
3	B	67	3	68	22	4	7	0.89	0.78	0.85	1,122	501	301	309	11
3	C	64	3	66	24	4	6	0.88	0.76	0.88	1,108	443	406	208	51
3	D	67	3	69	24	2	6	0.92	0.84	0.98	1,106	462	266	316	62
4	A	13	3	65	30	2	3	0.88	0.75	1.68	1,530	138	432	740	220
4	B	13	3	62	33	2	3	0.84	0.68	1.73	1,528	103	427	784	214
4	C	13	3	64	33	1	3	0.87	0.73	1.72	1,528	123	406	769	230
4	D	13	3	64	32	1	3	0.87	0.74	1.73	1,566	128	411	787	240
4	A	47	3	61	31	4	4	0.84	0.69	1.39	1,530	249	524	664	93
4	B	47	3	61	32	2	4	0.85	0.70	1.43	1,528	236	497	703	92
4	C	47	3	59	35	3	4	0.84	0.69	1.45	1,528	238	477	697	116
4	D	47	3	61	32	3	3	0.84	0.68	1.47	1,566	220	484	773	89
4	A	67	3	72	20	1	7	0.93	0.86	0.93	1,116	451	354	251	60
4	B	64	3	65	26	5	5	0.87	0.73	1.35	1,122	245	374	374	129
4	D	62	3	77	16	3	4	0.92	0.84	1.12	1,126	378	299	383	66
5	A	13	3	64	30	2	4	0.86	0.73	1.46	1,550	190	571	672	117
5	B	13	3	63	32	2	3	0.85	0.69	1.57	1,526	137	543	692	154
5	C	13	3	63	33	2	2	0.85	0.71	1.61	1,432	123	489	650	170

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-28 Cont'd
Inter-Rater Reliability Reading*

				Percentage Absolute Difference								Frequency			
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra. Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2	3
5	D	13	3	65	30	1	3	0.86	0.72	1.55	1,488	132	544	672	140
5	A	47	3	65	29	1	4	0.89	0.78	1.38	1,550	297	478	661	114
5	B	47	3	65	31	2	2	0.86	0.72	1.50	1,526	199	493	714	120
5	C	47	3	64	32	2	2	0.87	0.73	1.58	1,432	163	423	698	148
5	D	47	3	65	31	1	3	0.88	0.75	1.52	1,488	191	468	688	141
5	A	67	3	65	29	1	5	0.89	0.78	1.03	1,130	349	445	294	42
5	C	57	3	54	40	4	2	0.75	0.49	0.81	1,132	411	550	147	24
5	C	63	3	65	18	14	3	0.82	0.64	1.31	1,096	355	108	567	66
5	D	63	3	59	31	6	5	0.86	0.73	0.96	1,116	464	316	257	79
6	A	18	3	67	29	1	3	0.83	0.65	1.31	1,554	149	816	542	47
6	B	18	3	65	30	3	2	0.79	0.59	1.31	1,554	158	808	543	45
6	C	18	3	62	33	2	3	0.80	0.59	1.35	1,534	135	793	541	65
6	D	18	3	64	33	2	2	0.80	0.60	1.36	1,572	124	820	564	64
6	A	24	3	71	25	1	4	0.89	0.78	1.03	1,554	425	697	388	44
6	B	24	3	69	28	1	3	0.88	0.76	1.06	1,554	410	686	406	52
6	C	24	3	69	26	1	3	0.88	0.76	1.10	1,534	366	690	430	48
6	D	24	3	71	26	1	2	0.88	0.75	1.07	1,572	381	728	429	34
6	A	67	3	61	30	1	8	0.83	0.65	0.77	1,158	429	572	151	6
6	B	67	3	63	29	4	4	0.88	0.77	1.12	1,172	373	367	355	77
6	C	57	3	63	34	1	3	0.82	0.63	0.92	1,160	327	611	214	8
6	C	65	3	63	29	3	5	0.87	0.74	1.08	1,150	342	424	332	52
6	D	67	3	69	25	1	6	0.88	0.76	1.00	1,162	319	547	277	19
7	A	31	3	61	33	2	5	0.87	0.73	1.43	1,498	222	546	600	130
7	B	31	3	61	34	2	3	0.84	0.68	1.51	1,694	173	618	766	137
7	C	31	3	62	34	1	3	0.84	0.68	1.53	1,514	125	599	651	139
7	D	31	3	62	31	3	4	0.84	0.69	1.49	1,548	167	598	643	140

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-28 Cont'd
Inter-Rater Reliability Reading*

				Percentage Absolute Difference								Frequency			
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra. Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2	3
7	A	50	3	63	30	1	5	0.90	0.80	1.03	1,498	560	423	426	89
7	B	50	3	65	30	2	3	0.89	0.78	1.07	1,694	580	519	488	107
7	C	50	3	68	28	1	3	0.91	0.83	1.11	1,514	506	441	464	103
7	D	50	3	66	29	2	3	0.91	0.82	1.09	1,548	550	424	455	119
7	A	67	3	77	16	1	6	0.94	0.88	0.97	1,170	400	506	163	101
7	B	63	3	74	21	1	4	0.95	0.90	1.60	1,190	311	211	310	358
7	C	67	3	75	19	1	5	0.87	0.75	1.05	1,164	193	763	166	42
7	D	67	3	67	27	1	5	0.88	0.76	1.11	1,172	261	568	297	46
8	A	19	3	61	34	2	4	0.85	0.70	1.39	1,678	218	711	624	125
8	B	19	3	61	34	2	3	0.84	0.68	1.52	1,704	165	646	734	159
8	C	19	3	63	34	1	2	0.85	0.71	1.52	1,642	138	682	649	173
8	D	19	3	61	36	1	2	0.82	0.64	1.54	1,642	117	681	682	162
8	A	56	3	61	31	4	5	0.89	0.77	1.47	1,678	364	394	686	234
8	B	56	3	60	34	3	4	0.88	0.75	1.59	1,704	292	375	784	253
8	C	56	3	60	34	3	4	0.88	0.76	1.57	1,642	288	369	739	246
8	D	56	3	61	34	2	3	0.88	0.76	1.59	1,642	291	359	730	262
8	A	65	3	70	23	1	6	0.92	0.83	1.26	1,166	264	453	334	115
8	B	67	3	71	22	2	6	0.93	0.86	1.17	1,186	402	287	395	102
8	C	63	3	67	29	1	3	0.88	0.75	1.01	1,190	324	593	208	65
8	D	60	3	68	27	3	2	0.89	0.78	1.36	1,178	222	450	372	134
8	D	63	3	67	30	1	2	0.89	0.77	1.50	1,174	174	359	526	115
10	A	7	3	54	33	4	10	0.89	0.77	1.12	7,140	2,444	2,083	1,955	658
10	A	52	3	61	31	2	7	0.88	0.76	1.40	7,140	1,077	2,851	2,479	733

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-29
Inter-Rater Reliability Mathematics*

				Percentage Absolute Difference								Frequency		
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2
3	A	5	2	89	7	1	3	0.97	0.94	0.86	1,632	834	187	611
3	B	5	2	89	7	1	2	0.97	0.94	0.80	1,510	823	171	516
3	C	5	2	90	7	1	2	0.97	0.94	0.91	1,490	723	177	590
3	D	5	2	89	8	1	2	0.97	0.94	0.83	1,454	752	196	506
3	A	20A	1	88	2	5	6	0.98	0.94	0.44	1,632	915	717	0
3	A	20B	2	78	14	2	5	0.91	0.82	0.46	1,632	1,117	285	230
3	A	27A	1	93	2	1	4	0.98	0.95	0.59	1,632	672	960	0
3	A	27B	2	78	13	3	5	0.92	0.84	0.59	1,632	995	317	320
3	B	20A	1	86	2	4	7	0.97	0.94	0.45	1,510	826	684	0
3	B	20B	2	77	16	2	5	0.90	0.80	0.44	1,510	1,044	271	195
3	B	27A	1	93	2	0	5	0.98	0.96	0.63	1,510	561	949	0
3	B	27B	2	76	15	3	6	0.92	0.83	0.59	1,510	918	293	299
3	C	20A	1	91	1	4	4	0.99	0.97	0.50	1,490	743	747	0
3	C	20B	2	77	18	2	4	0.90	0.79	0.49	1,490	985	288	217
3	C	27A	1	94	2	1	3	0.97	0.95	0.63	1,490	553	937	0
3	C	27B	2	76	16	3	5	0.91	0.81	0.66	1,490	850	296	344
3	D	20A	1	87	2	2	8	0.97	0.94	0.45	1,454	795	659	0
3	D	20B	2	80	13	2	5	0.93	0.85	0.48	1,454	989	240	225
3	D	27A	1	91	2	1	6	0.98	0.96	0.60	1,454	581	873	0
3	D	27B	2	76	14	4	6	0.91	0.82	0.63	1,454	854	281	319
3	A	39A	1	91	1	1	7	0.99	0.97	0.30	1,632	1,147	485	0
3	A	39B	2	83	11	2	5	0.93	0.85	0.49	1,632	1,082	300	250
3	B	39A	1	92	1	1	6	0.99	0.98	0.33	1,510	1,013	497	0
3	B	39B	2	84	10	1	4	0.95	0.89	0.55	1,510	954	284	272
3	C	39A	1	93	0	1	6	1.00	0.99	0.36	1,490	949	541	0
3	C	39B	2	86	10	2	3	0.95	0.89	0.61	1,490	888	294	308
3	D	39A	1	91	1	1	7	0.99	0.98	0.32	1,454	993	461	0
3	D	39B	2	84	11	1	4	0.94	0.87	0.53	1,454	930	283	241

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-29 Cont'd
Inter-Rater Reliability Mathematics*

				Percentage Absolute Difference								Frequency		
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2
3	A	51	2	87	6	2	5	0.96	0.92	1.08	686	270	94	322
3	A	57A	1	97	1	0	3	0.99	0.98	0.82	1,134	203	931	0
3	A	57B	2	76	14	4	5	0.91	0.82	0.82	1,134	575	191	368
3	B	51	2	91	3	1	4	0.98	0.95	1.09	670	255	98	317
3	C	60	2	90	3	1	6	0.97	0.95	1.49	1,100	112	339	649
3	D	58A	1	97	1	0	2	0.99	0.97	0.82	1,098	201	897	0
3	D	58B	2	94	3	0	3	0.96	0.92	0.20	1,098	888	200	10
4	A	11A	1	96	2	0	2	0.98	0.95	0.22	1,530	1,200	330	0
4	A	11B	2	76	18	3	4	0.87	0.75	0.55	1,530	849	514	167
4	B	11A	1	94	2	0	4	0.97	0.93	0.23	1,528	1,181	347	0
4	B	11B	2	73	20	1	5	0.88	0.75	0.58	1,528	806	556	166
4	C	11A	1	95	2	0	2	0.96	0.93	0.19	1,528	1,231	297	0
4	C	11B	2	75	20	1	4	0.87	0.74	0.56	1,528	841	525	162
4	D	11A	1	97	1	0	2	0.98	0.96	0.23	1,566	1,205	361	0
4	D	11B	2	72	22	2	3	0.86	0.71	0.63	1,566	779	584	203
4	A	22	2	92	4	1	3	0.98	0.95	1.53	1,530	152	418	960
4	A	25A	1	97	1	0	2	0.99	0.98	0.41	1,530	905	625	0
4	A	25B	2	81	13	2	4	0.91	0.83	0.59	1,530	819	516	195
4	B	22	2	94	2	1	3	0.99	0.97	1.53	1,528	142	442	944
4	B	25A	1	96	1	0	3	0.99	0.98	0.43	1,528	878	650	0
4	B	25B	2	79	15	2	4	0.89	0.78	0.62	1,528	775	563	190
4	C	22	2	94	3	0	2	0.98	0.96	1.53	1,528	162	395	971
4	C	25A	1	97	1	0	2	0.99	0.98	0.45	1,528	843	685	0
4	C	25B	2	80	14	2	4	0.90	0.79	0.67	1,528	742	548	238
4	D	22	2	94	3	0	2	0.98	0.96	1.52	1,566	161	431	974
4	D	25A	1	96	1	0	2	0.99	0.97	0.45	1,566	865	701	0

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-29 Cont'd
Inter-Rater Reliability Mathematics*

				Percentage Absolute Difference								Frequency		
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2
4	D	25B	2	83	11	3	4	0.91	0.82	0.65	1,566	771	569	226
4	A	39A	1	98	0	0	2	1.00	1.00	0.54	1,530	708	822	0
4	A	39B	2	87	8	1	4	0.97	0.93	0.68	1,530	953	120	457
4	B	39A	1	96	1	0	2	0.99	0.97	0.56	1,528	674	854	0
4	B	39B	2	86	10	1	3	0.96	0.91	0.69	1,528	926	147	455
4	C	39A	1	98	1	0	2	0.99	0.99	0.52	1,528	729	799	0
4	C	39B	2	87	9	1	3	0.96	0.93	0.68	1,528	935	149	444
4	D	39A	1	97	1	0	2	0.99	0.97	0.54	1,566	728	838	0
4	D	39B	2	86	9	2	3	0.96	0.92	0.65	1,566	981	159	426
4	A	51	2	89	4	2	5	0.97	0.94	1.08	1,130	485	68	577
4	B	59A	1	97	1	0	2	0.97	0.95	0.88	1,128	134	994	0
4	B	59B	2	69	26	2	3	0.84	0.68	1.05	1,128	249	575	304
5	A	12	2	93	4	1	3	0.99	0.98	1.18	1,550	563	148	839
5	B	12	2	93	3	1	3	0.99	0.98	1.23	1,526	503	164	859
5	C	12	2	94	4	1	1	0.99	0.97	1.27	1,432	446	150	836
5	D	12	2	94	4	1	2	0.99	0.98	1.21	1,488	510	157	821
5	A	20A	1	97	1	0	2	0.98	0.97	0.83	1,550	269	1,281	0
5	A	20B	2	85	9	2	3	0.93	0.87	0.39	1,550	1,113	263	174
5	A	27A	1	96	1	0	3	0.99	0.97	0.55	1,550	704	846	0
5	A	27B	2	78	16	3	4	0.92	0.84	1.29	1,550	444	213	893
5	B	20A	1	97	0	0	2	0.99	0.99	0.84	1,526	252	1,274	0
5	B	20B	2	83	12	2	3	0.91	0.82	0.39	1,526	1,089	282	155
5	B	27A	1	96	1	0	3	0.99	0.98	0.59	1,526	629	897	0
5	B	27B	2	77	17	3	3	0.92	0.83	1.37	1,526	369	223	934
5	C	20A	1	98	0	0	1	0.99	0.99	0.85	1,432	222	1,210	0
5	C	20B	2	87	10	1	2	0.94	0.87	0.40	1,432	1,027	244	161

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-29 Cont'd
Inter-Rater Reliability Mathematics*

				Percentage Absolute Difference								Frequency		
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2
5	C	27A	1	97	1	0	2	1.00	0.98	0.58	1,432	608	824	0
5	C	27B	2	79	16	3	2	0.92	0.83	1.37	1,432	370	170	892
5	D	20A	1	97	1	0	2	0.99	0.97	0.83	1,488	252	1,236	0
5	D	20B	2	85	11	1	3	0.93	0.85	0.39	1,488	1,049	296	143
5	D	27A	1	98	0	0	2	1.00	0.99	0.59	1,488	611	877	0
5	D	27B	2	80	15	3	3	0.92	0.84	1.40	1,488	349	200	939
5	A	46A	1	97	0	0	3	1.00	0.99	0.33	1,550	1,036	514	0
5	A	46B	2	88	6	3	3	0.95	0.90	0.62	1,550	1,026	82	442
5	B	46A	1	97	0	0	2	1.00	0.99	0.36	1,526	979	547	0
5	B	46B	2	87	8	2	3	0.96	0.91	0.69	1,526	953	93	480
5	C	46A	1	98	0	0	1	1.00	0.99	0.40	1,432	866	566	0
5	C	46B	2	90	6	2	2	0.96	0.91	0.74	1,432	869	65	498
5	D	46A	1	97	1	0	2	1.00	0.99	0.36	1,488	956	532	0
5	D	46B	2	89	7	2	2	0.96	0.91	0.71	1,488	926	72	490
5	A	63	2	90	6	1	4	0.97	0.95	0.67	1,124	580	335	209
5	B	64A	1	90	1	0	9	1.00	0.99	0.59	1,112	456	656	0
5	B	64B	2	89	8	0	3	0.94	0.87	0.53	1,112	546	541	25
5	C	64	2	94	2	1	3	0.99	0.98	1.35	1,122	309	116	697
5	D	64A	1	96	1	0	3	0.99	0.98	0.43	1,122	637	485	0
5	D	64B	2	78	17	1	4	0.89	0.79	1.25	1,122	161	523	438
6	A	7A	1	92	2	1	6	0.98	0.97	0.49	1,554	792	762	0
6	A	7B	2	78	11	4	7	0.93	0.86	0.68	1,554	967	112	475
6	B	7A	1	94	1	1	5	0.99	0.98	0.50	1,554	780	774	0
6	B	7B	2	81	12	2	5	0.95	0.89	0.71	1,554	945	119	490
6	C	7A	1	92	1	1	6	0.99	0.97	0.47	1,534	809	725	0
6	C	7B	2	82	8	3	7	0.95	0.89	0.69	1,534	952	100	482

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-29 Cont'd
Inter-Rater Reliability Mathematics*

				Percentage Absolute Difference								Frequency		
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2
6	D	7A	1	94	1	0	5	0.99	0.98	0.48	1,572	812	760	0
6	D	7B	2	83	10	2	5	0.95	0.90	0.70	1,572	966	118	488
6	A	22A	1	97	1	0	3	0.99	0.97	0.83	1,554	259	1,295	0
6	A	22B	2	85	11	0	4	0.91	0.82	0.47	1,554	888	600	66
6	B	22A	1	98	0	0	2	1.00	1.00	0.84	1,554	243	1,311	0
6	B	22B	2	86	10	1	3	0.91	0.83	0.50	1,554	833	665	56
6	C	22A	1	97	1	0	3	0.99	0.98	0.86	1,534	222	1,312	0
6	C	22B	2	85	11	0	3	0.91	0.82	0.46	1,534	886	588	60
6	D	22A	1	97	1	0	2	0.98	0.97	0.87	1,572	212	1,360	0
6	D	22B	2	87	11	1	2	0.91	0.82	0.53	1,572	828	658	86
6	A	39A	1	94	2	0	3	0.98	0.96	0.42	1,554	898	656	0
6	A	39B	2	88	7	2	4	0.96	0.92	0.75	1,554	770	404	380
6	B	39A	1	97	2	0	1	0.98	0.97	0.41	1,554	920	634	0
6	B	39B	2	87	9	1	3	0.96	0.91	0.74	1,554	770	419	365
6	C	39A	1	95	2	0	4	0.98	0.97	0.41	1,534	900	634	0
6	C	39B	2	86	9	1	4	0.96	0.93	0.74	1,534	773	390	371
6	D	39A	1	96	2	0	2	0.98	0.97	0.42	1,572	907	665	0
6	D	39B	2	87	8	2	3	0.94	0.89	0.72	1,572	789	434	349
6	A	51	2	91	5	0	3	0.98	0.95	0.77	1,554	678	551	325
6	B	51	2	91	7	0	2	0.97	0.94	0.82	1,554	613	606	335
6	C	51	2	90	6	1	3	0.97	0.94	0.75	1,534	678	557	299
6	D	51	2	93	6	0	1	0.97	0.94	0.78	1,572	667	583	322
6	A	63A	1	96	1	0	4	1.00	0.99	0.59	1,170	480	690	0
6	A	63B	2	90	6	1	4	0.97	0.95	1.34	1,170	277	220	673
6	B	63A	1	95	0	0	4	1.00	0.99	0.36	1,170	747	423	0
6	B	63B	2	89	5	1	6	0.98	0.96	0.66	1,170	729	105	336
7	A	7A	1	96	1	0	3	1.00	0.99	0.65	1,498	528	970	0

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-29 Cont'd
Inter-Rater Reliability Mathematics*

				Percentage Absolute Difference								Frequency		
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2
7	A	7B	2	79	13	3	5	0.93	0.86	1.07	1,498	598	200	700
7	B	7A	1	96	2	0	2	0.98	0.96	0.70	1,694	509	1,185	0
7	B	7B	2	80	13	3	4	0.93	0.86	1.21	1,694	569	207	918
7	C	7A	1	96	1	0	3	0.99	0.98	0.68	1,514	484	1,030	0
7	C	7B	2	82	13	1	4	0.96	0.91	1.16	1,514	537	192	785
7	D	7A	1	96	1	0	3	0.99	0.97	0.67	1,548	510	1,038	0
7	D	7B	2	80	13	3	4	0.93	0.87	1.18	1,548	553	168	827
7	A	19A	1	94	1	0	4	0.99	0.98	0.47	1,498	794	704	0
7	A	19B	2	85	8	1	7	0.95	0.90	0.47	1,498	932	435	131
7	B	19A	1	95	2	0	3	0.98	0.97	0.49	1,694	869	825	0
7	B	19B	2	86	8	2	5	0.94	0.88	0.50	1,694	979	576	139
7	C	19A	1	95	2	0	3	0.98	0.96	0.53	1,514	712	802	0
7	C	19B	2	86	9	1	5	0.94	0.88	0.55	1,514	831	534	149
7	D	19A	1	93	2	1	4	0.98	0.96	0.49	1,548	798	750	0
7	D	19B	2	86	7	1	6	0.93	0.86	0.51	1,548	896	523	129
7	A	38A	1	87	2	1	10	0.98	0.96	0.51	1,498	740	758	0
7	A	38B	2	87	4	1	8	0.97	0.95	0.60	1,498	714	670	114
7	B	38A	1	87	2	2	8	0.98	0.96	0.52	1,694	815	879	0
7	B	38B	2	88	3	1	8	0.98	0.95	0.62	1,694	773	789	132
7	C	38A	1	87	3	1	8	0.98	0.95	0.54	1,514	690	824	0
7	C	38B	2	89	4	1	6	0.98	0.96	0.67	1,514	645	728	141
7	D	38A	1	86	2	2	10	0.98	0.96	0.52	1,548	745	803	0
7	D	38B	2	87	6	1	7	0.96	0.93	0.65	1,548	662	761	125
7	A	52	2	95	1	0	3	0.99	0.99	0.68	1,498	761	461	276
7	B	52	2	95	2	0	3	0.99	0.98	0.68	1,694	836	567	291
7	C	52	2	95	2	0	3	0.99	0.98	0.71	1,514	726	508	280

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-29 Cont'd
Inter-Rater Reliability Mathematics*

				Percentage Absolute Difference								Frequency		
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2
7	D	52	2	95	2	0	3	0.99	0.98	0.68	1,548	757	531	260
7	A	61	2	73	17	4	6	0.89	0.77	0.67	670	376	137	157
7	B	61A	1	89	2	1	8	0.98	0.96	0.35	1,172	761	411	0
7	B	61B	2	77	15	1	7	0.91	0.82	0.57	1,172	624	432	116
7	C	62	2	88	4	0	8	0.98	0.96	0.41	1,174	856	154	164
7	D	61	2	68	19	5	8	0.88	0.75	0.77	652	322	156	174
8	A	10	2	85	9	1	5	0.94	0.88	0.34	1,678	1,274	234	170
8	B	10	2	83	12	1	4	0.93	0.86	0.41	1,704	1,225	265	214
8	C	10	2	84	10	1	4	0.93	0.86	0.41	1,642	1,181	252	209
8	D	10	2	83	12	1	3	0.92	0.84	0.42	1,642	1,176	247	219
8	A	19A	1	94	2	1	4	0.98	0.96	0.61	1,678	649	1,029	0
8	A	19B	2	77	14	1	8	0.95	0.91	0.92	1,678	741	324	613
8	B	19A	1	95	1	1	3	0.99	0.97	0.68	1,704	543	1,161	0
8	B	19B	2	80	12	2	6	0.96	0.91	1.05	1,704	663	298	743
8	C	19A	1	95	1	0	3	0.99	0.97	0.66	1,642	562	1,080	0
8	C	19B	2	81	12	1	5	0.96	0.91	1.08	1,642	622	269	751
8	D	19A	1	96	1	0	2	0.99	0.97	0.65	1,642	579	1,063	0
8	D	19B	2	81	13	1	5	0.95	0.90	1.04	1,642	649	280	713
8	A	33A	1	78	12	1	9	0.87	0.74	0.34	1,678	1,109	569	0
8	A	33B	2	76	15	1	8	0.91	0.82	0.53	1,678	962	544	172
8	B	33A	1	80	12	1	7	0.88	0.75	0.40	1,704	1,021	683	0
8	B	33B	2	77	15	2	6	0.91	0.82	0.69	1,704	809	617	278
8	C	33A	1	77	15	1	8	0.85	0.69	0.41	1,642	977	665	0
8	C	33B	2	78	13	2	8	0.92	0.84	0.66	1,642	781	634	227
8	D	33A	1	81	12	1	6	0.87	0.74	0.39	1,642	1,000	642	0
8	D	33B	2	77	16	2	5	0.90	0.79	0.61	1,642	850	588	204

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-29 Cont'd
Inter-Rater Reliability Mathematics*

				Percentage Absolute Difference								Frequency		
Grade	Form	Item No.	Max Score	Perfect	Adjacent	Discrepant	Codes	Intra Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2
8	A	49A	1	95	1	0	4	0.98	0.97	0.22	1,678	1,317	361	0
8	A	49B	2	81	11	2	5	0.91	0.83	0.36	1,678	1,277	206	195
8	B	49A	1	95	1	0	4	0.99	0.98	0.25	1,704	1,282	422	0
8	B	49B	2	83	11	1	5	0.93	0.85	0.40	1,704	1,263	201	240
8	C	49A	1	96	1	0	3	0.99	0.98	0.27	1,642	1,195	447	0
8	C	49B	2	83	12	2	4	0.92	0.85	0.43	1,642	1,186	204	252
8	D	49A	1	96	1	0	2	0.98	0.96	0.25	1,642	1,233	409	0
8	D	49B	2	83	12	2	3	0.91	0.83	0.39	1,642	1,238	176	228
8	A	62	2	94	2	0	4	0.96	0.93	0.12	1,186	1,078	73	35
8	B	62	2	74	21	0	5	0.86	0.71	0.64	1,164	502	582	80
8	C	62A	1	91	5	0	4	0.95	0.90	0.62	1,180	446	734	0
8	C	62B	2	74	17	2	7	0.91	0.82	0.74	1,180	580	329	271
8	D	61A	1	97	1	0	2	0.99	0.98	0.58	1,178	493	685	0
8	D	61B	2	85	10	3	3	0.94	0.89	0.81	1,178	656	92	430
10	A	23	2	81	11	0	8	0.95	0.91	0.82	7,140	2,968	2,510	1,662
10	A	33	2	80	3	1	16	0.98	0.97	0.49	7,140	4,644	1,502	994
10	A	36	2	76	12	1	11	0.96	0.92	0.80	7,140	3,525	1,513	2,102
10	A	46	2	73	17	1	9	0.91	0.83	0.71	7,140	3,293	2,652	1,195

*The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 9-30
Inter-Rater Reliability, Writing Prompts* **

				Percentage Absolute Difference								Frequency						
Grade	Form	Item No.	Max Score	P	A	D	Codes	Intra. Corr.	Weighted Kappa	Mean	No. of Reads	0	1	2	3	4	5	6
4	A	1A	6	62	33	2	3	0.90	0.79	2.97	123,604	4,315	3,442	25,518	52,522	33,840	3,696	271
4	A	1B	3	92	6	0	2	0.88	0.78	1.92	123,604	2,839	5,037	114,505	1,223	0	0	0
8	A	1A	6	61	34	2	4	0.90	0.81	3.25	133,234	5,147	1,193	17,544	49,005	52,136	7,566	643
8	A	1B	3	96	2	0	3	0.96	0.93	1.94	133,234	3,609	1,374	127,850	401	0	0	0
10	A	1A	6	56	35	3	6	0.91	0.83	2.81	141,050	9,221	5,439	33,141	53,441	34,734	4,743	331
10	A	1B	3	88	6	0	5	0.93	0.86	1.90	141,050	7,405	3,291	126,318	4,036	0	0	0

*P is percent perfect agreement, A is percent adjacent agreement, and D is percent discrepant.
**The sum of the modes of agreement and codes may not equal exactly 100% due to rounding.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-1
Items Flagged for DIF, By Gender

Content	Grade	Form	Test Book Number	Item Type	Female			Male			SMD	Delta	LH Flag Female	LH Flag Male	Flag MH
					D+	D-	Z	D+	D-	Z					
RD	6	ABCD	18	CR	0.12	0.00	9.78	0.00	-0.13	-10.43	0.30		CC	-CC	CC
RD	7	ABCD	31	CR	0.11	0.00	8.23	0.02	-0.14	-8.79	0.27		CC	-CC	CC
RD	7	A	65	MC	0.03	0.00	4.84	0.00	-0.01	-1.62		1.65			C
RD	8	ABCD	19	CR	0.13	0.00	9.90	0.00	-0.13	-10.39	0.27		CC	-CC	CC
RD	8	ABCD	21	MC	0.00	-0.05	-4.74	0.03	0.00	5.78		-1.76			-C
RD	8	ABCD	23	MC	0.00	-0.03	-5.57	0.04	0.00	6.30		-2.15			-C
RD	8	ABCD	31	MC	0.00	-0.05	-6.92	0.05	0.00	6.91		-1.82			-C
RD	8	ABCD	56	CR	0.14	0.00	9.86	0.00	-0.13	-10.01	0.26		CC	-CC	CC
RD	10	A	7	CR	0.10	0.00	6.73	0.00	-0.13	-8.62	0.18			-CC	BB
RD	10	A	52	CR	0.12	-0.01	9.19	0.00	-0.14	-11.08	0.24		CC	-CC	BB
MA	5	ABCD	21	MC	0.00	-0.07	-8.61	0.07	-0.01	7.00		-1.57			-C
MA	6	ABCD	30	MC	0.00	-0.07	-8.22	0.06	0.00	7.19		-1.50			-C
MA	6	A	56	MC	0.00	-0.10	-21.00	0.01	-0.13	-11.03		-1.06	-C	-C	-B
MA	7	ABCD	10	MC	0.00	-0.07	-8.15	0.06	-0.01	7.95		-1.76			-C
MA	8	D	61B	CR	0.11	-0.04	3.45	0.00	-0.11	-4.92	0.20			-CC	BB
LA	10	E	5	MC	0.00	-0.05	-7.65	0.06	-0.03	6.59		-1.51			-C
WR	4	E	1A	CR	0.13	0.00	27.23	0.00	-0.14	-29.19	0.16		CC	-CC	
WR	8	E	1A	CR	0.10	-0.01	20.27	0.13	-0.13	-24.86	0.16			-CC	

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-2
Items Flagged for DIF, By Ethnicity, African American

Content	Grade	Form	Test Book Number	Item Type	D+	D-	Z	SMD	Delta	LH Flag	MH Flag
RD	3	D	62	MC	0.10	0.00	8.21		0.52	C	
MA	3	C	57	MC	0.00	-0.04	-5.29		-1.82		-C
MA	5	ABCD	35	MC	0.02	-0.02	-0.83		-1.86		-C
MA	5	C	61	MC	0.00	-0.08	-7.93		-1.86		-C
MA	6	A	56	MC	0.00	-0.15	-14.51		-1.06	-C	-B
MA	7	ABCD	18	MC	0.01	-0.02	-0.66		-1.65		-C
MA	8	ABCD	54	MC	0.02	-0.01	0.24		-2.05		-C
MA	10	A	36	CR	0.00	-0.11	-4.59	-0.04		-CC	
LA	4	E	26	MC	0.05	-0.06	-2.63		-2.31		-C
LA	8	E	26	MC	0.02	-0.05	-1.35		-2.31		-C
LA	10	E	26	MC	0.02	-0.08	-4.10		-2.30		-C
SS	8	E	14	MC	0.04	-0.10	-4.25		-1.60		-C
SS	8	E	22	MC	0.01	-0.03	-2.16		-1.74		-C
SS	10	E	2	MC	0.03	-0.05	-2.74		-2.02		-C

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-3
Items Flagged for DIF, By Ethnicity, Hispanic

Content	Grade	Form	Test Book Number	Item Type	D+	D-	Z	SMD	Delta	LH Flag	MH Flag
RD	6	ABCD	3	MC	0.00	-0.09	-4.92		-1.58		-C
RD	6	ABCD	37	MC	0.05	-0.09	-2.75		-2.01		-C
RD	7	ABCD	5	MC	0.00	-0.10	-5.74		-0.70	-C	
RD	7	ABCD	31	CR	0.11	-0.01	3.45	0.12		CC	
RD	8	ABCD	21	MC	0.01	-0.08	-2.65		-1.78		-C
MA	3	ABCD	21	MC	0.00	-0.07	-5.27		-1.67		-C
MA	4	BD	55	MC	0.00	-0.10	-10.71		-1.53		-C
MA	5	A	60	MC	0.00	-0.12	-10.40		-2.35	-C	-C
MA	6	A	56	MC	0.01	-0.13	-11.03		-0.64	-C	
MA	8	ABCD	19A	CR	0.12	0.00	6.52	0.14		CC	
MA	8	ABCD	54	MC	0.02	-0.04	-1.67		-1.96		-C
LA	4	E	26	MC	0.07	-0.05	0.10		-1.63		-C
LA	8	E	26	MC	0.06	-0.06	0.51		-1.63		-C
LA	10	E	26	MC	0.16	-0.12	-4.77		-1.66		-C
SC	4	ABCD	1	MC	0.01	-0.04	-1.97		-3.45		-C

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-4
Items Flagged for DIF, By Ethnicity, Asian

Content	Grade	Form	Test Book Number	Item Type	D+	D-	Z	SMD	Delta	LH Flag	MH Flag
RD	3	B	57	MC	0.04	-0.10	-4.39		-1.56		-C
RD	4	ABCD	13	CR	0.16	-0.02	3.82	0.18		CC	BB
RD	4	A	63	MC	0.11	-0.01	5.23		1.35	C	B
RD	4	D	62	CR	0.29	-0.18	3.05	0.26		CC	CC
RD	5	ABCD	47	CR	0.13	-0.08	3.77	0.30		CC	CC
RD	6	ABCD	18	CR	0.17	-0.18	4.73	0.27		CC	CC
RD	6	ABCD	24	CR	0.20	-0.08	5.96	0.21		CC	BB
RD	6	ABCD	37	MC	0.01	-0.11	-4.98		-2.41		-C
RD	6	ABCD	38	MC	0.01	-0.05	-2.44		-1.52		-C
RD	6	A	59	MC	0.09	-0.03	5.09		1.54		C
RD	6	A	67	CR	0.22	-0.02	5.12	0.29		CC	CC
RD	6	B	67	CR	0.22	-0.01	3.61	0.19		CC	BB
RD	6	C	57	CR	0.19	0.00	4.21	0.31		CC	CC
RD	7	ABCD	31	CR	0.16	0.00	5.01	0.21		CC	BB
RD	7	C	67	CR	0.18	-0.05	3.49	0.24		CC	BB
RD	8	ABCD	19	CR	0.22	-0.13	6.10	0.31		CC	CC
RD	8	ABCD	21	MC	0.00	-0.08	-5.93		-2.02		-C
RD	8	ABCD	56	CR	0.24	-0.04	6.00	0.33		CC	CC
RD	8	A	65	CR	0.32	-0.14	3.78	0.22		CC	BB
RD	8	B	67	CR	0.24	-0.10	3.77	0.23		CC	BB
RD	8	C	63	CR	0.31	-0.14	3.57	0.31		CC	CC
RD	8	D	60	CR	0.18	0.00	3.59	0.24		CC	BB
RD	8	D	63	CR	0.29	-0.07	2.94	0.34		CC	CC
RD	10	A	7	CR	0.25	-0.04	6.22	0.30		CC	CC
RD	10	A	21	MC	0.11	0.00	6.00		1.78	C	C
RD	10	A	52	CR	0.28	-0.03	7.97	0.29		CC	CC
MA	3	ABCD	16	MC	0.00	-0.12	-5.32		-1.91	-C	-C
MA	4	ABCD	34	MC	0.00	-0.05	-2.76		-1.55		-C

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-4 Cont'd
Items Flagged for DIF, By Ethnicity, Asian

Content	Grade	Form	Test Book Number	Item Type	D+	D-	Z	SMD	Delta	LH Flag	MH Flag
MA	4	A	52	MC	0.00	-0.13	-7.11		-1.75	-C	-C
MA	4	BD	55	MC	0.00	-0.12	-8.70		-1.79	-C	-C
MA	4	B	60	MC	0.01	-0.08	-3.90		-1.71		-C
MA	5	ABCD	18	MC	0.01	-0.12	-4.85		-1.69		-C
MA	5	ABCD	35	MC	0.01	-0.04	-4.17		-2.10		-C
MA	5	A	56	MC	0.08	-0.01	4.83		1.64		C
MA	5	A	60	MC	0.01	-0.16	-10.38		-2.80	-C	-C
MA	5	C	61	MC	0.01	-0.10	-6.21		-2.06		-C
MA	5	C	66	MC	0.01	-0.12	-5.59		-2.23		-C
MA	5	D	57	MC	0.04	0.00	2.92		1.60		C
MA	5	D	65	MC	0.01	-0.13	-5.61		-1.77	-C	-C
MA	6	ABCD	36	MC	0.01	-0.11	-5.77		-1.69		-C
MA	6	ABCD	39B	CR	0.15	-0.11	4.17	0.10		CC	
MA	6	ABCD	47	MC	0.00	-0.07	-5.79		-2.12		-C
MA	6	A	63A	CR	0.01	-0.15	-5.14	-0.35		-CC	-CC
MA	6	A	63B	CR	0.04	-0.16	-4.04	-0.27		-CC	-CC
MA	7	ABCD	18	MC	0.00	-0.06	-2.98		-1.82		-C
MA	7	ABCD	28	MC	0.01	-0.10	-3.51		-1.63		-C
MA	7	ABCD	37	MC	0.01	-0.04	-1.99		-1.70		-C
MA	7	ABCD	44	MC	0.00	-0.04	-1.73		-2.27		-C
MA	7	B	61A	CR	0.13	-0.03	4.29	0.29		CC	CC
MA	8	ABCD	54	MC	0.00	-0.14	-10.11		-2.60	-C	-C
MA	10	A	26	MC	0.00	-0.12	-7.37		-1.77	-C	-C
MA	10	A	42	MC	0.02	-0.12	-5.86		-1.07	-C	-B
LA	4	E	13	MC	0.00	-0.16	-9.63		-0.67	-C	
LA	4	E	26	MC	0.03	-0.05	-2.05		-3.66		-C
LA	8	E	17	MC	0.00	-0.12	-6.50		0.76	-C	
LA	8	E	26	MC	0.10	-0.04	3.93		-3.66		-C

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-4 Cont'd
Items Flagged for DIF, By Ethnicity, Asian

Content	Grade	Form	Test Book Number	Item Type	D+	D-	Z	SMD	Delta	LH Flag	MH Flag
LA	10	E	4	MC	0.00	-0.10	-5.77		-0.67	-C	
LA	10	E	26	MC	0.00	-0.26	-14.88		-3.75	-C	-C
SS	4	E	20	MC	0.02	-0.12	-5.27		-1.30	-C	-B
SS	8	E	3	MC	0.01	-0.15	-7.43		-2.29	-C	-C
SS	10	E	12	MC	0.13	0.00	6.67		1.39	C	B
SS	10	E	15	MC	0.12	-0.01	5.97		0.91	C	
SS	10	E	25	MC	0.01	-0.05	-0.95		-2.05		-C
SS	10	E	33	MC	0.13	0.00	6.60		1.07	C	B
SC	4	ABCD	1	MC	0.00	-0.03	-3.21		-4.85		-C
SC	8	ABCD	11	MC	0.01	-0.05	-1.40		-1.51		-C
SC	10	A	4	MC	0.11	0.00	5.83		0.68	C	
WR	8	E	1A	CR	0.15	-0.02	8.18	0.14		CC	
WR	10	E	1A	CR	0.18	0.00	4.77	0.16		CC	

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-5
Items Flagged for DIF, By Ethnicity, American Indian*

Content	Grade	Form	Test Book Number	Item Type	D+	D-	Z	SMD	Delta	LH Flag	MH Flag
RD	3	C	64	CR				0.33			CC
RD	4	ABCD	17	MC	0.00	-0.12	-2.80		0.28	-C	
RD	4	A	67	CR	0.11	-0.64	-1.12	-0.27			-CC
RD	4	B	64	CR				-0.40			-CC
RD	6	B	67	CR				-0.79			-CC
RD	6	C	57	CR				0.32			CC
RD	6	D	67	CR				0.37			CC
RD	7	ABCD	52	MC	0.10	-0.28	-2.82		0.20	-C	
RD	7	A	67	CR	0.29	-0.42	-2.15	-0.34			-CC
RD	7	C	67	CR				-0.34			-CC
RD	7	D	67	CR	0.43	-0.42	-1.83	-0.32			-CC
RD	8	ABCD	34	MC	0.20	-0.03	2.89		-0.02	C	
MA	3	AB	51	CR	0.19	-0.32	-1.92	-0.27			-CC
MA	3	A	57A	CR	1.64		1.30	-0.33			-CC
MA	3	A	57B	CR	1.74		1.38	-0.48			-CC
MA	3	D	58A	CR	0.12	-0.30	-1.56	-0.33			-CC
MA	4	ABCD	36	MC	0.02	0.00	1.06		1.94		C
MA	5	C	64	CR				-0.40			-CC
MA	5	D	64A	CR				-0.40			-CC
MA	5	D	64B	CR				-0.50			-CC
MA	6	B	63B	CR				-0.25			-CC
MA	7	B	61B	CR	0.04	-0.28	-1.55	-0.28			-CC
WR	8	E	1A	CR	0.05	-0.15	-4.66	-0.18		-CC	-BB

*Note: Linn-Harnisch DIF statistics can only be calculated for items with sufficient student N counts. In some cases here, the size of the tested population was too small here to include valid Linn-Harnisch DIF statistics.

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-6
Items Flagged for DIF, By English Language Proficiency

Content	Grade	Form	Test Book Number	Item Type	Not Proficient			Proficient			SMD	Delta	LH Flag Not Proficient	LH Flag Proficient	MH Flag
					D+	D-	Z	D+	D-	Z					
RD	4	ABCD	12	MC	0.00	-0.10	-4.16	0.03	-0.03	-0.01		-0.44	-C		
RD	4	ABCD	13	CR	0.17	-0.17	3.79	0.04	-0.06	-1.09	0.13		CC		
RD	4	ABCD	47	CR	0.19	-0.12	4.30	0.04	-0.05	-1.68	0.16		CC		
RD	4	A	63	MC	0.12	0.00	6.25	0.02	-0.01	-0.74		1.24	C		B
RD	4	B	64	CR	0.31	-0.19	3.19	0.04	-0.04	-1.36	0.23		CC		BB
RD	5	ABCD	13	CR	0.23	-0.14	4.35	0.04	-0.06	-1.39	0.11		CC		
RD	5	ABCD	47	CR	0.21	-0.04	3.67	0.05	-0.05	-1.52	0.19		CC		BB
RD	6	ABCD	3	MC	0.03	-0.16	-4.78	0.01	-0.01	0.97		-1.75	-C		-C
RD	6	ABCD	18	CR	0.16	-0.05	3.09	0.01	-0.04	-1.15	0.19		CC		BB
RD	6	ABCD	24	CR	0.18	-0.08	3.81	0.04	-0.03	-1.95	0.16		CC		
RD	6	ABCD	29	MC	0.00	-0.11	-4.45	0.02	-0.02	0.17		-0.88	-C		
RD	6	ABCD	37	MC	0.03	-0.17	-7.31	0.02	0.00	2.64		-1.88	-C		-C
RD	6	A	57	MC	0.07	-0.12	-5.70	0.01	0.00	2.34		-1.51			-C
RD	6	A	67	CR	0.21	-0.05	3.64	0.02	-0.04	-1.29	0.23		CC		BB
RD	7	ABCD	26	MC	0.00	-0.11	-4.06	0.01	-0.02	1.14		-1.07	-C		-B
RD	7	ABCD	31	CR	0.17	0.00	4.28	0.04	-0.04	-1.53	0.24		CC		BB
RD	8	ABCD	21	MC	0.03	-0.11	-3.07	0.01	-0.01	1.60		-1.64			-C
RD	8	ABCD	56	CR	0.22	-0.01	3.31	0.03	-0.04	-1.01	0.21		CC		BB
RD	10	A	7	CR	0.33	-0.07	4.49	0.02	-0.04	-1.91	0.19		CC		BB
RD	10	A	37	MC	0.00	-0.14	-4.80	0.02	-0.02	0.95		-0.56	-C		
RD	10	A	52	CR	0.23	-0.10	4.60	0.03	-0.05	-2.07	0.20		CC		BB
MA	3	ABCD	16	MC	0.05	-0.13	-5.44	0.02	-0.01	1.41		-1.53	-C		-C
MA	3	ABCD	21	MC	0.01	-0.09	-3.91	0.01	0.00	2.59		-1.73			-C
MA	3	C	55	MC	0.03	0.00	3.30	0.00	0.00	0.75		1.60			C
MA	4	ABCD	34	MC	0.03	-0.12	-5.15	0.01	-0.02	1.75		-1.17	-C		-B
MA	4	A	52	MC	0.00	-0.14	-7.29	0.01	-0.01	1.21		-1.58	-C		-C
MA	4	BD	55	MC	0.01	-0.15	-10.05	0.01	-0.01	0.85		-1.86	-C		-C
MA	5	ABCD	3	MC	0.12	-0.07	3.91	0.01	-0.01	-1.45		0.65	C		

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-6 Cont'd
Items Flagged for DIF, By English Language Proficiency

Content	Grade	Form	Test Book Number	Item Type	Not Proficient			Proficient			SMD	Delta	LH Flag Not Proficient	LH Flag Proficient	MH Flag
					D+	D-	Z	D+	D-	Z					
MA	5	A	60	MC	0.00	-0.15	-10.03	0.01	0.00	3.09		-2.27	-C		-C
MA	5	D	63	MC	0.08	-0.12	-5.36	0.01	-0.01	1.25		-1.52			-C
MA	5	D	64A	CR	0.16	-0.07	3.18	0.02	-0.02	-1.56	0.18		CC		BB
MA	5	D	65	MC	0.10	-0.15	-6.38	0.01	-0.02	0.91		-1.78	-C		-C
MA	6	ABCD	47	MC	0.02	-0.04	-0.55	0.01	-0.01	0.51		-1.57			-C
MA	6	A	56	MC	0.00	-0.14	-9.18	0.01	-0.10	-21.38		-1.28	-C		-B
MA	6	A	57	MC	0.13	-0.10	6.36	0.01	-0.01	-0.81		1.36	C		B
MA	6	B	63A	CR	0.14	0.00	4.31	0.01	-0.03	-1.40	0.30		CC		CC
MA	6	B	63B	CR	0.21	-0.05	3.11	0.04	-0.06	-1.62	0.21		CC		BB
MA	8	ABCD	45	MC	0.01	-0.12	-4.43	0.03	-0.01	0.82		-0.67	-C		
MA	8	ABCD	54	MC	0.03	-0.13	-6.23	0.01	0.00	3.26		-1.70	-C		-C
MA	10	A	42	MC	0.03	-0.11	-3.85	0.01	-0.01	0.75		-1.07	-C		-B
MA	10	A	48	MC	0.12	-0.02	4.33	0.03	-0.02	-0.92		0.50	C		
LA	4	E	26	MC	0.02	-0.06	-1.82	0.01	-0.01	-0.91		-2.28			-C
LA	8	E	17	MC	0.01	-0.13	-4.65	0.01	-0.03	1.07		0.99	-C		
LA	8	E	26	MC	0.08	-0.04	2.41	0.01	-0.02	-2.82		-2.28			-C
LA	10	E	6	MC	0.00	-0.13	-4.02	0.01	-0.03	-0.30		-0.60	-C		
LA	10	E	26	MC	0.12	-0.30	-9.59	0.01	-0.01	1.69		-2.35	-C		-C
SS	8	E	3	MC	0.03	-0.14	-4.99	0.01	0.00	0.65		-1.38	-C		-B
SS	10	E	4	MC	0.17	-0.04	3.81	0.02	-0.03	-1.60		0.50	C		
SS	10	E	15	MC	0.14	-0.04	4.16	0.01	-0.02	-2.12		0.69	C		
SC	4	ABCD	1	MC	0.01	-0.08	-2.80	0.01	0.00	2.29		-3.89			-C
SC	8	ABCD	15	MC	0.00	-0.12	-5.14	0.02	-0.01	0.62		-0.66	-C		
SC	10	A	4	MC	0.17	0.00	6.10	0.00	-0.01	-2.05		1.07	C		B
SC	10	A	47	MC	0.12	-0.12	-3.46	0.02	-0.01	0.19		-0.40	-C		

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-7
Items Flagged for DIF, By Socioeconomic Status

Content	Grade	Form	Test Book Number	Item Type	Not Economically Disadv.			Economically Disadv.			SMD	Delta	LH Flag Not Econ Disadv	LH Flag Econ Disadv	MH Flag
					D+	D-	Z	D+	D-	Z					
MA	6	A	56	MC	0.01	-0.07	-12.55	0.00	-0.12	-21.14		-0.72		-C	

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-8
Items Flagged for DIF, By Disability Status

Content	Grade	Form	Test Book Number	Item Type	Not Disabled			Disabled			SMD	Delta	LH Flag Not Disabled	LH Flag Disabled	MH Flag
					D+	D-	Z	D+	D-	Z					
RD	4	ABCD	47	CR	0.06	-0.04	0.68	0.08	-0.15	-3.98	-0.10			-CC	
RD	5	ABCD	13	CR	0.07	-0.06	1.40	0.04	-0.15	-5.51	-0.17			-CC	-BB
RD	5	C	57	CR	0.03	-0.02	-0.07	0.22	-0.13	-2.87	-0.18			-CC	-BB
RD	5	C	63	CR	0.04	-0.07	0.13	0.07	-0.20	-2.81	-0.19			-CC	-BB
RD	6	ABCD	18	CR	0.03	-0.04	0.97	0.05	-0.12	-4.10	-0.16			-CC	
RD	7	ABCD	31	CR	0.04	-0.04	1.30	0.02	-0.19	-5.61	-0.13			-CC	
RD	8	ABCD	19	CR	0.05	-0.03	1.15	0.08	-0.13	-4.21	-0.15			-CC	
RD	8	ABCD	56	CR	0.05	-0.03	1.68	0.15	-0.17	-4.97	-0.17			-CC	
RD	8	C	63	CR	0.04	-0.04	-0.70	0.05	-0.17	-3.00	-0.23			-CC	-BB
RD	10	A	52	CR	0.04	-0.03	0.33	0.00	-0.12	-5.12	-0.13			-CC	
MA	3	B	52	MC	0.02	-0.02	-5.93	0.09	0.00	10.98		1.60			C
MA	6	A	56	MC	0.01	-0.09	-17.18	0.03	-0.15	-16.42		-0.15		-C	
MA	6	D	58	MC	0.01	0.00	3.25	0.01	-0.05	-4.26		-1.56			-C
MA	7	C	62	CR	0.02	-0.05	-2.12	0.15	-0.01	4.61	0.19			CC	BB
SC	8	ABCD	2	MC	0.01	0.00	1.36	0.03	-0.02	-3.48		-1.90			-C
SC	8	ABCD	3	MC	0.02	0.00	-0.08	0.01	-0.02	-1.18		-1.64			-C
WR	4	E	1A	CR	0.07	-0.04	7.78	0.00	-0.25	-24.94	-0.33			-CC	-CC
WR	4	E	1B	CR	0.02	-0.02	7.10	0.00	-0.03	-13.74	-0.27				-CC
WR	8	E	1A	CR	0.06	-0.03	2.40	0.13	-0.22	-17.09	-0.33			-CC	-CC
WR	8	E	1B	CR	0.01	-0.01	4.56	0.00	-0.01	-6.94	-0.27				-CC

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-9
Correlations among Reading Objectives

Grade	CS	1	2	3
3	2	0.75		
	3	0.76	0.84	
	4	0.58	0.62	0.65
4	2	0.74		
	3	0.76	0.79	
	4	0.60	0.61	0.64
5	2	0.74		
	3	0.75	0.76	
	4	0.66	0.67	0.68
6	2	0.65		
	3	0.69	0.71	
	4	0.65	0.66	0.72
7	2	0.65		
	3	0.69	0.72	
	4	0.63	0.66	0.71
8	2	0.65		
	3	0.68	0.71	
	4	0.66	0.71	0.73
10	2	0.58		
	3	0.67	0.72	
	4	0.66	0.70	0.80

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-10
Correlations among Mathematics Objectives

Grade	CS	A	B	C	D	E
3	B	0.62				
	C	0.53	0.59			
	D	0.50	0.58	0.53		
	E	0.63	0.64	0.59	0.56	
	F	0.54	0.67	0.55	0.52	0.57
4	B	0.53				
	C	0.50	0.48			
	D	0.58	0.61	0.51		
	E	0.54	0.51	0.46	0.55	
	F	0.62	0.65	0.54	0.63	0.54
5	B	0.55				
	C	0.50	0.49			
	D	0.56	0.61	0.51		
	E	0.64	0.57	0.51	0.61	
	F	0.64	0.67	0.52	0.68	0.65
6	B	0.62				
	C	0.56	0.56			
	D	0.62	0.67	0.60		
	E	0.68	0.62	0.59	0.63	
	F	0.66	0.66	0.54	0.63	0.60
7	B	0.64				
	C	0.64	0.63			
	D	0.61	0.68	0.61		
	E	0.64	0.68	0.64	0.65	
	F	0.66	0.68	0.60	0.63	0.64
8	B	0.58				
	C	0.62	0.54			
	D	0.64	0.64	0.57		
	E	0.65	0.59	0.57	0.62	
	F	0.67	0.66	0.64	0.67	0.66
10	B	0.70				
	C	0.71	0.66			
	D	0.71	0.67	0.71		
	E	0.69	0.65	0.65	0.67	
	F	0.72	0.68	0.72	0.74	0.67

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-11
Correlations among Language Arts Objectives

Grade	CS	B	D
4	D	0.52	
	F	0.52	0.38
8	D	0.67	
	F	0.48	0.41
10	D	0.68	
	F	0.53	0.53

Table 10-12
Correlations among Social Studies Objectives

Grade	CS	A	B	C	D
4	B	0.65			
	C	0.53	0.52		
	D	0.55	0.54	0.48	
	E	0.64	0.61	0.54	0.56
8	B	0.62			
	C	0.54	0.56		
	D	0.59	0.60	0.54	
	E	0.51	0.54	0.50	0.51
10	B	0.62			
	C	0.61	0.69		
	D	0.58	0.65	0.68	
	E	0.54	0.59	0.62	0.59

Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-13
Correlations among Science Objectives

Grade	CS	A	B	C	D	E	F	G
	B	0.36						
	C	0.41	0.54					
	D	0.26	0.32	0.35				
4	E	0.31	0.40	0.43	0.32			
	F	0.37	0.49	0.52	0.36	0.46		
	G	0.35	0.50	0.51	0.32	0.39	0.49	
	H	0.31	0.45	0.46	0.30	0.37	0.45	0.44
	B	0.40						
	C	0.51	0.47					
	D	0.38	0.40	0.46				
8	E	0.41	0.43	0.48	0.43			
	F	0.41	0.44	0.48	0.44	0.46		
	G	0.43	0.41	0.51	0.39	0.39	0.42	
	H	0.42	0.41	0.50	0.38	0.41	0.43	0.43
	B	0.48						
	C	0.50	0.52					
	D	0.51	0.52	0.56				
10	E	0.39	0.39	0.41	0.42			
	F	0.40	0.40	0.42	0.41	0.32		
	G	0.50	0.51	0.51	0.49	0.40	0.40	
	H	0.50	0.53	0.56	0.53	0.40	0.42	0.55

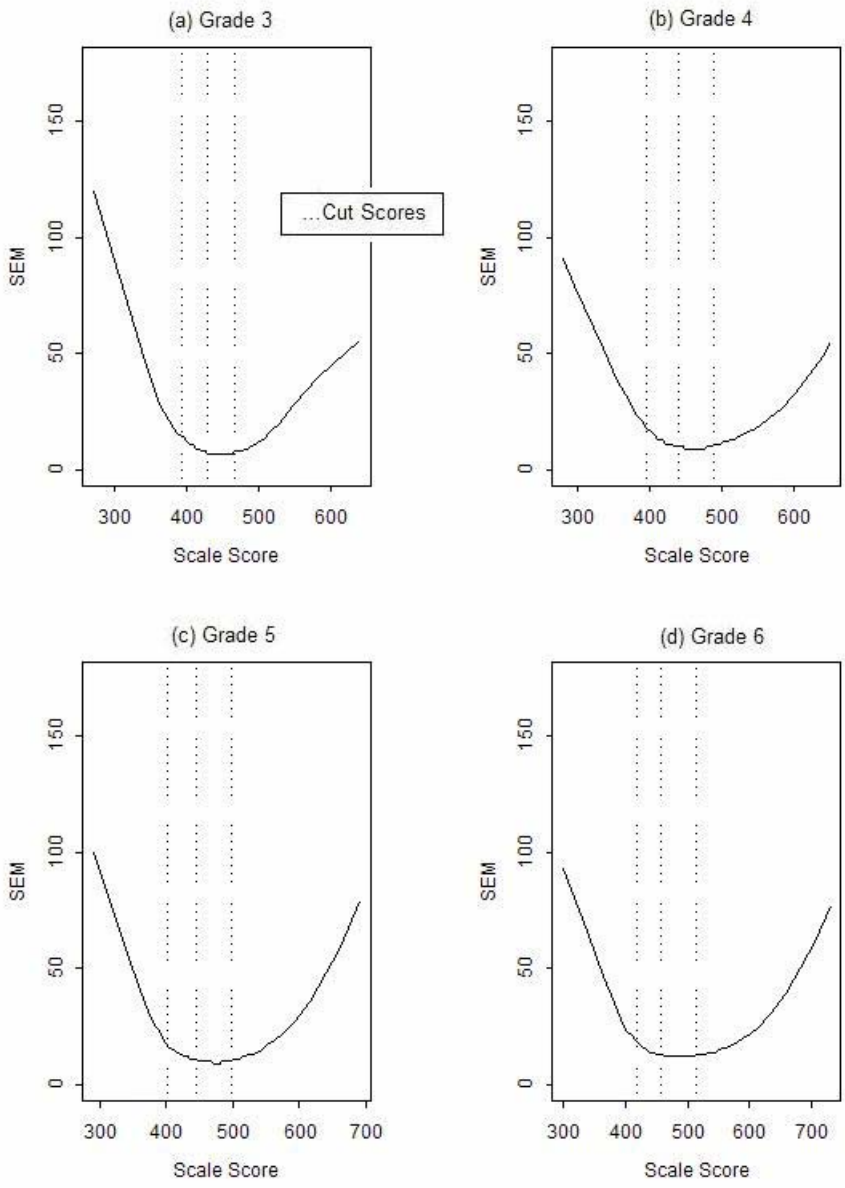
Copyright © 2008 by the Wisconsin Department of Public Instruction

Table 10-14
Factor Analysis

Content Area	Grade	First Eigenvalue	Second Eigenvalue	Ratio of First Two Eigenvalues
Reading	3	16.17	1.08	14.97
	4	14.29	0.96	14.89
	5	13.11	1.14	11.50
	6	10.79	1.03	10.48
	7	10.86	1.15	9.44
	8	11.69	0.86	13.59
	10	13.73	1.55	8.86
Mathematics	3	10.95	1.59	6.89
	4	10.00	1.55	6.45
	5	14.05	4.92	2.86
	6	13.66	2.38	5.74
	7	14.26	1.67	8.54
	8	13.53	2.25	6.01
	10	15.96	1.33	12.00
Language Arts	4	5.16	0.55	9.38
	8	6.41	0.46	13.93
	10	6.75	0.71	9.51
Social Studies	4	7.64	0.96	7.96
	8	7.85	0.86	9.13
	10	10.31	1.13	9.12
Science	4	6.93	0.86	8.06
	8	6.93	0.90	7.70
	10	8.51	0.59	14.42

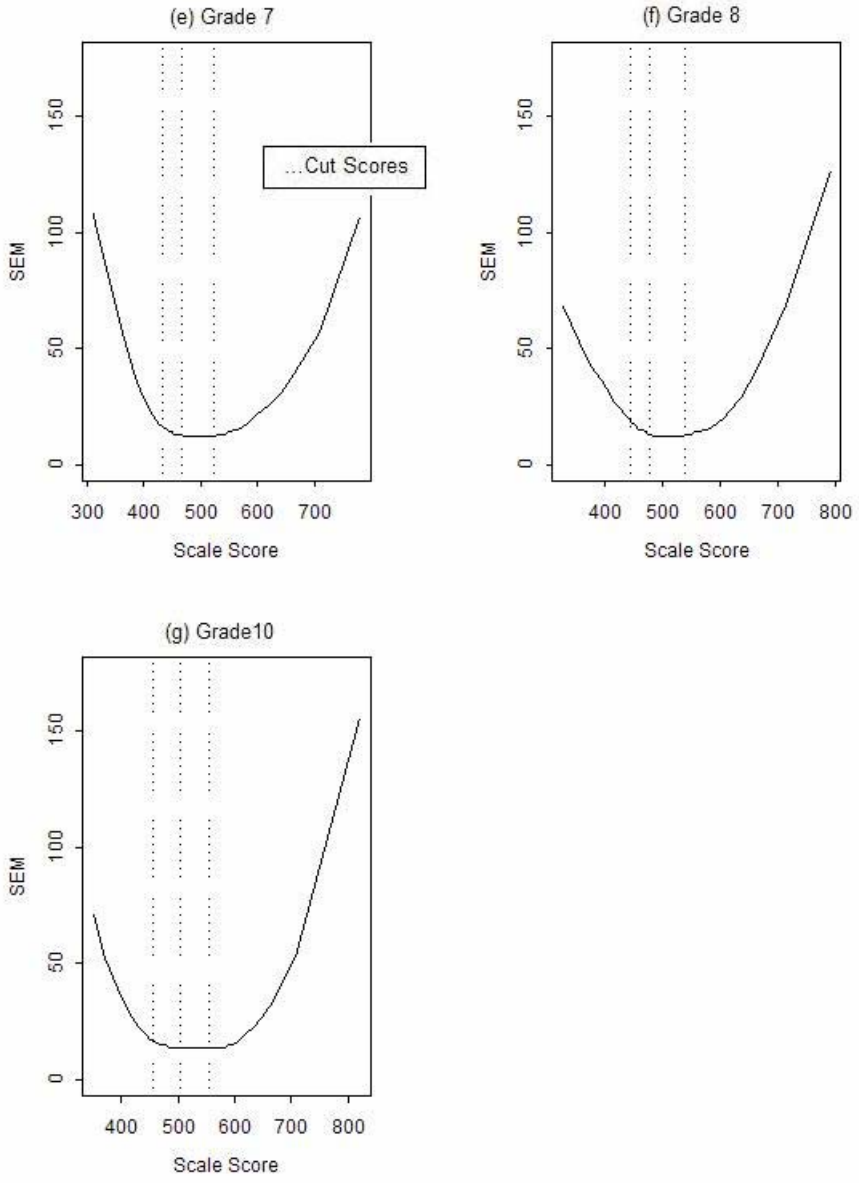
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-1
SEM Curves, Reading Grades 3-6



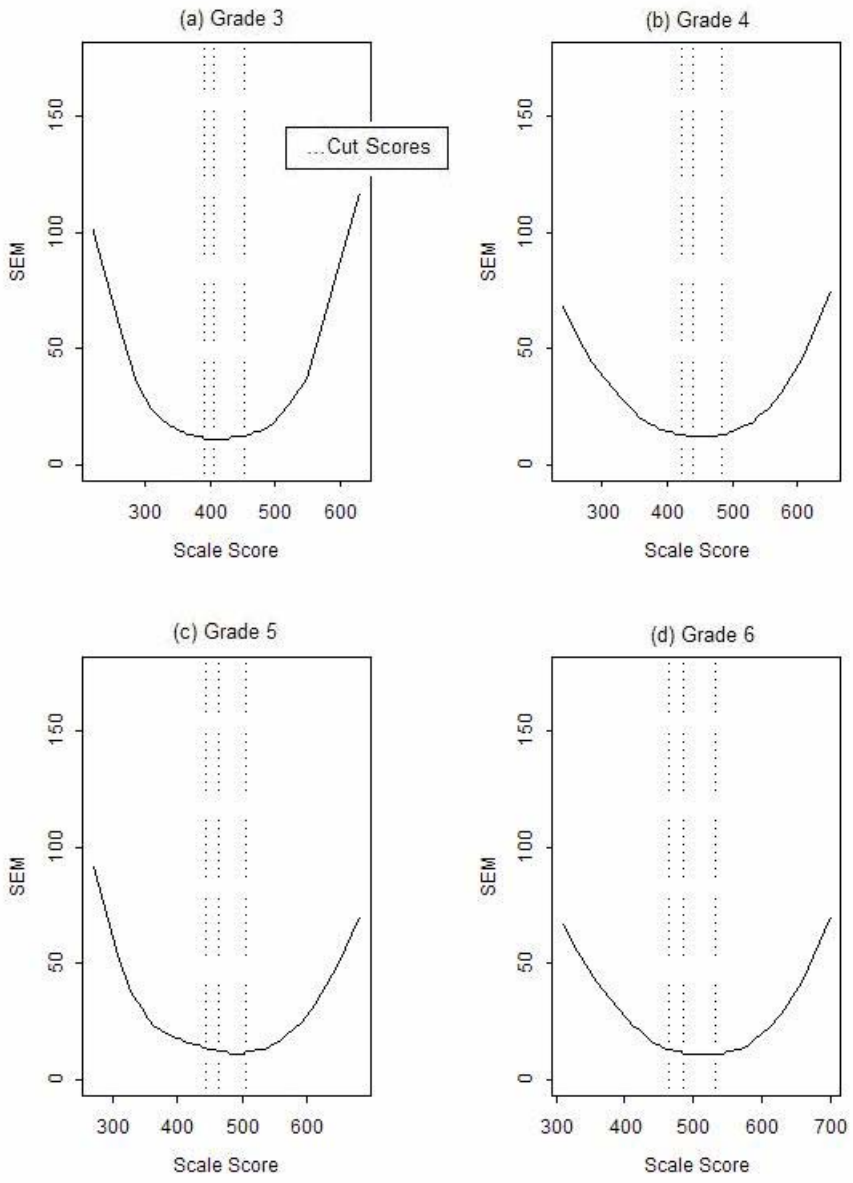
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-1 Cont'd
SEM Curves, Reading Grades 7, 8, 10



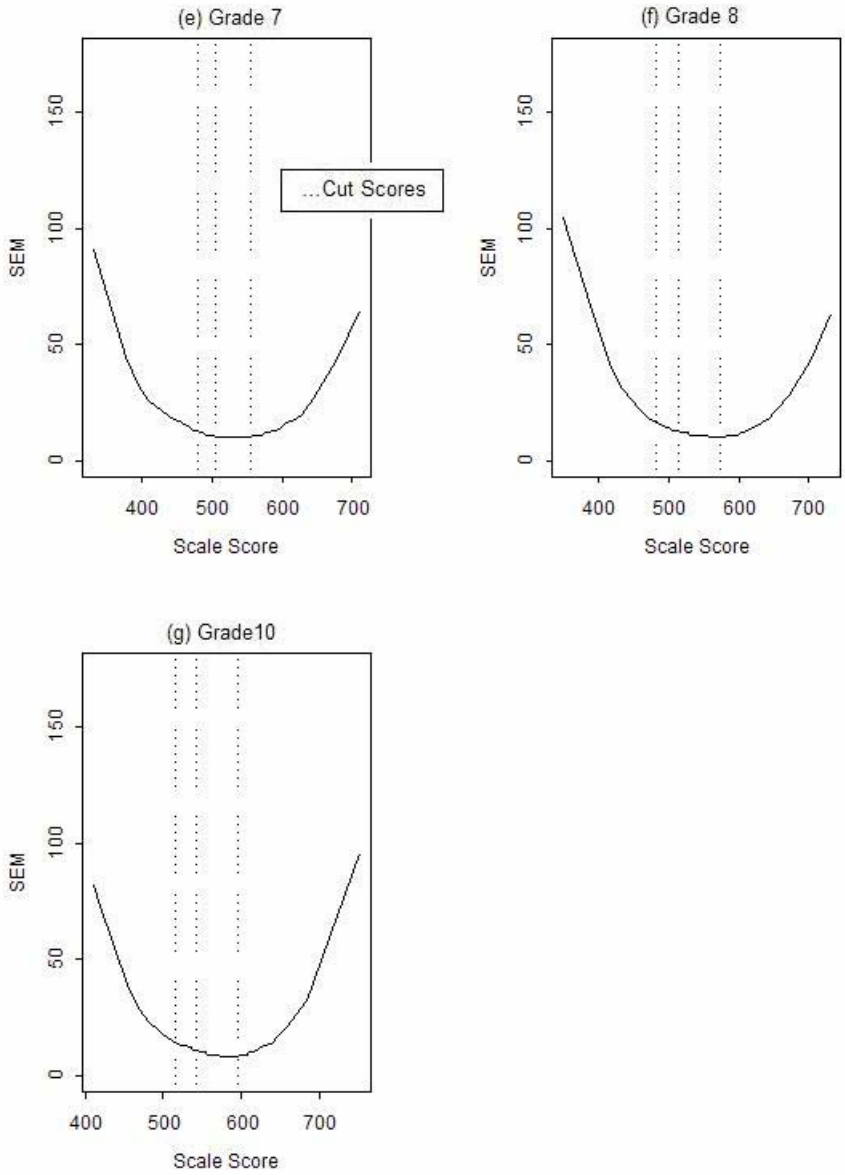
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-2
SEM Curves, Mathematics Grades 3-6



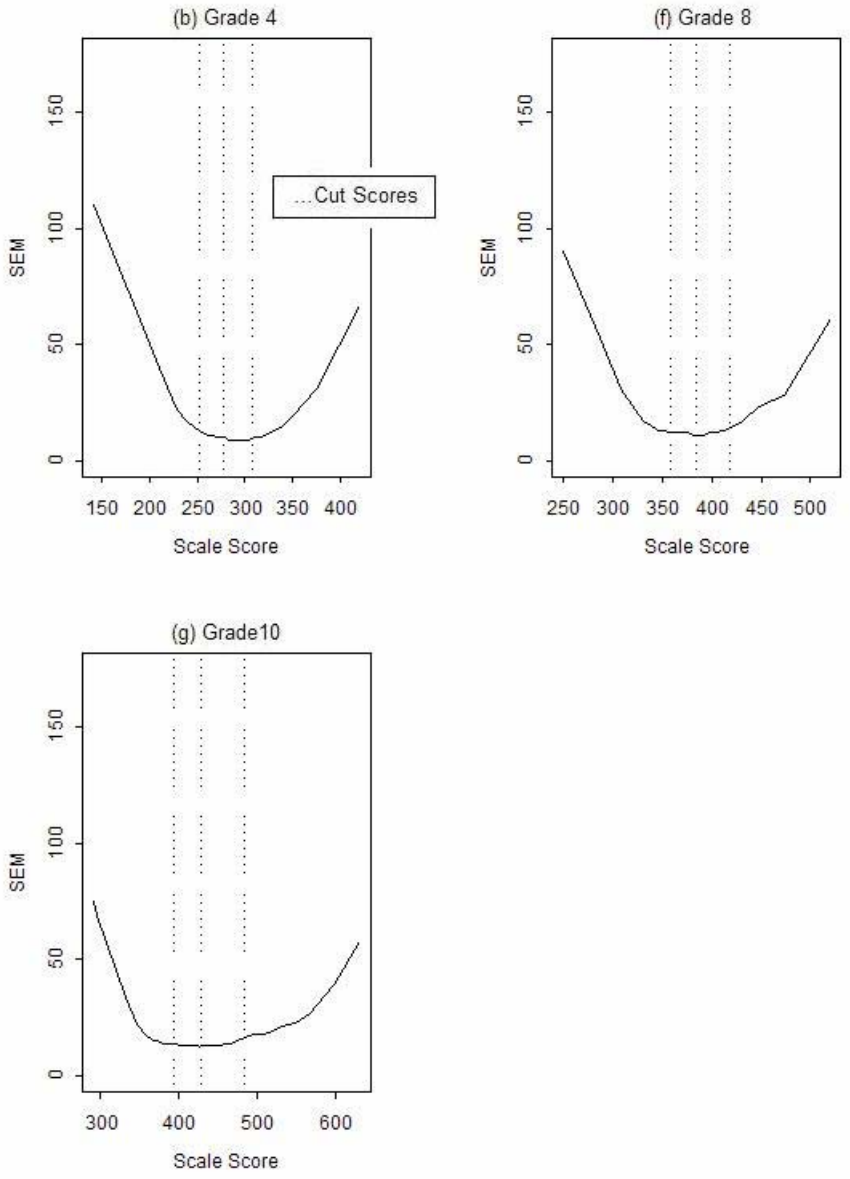
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-2 Cont'd
SEM Curves, Mathematics Grades 7, 8, 10



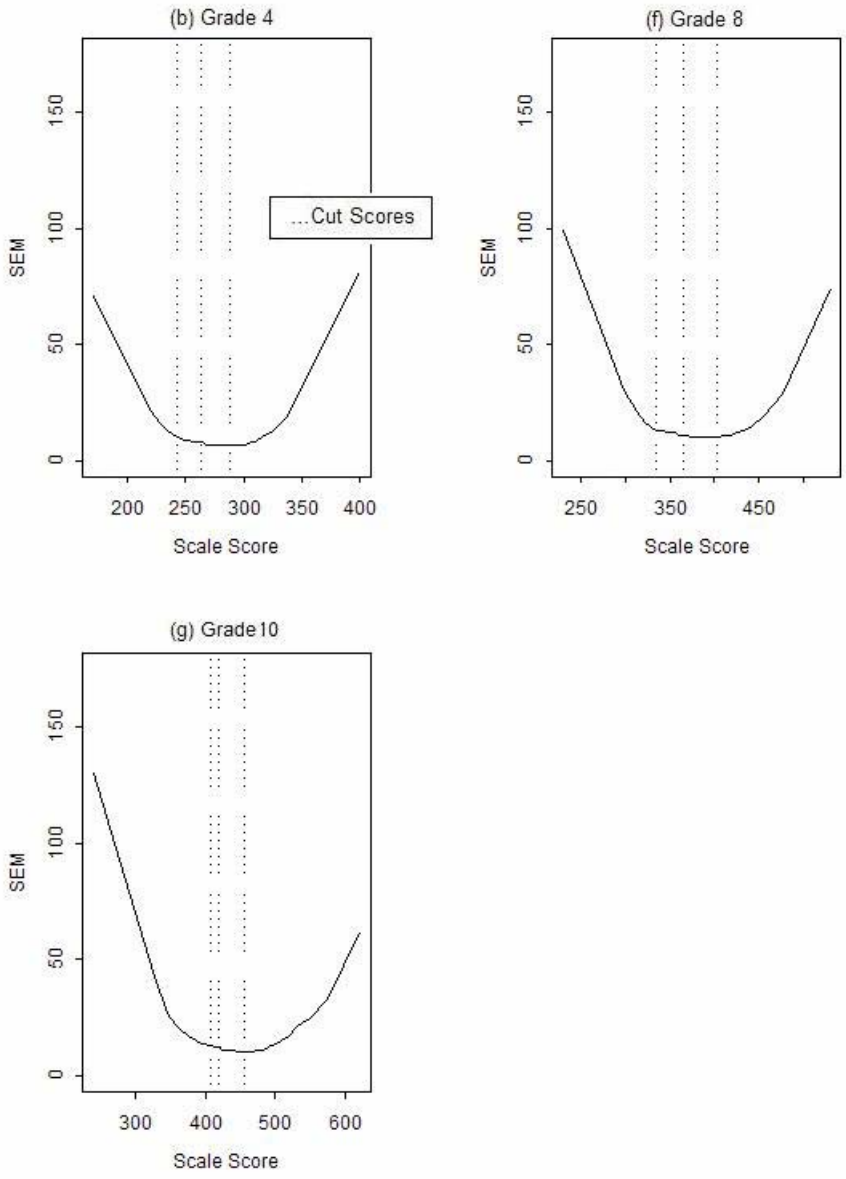
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-3
SEM Curves, Language Arts Grades 4, 8, 10



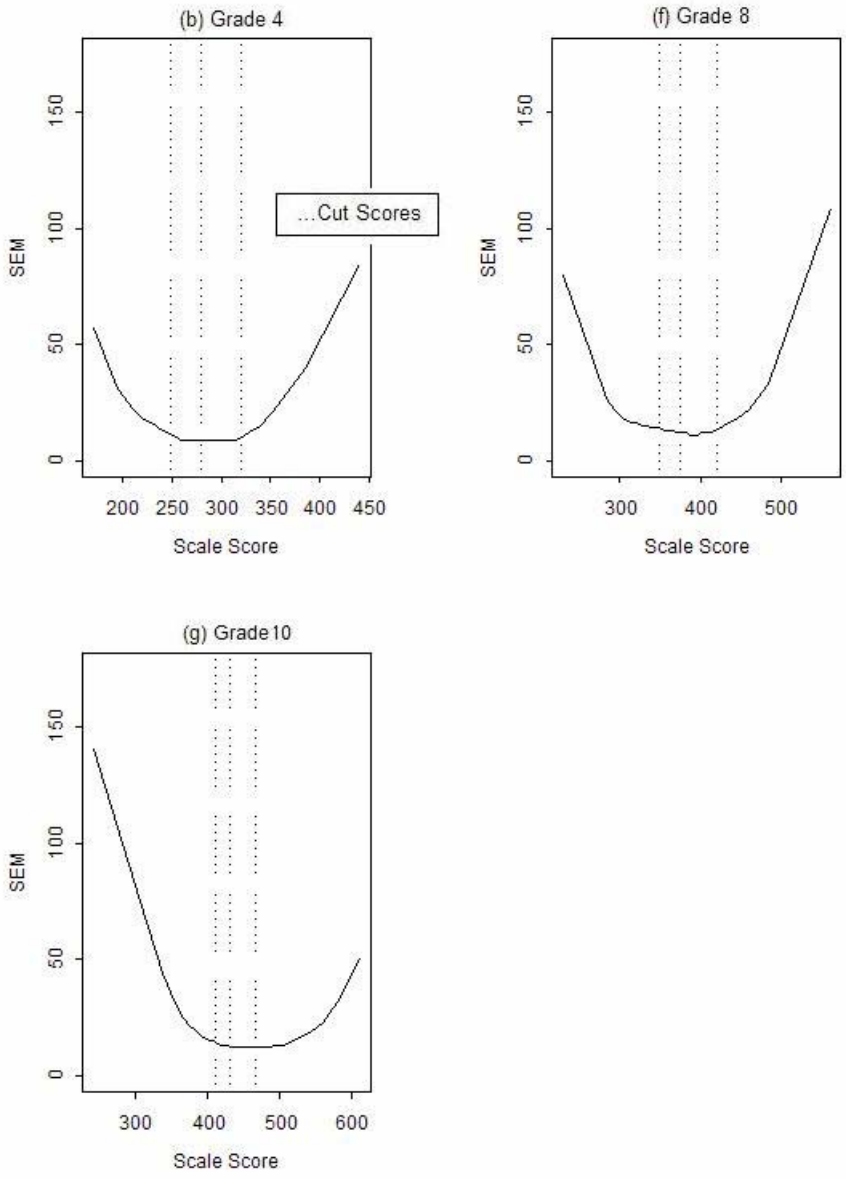
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-4
SEM Curves, Social Studies Grades 4, 8, 10



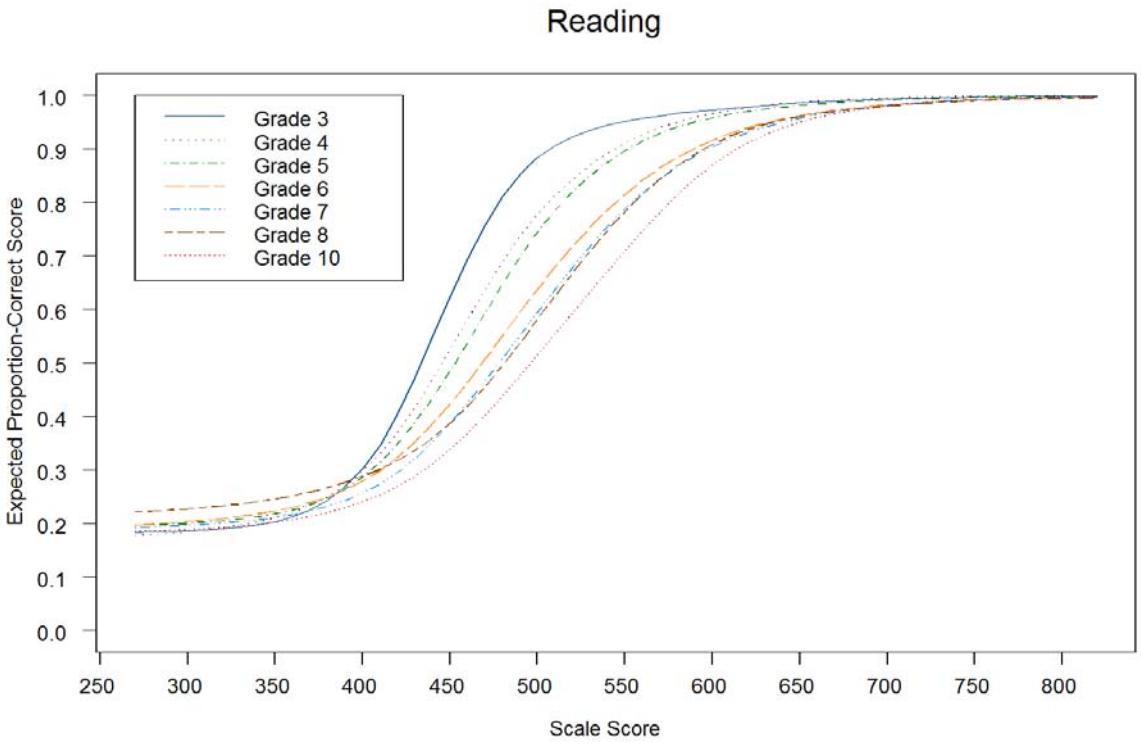
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-5
SEM Curves, Science Grades 4, 8, 10



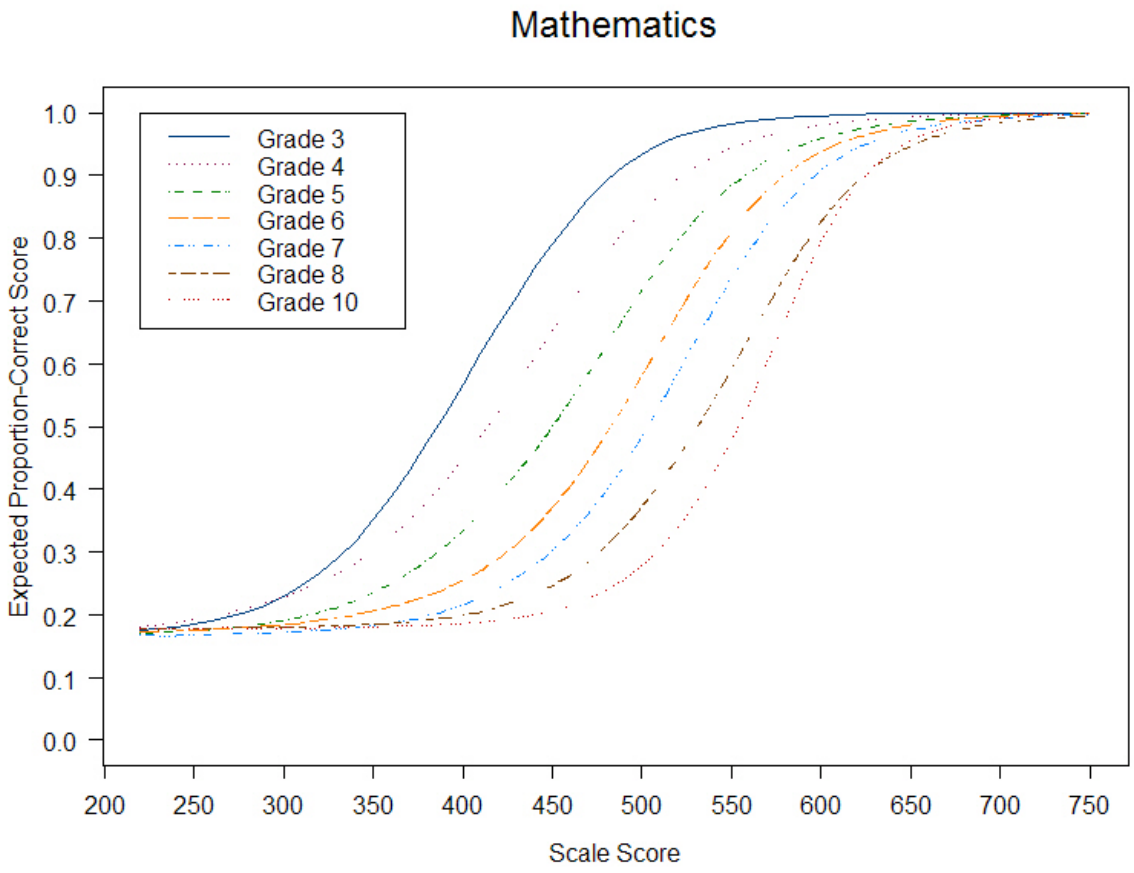
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-6
TCC Curve for Reading Grades 3-8, 10



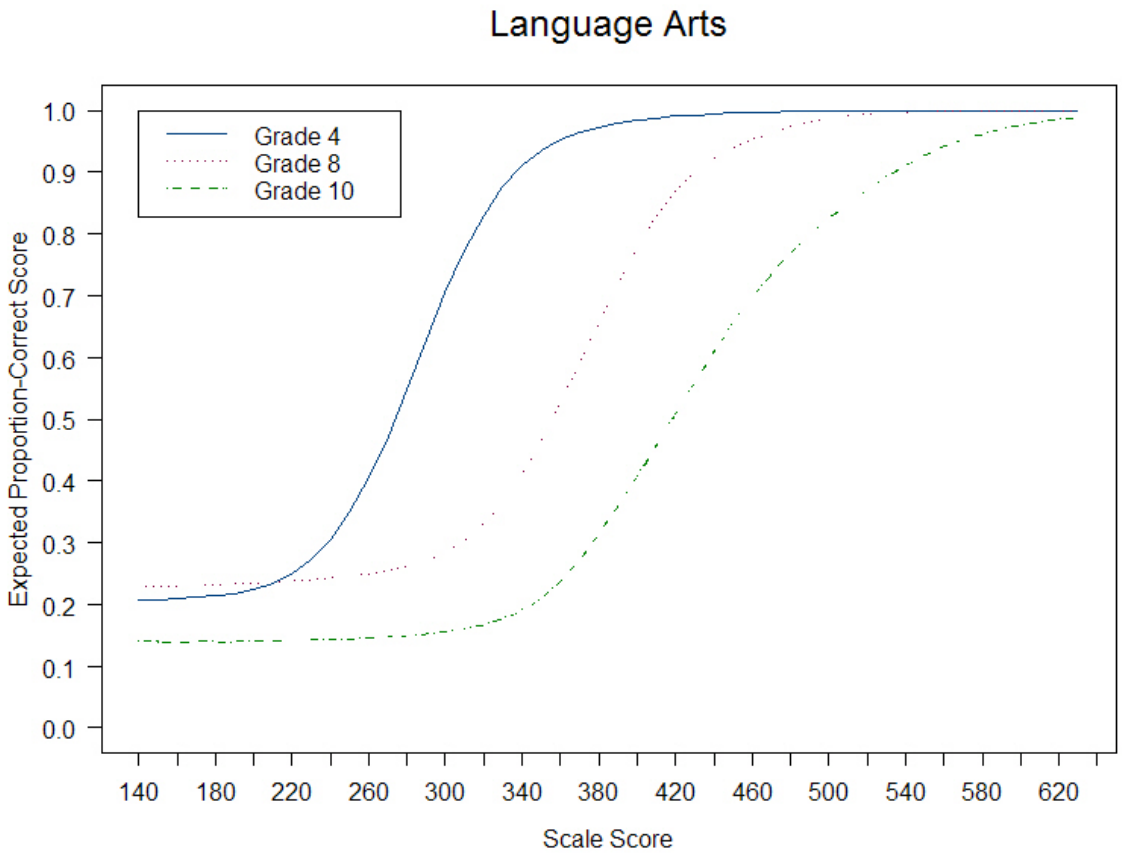
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-7
TCC Curve for Mathematics Grades 3-8, 10



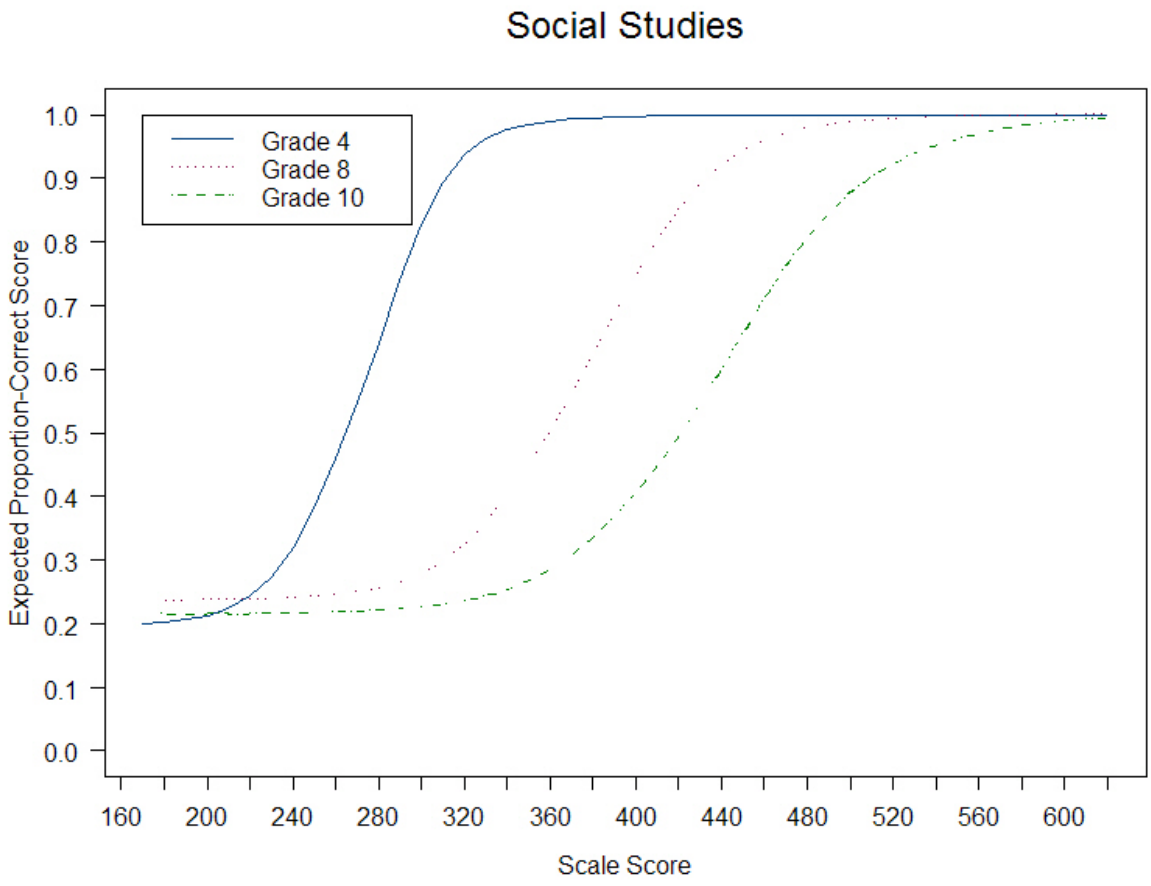
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-8
TCC Curve for Language Arts Grades 4, 8, 10



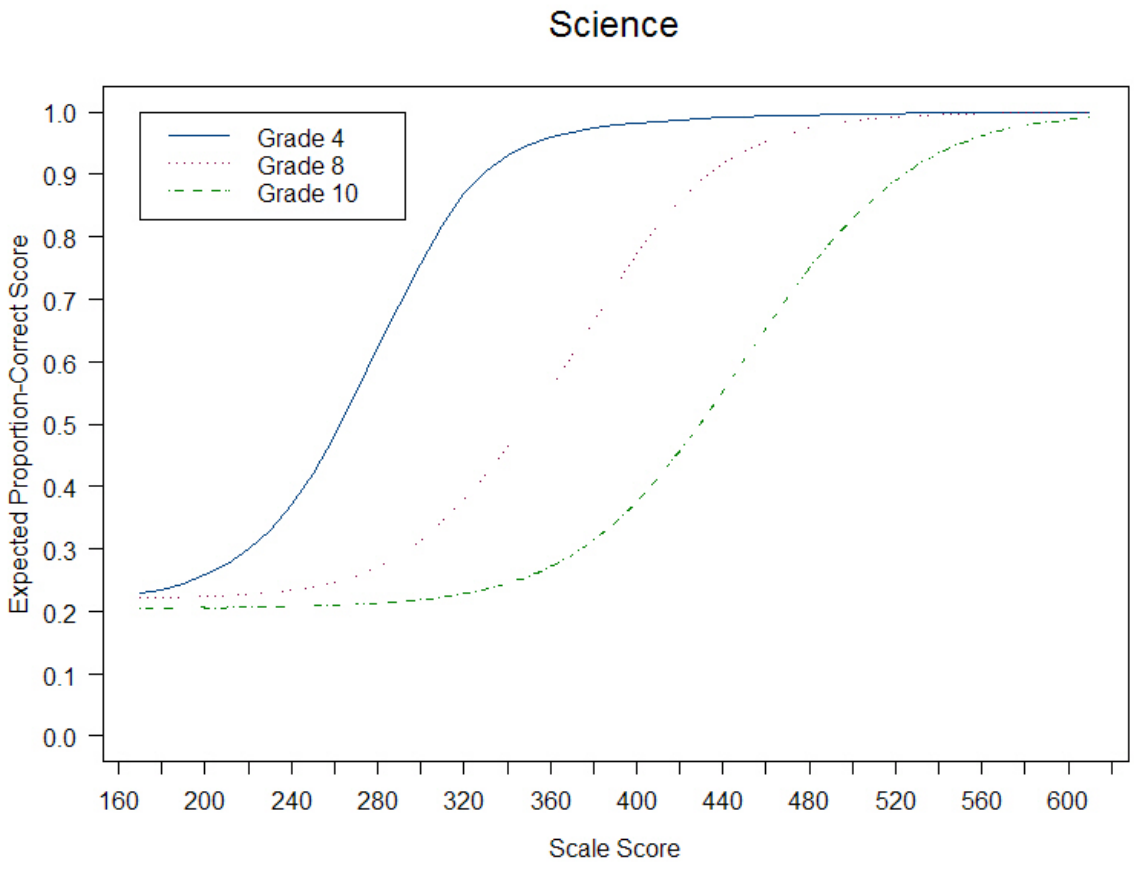
Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-9
TCC Curve for Social Studies Grades 4, 8, 10



Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 7-10
TCC Curve for Science Grades 4, 8, 10



Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 9-1
Reading Indices for Classification Consistency and Classification Accuracy

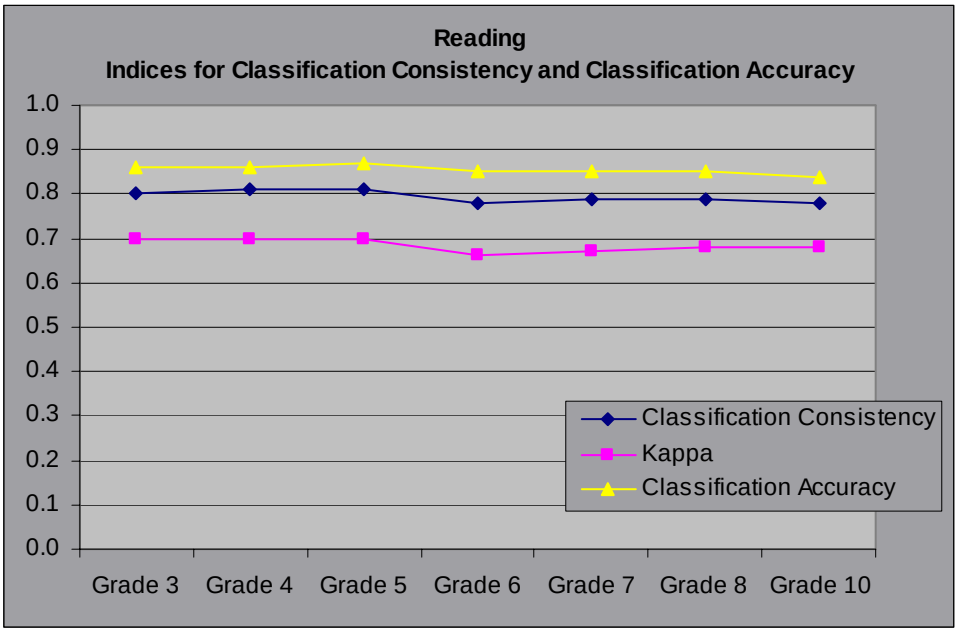
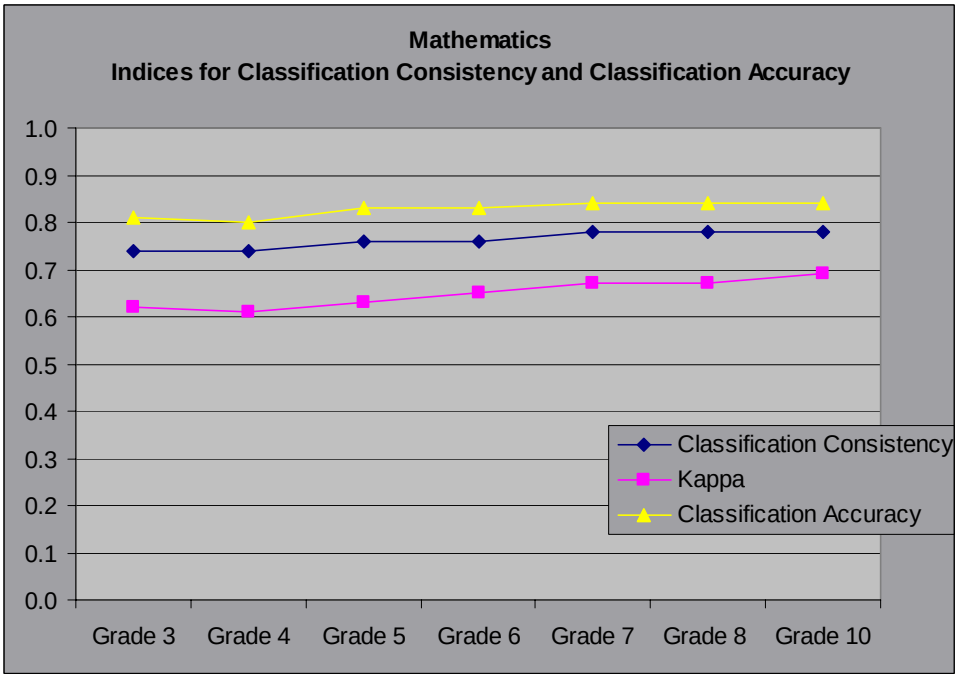


Figure 9-2
Mathematics Indices for Classification Consistency and Classification Accuracy



Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 9-3
Language Arts Indices for Classification Consistency and Classification Accuracy

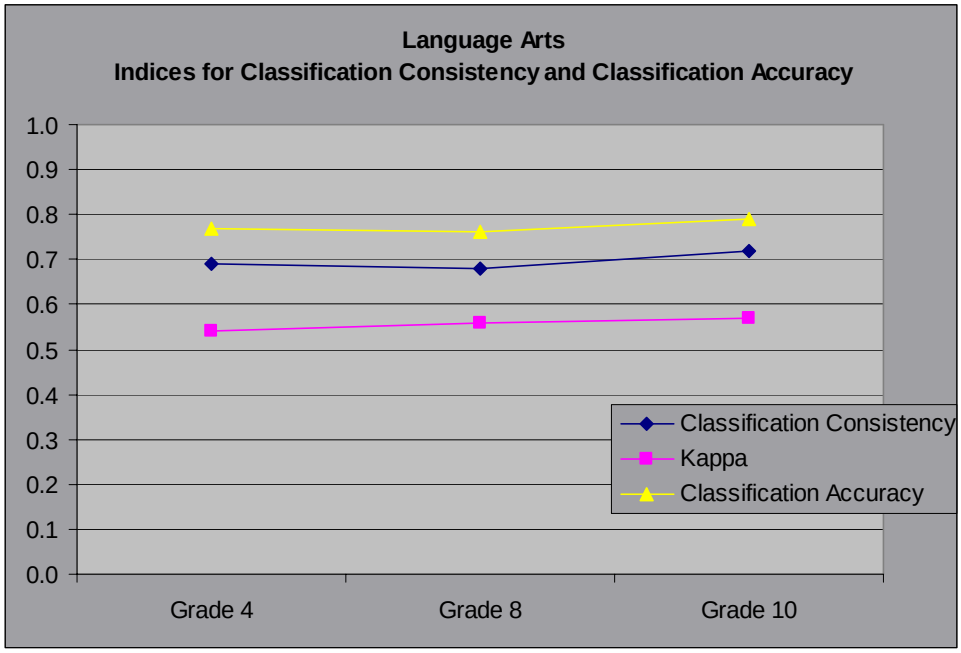
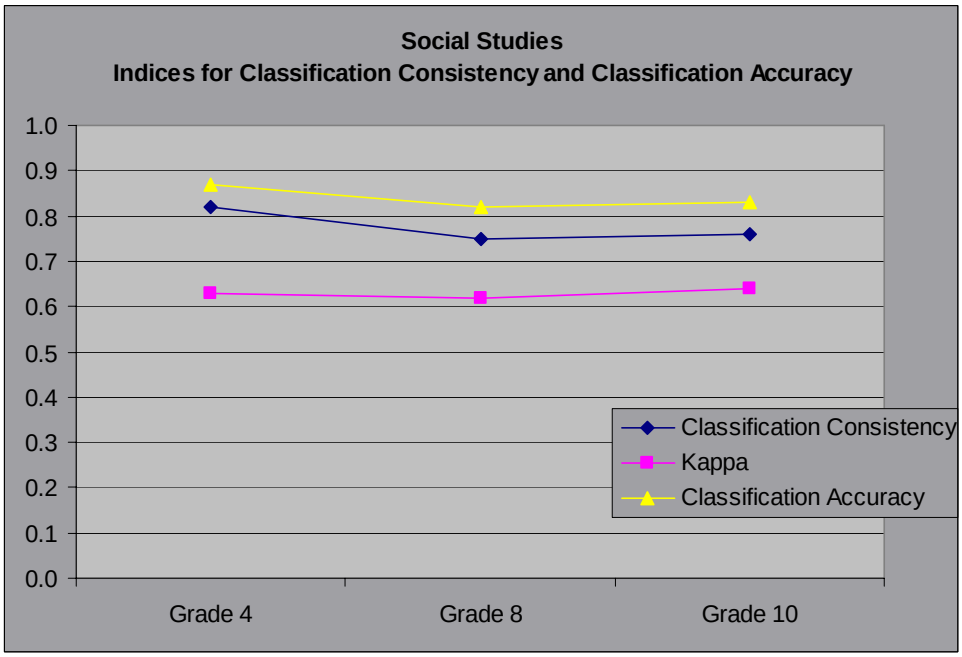
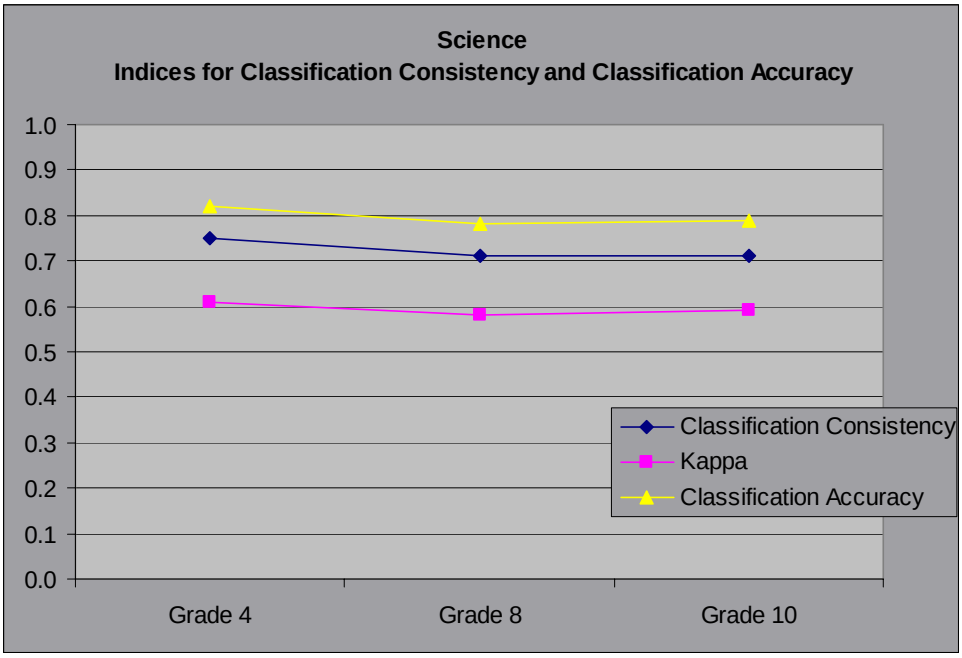


Figure 9-4
Social Studies Indices for Classification Consistency and Classification Accuracy



Copyright © 2008 by the Wisconsin Department of Public Instruction

Figure 9-5
Science Indices for Classification Consistency and Classification Accuracy



Copyright © 2008 by the Wisconsin Department of Public Instruction

Appendix 1: The Form Selection Summary Document

WKCE-CRT Form Selection Summary Document

Content Area:

Grade Level:

Test Description

	Form 2006				Anchor Items				Form 2007			
	No. Items	% No. Items	No. Points	% No. Points	No. Items	% No. Items	No. Points	% No. Points	No. Items	% No. Items	No. Points	% No. Points
SR												
CR												
Total												

Blueprint Comparison (Number of items)

	2006 Blueprint Requirement		2007 Blueprint Requirement		Form 2006		Anchors		Form 2007	
Reporting Category	SR	CR	SR	CR	SR	CR	SR	CR	SR	CR
A										
B										
C										
D										
E										
F										

Blueprint Comparison (% Number of items)

	2006 Blueprint Requirement		2007 Blueprint Requirement		Form 2006		Anchors 07		Form 2007	
Reporting Category	SR	CR	SR	CR	SR	CR	SR	CR	SR	CR
A										
B										
C										
D										
E										
F										

Copyright © 2008 by the Wisconsin Department of Public Instruction

WKCE-CRT Form Selection Summary Document, Cont'd

Form 2007 Number of Items by DOK & Objective

Objective	Obj DoK	Dok Level 1	Dok Level 2	Dok Level 3	Dok Level 4
A					
B					
C					
D					
E					
F					

*Combine SR & CR items

Number of Items on DPI Watch List

	Anchor 07	Form 2007
Number of items		

Number of easy and difficult items for preventing ceiling and floor effect

	Form 2006		**Anchors 07 (SR only)**		**Form 2007**	
	SR	CR	SR	CR	SR	CR
Mean *p*-value						
No. of items: *p*-value < 0.30						
No. of items: 0.30 < *p*-value < 0.40						
No. of items: *p*-value > 0.80						
No. of items: 0.80 < *p*-value < 0.90						

TCCs overlay each other closely

	2006 & Anchor	**Anchor & 2007**	**2006 & 2007**
TCCs of Selected Form			

SE curves are smoothly bow-shaped without dips, bumps, and twists

SE curves of Selected Form	

Expected % Max. RS difference between any two Selected Forms < or = 0.05:

	2006 & Anchor
Max Raw Score Difference	

Copyright © 2008 by the Wisconsin Department of Public Instruction

WKCE-CRT Form Selection Summary Document, Cont'd

Number of Items with DIF

	DIF Code (Ag/Fav)	**2006** Against/favor	**Anchor** Against/favor	**2007** Against/favor
Gender	F			
	M			
Ethnicity	White			
	African American			
	Hispanic			
	Asian			
	American Indian			
ELL	Proficient			
	Not Proficient			
SES	Disadvantaged			
	Not Disadvantaged			
Disability	Disabled			
	Not Disabled			

Number of Items with Poor Fit

	2006	**Anchor**	**2007**
Fit = 3			

Approval

Content Editor	Content Supervisor/Lead	Date	Project Manager	Date	Research Monitor	Date

Copyright © 2008 by the Wisconsin Department of Public Instruction

WKCE-CRT Form Selection Summary Document, Cont'd

Plots

- **2006 Form (Reference) vs. Anchor (Working) Set:**
  (TCC & SE curves & Summary Report)

- **Anchor (Reference) vs. 2007(Working) Set:**
  (TCC & SE curves & Summary Report)

- **2006 (Reference) vs. 2007 (Working) Set:**
  (TCC & SE curves & Summary Report)

- **2006 Form, Anchor Set, 2007 Form TCC & SE Curves:**

Copyright © 2008 by the Wisconsin Department of Public Instruction