# MONTCAS, PHASE 2
# Criterion-Referenced Test

# 2006
# TECHNICAL REPORT

Linda McCulloch, Superintendent
Montana Office of Public Instruction
PO Box 202501
Helena, Montana 59620-2501
www.opi.mt.gov

# TABLE OF CONTENTS

**SECTION I: ASSESSMENT DEVELOPMENT** ..... **3**

CHAPTER 1—BACKGROUND AND OVERVIEW ..... 3
*Purpose of This Report* ..... *3*
*Overview of the Assessment System* ..... *4*
*Options for Participation* ..... *5*
*Brief Summary of Technical Evidence in This Report* ..... *6*
CHAPTER 2—OVERVIEW OF TEST DESIGN ..... 9
*Criterion-Referenced Test (CRT)* ..... *9*
*Item Types* ..... *9*
*Common-Matrix Design* ..... *10*
CHAPTER 3—TEST DEVELOPMENT PROCESS ..... 11
*Criterion-Referenced Test (CRT) Item Development* ..... *11*
*MPSSIP Item Development* ..... *11*
*Item Development Process Overview* ..... *12*
*Montana-Augmented Item Development* ..... *13*
*Montana-Augmented Item Development Process Overview* ..... *14*
*Internal Item Review* ..... *15*
*External Item and Bias Reviews* ..... *15*
*Item Editing* ..... *16*
*Operational Test Assembly* ..... *16*
*Editing Drafts of Operational Tests* ..... *17*
*Test Booklet Design* ..... *18*
*Braille and Large-Print Translation* ..... *18*
CHAPTER 4—DESIGN OF THE READING ASSESSMENT ..... 19
*Reading Test Specifications* ..... *19*
*Item Types* ..... *22*
*Test Design* ..... *23*
CHAPTER 5—DESIGN OF THE MATHEMATICS ASSESSMENT ..... 24
*Mathematics Specifications* ..... *24*
*Content Specifications* ..... *26*
*Item Types* ..... *27*
*Test Design* ..... *27*
*The Use of Calculators in the CRT* ..... *28*

**SECTION II: TEST ADMINISTRATION** ..... **29**

CHAPTER 6—TEST ADMINISTRATION ..... 29
*Responsibility for Administration* ..... *29*
*Procedures* ..... *29*
*Test Administrator Training* ..... *29*
*Participation Requirements* ..... *30*
*Test Scheduling* ..... *31*
*Help Desk* ..... *35*

**SECTION III: DEVELOPMENT AND REPORTING OF SCORES** ..... **36**

CHAPTER 7—SCORING ..... 36

*Scanning* ..... *36*
*Scanning Quality Control* ..... *37*
*Electronic Data Files* ..... *38*
*Items Scored by Readers* ..... *39*
*I-Score* ..... *39*
*Preliminary Activities* ..... *41*
*Planning and Designing Documents* ..... *41*
*Benchmarking* ..... *42*
*Selecting and Training Scoring Staff* ..... *42*
Quality Assurance Coordinators (QACs) and Verifiers ..... 42
Training Quality Assurance Coordinators and Verifiers ..... 43
Selecting Readers ..... 43
Training Readers ..... 45
*Scoring Activities* ..... *45*
*Monitoring Readers* ..... *46*
*General Scoring Guides* ..... *48*
CHAPTER 8—ITEM ANALYSES ..... 50
*Difficulty Indices (p)* ..... *50*
*Item-Test Correlations (Item Discrimination)* ..... *51*
*Summary of Item Analysis Results* ..... *52*
*Differential Item Functioning (DIF)* ..... *58*
*Item Response Theory Analyses* ..... *61*
CHAPTER 9—RELIABILITY ..... 62
*Reliability and Standard Errors of Measurement* ..... *64*
*Reliability of Performance Level Categorization* ..... *68*
*Accuracy* ..... *68*
*Consistency* ..... *68*
*Calculating Accuracy* ..... *68*
*Calculating Consistency* ..... *69*
*Kappa* ..... *69*
*Results of Accuracy, Consistency, and Kappa Analyses* ..... *70*
CHAPTER 10— SCALING AND EQUATING ..... 78
*General Rules* ..... *78*
*IRT Equating* ..... *79*
*Translating Raw Scores to Scaled Scores and Performance Levels* ..... *81*
CHAPTER 11—REPORTING ..... 84
*Local Data Analysis and Interpretation* ..... *85*
CHAPTER 12—VALIDITY SUMMARY ..... 86

**SECTION IV—REFERENCES ..... 89**

APPENDIX A: ITEM PARAMETER FILES ..... 90
APPENDIX B: TECHNICAL ADVISORY COMMITTEE ..... 115
APPENDIX C: CRT STANDARD SETTING REPORT ..... 116
APPENDIX D: CRT PERFORMANCE LEVEL DESCRIPTORS, SCALED SCORES, AND RAW SCORES ..... 177
APPENDIX E: REPORT SHELLS ..... 210
APPENDIX F: REPORTING DECISION RULES ..... 225

# SECTION I: ASSESSMENT DEVELOPMENT

## CHAPTER 1—BACKGROUND AND OVERVIEW

### PURPOSE OF THIS REPORT

In the spring of 2006, Montana students in grades 3 through 8 and 10 participated in the MontCAS, Phase 2 Criterion Referenced Test (CRT) in reading and mathematics in order to measure their reading and mathematics achievement as articulated by the Montana Content Standards and Grade Level Expectations. This represents the third year of the operational CRT program, which will expand next year to include field tests in science (grades 4, 8 and 10).

The purpose of this report is to describe several technical aspects of the CRT in an effort to contribute to the accumulation of validity evidence to support CRT score interpretations. Because it is the interpretations of test scores that are evaluated for validity, not the test itself, this report presents documentation to substantiate intended interpretations (American Educational Research Association (AERA), American Psychological Association & National Council on Measurement in Education, 1999). Subsequent chapters of this report discuss test development, test alignment, test administration, scoring, equating, item analyses, reliability, scaled scores, performance levels and reporting. Each of these topics contributes important information to the validity argument. However, note that certain aspects of a comprehensive validity argument are not included in this report, but could also be important to consider when drawing conclusions about validity. Additional sources of validity evidence might speak to the extent to which scores from the CRT assessments converge with other measures of the same or similar constructs and diverge from measures of different constructs, as well as additional consequences arising from scores at the student, school, district and state levels.

Historically, while some parts of a technical report may have been used by educated laypersons, the intended audience was experts in psychometrics and educational research. This edition of the CRT technical report is an attempt to make the information contained herein more accessible to educated lay people by providing richer descriptions of general categories of information. In making some of the information more accessible we have purposefully preserved the depth of technical information that has historically been provided in our technical reports. The reader will find that some of the discussion and tables continue to require a working knowledge of measurement concepts such as "reliability" and "validity", and statistical concepts such as "correlation" and "central tendency." To fully understand some data, the reader will also have to possess basic familiarity with advanced topics in measurement and statistics.

## Overview of the Assessment System

The MONTCAS Phase 2 assessment program CRTs are designed to measure student acquisition of the knowledge and skills in Montana's content standards for reading and mathematics. The assessments were developed to provide information at the student, class, school, and system level. These assessments are part of NCLB accountability for Montana schools. The results are meant to be useful for program and instructional improvement.

The CRTs are based on, and aligned to, Montana's Content Standards and Grade Level Expectations in Reading and Mathematics. Montana educators worked with OPI and its contractor, Measured Progress, in the development and review (content and bias) of these tests to assess how well students have learned the Montana content standards for their grade. In addition, an independent alignment study was performed by Northwest Regional Educational Laboratory (NWREL) in fall 2005 prior to 2006 test form production. NWREL's alignment study may be found on OPI's Web site www.opi.mt.gov/assessment.

CRT scores are intended to be useful indicators of the extent to which students have mastered material outlined in the Montana reading and mathematics content standards. For a particular student, his/her CRT score should be used as part of a body of evidence regarding mastery and should not be used in isolation to make high stakes decisions. CRT scores, when aggregated to school, system or state levels, are more reliable indicators of program success, particularly when monitored over the course of several years.

**TABLE 1-1: TIMELINE OF MAJOR PROGRAM MILESTONES**

| Milestone | Year | Subjects |
|---|---|---|
| Montana Content Standards adopted by Montana's Board of Education | 1998 | Reading and Mathematics |
| Item development and field test administration of the grades 3 through 8 and 10 CRT Montana-specific items | 2003 | Reading and Mathematics |
| First operational administration of the CRT in grades 4, 8 & 10 | 2004 | Reading and Mathematics |
| Standard Setting for grades 4, 8 and 10 | 2004 | Reading and Mathematics |
| Second operational administration of the CRT in grades 4, 8 & 10 | 2005 | Reading and Mathematics |
| Field test administration in grades 3, 5, 6 and 7 | 2005 | Reading and Mathematics |
| Third operational administration of the CRT in grades 4, 8 & 10; First operational administration of the CRT in grades 3, 5 6 and 7 | 2006 | Reading and Mathematics |
| Standard Setting for grades 3 through 8 and 10 | 2006 | Reading and Mathematics |
| Item development and bias review by Montana educators to prepare for science field test in spring 2007 | 2006 | Science |

## OPTIONS FOR PARTICIPATION

All Montana students enrolled in accredited schools are expected to participate in either the CRT or the CRT Alternate assessment (CRT-ALT). The vast majority of students will participate in the CRT, and most of them will participate under standard administration procedures. However, there is an array of standard accommodations which are available to any student, with or without disabilities, when such accommodations are necessary to allow the student to demonstrate his/her skills and competencies. Standard accommodations are not considered to change the construct being measured and may be

provided to students for either the reading or math portions of the assessment, or both, as necessary. Student's tests are scored the same way regardless of whether or not they took the test using standard accommodations.

In addition to standard accommodations, other accommodations for the CRT are available to a student when specified in his/her IEP, 504, or LEP plan. These other accommodations are referred to as non-standard accommodations and, because they are considered to alter the construct being measured, affect the student's score on the CRT. When a non-standard accommodation is used, the student's score will be reported as the lowest possible score (i.e., a scaled score of 200 which falls into the Novice performance level) for that content area. Non-standard accommodations on the CRT may be provided in reading or math, or both, as dictated by the student's IEP, 504, or LEP plan.

For a very small percentage of students, participation in the statewide assessment program will be achieved by participating in the CRT-ALT. Students with significant cognitive disabilities who are working toward alternate academic achievement standards, as documented in their IEP plans, are eligible to take the CRT-ALT. Technical characteristics of the CRT-ALT program are described in a companion technical report.

## Brief Summary of Technical Evidence in This Report

The *Standards for Educational and Psychological Testing* (AERA et al, 1999) provides a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five general areas: test content, response processes, internal structure, relationship to other variables, and consequences of testing. Although each of these sources may speak to a different *aspect* of validity, they are not distinct *types* of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

This manual does not attempt to give a thorough treatment regarding all evidence of validity for the MONTCAS assessments. Rather this manual provides discussion of the

work done by the testing contractor that contributes to the collection of validity evidence. Other evidence of validity that could further contribute to the validity argument is outside of the scope of the contract between Measured Progress and Montana OPI.

Viewed through the lens provided by the Standards, evidence based on test content is extensively described in Chapters 2 through 6. Item alignment with Montana content standards; item bias, sensitivity and content appropriateness review processes; adherence to the test blueprint; use of multiple item types; use of standardized administration procedures, with accommodated options for participation; and appropriate test administration training are all components of validity evidence based on test content.

The scoring information in Chapter 7 describes the steps taken to train and monitor hand-scorers, as well as quality control procedures related to scanning and machine scoring. This evidence describes the quality assurance measures taken during scoring of constructed response items to ensure that results are valid.

Technical characteristics of the internal structure of the assessments are presented in the discussions of item analyses in Chapter 8 in terms of classical item statistics (item difficulty, item-test correlations), differential item functioning analyses, a variety of reliability coefficients, standard errors of measurement, and item response theory parameters and procedures. This section will contribute to the evidence of strong internal structure of the assessments.

Some evidence based on the consequences of testing is addressed in the scaled scores, equating, and reporting information in Chapters 10 and 11, as well as in the test interpretation guide, which is a separate document that is referenced in the discussion of reporting. Each of these chapters speaks to the efforts undertaken to promote accurate and clear information provided to the public regarding test scores.

With this introduction to a conceptual understanding of how the information presented in this report contributes to an overarching validity argument in mind, the reader should be

in position to organize the extensive detail contained in the following chapters. The organization of this report is based on the conceptual flow of an assessment cycle. The report begins with the initial test specification and addresses all the intermediate steps that lead to final score reporting.

# CHAPTER 2—OVERVIEW OF TEST DESIGN

## CRITERION-REFERENCED TEST (CRT)

Items on the CRT originate from the Measured Progress State Secure Item Pool (MPSSIP) and Montana-augmented item banks (see Chapter 3) and are directly linked to **Montana's Content Standards**. The content standards are the basis for the reporting categories developed for each subject area and are used to help guide the development of test items. No other content or process is subject to statewide assessment. An item may address part, all, or several of the benchmarks within a standard.

## ITEM TYPES

Montana's educators and students were familiar with most of the item types that were used in the assessment program. The types of items used and the functions of each are described below.

**Multiple-choice items** were used, in part, to provide breadth of coverage of a content area. Because they require no more than a minute for most students to answer, these items make efficient use of limited testing time and allow coverage of a wide range of knowledge and skills.

**Short-answer items** were used to assess students' skills and their abilities to work with brief, well-structured problems that had one or a very limited number of solutions (e.g., mathematical computations). Short-answer items require approximately two minutes for most students to answer. The advantage of this type of item is that it requires students to demonstrate knowledge and skills by generating, rather than merely selecting, an answer.

**Constructed-response items** typically require students to use higher-order thinking skills—evaluation, analysis, summarization, and so on—in constructing a satisfactory response. Constructed-response items should take most students approximately five to ten minutes to complete. It should be noted that the use of released CRT items to prepare students to answer this kind of item is appropriate and encouraged.

## COMMON-MATRIX DESIGN

The CRT measures what students know and are able to do by using a variety of item types. The tests are structured using both **common** and **matrix-sampled** items. Common items are those taken by all students at a given grade level. Students' scores are based only on common items. In addition, a larger pool of matrix-sampled items is divided among the sixteen forms of the test at each grade level. Each student takes only one form of the test and so answers a fraction of the matrix-sampled items in the entire pool. The matrix-sampled items (field test items) were invisible to test takers and had a negligible impact on testing time. Because the field test were randomly distributed, it provided the sample size needed to produce reliable data (750-1500 students per item as some items were repeated across forms) on which to inform item selection for future tests.

The CRT reports were delivered to schools on September 15, 2006. In addition, common items were released via OPI's assessment Web site and with a data management tool called *iAnalyze* (see Chapter 11: "Reporting" and Appendix E: Report Shells.)

# CHAPTER 3—TEST DEVELOPMENT PROCESS

## CRITERION-REFERENCED TEST (CRT) ITEM DEVELOPMENT

As previously mentioned, items in the CRT are derived from either the Measured Progress State Secure Item Pool (MPSSIP) or a Montana-augmented item bank. The item development process for both item banks is similar and is discussed in greater detail in this chapter.

## MPSSIP ITEM DEVELOPMENT

The items developed for the Measured Progress State Secure Item Pool (MPSSIP) and forms were consistent with national and Montana Content Standards. Measured Progress curriculum and assessment specialists worked with Montana educators to verify the alignment of items to the appropriate Montana Content Standards. As an additional quality control check, Northwest Regional Educational Laboratory (NWREL) performed an independent alignment study to verify item alignment to Montana Content Standards.

The development process Measured Progress followed combined the expertise of the item development team and a panel of educators nationwide to help ensure that these items met the needs of the core MPSSIP program and the CRT program. All items used in the MPSSIP common and matrix portions of the CRT program underwent review by a national panel of content and bias reviewers. This panel included numerous Montana educators. Annual MPSSIP item development is depicted in the following tables:

**TABLE 3-1: TOTAL NUMBER OF MPSSIP ITEMS DEVELOPED PER YEAR**

| GRADE | READING | MATH |
|---|---|---|
| 3 | 168 | 78 |
| 4 | 168 | 78 |
| 5 | 168 | 78 |
| 6 | 168 | 78 |
| 7 | 168 | 78 |
| 8 | 168 | 78 |
| 10 | 168 | 78 |

**TABLE 3-2: ANNUAL MPSSIP READING ITEM DEVELOPMENT GRADES 3 - 8 & 10**

| Passages | Multiple Choice | Constructed Response |
|---|---|---|
| 2 long literary passages | 40 | 4 |
| 2 long informational passages | 40 | 4 |
| 4 short literary passages | 40 | 0 |
| 4 short informational passages | 40 | 0 |
| **12** | **160** | **8** |

**TABLE 3-3: ANNUAL MPSSIP MATH ITEM DEVELOPMENT GRADES 3 - 8 & 10**

| Multiple Choice | Short Answer | Constructed Response |
|---|---|---|
| 68 | 4 | 6 |

## ITEM DEVELOPMENT PROCESS OVERVIEW

An overview of the test development process for the common and matrix items, including conducting the field tests, follows.

**TABLE 3-4: DEVELOPMENT PROCESS OVERVIEW**

| DEVELOPMENT STEP | DESCRIPTION |
|---|---|
| Select reading passages and conduct external review for bias and sensitivity issues (2005) | • Measured Progress Curriculum and Assessment Specialists located potential reading passages.<br>• Reading passages were reviewed for bias and sensitivity issues before the development of reading item sets. |
| Develop items (January through May 2005) | • Measured Progress Curriculum and Assessment Specialists developed reading item sets and mathematics items. |
| National item review for bias and sensitivity issues and content appropriateness (summer 2005) | • Panels of national educators reviewed newly-developed reading and mathematics items for bias and sensitivity issues and content appropriateness. |

| | |
|---|---|
| Edit items (summer 2005) | All items reviewed by national committee members were edited to assure<br>• clarity and unambiguousness of items.<br>• correct grammar, punctuation, usage, and spelling.<br>• technical quality with respect to stems, options, and scoring guides. |
| Montana educators review items for bias and sensitivity issues and content appropriateness (Sept/Oct. 2005) | • Panels of Montana educators reviewed and selected 2006 common items and reviewed and edited field test items for 2006 forms. |
| Field test items (spring 2006) | • Embedded matrix (field test) items were administered to a sample of students (minimum of 1,500 students per item/16 forms per grade and content). |
| Item Review and Selection Meeting (summer 2006) | • Measured Progress test developers and Montana educators reviewed the results of the Spring 2006 field test and selected MPSSIP common items for the Spring 2007 operational CRT forms. |

## MONTANA-AUGMENTED ITEM DEVELOPMENT

The items developed for the augmented CRT item bank were aligned with Montana's content standards. Measured Progress's development specialists worked with OPI and Montana educators to align the items developed to augment the CRT to appropriate Montana content standards. As an additional quality control check, lead developers in each content area checked for their agreement that each item was appropriately aligned. Where there were any apparent discrepancies, lead Curriculum and Assessment specialists resolved them with OPI personnel.

The development process Measured Progress followed, combining the expertise of the item development team and Montana educators, helped ensure that these items met the needs of the CRT program. The item specifications were built on the Montana content standards, thus assuring complete alignment between the content standards and the augmented portion of the CRT. In addition to internal review, all test materials and items used in the CRT program underwent review by Montana educators and bias review committees prior to print. Table 3-5 depicts the number of items developed and field tested in 2002-2003 to support the program's item bank 2004 through 2007.

**TABLE 3-5: TOTAL NUMBER OF MONTANA-AUGMENTED ITEMS DEVELOPED AND FIELD TESTED BY GRADE AND CONTENT (ALL MULTIPLE CHOICE ITEMS)**

| GRADE | READING | MATH |
|---|---|---|
| 3 | 60 | 60 |
| 4 | 100 | 100 |
| 5 | 60 | 60 |
| 6 | 60 | 60 |
| 7 | 60 | 60 |
| 8 | 100 | 100 |
| 10 | 150 | 150 |

## MONTANA-AUGMENTED ITEM DEVELOPMENT PROCESS OVERVIEW

An overview of the test development process for the Montana-augmented item bank, including conducting the field tests, follows.

**TABLE 3-6: DEVELOPMENT PROCESS OVERVIEW**

| DEVELOPMENT STEP | DESCRIPTION |
|---|---|
| Review by Montana educators of passages for the reading tests (Aug. 2002) | • Measured Progress Curriculum and Assessment reading specialists located potential reading passages.<br>• Montana educators approved the passages in consultation with a Montana Bias Review Committee prior to item writing.<br>• Measured Progress Permissions staff secured permissions to use the passages prior to item writing meetings. |
| Item drafting/editing meetings (Sept. 2002) | Measured Progress Curriculum and Assessment specialists<br>• provided item development training to Montana participants;<br>• facilitated the development of item ideas by the participants. |
| Editorial review of items (Oct. 2002) | All items were reviewed by members of Measured Progress's Publications staff to ensure<br>• clarity and unambiguousness of items;<br>• correct grammar, punctuation, usage, and spelling;<br>• technical quality with respect to stems, options, and scoring guides;<br>• compliance with OPI sensitivity standards and style guidelines. |
| Item review meetings (Nov. 2002) | Curriculum and Assessment Specialists facilitated the review of all items with Montana educators and selected appropriate items for field testing in 2003. |
| Bias Review Committee meetings (Nov. 2002) | Measured Progress staff facilitated the review of all test items for sensitivity and bias considerations based on OPI guidelines. Members of this committee were selected by OPI. Measured Progress provided OPI with guidelines for committee membership. |
| Field Test of MT-Augmented Items (April 2003) | Measured Progress provided field test forms which were administered to a sample of students in Montana prior to use of the items in operational assessment to assure quality of items. |

| | |
|---|---|
| Final Item Selection (August 2003) | Measured Progress provided the reports necessary for Montana educators to review the results of field-testing, revise as necessary, and select items for the augmented portion of the assessment. |

## INTERNAL ITEM REVIEW

The lead or peer Curriculum and Assessment Specialist within the content specialty reviewed each item for:

- item "integrity", item content and structure, appropriateness to designated content area, item format, clarity, possible ambiguity, keyability, single "keyness", appropriateness and quality of reading selections and graphics, and appropriateness of scoring guide descriptions and distinctions (as correlated to the item and within the guide itself).
- scorability and evaluated whether the scoring guide adequately addressed performance on the item.
- fundamental issues including the following:
  - What is the item asking?
  - Is the key the only possible key?
  - Is the constructed-response item scorable as written (are the correct words used to elicit the response defined by the guide)?
  - Is the wording of the scoring guide appropriate and parallel to the item wording?
  - Is the item complete (i.e., with scoring guide, content codes, key, grade level, and contract identified)?
  - Is the item appropriate for the designated grade level?

## EXTERNAL ITEM AND BIAS REVIEWS

All MPSSIP and Montana-augmented items undergo the following external reviews:

- In July 2005, MPSSIP National Bias and Content Review Committees reviewed common and matrix passages and items used for the 2006 administration during two, two-day meetings, held in Chicago, IL.
- In early October 2005, common item sets were reviewed by Measured Progress content specialists and Montana educators. Feedback from the Montana content and bias reviews were incorporated into the final editing processes.

## ITEM EDITING

Editors reviewed and edited the items to ensure uniform style (based on *The Chicago Report of Style, 15th Edition*) and adherence to sound testing principles. These principles included the stipulation that items

- were correct with regard to grammar, punctuation, usage, and spelling;
- were written in a clear, concise style;
- contained unambiguous explanations for students as to what was required to attain a maximum score;
- were written at a reading level that would allow the student to demonstrate his or her knowledge of the tested subject matter regardless of reading ability;
- exhibited high technical quality regarding psychometric characteristics;
- had appropriate answer options or score-point descriptors; and
- were free of potentially insensitive content.

## OPERATIONAL TEST ASSEMBLY

Test assembly is the sorting and laying out of item sets into test forms. Criteria considered during this process included the following:

- **Content coverage/match to test design**. The curriculum specialist completed an initial sorting of items into sets based on a balance of content categories across sessions and forms, as well as a match to the test design (e.g., number of multiple-choice, short-answer, and constructed-response items).
- **Item difficulty and complexity**. Item statistics drawn from the data analysis of previously tested items were used to ensure that there were similar levels of difficulty and complexity across forms.
- **Visual balance**. Item sets were reviewed to ensure that each reflected a similar length and "density" of selected items (e.g., length/complexity of reading selections or number of graphics).
- **Option balance**. Each item set was checked to verify that it contained a roughly equivalent number of key options (As, Bs, Cs, and Ds).
- **Name balance**. Item sets were reviewed to ensure that a diversity of names was used.

- **Bias**. Each item set was reviewed to ensure fairness and balance based on gender, ethnicity, religion, socioeconomic status, and other factors.
- **Page fit**. Item placement was modified to ensure the best fit and arrangement of items on any given page.
- **Facing-page issues**. For multiple items associated with a single stimulus (a graphic or a reading selection), consideration was given to whether those items needed to begin on a left- or right-hand page, as well as to the nature and the amount of material that needed to be placed on facing pages. These considerations served to minimize the amount of page flipping required of the students.
- **Relationships between forms**. Sets of common items were placed identically in each version of the forms. Although matrix-sampled item sets differed from form to form, they took up the same number of pages in each form so that sessions and content areas began on the same page in every form. Therefore, the number of pages needed for the longest form often determined the layout of each form.
- **Visual appeal**. The visual accessibility of each page of the form was always taken into consideration, including such aspects as the amount of white space, the density of the text, and the number of graphics.

## EDITING DRAFTS OF OPERATIONAL TESTS

Any changes made during the test construction had to be reviewed and approved by the Curriculum and Assessment Specialist. Once a form had been laid out in what was considered its final form, it was reread to identify any final considerations, including the following:

- **Editorial changes**. All text was scrutinized for editorial accuracy, including consistency of instructional language, grammar, spelling, punctuation, and layout. Measured Progress's publishing standards are based on *The Chicago Report of Style, 15th Edition*.
- **Keying items**. Items were reviewed for any information that might "key" or provide information that would help students answer another item. Decisions about moving keying items were based on the severity of the key-in and the placement of the items in relation to each other within the form.
- **Key patterns**. The final sequence of keys was reviewed to ensure that the order appeared random (i.e., no recognizable pattern and no more than three of the same key in a row).

## TEST BOOKLET DESIGN

In order to accommodate the embedded field test design, sixteen versions of the test were administered in grade 3 through 8 and 10.

## BRAILLE AND LARGE-PRINT TRANSLATION

*Form One* for grades 3 through 8, and 10 tests was translated into Braille by National Braille Press, a subcontractor that specializes in test materials for blind and visually impaired students. In addition, *Form One* for each grade was adapted into a large-print version.

# CHAPTER 4—DESIGN OF THE READING ASSESSMENT

## READING TEST SPECIFICATIONS

As indicated earlier, the test blueprint/specifications for reading was based on MPSSIP and Montana's reading content standards, which identifies five **Montana Content Standards** that apply specifically to reading and reading comprehension. Those content standards follow:

- **Reading Standard 1:** Students construct meaning as they comprehend, interpret, and respond to what they read.
- **Reading Standard 2:** Students apply a range of skills and strategies to read.
- **Reading Standard 3:** Students set goals, monitor, and evaluate their reading progress. (Cannot measure this benchmark with traditional paper/pencil test.)
- **Reading Standard 4:** Students select, read, and respond to print and non-print material for a variety of purposes.
- **Reading Standard 5:** Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audiences.

The reading assessments are designed to measure students reading achievement using a context of passages. Each test has a combination of long and short passages, as shown in Tables 4-1 and 4-2 on the following page.

### TABLE 4-1: GRADES 3-8 READING TEST SPECIFICATIONS

| READING GRADES 3-8 (PER FORM) | |
|---|---|
| **Passages** | **Number of items** |
| Session 1 Common | |
| Short passage | 5 MC |
| Short passage | 5 MC |
| Long passage | 11 MC, 1 CR |
| **Session total** | **21 MC, 1 CR** |
| Session 2 Montana-specific common and embedded matrix field test | |
| Montana-specific passage (common) | 10 MC |
| Embedded long passage (field test) | 6 MC, 1 CR |
| Embedded short passage (field test) | 6 MC |
| **Session total** | **22 MC, 1 CR** |
| Session 3 Common | |
| Short passage | 5 MC |
| Short passage | 5 MC |
| Long passage | 11 MC, 1 CR |
| **Session total** | **21 MC, 1 CR** |
| **Common total** | **52 MC, 2 CR** |

### TABLE 4-2: GRADE 10 READING SPECIFICATIONS

| READING GRADE 10 (PER FORM) | |
|---|---|
| **Passages** | **Number of items** |
| Session 1 Common | |
| Short passage | 5 MC |
| Short passage | 5 MC |
| Long passage | 11 MC, 1 CR |
| **Session total** | **21 MC, 1 CR** |
| Session 2 Montana-specific common and embedded matrix (field test) | |
| Montana-specific passage (common) | 15 MC |
| Embedded long passage (field test) | 6 MC, 1 CR |
| Embedded short passage (field test) | 6 MC |
| **Session total** | **27 MC, 1 CR** |
| Session 3 Common | |
| Short passage | 5 MC |
| Short passage | 5 MC |
| Long passage | 11 MC, 1 CR |
| **Session total** | **21 MC, 1 CR** |
| **Common total** | **57 MC, 2 CR** |

**Key**

- MC = multiple-choice items
- CR = constructed-response items

Passages included both long and short texts selected from reading sources that students at each grade level would be likely to encounter in their classroom and in their independent reading. No passages were written specifically for the assessment, but instead were collected from published works. Each passage is classified as one of three types described below.

- **Literary passages** are represented by a variety of genres—modern narratives; diary entries; drama; poetry; biographies; essays; excerpts from novels; short stories; and traditional narratives, such as fables, myths, and folktales.
- **Content passages** are primarily informational and often deal with the areas of science and social studies. They are drawn from such sources as newspapers, magazines, and books.
- **Practical passages** are functional materials that instruct or advise the reader—for example, directions, reference tools, or reports.

The main difference in the passages used for grades 3 – 8, and 10 was their degree of difficulty. All passages were selected to be appropriate for the intended audience; however, the ideas expressed became increasingly more complex at grade levels 8 and 10.

The items related to these passages required students to demonstrate their skills in both literal comprehension, where the answer is stated explicitly in the text, and inferential comprehension, where the answer is implied by the text and/or the text must be connected to relevant prior knowledge to determine an answer. All items focused reading skills reflected in Montana's Reading Content Standards. Items of this type required students to use reading skills and strategies to answer items—for example, how to identify the author's principal purpose, such as to persuade, entertain, or inform—and to demonstrate their understanding of how words and images communicate to readers. Table 4-3 depicts passage distribution, length, and reporting categories.

**TABLE 4-3: PASSAGE DISTRIBUTION**

| Reading Passage Distribution | | | |
|---|---|---|---|
| Literary | | 50% | 25 points |
| Informational | Comprised of both content and practical passages | 50% | 25 points |
| | | 100% | 50 points |
| **Reading Passage Length** | | | |
| Long* | Either a literary or informational per session | 50% | 25 points |
| Short* | At least one literary and informational per session | 50 % | 25 points |
| | | 100% | 50 points |
| **Reporting Categories** | | | |
| Comprehension and Analysis | | 70% | 35 points |
| Reading Process and Skills | | 30 % | 15 points |
| | | 100 % | 50 points |

## ITEM TYPES

The CRT assessments in reading include a mix of multiple-choice and constructed-response items. Constructed-response items required students to write an answer consisting of several phrases or short sentences. Each type of item was worth a specific number of points in the student's total reading score as shown in Table 4-4.

**TABLE 4-4: ITEM TYPES**

| Type of Item | Possible Score Points |
|---|---|
| Multiple-Choice | 0 or 1 |
| Constructed-Response | 0,1, 2, 3, or 4 |

## TEST DESIGN

Table 4-5 summarizes the number and types of common reading items and shows the placement of the common portions of the assessment.

### TABLE 4-5: COMMON READING ITEMS

| | | | | TOTAL | |
|---|---|---|---|---|---|
| **Grade** | **Session 1** | **Session 2** | **Session 3** | **MC** | **CRs** |
| 3 | 21 MC, 1 CR | 10 MC | 21 MC, 1 CR | 52 | 2 |
| 4 | 21 MC, 1 CR | 10 MC | 21 MC, 1 CR | 52 | 2 |
| 5 | 21 MC, 1 CR | 10 MC | 21 MC, 1 CR | 52 | 2 |
| 6 | 21 MC, 1 CR | 10 MC | 21 MC, 1 CR | 52 | 2 |
| 7 | 21 MC, 1 CR | 10 MC | 21 MC, 1 CR | 52 | 2 |
| 8 | 21 MC, 1 CR | 10 MC | 21 MC, 1 CR | 52 | 2 |
| 10 | 21 MC, 1 CR | 15 MC | 21 MC, 1 CR | 57 | 2 |

**Key**

- MC = multiple-choice items
- CR = constructed-response items

# CHAPTER 5—DESIGN OF THE MATHEMATICS ASSESSMENT

## MATHEMATICS SPECIFICATIONS

Mathematics specifications/blueprint is based on Montana's Mathematics Content Standards, which identifies seven standards:

- **Mathematics Standard 1:** Problem Solving
- **Mathematics Standard 2:** Numbers and Operations
- **Mathematics Standard 3:** Algebra
- **Mathematics Standard 4:** Geometry
- **Mathematics Standard 5:** Measurement
- **Mathematics Standard 6:** Data Analysis, Statistics, and Probability
- **Mathematics Standard 7:** Patterns, Relations, and Functions

### TABLE 5-1: MATHEMATICS SPECIFICATIONS/BLUEPRINT

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Test Design:** | **55** multiple-choice items<br>3 1-point short-answer items<br>2 4-point constructed-response items<br>Total points: 66 | | | | | | |
| **Percent Point distribution by content standard*** | | | | | | | |
| **MPSSIP Standards** | **Grade 3** | **Grade 4** | **Grade 5** | **Grade 6** | **Grade 7** | **Grade 8** | **Grade 10** |
| **Number and Operations** | 32% | 32% | 32% | 32% | 30% | 20% | 20% |
| **Algebra** | 20% | 20% | 20% | 20% | 20% | 29% | 27% |
| **Geometry** | 16% | 16% | 16% | 16% | 16% | 18% | 23% |
| **Measurement** | 13% | 13% | 13% | 13% | 14% | 14% | 11% |
| **Data Analysis/Probability** | 20% | 20% | 20% | 20% | 20% | 20% | 20% |
| *Because percents are rounded to the nearest whole number, not all sums add to 100%.<br>**Note:** Geometry and Measurement comprise a single reporting category. | | | | | | | |
| | | | | | | | |
| **Point distribution by content standard** | | | | | | | |
| | **Grade 3** | **Grade 4** | **Grade 5** | **Grade 6** | **Grade 7** | **Grade 8** | **Grade 10** |
| **Number and Operations** | 18 | 18 | 18 | 18 | 17 | 11 | 11 |
| **Algebra** | 11 | 11 | 11 | 11 | 11 | 16 | 15 |
| **Geometry** | 9 | 9 | 9 | 9 | 9 | 10 | 13 |
| **Measurement** | 7 | 7 | 7 | 7 | 8 | 8 | 6 |
| **Data Analysis/Probability** | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| **Four-point items:** Each test contains two 4-point constructed-response items. In any given year, the two items will measure two different standards. From year to year, those standards may change. | | | | | | | |

**One-point items:** There are two types of one-point items: multiple-choice and short answer items. Each test contains 55 multiple-choice items and three short-answer items. The number of one-point items per standard will vary from year to year depending on which two standards are measured by the four-point items. (The number of total points per standard is kept constant from year to year.)

| Number of 1-point items per content standard | | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Grade 3** | **Grade 4** | **Grade 5** | **Grade 6** | **Grade 7** | **Grade 8** | **Grade 10** |
| **Number and Operations** | 14 or 18 | 14 or 18 | 14 or 18 | 14 or 18 | 13 or 17 | 7 or 11 | 7 or 11 |
| **Algebra** | 7 or 11 | 7 or 11 | 7 or 11 | 7 or 11 | 7 or 11 | 12 or 16 | 11 or 15 |
| **Geometry** | 5 or 9 | 5 or 9 | 5 or 9 | 5 or 9 | 5 or 9 | 6 or 10 | 9 or 13 |
| **Measurement** | 3 or 7 | 3 or 7 | 3 or 7 | 3 or 7 | 4 or 8 | 4 or 8 | 2 or 6 |
| **Data Analysis/Probability** | 7 or 11 | 7 or 11 | 7 or 11 | 7 or 11 | 7 or 11 | 7 or 11 | 7 or 11 |

**Distribution of One-Point Items by Standard**

The distribution of one-point items within a standard is partially dependent on the specific items selected for a given test. However, a minimal number of one-point items per standard has been established. Those numbers are shown in the table below.

| Minimum Number of 1-Point Items Per Standard | | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Grade 3** | **Grade 4** | **Grade 5** | **Grade 6** | **Grade 7** | **Grade 8** | **Grade 10** |
| **Number and Operations** | | | | | | | |
| ***Total Number of points*** | ***18*** | ***18*** | ***18*** | ***18*** | ***17*** | ***11*** | ***11*** |
| Number concepts | 4 | 3 | 2 | 3 | 3 | 2 | 2 |
| Meanings of operations | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Computation/estimation | 4 | 5 | 6 | 5 | 4 | 2 | 2 |
| *Floating points* | *5 or 9* | *5 or 9* | *5 or 9* | *5 or 9* | *5 or 9* | *2 or 6* | *2 or 6* |
| | | | | | | | |
| **Algebra** | | | | | | | |
| ***Total Number of points*** | ***11*** | ***11*** | ***11*** | ***11*** | ***11*** | ***16*** | ***15*** |
| Patterns | 3 | 2 | 2 | 1 | 1 | 1 | 1 |
| Algebraic symbols | 1 | 1 | 1 | 2 | 2 | 4 | 4 |
| Mathematical models | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Change | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *Floating points* | *1 or 5* | *2 or 6* | *2 or 6* | *2 or 6* | *2 or 6* | *5 or 9* | *4 or 8* |
| | | | | | | | |
| **Geometry** | | | | | | | |
| ***Total Number of points*** | ***9*** | ***9*** | ***9*** | ***9*** | ***9*** | ***10*** | ***13*** |
| Properties of 2-and 3-d shapes | 2 | 2 | 2 | 2 | 2 | 2 | 3 |
| Coordinate Geometry | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Transformations/symmetry | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Visualization/spatial reasoning | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *Floating points* | *0 or 4* | *0 or 4* | *0 or 4* | *0 or 4* | *0 or 4* | *1 or 5* | *3 or 7* |
| | | | | | | | |
| **Measurement** | | | | | | | |
| ***Total Number of points*** | ***7*** | ***7*** | ***7*** | ***7*** | ***8*** | ***8*** | ***6*** |
| Concepts of measurement | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Techniques, tools, formulas | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *Floating points* | *1 or 5* | *1 or 5* | *1 or 5* | *1 or 5* | *2 or 6* | *2 or 6* | *0 or 4* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Data Analysis/Probability** | | | | | | | |
| ***Total Number of points*** | **11** | **11** | **11** | **11** | **11** | **11** | **11** |
| Collect/organize/display data | 2 | 2 | 2 | 1 | 1 | 1 | 1 |
| Statistical methods | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Inferences/predictions | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Probability | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *Floating points* | *2 or 6* | *2 or 6* | *2 or 6* | *3 or 7* | *3 or 7* | *3 or 7* | *3 or 7* |

## CONTENT SPECIFICATIONS

For students to function effectively as mathematical problem solvers, they must be taught how to apply and communicate basic concepts and procedures, as well as how to do the procedures themselves.

**Content items** measure what students have been taught directly. Included in these are the basic concepts and procedural skills from all the content standards. For example, in the numbers and number sense standard and the computation standard, conceptual and procedural knowledge includes understanding of place value in our number system; the computational algorithms as applied to whole numbers, fractions, and decimals; and the concepts of ratio, proportion, and percent. In the data analysis and statistics standard, conceptual and procedural knowledge includes the ability to read charts and graphs as well as to understand concepts of averages (means, medians, and modes) and the methods for computing them. Contextual settings used in items measuring this category were very simple and were directly related to those used in the teaching of the concepts and the procedures.

**Application items** measure what the students can do with the content they have learned. Included are items requiring students to combine the basic concepts and procedures to solve real-life and mathematical problems, to evaluate their own ideas and the ideas of others using mathematical reasoning, and to communicate their ideas using the wealth of symbolic, pictorial, graphic, and verbal representations available in mathematics.

It is important to understand that application items also measure mastery of the basic concepts and procedures. For example, in mathematics, items were either short-answer or constructed-response items (see "Item Types" in the table below), which were worth up to four score points. In most cases, portions of these items required the student to perform some problem solving, reasoning, and/or communicating. At the same time, however, the items required the students to demonstrate their

understanding of mathematics content. If a student did not show mastery of all aspects of a constructed-response item, or if he/she made careless errors, the student did not earn the highest score for that item. Thus, it can be said that **all** mathematics items in the CRT measured content; some items went beyond that realm (short-answer and constructed-response), however, and were classified as application.

**TABLE 5-2: DISTRIBUTION OF MATHEMATICS PROCESS CATEGORIES**

| Grade | 3 | 4 | 5 | 6 | 7 | 8 | HS |
|---|---|---|---|---|---|---|---|
| Procedures/Concepts | 65% | 65% | 60% | 60% | 55% | 55% | 55% |
| Problem Solving/ Reasoning | 35% | 35% | 40% | 40% | 45% | 45% | 45% |

## ITEM TYPES

The CRT mathematics assessment included multiple-choice, short-answer, and constructed-response items. Short-answer items required students to perform a computation or solve a simple problem. Constructed-response items were more complex, requiring 8-10 minutes of response time. Each type of item was worth a specific number of points in the student's total mathematics score, as shown below.

**TABLE 5-3: ITEM TYPES**

| Type of Item | Possible Score Points |
|---|---|
| Multiple-Choice | 0 or 1 |
| Short-Answer | 0 or 1 |
| Constructed-Response | 0,1, 2, 3, or 4 |

## TEST DESIGN

Table 5-4 summarizes the number and types of items that were used in the CRT mathematics assessment for 2006, and shows the construction of the common portions of the assessment.

**TABLE 5-4: COMMON MATH ITEMS**

| | | | | | TOTAL | |
|---|---|---|---|---|---|---|
| **Gr.** | **Session 1 Cal** | **Session 2A Cal** | **Session 2B No Cal** | **Session 3 No Cal** | **M C** | **SA & CRs** |
| 3 | 24 MC, 1 CR | 5 MC | 5 MC | 21 MC, 3 SA, 1 CR | 55 | 3 SA, 2 CRs |
| 4 | 24 MC, 1 CR | 5 MC | 5 MC | 21 MC, 3 SA, 1 CR | 55 | 3 SA, 2 CRs |
| 5 | 24 MC, 1 CR | 5 MC | 5 MC | 21 MC, 3 SA, 1 CR | 55 | 3 SA, 2 CRs |
| 6 | 24 MC, 1 CR | 5 MC | 5 MC | 21 MC, 3 SA, 1 CR | 55 | 3 SA, 2 CRs |
| 7 | 24 MC, 1 CR | 5 MC | 5 MC | 21 MC, 3 SA, 1 CR | 55 | 3 SA, 2 CRs |
| 8 | 24 MC, 1 CR | 5 MC | 5 MC | 21 MC, 3 SA, 1 CR | 55 | 3 SA, 2 CRs |
| 10 | 24 MC, 1 CR | 8 MC | 7 MC | 21 MC, 3 SA, 1 CR | 60 | 3 SA, 2 CRs |

**Key**

- Cal = calculator use allowed
- No Cal = no calculator use allowed
- MC = multiple-choice items
- SA = short-answer items
- CR = constructed-response items

## THE USE OF CALCULATORS IN THE CRT

The Montana educators who helped develop the CRT acknowledged the importance of mastering arithmetic algorithms. At the same time, they understood that the use of calculators is a necessary and important skill in society today. Calculators can save time and prevent error in the measurement of some higher-order thinking skills and allow students to do more sophisticated and intricate problems. For these reasons, calculators were permitted on some parts of the CRT mathematics assessment and prohibited on others. (Students were allowed to use any calculator with which they were familiar.)

# SECTION II: TEST ADMINISTRATION

## CHAPTER 6—TEST ADMINISTRATION

### RESPONSIBILITY FOR ADMINISTRATION

As indicated in the *Test Coordinator's Manual*, principals and/or their designated School Test Coordinators were responsible for the proper administration of the CRT. This report was used to ensure the uniformity of administration procedures from school to school.

### PROCEDURES

School Test Coordinators were instructed to read the *Test Coordinator's Manual* prior to testing, and to be familiar with the instructions given in the *Test Administrator's Manual*. The *Test Coordinator's Manual* provided each school with checklists to help prepare for testing. The checklists outlined tasks to be performed before, during, and after test administration. Along with providing these checklists, the *Test Coordinator's Manual* outlined the nature of the testing material being sent to each school, how to inventory the material, how to track it during administration, and how to return the material once testing was complete. It also contained information about including or excluding students. The *Test Administrator's Manual* included checklists for the administrators to prepare themselves, their classrooms, and their students for the administration of the test. The *Test Administrator's Manual* contained sections that detailed the procedure to be followed for each test session, and it contained instructions on preparing the material prior to giving it to the School Test Coordinator for its return to Measured Progress.

### TEST ADMINISTRATOR TRAINING

In addition to distributing the *2006 Test Coordinator's Manuals* and *Test Administrator's Manuals*, OPI and Measured Progress produced and distributed two audio PowerPoint presentations, "Spring 2006: CRT and CRT-ALT Overview and Update of System and School Test Coordinators" and "CRT-ALT Test Administrator Training CD" to each system and school test coordinator. Training materials and the audio PowerPoint presentations were also posted on OPI's Web site. OPI conducted a survey requesting test coordinators to provide feedback to this form of training (audio CD vs. live

presentations). Positive feedback was received for the audio PowerPoints. System and school test coordinators were not required to travel long distances to attend pre-administration workshops and they could share the training CD with other educators within their buildings.

## PARTICIPATION REQUIREMENTS

All public school students participated in the CRT; however, scores of students in the following categories were excluded from the calculation of averages (CRT reports only):

- Foreign exchange students
- Students not enrolled in an accredited Montana school (for example: home-schooled student)
- Students enrolled in a private accredited school
- Students enrolled in a private nonaccredited school
- Students enrolled in a private nonaccredited Title 1 school
- Students enrolled part-time (less than 180 hours) taking a mathematics or reading course
- First year in US LEP students **were required** to participate in the math assessment only.

A summary of this information is shown in the table below which was published in the *Test Administrator's Manual* and *Test Coordinator's Manual*.

**TABLE 6-1: SUMMARY OF ELIGIBILITY FOR EXCLUSION**

| EXCLUDED FROM AVERAGES | MUST PARTICIPATE | MAY PARTICIPATE |
|---|---|---|
| FOREIGN EXCHANGE STUDENTS | **YES** | |
| STUDENTS NOT ENROLLED IN AN ACCREDITED MONTANA SCHOOL | | **YES** |
| STUDENTS ENROLLED IN A PRIVATE ACCREDITED SCHOOL | **YES** | |
| STUDENTS ENROLLED IN A PRIVATE NONACCREDITED SCHOOL | | **YES** |
| STUDENTS ENROLLED IN A PRIVATE NONACCREDITED TITLE I SCHOOL | | **YES** |

| | | |
|---|---|---|
| STUDENTS ENROLLED PART-TIME (LESS THAN 180 HRS.) TAKING A MATHEMATICS OR READING COURSE | | **YES** |
| READING: FIRST YEAR IN US LEP STUDENTS | | **YES** |
| MATHEMATICS: FIRST YEAR IN US LEP STUDENTS | **YES** | |

Information about the exclusion was coded by staff after testing was completed in the Student Response Booklet, if applicable. The *Test Coordinator's Manual* and *Test Administrator's Manual* provided detailed instructions for coding exclusions. In addition, testing exclusions were discussed thoroughly in the pre-administration training audio CD. Please refer to Appendix G: Reporting Decision Rules.

## TEST SCHEDULING

The CRTs were given during the spring: **reading** and **mathematics** were administered to grades 3 through 8 and 10 during the four-week period, March 6–29, 2006. Schools were able to schedule testing sessions at any time during this period, provided they followed the sequence in the scheduling guidelines detailed in *Test Administrator's Manual.* Schools were asked to schedule makeup testing of students who were absent from initial test sessions during this testing window.

The CRT is an untimed assessment; however, guidelines or ranges were provided in the *2006 Test Coordinator's Manual* and *2006 Test Administrator's Manual* based on estimates of the time it would take an average student to respond to each type of item that made up the test:

- multiple-choice items – 1 minute per item
- short-answer items – 2 minutes per item
- constructed-response items – 10 minutes per item

While the guidelines for scheduling were based on the assumption that most students would complete the test within the time estimated, each test administrator was asked to allow additional time for students who needed it (see Tables 6-2 through 6-5). If additional classroom space was not available for students who required additional time to complete the tests, schools were encouraged to consider

using another space, such as the guidance office, for this purpose. If additional areas were not available, it was recommended that each classroom being used for test administration be scheduled for the maximum amount of time.

### TABLE 6-2: GRADES 3 THROUGH 8 RECOMMENDED READING SCHEDULE

| Grades 3 through 8 Recommended Testing Schedule—Reading | | |
|---|---|---|
| **DAY 1 Reading** | **Test Activity** | Time Range (in minutes) |
| | General Instructions | 5-10 |
| | | |
| Session 1 | Reading Session 1 | 45-55 |
| **DAY 2 Reading** | | |
| Session 2 | Reading Session 2 | 45-55 |
| | Break | |
| Session 3 | Reading Session 3 | 45-55 |

### TABLE 6-3: GRADES 3 THROUGH 8 RECOMMENDED MATHEMATICS SCHEDULE

| Grades 3 through 8 Recommended Testing Schedule—Mathematics | | |
|---|---|---|
| **DAY 3 Mathematics** | **Calculators ARE allowed** | Time Range (in minutes) |
| Session 1 | Mathematics Session 1 | 45-55 |
| | Break | |
| Session 2A | Mathematics Session 2A | 20-30 |
| **DAY 4 Mathematics** | **Calculators are NOT allowed** | |
| Session 2B | Mathematics Session 2B | 20-30 |
| | Break | |
| Session 3 | Mathematics Session 3 | 45-55 |

### TABLE 6-4: GRADE 10
### RECOMMENDED READING SCHEDULE

| Grade 10 Recommended Testing Schedule—Reading | | |
|---|---|---|
| **DAY 1 Reading** | **Test Activity** | Time Range (in minutes) |
| | General Instructions | 10-20 |
| | Break | |
| Session 1 | Reading Session 1 | 50-60 |
| **DAY 2 Reading** | | |
| Session 2 | Reading Session 2 | 50-60 |
| | Break | |
| Session 3 | Reading Session 3 | 50-60 |

### TABLE 6-5: GRADE 10
### RECOMMENDED MATHEMATICS SCHEDULE

| Grade 10 Recommended Testing Schedule—Mathematics | | |
|---|---|---|
| **DAY 3 Mathematics** | **Calculators ARE allowed** | Time Range (in minutes) |
| Session 1 | Mathematics Session 1 | 50-60 |
| | Break | |
| Session 2A | Mathematics Session 2A | 20-30 |
| **DAY 4 Mathematics** | **Calculators are NOT allowed** | |
| Session 2B | Mathematics Session 2B | 20-30 |
| | Break | |
| Session 3 | Mathematics Session 3 | 50-60 |

## HELP DESK

To address testing concerns, Measured Progress established a help desk dedicated to the State of Montana. Help desk support is an essential element to the successful administration of large-scale assessments. It provides a centralized location where individuals in the field can call a toll-free number to request assistance, report problems they are experiencing, or ask specific questions.

The Measured Progress help desk provided support during all phases of the testing window. It was staffed at varying levels based on need and volume and was available from 8:00 A.M. to 4:00 P.M. MST during the testing window. At a minimum, the help desk consisted of a product support specialist who was responsible for receiving, responding to, and tracking calls and e-mails, and routing issues to the appropriate person(s) for resolution. In addition, communications requiring a higher level of program support were routed to the program manager and/or program assistant

When possible, all calls and e-mails received during business hours were responded to immediately with resolution or updated within hours of receipt.

# SECTION III: DEVELOPMENT AND REPORTING OF SCORES

## CHAPTER 7—SCORING

This chapter describes the scoring processes for the multiple-choice, short-answer, and constructed-response items for Montana's Criterion-Referenced Test (CRT) program.. Responses to multiple-choice items are first captured electronically from the student response booklets through a scanning process and then scored through an electronic process in data processing. The short-answer and constructed-response items begin the scoring process in scanning as well, but after scanning, the student responses are scored by human readers.

### SCANNING

Months prior to test administration and subsequent scanning activities, the scanning department met with the program management team to determine decision rules and required scanning and imaging specifications. The information gathered at these meetings was then used to develop a customized scanning program for Montana.

At the conclusion of testing, Montana schools shipped all test materials back to Measured Progress. To expedite the scanning and scoring process, used student response booklets were express-shipped separately from other test materials. Once the approximately 80,000 used student response booklets were logged in, identified with appropriate scannable, preprinted school information sheets, examined for extraneous materials, and batched, they were moved into the scanning area.

The first step in that conversion was the removal of the booklet bindings so that the individual pages could pass through the scanners one at a time. Once cut, the sheets were put back in their proper boxes and placed in storage until needed for the scanning/imaging process.

Customized scanning programs for all scannables were prepared to selectively read the student response booklets and to format the scanned information electronically according to predetermined requirements. Any information (including multiple-choice response data) that had been designated time-critical or process-critical was handled first.

All student response documents and other scannable information necessary to produce the required reports were captured and converted into an electronic format, including all student identification and demographics, and digital image clips of short-answer and constructed-response student responses. The digital image clip information allowed Measured Progress to replicate student responses on the readers' monitors just as they had appeared on the originals. From that point on, the entire process—data processing, scoring, benchmarking data analysis, and reporting—was accomplished without further reference to the originals.

## SCANNING QUALITY CONTROL

Throughout the scanning process, quality control measures were implemented by Measured Progress staff to assure that the choices that the students marked for multiple-choice questions and the responses that the students wrote for short-answer and constructed-response items were captured accurately.

The scanners used for the Montana CRT program are equipped with many built-in safeguards that prevent data errors. The scanning hardware was continually monitored for conditions that would cause the machine to shut down if standards were not maintained.

When a scanning error occurred (such as a mis-fed document), the scanner displayed an error message and prevented further scanning until the condition was corrected. Things monitored by the scanners included document page and integrity checks, user-designed on-line edits, and many internal checks of electronic functions.

Before every scanning shift began, Measured Progress operators performed a daily diagnostic routine. In the rare event that the routine detected a photocell that appeared to be out of range, that machine was calibrated and the test performed again. If the read was still not up to standard, for a service call was made to the field service engineer for correction.

As a final safeguard, spot checks of scanned files, bubble by bubble and image by image, were routinely made throughout scanning runs. The result of these precautions, from the original layout of the scanning form to the daily vigilance of our operators, was a scan error rate well below 1 per 1000.

## Electronic Data Files

Once the scanning process was completed, the booklets themselves were put into storage (where they stayed for at least 180 days beyond the close of the fiscal year). When it had been determined that the files were complete and accurate, those files were duplicated electronically and made available for many other processing options. Completed files were loaded onto our local area network (LAN) for transfer to Measured Progress's proprietary I-Score system for scoring. Those files were then used to identify (and print out) papers to be used in the benchmarking processes, and the data made transferable via the Internet, CD-ROM, or optical disk.

| Table 7-1: Number of Responses Scanned and Scored**Grade/Content** | **Number of Responses Scanned and Scored** |
|---|---|
| 3 Math | 80,035 |
| 4 Math | 85,861 |
| 5 Math | 86,017 |
| 6 Math | 89,367 |
| 7 Math | 91,611 |
| 8 Math | 97,611 |
| 10 Math | 97,820 |
| 3 Reading | 25,818 |
| 4 Reading | 26,948 |
| 5 Reading | 27,107 |
| 6 Reading | 28,260 |
| 7 Reading | 28,856 |
| 8 Reading | 31,016 |
| 10 Reading | 31,273 |

**NOTE:** Common math items scanned and scored consisted of three short-answer and two constructed-response items per student compared to two constructed-response common items per student in reading.

## ITEMS SCORED BY READERS

Test and answer materials were handled as little as possible to minimize the possibility of loss, mishandling, or breach of security. Once scanned, either by optical mark reader or the I-Score system, papers were stored securely in areas with limited personnel access.

As explained in the following sections on scoring, the I-Score system itself ensures the security of responses and test items: all scoring is "blind"; that is, no student names are associated with viewed responses or raw scores and all scoring personnel are subject to the same nondisclosure requirements and supervision as regular Measured Progress staff.

## I-SCORE

All of Measured Progress's scoring facilities use the *iScore* process. *iScore* is Measured Progress's Web-based proprietary software used to score short-answer and constructed response items. Images of student responses are transferred electronically via a secure

Web site to a scorer's computer screen at any one of Measured Progress's scoring facilities. For Montana's CRT program, scoring took place in Dover, New Hampshire, Albany, New York, and Denver, Colorado.

After the 2006 test material had been loaded into the LAN, I-Score sent electronically scanned images of student work to individual readers at computer terminals, who evaluated each response and recorded each student's score via keypad or mouse entry. When the reader had finished with one response, the next response appeared immediately on the computer screen. In that way, the system guaranteed complete anonymity of individual students and ensured the randomization of responses during scoring.

Although I-Score is based on conventional scoring techniques, it also offers numerous benefits, not the least of which is raising the bar on scoring process capability. Some of the benefits are

- real-time information on scorer reliability, read-behinds, and overall process monitoring;
- early access to subsets of data for tasks such as standard setting;
- reduced material handling, which not only saves time and labor, but also enhances the security of materials; and
- immediate access to samples of student responses and scores for reporting and analysis through electronic media.

Scoring operations, directed by the manager of scoring services, were carried out by a highly qualified staff. The staff included

- chief readers, who oversaw all training and scoring within particular subject areas;
- quality assurance coordinators (QACs), who led benchmarking and training activities and monitored scoring consistency and rates;
- verifiers, who performed read-behinds of readers and assisted at scoring tables as necessary; and
- readers, who performed the bulk of the scoring.

Table 7-2, below, summarizes the qualifications of the 2006 CRT quality assurance coordinators and readers.

**TABLE 7-2: EDUCATIONAL CREDENTIALS**

**Montana Reader Education Credentials**

| Description | Albany, NY | Denver, CO | Dover, NH | Total | Pct |
|---|---|---|---|---|---|
| Less then 48 college credits | 0 | 0 | 0 | 0 | 0.00% |
| 48+ college credits | 3 | 1 | 3 | 7 | 3.00% |
| Associate's degree | 14 | 0 | 7 | 21 | 9.00% |
| Bachelor's degree | 82 | 18 | 33 | 133 | 55.00% |
| Master's degree | 48 | 9 | 15 | 72 | 30.00% |
| Doctorate | 6 | 0 | 1 | 7 | 3.00% |
| **Total** | **153** | **28** | **59** | **240** | |

**Montana Quality Assurance Coordinators Education Credentials**

| Description | Albany, NY | Denver, CO | Dover, NH | Total | Pct |
|---|---|---|---|---|---|
| Less then 48 college credits | 0 | 0 | 0 | 0 | 0.00% |
| 48+ college credits | 0 | 0 | 0 | 0 | 0.00% |
| Associate's degree | 1 | 0 | 0 | 1 | 5.00% |
| Bachelor's degree | 4 | 0 | 4 | 8 | 40.00% |
| Master's degree | 5 | 1 | 3 | 9 | 45.00% |
| Doctorate | 1 | 1 | 0 | 2 | 10.00% |
| **Total** | **11** | **2** | **7** | **20** | |

## PRELIMINARY ACTIVITIES

Preliminary activities for scoring included (1) participating in the planning and design of documents to be used for scoring, (2) reviewing items and score guides for benchmarking and training and the creation of benchmarking packets, and (3) selecting scoring staff and training them for scoring.

## PLANNING AND DESIGNING DOCUMENTS

At the request of the project manager, scoring personnel advised project management and OPI staff on the program design in order to support an efficient and effective scoring process. Scoring staff also contributed to the design of

- response documents and the image-capture process to yield acceptable image clips (also defining file format and layout); and
- scoring benchmarks composed of the guide, subject background information, and anchor papers.

## BENCHMARKING

Before the scheduled start of scoring activities, scoring center staff and Montana educators reviewed test items and scoring guides for benchmarking. At that point, chief readers and selected QACs prepared scorer training materials.

Scoring staff from Measured Progress (including test developers) and Montana educators selected one or two anchor examples for each item score point. An additional six to ten responses per item were chosen as part of the training pack. The anchor pack consisted of midrange exemplars, while the training pack exemplars illustrated the range within each score point. The chief readers, who worked closely with QACs for each content area, facilitated the selection of response exemplars.

## SELECTING AND TRAINING SCORING STAFF

As mentioned earlier in this chapter, staff assigned to scoring activities for the Montana CRT program are highly qualified, both in terms of their education and their scoring capability. Each scoring employee is required to have a minimum of 48 college credits, and at least two college courses in the content area they are scoring.

### QUALITY ASSURANCE COORDINATORS (QACS) AND VERIFIERS

Because the read-behinds performed by the QACs and verifiers moderated the scoring process and thus maintained the integrity of the scores, individuals chosen to fill those positions were selected for their accuracy (see below). In addition, QACs, who train readers to score each item in their content areas, were selected for their ability to instruct and for their level of expertise in their content areas. For this reason, QACs typically are

retired teachers who have demonstrated a high level of expertise in their respective disciplines. The average ratio of QACs and verifiers to readers was approximately 1:11.

### TRAINING QUALITY ASSURANCE COORDINATORS AND VERIFIERS

To ensure that all QACs provided consistent training and feedback, the chief readers spent two days training and qualifying the QACs, and the QACs reviewed all items with the verifiers before scoring. In addition, QACs rotated among tables, supervising readers and reading behind verifiers, who in turn read behind a different table of readers each day.

### SELECTING READERS

Applicants were required to demonstrate their ability by participating in a preliminary scoring evaluation. The I-Score system enables Measured Progress to efficiently measure a prospective reader's ability to score student responses accurately. After participating in a training session, applicants were required to achieve at least 80% exact scoring agreement for a qualifying pack consisting of 20 responses to a predetermined item in their content area. Those 20 responses were randomly selected from a bank of approximately 150, all of which had been selected by QACs and approved by the chief readers and developers. Table 7-3 depicts the accuracy and qualification percentages of the readers.

**TABLE 7-3**

**MONTANA SCORING ACCURACY AND QUALIFICATION STATISTICS 2006**

| Content | Grade | Item | Average % Exact Agreement for Embedded CR sets | Average % Exact Agreement for Double Blind Scoring | Number of Readers taking Qualification Sets | Number Successfully Qualifying | Percent Successfully Qualifying |
|---|---|---|---|---|---|---|---|
| Math | 3 | 25 | 87.8 | 82.3 | NA | NA | NA |
| Math | 3 | 65 | 71.2 | 96.3 | NA | NA | NA |
| Math | 3 | 66 | NA | 91.1 | NA | NA | NA |
| Math | 3 | 67 | NA | 94.5 | NA | NA | NA |
| Math | 3 | 68 | NA | 82.1 | NA | NA | NA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Math | 4 | 25 | 92.6 | 35.9 | 9 | 6 | 66.7 |
| Math | 4 | 65 | NA | 90.7 | NA | NA | NA |
| Math | 4 | 66 | NA | 90.2 | NA | NA | NA |
| Math | 4 | 67 | NA | 96.1 | NA | NA | NA |
| Math | 4 | 68 | 96.0 | 93.8 | 14 | 14 | 100.0 |
| Math | 5 | 25 | 86.3 | 75.2 | 10 | 7 | 70.0 |
| Math | 5 | 65 | NA | 96.9 | NA | NA | NA |
| Math | 5 | 66 | NA | 96.4 | NA | NA | NA |
| Math | 5 | 67 | NA | 96.8 | NA | NA | NA |
| Math | 5 | 68 | 95.6 | 91.0 | 13 | 13 | 100.0 |
| Math | 6 | 25 | 84.7 | 84.2 | 12 | 10 | 83.3 |
| Math | 6 | 65 | NA | 98.1 | NA | NA | NA |
| Math | 6 | 66 | NA | 80.2 | NA | NA | NA |
| Math | 6 | 67 | NA | 94.0 | NA | NA | NA |
| Math | 6 | 68 | 92.9 | 95.9 | 14 | 14 | 100.0 |
| Math | 7 | 25 | 79.8 | 77.4 | 25 | 18 | 72.0 |
| Math | 7 | 65 | NA | 96.2 | NA | NA | NA |
| Math | 7 | 66 | NA | 93.6 | NA | NA | NA |
| Math | 7 | 67 | NA | 97.4 | NA | NA | NA |
| Math | 7 | 68 | 91.4 | 92.4 | 24 | 20 | 83.3 |
| Math | 8 | 25 | 88.3 | 91.4 | 24 | 17 | 70.8 |
| Math | 8 | 65 | NA | 98.2 | NA | NA | NA |
| Math | 8 | 66 | NA | 95.4 | NA | NA | NA |
| Math | 8 | 67 | NA | 95.3 | NA | NA | NA |
| Math | 8 | 68 | 94.3 | 91.6 | 12 | 11 | 91.7 |
| Math | 10 | 25 | 59.3 | 88.8 | 22 | 14 | 63.6 |
| Math | 10 | 70 | NA | 96.3 | NA | NA | NA |
| Math | 10 | 71 | NA | 98.3 | NA | NA | NA |
| Math | 10 | 72 | NA | 98.1 | NA | NA | NA |
| Math | 10 | 73 | 89.0 | 94.4 | 25 | 25 | 100.0 |
| Reading | 3 | 22 | 78.2 | 75.7 | NA | NA | NA |
| Reading | 3 | 67 | 71.0 | 78.5 | NA | NA | NA |
| Reading | 4 | 22 | 89.7 | 73.4 | 31 | 31 | 100.0 |
| Reading | 4 | 67 | 84.2 | 72.9 | 28 | 25 | 89.3 |
| Reading | 5 | 22 | 87.6 | 78.4 | 33 | 32 | 96.7 |
| Reading | 5 | 67 | 91.2 | 80.3 | 28 | 28 | 100.0 |
| Reading | 6 | 22 | 88.3 | 78.8 | 33 | 31 | 93.9 |
| Reading | 6 | 67 | 87.4 | 77.0 | 28 | 27 | 96.4 |
| Reading | 7 | 22 | 93.1 | 77.1 | 33 | 32 | 96.7 |
| Reading | 7 | 67 | 91.7 | 77.5 | 29 | 26 | 89.7 |
| Reading | 8 | 22 | 83.7 | 77.1 | 60 | 56 | 93.3 |
| Reading | 8 | 67 | 89.1 | 89.1 | 37 | 37 | 100.0 |
| Reading | 10 | 22 | 81.8 | 77.5 | 42 | 38 | 90.4 |
| Reading | 10 | 72 | 78.4 | 82.5 | 34 | 30 | 88.2 |

### Training Readers

The QACs first applied the language of the scoring guide for an item to its anchor pack exemplars. Once discussion of the anchor pack had concluded, readers attempted to score the training pack exemplars correctly. The QACs then reviewed the training pack and answered any items readers had before actual scoring began. With this system, two aspects of scoring efficiency are in conflict. First, in order to minimize training expense, it is desirable to train each reader on as few items as possible. Second, to prevent reader drift and to minimize retraining requirements, it is desirable to score a given item in a brief period of time. But the lower the number of unique items each reader scores, the greater the number of readers required to score that item quickly. To minimize that conflict, we divided each subject area's readers into two or more groups. On the first day of scoring, each group was trained to score a different item. When a group had completed all of an item's responses, those readers were trained on another item (or set).

## Scoring Activities

Student test booklets at grade levels 3 through 8 and 10 were digitally scanned and scored on a file server for a dedicated, secure LAN. I-Score then distributed digital images of student responses to readers. Training and scoring took place over a period of approximately two weeks.

Items were randomly assigned to readers; thus, each item in a student's response booklet was more than likely scored by a different reader. By using the maximum possible number of readers for each student, the procedure effectively minimized error variance due to reader sampling. All common and matrix constructed-response items were scored once with a 2% read-behind to ensure consistency among readers and accuracy of individual readers.

**TABLE 7-4: MONTANA 2006 SUMMARY STATISTICS**

| Grade/Content | Number of Responses Scored | Total Number of Responses Scored in Double-Blind | Total Number of Arbitrations Required | Percentage of Double-Blinds Arbitrated |
|---|---|---|---|---|
| 3 Math | 80,035 | 3,060 | 141 | 4.96% |
| 4 Math | 85,861 | 3,900 | 199 | 4.49% |
| 5 Math | 86,017 | 3,964 | 157 | 4.54% |
| 6 Math | 89,367 | 4,856 | 154 | 4.25% |
| 7 Math | 91,611 | 6,150 | 233 | 4.52% |
| 8 Math | 97,611 | 5,733 | 184 | 3.49% |
| 10 Math | 97,820 | 7,813 | 239 | 3.47% |
| 3 Reading | 25,818 | 877 | 31 | 5.25% |
| 4 Reading | 26,948 | 1,015 | 39 | 3.52% |
| 5 Reading | 27,107 | 1,100 | 30 | 2.54% |
| 6 Reading | 28,260 | 1,042 | 37 | 3.00% |
| 7 Reading | 28,856 | 1,223 | 43 | 3.21% |
| 8 Reading | 31,016 | 1,393 | 38 | 2.47% |
| 10 Reading | 31,273 | 1,414 | 42 | 2.86% |

## MONITORING READERS

To ensure high inter-rater reliability and to prevent scoring drift after a reader scored a student response. *iScore* determined whether the reader met the standard accuracy requirement that states that a reader's scoring, based on double-scored responses, must be exact more than 90% of the time and that up to the 10% that are not exact, their score is adjacent at least 80% of the time. If a reader's scores do not meet these three standards, *iScore* will freeze or block the reader's screen and alert the senior reader. The senior reader will then determine whether responses should also be scored by another reader, scored by a QAC, or routed for special attention. QAC's and senior readers were able to obtain current reader accuracy reports and speed reports online at any time. Table 7-4 summarizes how often readers screens were blocked through the process and the resolutions.

## TABLE 7-5: MONTANA BLOCKED READER STATISTICS 2006

| Content | Grade/Item | Number of Readers Blocked From Scoring by *iScore* | Number of Readers Allowed to Continue Scoring Based upon Other Quality Monitoring (Read-Behinds and Double Blinds) | Number of Readers NOT Allowed To Continue Scoring Item and Reassigned to Other Items or Dismissed from Project |
|---|---|---|---|---|
| Math | 3, 25 | NA | NA | NA |
| Math | 3, 65 | NA | NA | NA |
| Math | 3, 66 | NA | NA | NA |
| Math | 3, 67 | NA | NA | NA |
| Math | 3, 68 | NA | NA | NA |
| Math | 4, 25 | 3 | 3 | 0 |
| Math | 4, 65 | NA | NA | NA |
| Math | 4, 66 | NA | NA | NA |
| Math | 4, 67 | NA | NA | NA |
| Math | 4, 68 | 0 | 0 | 0 |
| Math | 5, 25 | 3 | 3 | 0 |
| Math | 5, 65 | NA | NA | NA |
| Math | 5, 66 | NA | NA | NA |
| Math | 5, 67 | NA | NA | NA |
| Math | 5, 68 | 0 | 0 | 0 |
| Math | 6, 25 | 2 | 0 | 2 |
| Math | 6, 65 | NA | NA | NA |
| Math | 6, 66 | NA | NA | NA |
| Math | 6, 67 | NA | NA | NA |
| Math | 6, 68 | 0 | 0 | 0 |
| Math | 7, 25 | 7 | 6 | 1 |
| Math | 7, 65 | NA | NA | NA |
| Math | 7, 66 | NA | NA | NA |
| Math | 7, 67 | NA | NA | NA |
| Math | 7, 68 | 4 | 4 | 0 |
| Math | 8, 25 | 7 | 7 | 0 |
| Math | 8, 65 | NA | NA | NA |
| Math | 8, 66 | NA | NA | NA |
| Math | 8, 67 | NA | NA | NA |
| Math | 8, 68 | 1 | 0 | 1 |
| Math | 10, 25 | 8 | 0 | 8 |
| Math | 10, 70 | NA | NA | NA |
| Math | 10,71 | NA | NA | NA |
| Math | 10, 72 | NA | NA | NA |
| Math | 10, 73 | 0 | 0 | 0 |
| Reading | 3, 22 | NA | NA | NA |
| Reading | 3, 67 | NA | NA | NA |
| Reading | 4, 22 | 0 | 0 | 0 |
| Reading | 4, 67 | 3 | 3 | 3 |

| | | | | |
|---|---|---|---|---|
| Reading | 5, 22 | 1 | 1 | 0 |
| Reading | 5, 67 | 0 | 0 | 0 |
| Reading | 6, 22 | 2 | 2 | 0 |
| Reading | 6, 67 | 1 | 1 | 0 |
| Reading | 7, 22 | 1 | 1 | 0 |
| Reading | 7, 67 | 3 | 3 | 0 |
| Reading | 8, 22 | 4 | 4 | 4 |
| Reading | 8, 67 | 0 | 0 | 0 |
| Reading | 10, 22 | 4 | 4 | 4 |
| Reading | 10, 72 | 4 | 4 | 4 |

**NOTE:** All readers who were allowed to continue scoring did so under increased quality screening/additional read-behinds were conducted on these readers.

## GENERAL SCORING GUIDES

Tables 7-6 and 7-7 are examples of general CRT short-answer and constructed-response scoring guides. Item-specific rubrics are prepared for each short-answer and constructed-response item, and are derived from the general rubrics provided here.

### TABLE 7-6: SHORT-ANSWER ITEMS

| Score Point | Description |
|---|---|
| 1 | The student's response provides a complete and correct answer. |
| 0 | The student's response is totally incorrect or too minimal to evaluate. |
| 0 | Blank/no response. |

### TABLE 7-7: CONSTRUCTED- RESPONSE ITEMS

| Score Point | Description |
|---|---|
| 4 | • The student completes all important components of the task and communicates ideas clearly.<br>• The student demonstrates in-depth understanding of the relevant concepts and/or processes.<br>• When instructed to do so, the student chooses more efficient and/or sophisticated processes.<br>• When instructed to do so, the student offers insightful interpretations or extensions (e.g., generalizations, applications, and analogies). |
| 3 | • The student completes the most important components of the task and communicates clearly. |

| | |
|---|---|
| | • The student demonstrates understanding of major concepts even though he/she overlooks or misunderstands some less important ideas or details. |
| 2 | • The student completes most important components of the task and communicates those clearly.<br>• The student demonstrates that there are gaps in his/her conceptual understanding. |
| 1 | • The student shows minimal understanding.<br>• The student addresses only a small portion of the required task(s). |
| 0 | • The student's response is totally incorrect or irrelevant. |
| B | • Blank/no response. |

# CHAPTER 8—ITEM ANALYSES

As noted in Brown (1983), "A test is only as good as the items it contains." A complete evaluation of a test's quality must include an evaluation of each item. Both the *Standards for Educational and Psychological Testing (AERA et al.,1999)* and the *Code of Fair Testing Practices in Education (2004)* include standards for identifying quality items. Items should assess only knowledge or skills that are identified as part of the domain being tested and should avoid assessing irrelevant factors. They should also be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. Further, items must not unfairly disadvantage test takers from particular racial, ethnic, or gender groups.

Both qualitative and quantitative analyses are conducted to ensure that Montana CRT items meet these standards. Qualitative analyses are described in earlier sections of this report; this section focuses on the more quantitative evaluations. The statistical evaluations are presented in three parts: 1) difficulty indices, 2) item-test correlations, and 3) differential item functioning (DIF) statistics. The item analyses presented here are based on the statewide administration of the Montana CRT in spring 2006. The numbers of students who participated in the assessment at each grade level were about 10,030 in grade 3, 10,340 in grade 4, 10,350 in grade 5, 10,770 in grade 6, 10,993 in grade 7, 11,685 in grade 8, and 11,485 in grade 10. Note that the information presented in this chapter is based on the items common to all forms since those are the items on which student scores are calculated. Item analyses are also performed for field test items; the statistics are then used in the item review process, as well as during form assembly for future administrations.

## DIFFICULTY INDICES (P)

All multiple-choice and constructed-response (constructed-response and short-answer) items were evaluated in terms of item difficulty according to standard classical test theory practices. Difficulty was defined as the average proportion of points achieved on an item, and was measured by obtaining the average score on an item and dividing by the maximum possible score for the item. Multiple-choice items were scored dichotomously (correct vs. incorrect), so for

those items, the difficulty index is simply the proportion of students who correctly answered the item. Constructed-response items (two on each math form and two on each reading form) were scored polytomously, where a student can achieve a score of 0, 1, 2, 3, or 4. Short-answer items (three computation items on each math form) were scored 0 or 1. By computing the difficulty index as the average proportion of points achieved, the indices for the different item types are placed on a similar scale; the index ranges from 0.0 to 1.0 regardless of the item type. Although this index is traditionally described as a measure of difficulty, it is properly interpreted as an "easiness index" because larger values indicate easier items. An index of 0.0 indicates that all students received no credit for the item, and an index of 1.0 indicates that all students received full credit for the item.

Items that are answered correctly by almost all students provide little information about differences in student ability, but they do indicate knowledge or skills that have been mastered by most students. Similarly, items that are correctly answered by very few students may indicate knowledge or skills that have not yet been mastered by most students, but such items provide little information about differences in student ability. In general, to provide best measurement, difficulty indices should range from near-chance performance (.25 for four-option, multiple-choice items or essentially zero for constructed-response or short-answer items) to .90. However, on a standards-referenced assessment such as the Montana CRT, it may be appropriate to include some items with very low or very high item difficulty values to ensure sufficient content coverage (minimum of six items/points per standard).

## ITEM-TEST CORRELATIONS (ITEM DISCRIMINATION)

A desirable feature of an item is that the higher-ability students perform better on the item than lower-ability students. The correlation between student performance on a single item and total test score is a commonly used measure of this characteristic of an item. Within classical test theory, the item-test correlation is referred to as the item's discrimination because it indicates the extent to which successful performance on an item discriminates between high and low scores on the test. For constructed-response items, the item discrimination index used was the Pearson product-moment correlation; for dichotomous items (multiple-choice and short-answer), the

corresponding statistic is commonly referred to as a point-biserial correlation. The theoretical range of these statistics is –1 to +1, with a typical range from .2 to .6.

Discrimination indices can be thought of as measures of how closely an item assesses the same knowledge and skills assessed by other items contributing to the criterion total score. That is, the discrimination index can be thought of as a measure of construct consistency. In light of this interpretation, the selection of an appropriate criterion total score is crucial to the interpretation of the discrimination index. Because each form of the Montana CRT was constructed to be parallel in content, the criterion score selected for each item was the raw score total for each form. The analyses were conducted for each form separately.

## SUMMARY OF ITEM ANALYSIS RESULTS

Summary statistics of the difficulty and discrimination indices for each item are provided in Tables 8-1 through 8-7 for grades 3 through 8 and 10. Mean difficulty and discrimination indices, broken down by item type – multiple-choice, constructed-response (constructed-response and short-answer), and all items – are shown in Table 8-8 (standard deviations are shown in parentheses). In general, the item difficulty and discrimination indices are within generally acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that students who performed well on individual items tended to perform well overall. There were a small number of items with near-zero discrimination indices, but none were reliably negative. While it is not inappropriate to include items with low discrimination values or with very high or very low item difficulty values to ensure that content is appropriately covered, there were very few such cases on the Montana CRT.

A comparison of indices across grade levels is complicated because these indices are population dependent. Direct comparisons would require that either the items or students were common across groups. Since that is not the case, it can not be determined whether differences in performance across grade levels are due to differences in student ability or differences in item

difficulty or both. However, one can say that for math, students in grade 3 and 4 found their items somewhat less difficult than students in higher grades found their items.

Comparing the difficulty indices of multiple-choice and constructed-response (constructed-response or short-answer) items is inappropriate because multiple-choice items can be answered correctly by guessing. Thus, it is not surprising that the difficulty indices for multiple-choice items tend to be higher (indicating that students performed better on these items) than the difficulty indices for constructed-response items. Similarly, the partial credit allowed by constructed-response items is advantageous in the computation of item-test correlations, so the discrimination indices for these items tend to be larger than the discrimination indices of multiple-choice items.

The statistics in Tables 8-1 through 8-7 and those calculated for the full set of items in Table 8-8 are weighted according to the number of points contributed by each item. In the event that an item's statistics indicate it is flawed, the item is dropped from the operational form. An item may be dropped, for example, if more than one of the response options is a defensible answer, or if the item is misleading or unclear in some way. No flawed items were found for the 2006 MontCAS, Phase 2 CRT test administration.

**TABLE 8-1**
**ITEM ANALYSIS: GRADE 3**

| Content Area | | Difficulty | Discrimination |
|---|---|---|---|
| **Math** | Mean | 0.70 | 0.35 |
| | StDev | 0.14 | 0.08 |
| | Min | 0.36 | 0.21 |
| | Max | 0.92 | 0.52 |
| | Range | 0.56 | 0.31 |
| **Reading** | Mean | 0.69 | 0.38 |
| | StDev | 0.14 | 0.09 |
| | Min | 0.36 | 0.15 |
| | Max | 0.90 | 0.52 |
| | Range | 0.54 | 0.37 |

**TABLE 8-2**
**ITEM ANALYSIS: GRADE 4**

| Content Area | | Difficulty | Discrimination |
|---|---|---|---|
| **Math** | Mean | 0.68 | 0.36 |
| | StDev | 0.14 | 0.08 |
| | Min | 0.32 | 0.24 |
| | Max | 0.93 | 0.58 |
| | Range | 0.61 | 0.34 |
| **Reading** | Mean | 0.70 | 0.36 |
| | StDev | 0.15 | 0.09 |
| | Min | 0.39 | 0.14 |
| | Max | 0.94 | 0.56 |
| | Range | 0.55 | 0.42 |

**TABLE 8-3**
**ITEM ANALYSIS: GRADE 5**

| Content Area | | Difficulty | Discrimination |
|---|---|---|---|
| **Math** | Mean | 0.59 | 0.36 |
| | StDev | 0.14 | 0.08 |
| | Min | 0.27 | 0.20 |
| | Max | 0.90 | 0.55 |
| | Range | 0.63 | 0.35 |
| **Reading** | Mean | 0.65 | 0.34 |
| | StDev | 0.15 | 0.09 |
| | Min | 0.17 | 0.09 |
| | Max | 0.89 | 0.49 |
| | Range | 0.72 | 0.40 |

**TABLE 8-4**
**ITEM ANALYSIS: GRADE 6**

| Content Area | | Difficulty | Discrimination |
|---|---|---|---|
| **Math** | Mean | 0.52 | 0.34 |
| | StDev | 0.17 | 0.10 |
| | Min | 0.16 | 0.16 |
| | Max | 0.92 | 0.64 |
| | Range | 0.76 | 0.48 |
| **Reading** | Mean | 0.70 | 0.33 |
| | StDev | 0.15 | 0.08 |
| | Min | 0.30 | 0.14 |
| | Max | 0.95 | 0.55 |
| | Range | 0.65 | 0.41 |

**TABLE 8-5**
**ITEM ANALYSIS: GRADE 7**

| Content Area | | Difficulty | Discrimination |
|---|---|---|---|
| **Math** | Mean | 0.52 | 0.33 |
| | StDev | 0.17 | 0.11 |
| | Min | 0.08 | 0.18 |
| | Max | 0.95 | 0.60 |
| | Range | 0.87 | 0.66 |
| **Reading** | Mean | 0.69 | 0.37 |
| | StDev | 0.12 | 0.09 |
| | Min | 0.45 | 0.18 |
| | Max | 0.90 | 0.54 |
| | Range | 0.45 | 0.36 |

**TABLE 8-6**
**ITEM ANALYSIS: GRADE 8**

| Content Area | | Difficulty | Discrimination |
|---|---|---|---|
| **Math** | Mean | 0.56 | 0.38 |
| | StDev | 0.15 | 0.09 |
| | Min | 0.24 | 0.22 |
| | Max | 0.86 | 0.62 |
| | Range | 0.62 | 0.40 |
| **Reading** | Mean | 0.71 | 0.34 |
| | StDev | 0.14 | 0.08 |
| | Min | 0.40 | 0.19 |
| | Max | 0.96 | 0.56 |
| | Range | 0.56 | 0.37 |

**TABLE 8-7**
**ITEM ANALYSIS: GRADE 10**

| Content Area | | Difficulty | Discrimination |
|---|---|---|---|
| **Math** | Mean | 0.57 | 0.39 |
| | StDev | 0.13 | 0.11 |
| | Min | 0.29 | 0.15 |
| | Max | 0.82 | 0.70 |
| | Range | 0.53 | 0.55 |
| **Reading** | Mean | 0.70 | 0.34 |
| | StDev | 0.13 | 0.09 |
| | Min | 0.18 | 0.12 |
| | Max | 0.94 | 0.59 |
| | Range | 0.76 | 0.47 |

**TABLE 8-8**
**AVERAGE DIFFICULTY AND DISCRIMINATION OF DIFFERENT ITEM TYPES FOR EACH GRADE/CONTENT AREA COMBINATION**

| Grade | Content Area | | Item Type | | |
|---|---|---|---|---|---|
| | | | All | MC | Constructed-Response |
| 3 | Reading | Difficulty | 0.69 ( 0.14) | 0.70 ( 0.13) | 0.41 ( 0.04) |
| | | Discrimination | 0.38 ( 0.09) | 0.38 ( 0.09) | 0.39 ( 0.05) |
| | | Number of Items | 54 | 52 | 2 |
| | Mathematics | Difficulty | 0.70 ( 0.14) | 0.71 ( 0.15) | 0.68 ( 0.07) |
| | | Discrimination | 0.35 ( 0.08) | 0.35 ( 0.07) | 0.43 ( 0.08) |
| | | Number of Items | 60 | 55 | 5 |
| 4 | Reading | Difficulty | 0.70 ( 0.15) | 0.71 ( 0.14) | 0.48 ( 0.12) |
| | | Discrimination | 0.36 ( 0.09) | 0.35 ( 0.08) | 0.49 ( 0.10) |
| | | Number of Items | 54 | 52 | 2 |
| | Mathematics | Difficulty | 0.68 ( 0.14) | 0.69 ( 0.14) | 0.59 ( 0.07) |
| | | Discrimination | 0.36 ( 0.08) | 0.36 ( 0.07) | 0.46 ( 0.09) |
| | | Number of Items | 60 | 55 | 5 |
| 5 | Reading | Difficulty | 0.65 ( 0.15) | 0.66 ( 0.15) | 0.42 ( 0.01) |
| | | Discrimination | 0.34 ( 0.09) | 0.34 ( 0.09) | 0.46 ( 0.04) |
| | | Number of Items | 54 | 52 | 2 |
| | Mathematics | Difficulty | 0.59 ( 0.14) | 0.60 ( 0.13) | 0.42 ( 0.14) |
| | | Discrimination | 0.36 ( 0.08) | 0.35 ( 0.08) | 0.47 ( 0.06) |
| | | Number of Items | 60 | 55 | 5 |
| 6 | Reading | Difficulty | 0.70 ( 0.15) | 0.71 ( 0.15) | 0.44 ( 0.01) |
| | | Discrimination | 0.33 ( 0.08) | 0.33 ( 0.08) | 0.44 ( 0.02) |
| | | Number of Items | 54 | 52 | 2 |
| | Mathematics | Difficulty | 0.52 ( 0.17) | 0.54 ( 0.16) | 0.27 ( 0.09) |
| | | Discrimination | 0.34 ( 0.10) | 0.33 ( 0.09) | 0.45 ( 0.13) |
| | | Number of Items | 60 | 55 | 5 |
| 7 | Reading | Difficulty | 0.69 ( 0.12) | 0.70 ( 0.12) | 0.47 ( 0.00) |
| | | Discrimination | 0.37 ( 0.09) | 0.37 ( 0.08) | 0.50 ( 0.02) |
| | | Number of Items | 54 | 52 | 2 |
| | Mathematics | Difficulty | 0.52 ( 0.17) | 0.54 ( 0.17) | 0.38 ( 0.11) |
| | | Discrimination | 0.33 ( 0.11) | 0.32 ( 0.10) | 0.47 ( 0.08) |
| | | Number of Items | 60 | 55 | 5 |
| 8 | Reading | Difficulty | 0.71 ( 0.14) | 0.72 ( 0.14) | 0.56 ( 0.04) |
| | | Discrimination | 0.34 ( 0.08) | 0.34 ( 0.07) | 0.53 ( 0.04) |
| | | Number of Items | 54 | 52 | 2 |
| | Mathematics | Difficulty | 0.56 ( 0.15) | 0.57 ( 0.14) | 0.40 ( 0.12) |
| | | Discrimination | 0.38 ( 0.09) | 0.37 ( 0.07) | 0.55 ( 0.05) |
| | | Number of Items | 60 | 55 | 5 |
| 10 | Reading | Difficulty | 0.70 ( 0.13) | 0.71 ( 0.13) | 0.59 ( 0.03) |
| | | Discrimination | 0.34 ( 0.09) | 0.33 ( 0.09) | 0.56 ( 0.04) |
| | | N | 59 | 57 | 2 |
| | Mathematics | Difficulty | 0.57 ( 0.13) | 0.57 ( 0.13) | 0.50 ( 0.16) |
| | | Discrimination | 0.39 ( 0.11) | 0.38 ( 0.10) | 0.54 ( 0.15) |
| | | N | 65 | 60 | 5 |

*Note: Numbers shown in parentheses are standard deviations.

## DIFFERENTIAL ITEM FUNCTIONING (DIF)

The *Code of Fair Testing Practices in Education* (2004) explicitly states that subgroup differences in performance should be examined when sample sizes permit, and actions should be taken to make certain that differences in performance are due to construct-relevant, rather than irrelevant, factors. The *Standards for Educational and Psychological Testing* (AERA et al., 1999) includes similar guidelines. As part of the effort to identify such problems, Montana CRT items were evaluated in terms of differential item functioning (DIF) statistics.

DIF procedures are designed to identify items for which subgroups of interest perform differently beyond the impact of differences in overall achievement. For the Montana CRT, the standardization DIF procedure (Dorans and Kulick, 1986) was employed to evaluate subgroup differences for three comparison groups: male/female, white/Native American, and white/Hispanic. This procedure calculates the difference in item performance for groups of students matched for achievement on the total test. That is, the average item performance is calculated for students at every total score, then an overall average is calculated weighting by the total score distribution so the weighting is the same for the two groups. The index ranges from –1 to 1 for multiple-choice items and is adjusted to the same scale for constructed-response items. Negative numbers indicate that the item was more difficult for female or non-white students. Dorans and Holland (1993) suggested that index values between –0.05 and 0.05 should be considered negligible. Most Montana CRT items fall within this range. Dorans and Holland further stated that items with values between –0.10 and –0.05 and between 0.05 and 0.10 (i.e., "low" DIF) should be inspected to ensure that no possible effect is overlooked, and that items with values outside the [–0.10, 0.10] range (i.e., "high" DIF) are more unusual and should be examined very carefully.

DIF indices indicate the degree of differential performance between two groups. That differential performance may or may not be indicative of bias in the test. Course-taking patterns, group differences in interests, or differences in school curricula can lead to DIF. If subgroup differences in performance are related to construct-relevant factors, the items should be considered for inclusion on a test.

Each item was categorized according to the guidelines adapted from Dorans and Holland (1993). Table 8-9 shows the number of items classified into each category separately by item type (multiple choice versus constructed response). Results are shown for male/female, white/Native American, and white/Hispanic comparisons. Table 8-10 provides the number of items in each of the three DIF categories that favor males or females, also separately by item type (multiple-choice and constructed-response). There are some Montana CRT items categorized as "low" or "high" DIF. These indices must not be interpreted as indisputable evidence of bias. Both the *Code of Fair Testing Practices in Education* (2004) and the *Standards for Educational and Psychological Testing* (AERA et al., 1999) assert that test items must be free from construct-irrelevant sources of differential difficulty. If subgroup differences in performance can be plausibly attributed to construct-relevant factors, the items may be included on a test. What is important is to determine if the cause of this differential performance is construct relevant.

For the Montana CRT, there were relatively few items (less than five) flagged as having low or high DIF. The items that were flagged were reviewed for potential bias, and no obvious biases were detected. For this reason, and in order to ensure sufficient content coverage, no items were excluded from the test as a result of the DIF analyses.

## TABLE 8-9
## DIF ANALYSIS – ALL GRADES

| Grade | Content Area | Male/Female DIF Class | | | | | | | | | White/Native American DIF Class | | | | | | | | | White/Hispanic DIF Class | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | All | | | MC | | | CR | | | All | | | MC | | | CR | | | All | | | MC | | | CR | | |
| | | A | B | C | A | B | C | A | B | C | A | B | C | A | B | C | A | B | C | A | B | C | A | B | C | A | B | C |
| 3 | Reading | 51 | 3 | 0 | 49 | 3 | 0 | 2 | 0 | 0 | 48 | 5 | 1 | 46 | 5 | 1 | 2 | 0 | 0 | 49 | 3 | 2 | 47 | 3 | 2 | 2 | 0 | 0 |
| | Math | 57 | 2 | 1 | 52 | 2 | 1 | 5 | 0 | 0 | 57 | 3 | 0 | 52 | 3 | 0 | 5 | 0 | 0 | 53 | 7 | 0 | 48 | 7 | 0 | 5 | 0 | 0 |
| 4 | Reading | 52 | 2 | 0 | 50 | 2 | 0 | 2 | 0 | 0 | 52 | 2 | 0 | 50 | 2 | 0 | 2 | 0 | 0 | 49 | 5 | 0 | 47 | 5 | 0 | 0 | 2 | 0 |
| | Math | 51 | 9 | 0 | 46 | 9 | 0 | 5 | 0 | 0 | 58 | 2 | 0 | 54 | 1 | 0 | 4 | 1 | 0 | 55 | 5 | 0 | 51 | 4 | 0 | 0 | 4 | 1 |
| 5 | Reading | 48 | 6 | 0 | 48 | 4 | 0 | 0 | 2 | 0 | 53 | 1 | 0 | 51 | 1 | 0 | 2 | 0 | 0 | 45 | 9 | 0 | 43 | 9 | 0 | 0 | 2 | 0 |
| | Math | 44 | 15 | 1 | 40 | 14 | 1 | 4 | 1 | 0 | 55 | 5 | 0 | 50 | 5 | 0 | 5 | 0 | 0 | 48 | 12 | 0 | 45 | 10 | 0 | 0 | 3 | 2 |
| 6 | Reading | 47 | 7 | 0 | 47 | 5 | 0 | 0 | 2 | 0 | 49 | 5 | 0 | 47 | 5 | 0 | 2 | 0 | 0 | 52 | 2 | 0 | 50 | 2 | 0 | 0 | 2 | 0 |
| | Math | 49 | 11 | 0 | 46 | 9 | 0 | 3 | 2 | 0 | 56 | 4 | 0 | 52 | 3 | 0 | 4 | 1 | 0 | 54 | 6 | 0 | 50 | 5 | 0 | 0 | 4 | 1 |
| 7 | Reading | 44 | 10 | 0 | 44 | 8 | 0 | 0 | 2 | 0 | 47 | 5 | 2 | 45 | 5 | 2 | 2 | 0 | 0 | 49 | 5 | 0 | 47 | 5 | 0 | 0 | 2 | 0 |
| | Math | 46 | 12 | 2 | 43 | 10 | 2 | 3 | 2 | 0 | 56 | 4 | 0 | 52 | 3 | 0 | 4 | 1 | 0 | 51 | 9 | 0 | 47 | 8 | 0 | 0 | 4 | 1 |
| 8 | Reading | 41 | 10 | 3 | 41 | 8 | 3 | 0 | 2 | 0 | 50 | 4 | 0 | 48 | 4 | 0 | 2 | 0 | 0 | 48 | 6 | 0 | 46 | 6 | 0 | 0 | 2 | 0 |
| | Math | 46 | 13 | 1 | 42 | 13 | 0 | 4 | 0 | 1 | 53 | 7 | 0 | 49 | 6 | 0 | 4 | 1 | 0 | 51 | 8 | 1 | 46 | 8 | 1 | 0 | 5 | 0 |
| 10 | Reading | 48 | 9 | 2 | 48 | 7 | 2 | 0 | 2 | 0 | 53 | 5 | 1 | 51 | 5 | 1 | 2 | 0 | 0 | 53 | 6 | 0 | 51 | 6 | 0 | 0 | 2 | 0 |
| | Math | 57 | 7 | 1 | 53 | 6 | 1 | 4 | 1 | 0 | 63 | 2 | 0 | 58 | 2 | 0 | 5 | 0 | 0 | 56 | 9 | 0 | 51 | 9 | 0 | 0 | 5 | 0 |

A = negligible DIF, B = low DIF, C = high DIF

## TABLE 8-10
## MALE VS. FEMALE DIFFERENTIAL ITEM FUNCTIONING (DIF) CATEGORIZATION BY ITEM TYPE (MULTIPLE-CHOICE AND CONSTRUCTED-RESPONSE)

| Grade | Content Area | Item Type | Negligible DIF (A) | | | | Low DIF (B) | | | | High DIF (C) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Favor Female | Favor Male | N | % | Favor Female | Favor Male | N | % | Favor Female | Favor Male | N | % |
| 3 | Reading | MC | 30 | 19 | 49 | 94 | 1 | 2 | 3 | 6 | 0 | 0 | 0 | 0 |
| | | CR | 2 | 0 | 2 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Math | MC | 28 | 24 | 52 | 95 | 0 | 2 | 2 | 4 | 0 | 1 | 1 | 2 |
| | | CR | 4 | 1 | 5 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | Reading | MC | 28 | 22 | 50 | 96 | 1 | 1 | 2 | 4 | 0 | 0 | 0 | 0 |
| | | CR | 2 | 0 | 2 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Math | MC | 22 | 24 | 46 | 84 | 3 | 6 | 9 | 16 | 0 | 0 | 0 | 0 |
| | | CR | 5 | 0 | 5 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | Reading | MC | 23 | 25 | 48 | 92 | 1 | 3 | 4 | 08 | 0 | 0 | 0 | 0 |
| | | CR | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 100 | 0 | 0 | 0 | 0 |
| | Math | MC | 19 | 21 | 40 | 73 | 7 | 7 | 14 | 25 | 0 | 1 | 1 | 2 |
| | | CR | 3 | 1 | 4 | 80 | 1 | 0 | 1 | 20 | 0 | 0 | 0 | 0 |
| 6 | Reading | MC | 27 | 20 | 47 | 90 | 1 | 4 | 5 | 10 | 0 | 0 | 0 | 0 |
| | | CR | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 100 | 0 | 0 | 0 | 0 |
| | Math | MC | 22 | 24 | 46 | 84 | 3 | 6 | 9 | 16 | 0 | 0 | 0 | 0 |
| | | CR | 3 | 0 | 3 | 60 | 2 | 0 | 2 | 40 | 0 | 0 | 0 | 0 |
| 7 | Reading | MC | 28 | 16 | 44 | 85 | 0 | 8 | 8 | 15 | 0 | 0 | 0 | 0 |
| | | CR | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 100 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Math** | **MC** | 26 | 17 | 43 | 78 | 4 | 6 | 10 | 18 | 0 | 2 | 2 | 4 |
| | | **CR** | 3 | 0 | 3 | 60 | 2 | 0 | 2 | 40 | 0 | 0 | 0 | 0 |
| **8** | **Reading** | **MC** | 23 | 18 | 41 | 79 | 3 | 5 | 8 | 15 | 1 | 2 | 3 | 6 |
| | | **CR** | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 100 | 0 | 0 | 0 | 0 |
| | **Math** | **MC** | 23 | 19 | 42 | 76 | 4 | 9 | 13 | 24 | 0 | 0 | 0 | 0 |
| | | **CR** | 2 | 2 | 4 | 80 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 20 |
| **10** | **Reading** | **MC** | 29 | 19 | 48 | 84 | 2 | 5 | 7 | 12 | 0 | 2 | 2 | 4 |
| | | **CR** | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 100 | 0 | 0 | 0 | 0 |
| | **Math** | **MC** | 29 | 24 | 53 | 88 | 1 | 5 | 6 | 10 | 0 | 1 | 1 | 2 |
| | | **CR** | 4 | 0 | 4 | 80 | 1 | 0 | 1 | 20 | 0 | 0 | 0 | 0 |

## ITEM RESPONSE THEORY ANALYSES

In addition to the classical test theory item analyses previously described, the Montana CRT tests were analyzed according to item response theory (IRT) models. IRT analyses were used, first, to place all 2006 forms on the same scale, and second, to equate the 2006 test to the previous year's test. Details on the IRT calibration and equating procedures for the Montana CRT are provided in Chapter 10.

# CHAPTER 9—RELIABILITY

Although an individual item's performance is an important focus for evaluation, a complete evaluation of an assessment must also address the way items function together and complement one another. Tests that function well provide an accurate assessment of the student's level of ability. Unfortunately, no test can do this perfectly. A variety of factors can contribute to a given student's score being either higher or lower than his or her true ability. For example, a student may mis-read an item, or mistakenly fill in the wrong bubble when he or she knew the answer; similarly a student may get an item correct by guessing, even though he or she did not know the answer. Collectively, these extraneous factors that impact a student's score are referred to as measurement error. Any assessment includes some amount of measurement error; that is, no measurement can be perfectly accurate. This is true of academic assessments—no assessment can measure students perfectly accurately; some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. When tests have a high amount of measurement error, student scores are very unstable. Students with high ability may get low scores or vice versa. Consequently, one cannot reliably tell a student's true level of ability with such a test. Assessments that have less measurement error (i.e., errors made are small on average and student scores on such a test will consistently represent their ability) are described as reliable.

There are a number of ways to estimate an assessment's reliability. One possible approach is to give the same test to the same students at two different points in time. If students receive the same scores on each test, then the extraneous factors affecting performance are small and the test is reliable. (This is referred to as test-retest reliability.) A potential problem with this approach is that students may remember items from the first administration or may have gained (or lost) knowledge or skills in the interim between the two administrations. A solution to the 'remembering items' problem is to give a different, but parallel test at the second administration. If student scores on each test correlate highly the test is considered reliable. (This is known as alternate forms reliability, because an alternate form of the test is used in each administration.) This approach, however, does not address the problem that students may have gained (or lost) knowledge or skills in the interim between the two administrations. In addition, the practical

challenges of developing and administering parallel forms generally preclude the use of parallel forms reliability indices. One way to address these problems is to split the test in half and then correlate students' scores on the two half-tests; this in effect treats each half-test as a complete test. By doing this, the problems associated with an intervening time interval, and of creating and administering two parallel forms of the test, are alleviated. This is known as a split-half estimate of reliability. If the two half-test scores correlate highly, items on the two half-tests must be measuring very similar knowledge or skills. This is evidence that the items complement one another and function well as a group. This also suggests that measurement error will be minimal.

The split-half method requires a judgment regarding the selection of which items contribute to which half-test score. This decision may have an impact on the resulting correlation; different splits will give different estimates of reliability. Cronbach (1951) provided a statistic, α, that avoids this concern about the split-half method. Cronbach's α gives an estimate of the average of all possible splits for a given test. Cronbach's α is often referred to as a measure of internal consistency because it provides a measure of how well all the items in the test measure one single underlying ability. Cronbach's a is computed using the following formula:

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma^2\left(Y_i\right)}{\sigma_x^2}\right]$$

where $i$ indexes the item
$n$ is the total number of items,
$\sigma^2\left(Y_i\right)$ represents individual item variance, and
$\sigma_x^2$ represents the total test variance

In addition to Cronbach's α, another approach to estimating the reliability for a test with differing item types (i.e., multiple-choice and constructed-response) is to assume that at least a small, but important, degree of unique variance is associated with item type (Feldt and Brennan, 1989). In contrast, Cronbach's coefficient α is built upon the assumption that there are no such

local or clustered dependencies. A stratified version of coefficient α corrects for this problem by using the following formula:

$$\alpha_{strat} = 1 - \frac{\sum_{j=1}^{k} \sigma_{x_j}^2 (1 - \alpha_j)}{\sigma_x^2}$$

where j indexes the subtests or categories,
$\sigma_{x_j}^2$ represents the variance of each of the k individual subtests or categories,
$\alpha_j$ is the unstratified Cronbach's $\alpha$ coefficient for each subtest, and
$\sigma_x^2$ represents the total test variance.

## RELIABILITY AND STANDARD ERRORS OF MEASUREMENT

Table 9-1 provides descriptive statistics, the overall Cronbach's α coefficient for each grade/content combination, and raw score standard errors of measurement. Tables 9-2 through 9-8 present Cronbach's α for each test form in each subject area (reading and mathematics), separately for each grade level. The tables also show reliability coefficients separately for multiple-choice and constructed-response (constructed-response and short-answer) items, and stratified reliability coefficients that adjust for the fact that different item formats are included in the test.

Across the grades and content areas, the overall a coefficients, multiple-choice a coefficients, and stratified a coefficients range from the mid-.80s to the low-.90s. There are little or no differences between the overall a and stratified a coefficients. The a coefficients for the constructed-response items are substantially lower, ranging from around 0.50 to around 0.75. These lower values can be explained, at least to some extent, by the fact that there are greater scoring inconsistencies for constructed-response items, as well as the relatively small numbers of these items on the test. Note that, for reading, it is possible that the reliability coefficients are inflated as a result of passage-based item dependency.

## TABLE 9-1
## RELIABILITIES, STANDARD ERRORS OF MEASUREMENT, AND DESCRIPTIVE STATISTICS

| Grade | Content Area | N | Total Points | Mean | SD | Rel | SEM |
|---|---|---|---|---|---|---|---|
| 3 | Reading | 10019 | 60 | 39.53 | 10.13 | 0.90 | 3.16 |
| | Mathematics | 10043 | 66 | 46.03 | 10.76 | 0.90 | 3.37 |
| 4 | Reading | 10326 | 60 | 40.83 | 9.77 | 0.89 | 3.21 |
| | Mathematics | 10349 | 66 | 44.70 | 11.63 | 0.91 | 3.57 |
| 5 | Reading | 10333 | 60 | 37.75 | 9.70 | 0.89 | 3.26 |
| | Mathematics | 10368 | 66 | 37.60 | 11.90 | 0.91 | 3.65 |
| 6 | Reading | 10764 | 60 | 40.31 | 9.02 | 0.88 | 3.12 |
| | Mathematics | 10774 | 66 | 33.09 | 11.73 | 0.90 | 3.73 |
| 7 | Reading | 10993 | 60 | 40.13 | 10.31 | 0.90 | 3.21 |
| | Mathematics | 10994 | 66 | 33.89 | 11.39 | 0.89 | 3.76 |
| 8 | Reading | 11692 | 60 | 41.70 | 9.49 | 0.89 | 3.21 |
| | Mathematics | 11681 | 66 | 35.69 | 12.81 | 0.92 | 3.72 |
| 10 | Reading | 11496 | 65 | 44.88 | 10.18 | 0.89 | 3.37 |
| | Mathematics | 11472 | 71 | 38.91 | 14.10 | 0.93 | 3.77 |

## TABLE 9-2
## RELIABILITY ANALYSIS – GRADE 3

| Content Area | Reliability | Form | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Reading | Coeff α | 0.91 | 0.91 | 0.90 | 0.90 | 0.91 | 0.90 | 0.90 | 0.90 | 0.91 | 0.90 | 0.91 | 0.91 | 0.90 | 0.90 | 0.90 | 0.89 |
| | MC α | 0.91 | 0.91 | 0.90 | 0.90 | 0.91 | 0.90 | 0.90 | 0.90 | 0.91 | 0.91 | 0.91 | 0.91 | 0.90 | 0.90 | 0.90 | 0.89 |
| | CR α | 0.58 | 0.54 | 0.58 | 0.58 | 0.55 | 0.49 | 0.60 | 0.53 | 0.54 | 0.55 | 0.58 | 0.57 | 0.49 | 0.51 | 0.58 | 0.53 |
| | Strat α | 0.91 | 0.91 | 0.90 | 0.90 | 0.91 | 0.90 | 0.90 | 0.90 | 0.91 | 0.91 | 0.92 | 0.91 | 0.90 | 0.90 | 0.90 | 0.90 |
| Mathe-matics | Coeff α | 0.91 | 0.91 | 0.89 | 0.90 | 0.90 | 0.90 | 0.90 | 0.91 | 0.90 | 0.90 | 0.91 | 0.90 | 0.89 | 0.91 | 0.90 | 0.90 |
| | MC α | 0.90 | 0.89 | 0.88 | 0.89 | 0.90 | 0.89 | 0.89 | 0.90 | 0.89 | 0.89 | 0.90 | 0.89 | 0.88 | 0.90 | 0.89 | 0.89 |
| | CR α | 0.59 | 0.56 | 0.48 | 0.54 | 0.55 | 0.57 | 0.56 | 0.58 | 0.57 | 0.56 | 0.59 | 0.56 | 0.50 | 0.60 | 0.54 | 0.59 |
| | Strat α | 0.91 | 0.91 | 0.89 | 0.90 | 0.91 | 0.91 | 0.90 | 0.91 | 0.90 | 0.90 | 0.91 | 0.90 | 0.89 | 0.91 | 0.90 | 0.90 |

## TABLE 9-3
## RELIABILITY ANALYSIS – GRADE 4

| Content Area | Reliability | Form | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Reading | Coeff α | 0.90 | 0.90 | 0.90 | 0.89 | 0.90 | 0.89 | 0.89 | 0.88 | 0.89 | 0.89 | 0.89 | 0.89 | 0.89 | 0.89 | 0.88 | 0.90 |
| | MC α | 0.89 | 0.89 | 0.90 | 0.89 | 0.89 | 0.88 | 0.88 | 0.88 | 0.89 | 0.89 | 0.88 | 0.89 | 0.89 | 0.89 | 0.88 | 0.89 |
| | CR α | 0.54 | 0.54 | 0.53 | 0.51 | 0.51 | 0.56 | 0.55 | 0.50 | 0.52 | 0.53 | 0.53 | 0.57 | 0.48 | 0.41 | 0.50 | 0.57 |
| | Strat α | 0.90 | 0.90 | 0.91 | 0.90 | 0.90 | 0.89 | 0.89 | 0.89 | 0.90 | 0.90 | 0.89 | 0.90 | 0.90 | 0.89 | 0.89 | 0.90 |
| Mathe-matics | Coeff α | 0.92 | 0.91 | 0.91 | 0.91 | 0.91 | 0.90 | 0.90 | 0.90 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 | 0.90 | 0.90 | 0.89 |
| | MC α | 0.91 | 0.90 | 0.91 | 0.90 | 0.90 | 0.89 | 0.89 | 0.89 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.89 | 0.89 | 0.88 |
| | CR α | 0.63 | 0.60 | 0.60 | 0.57 | 0.60 | 0.55 | 0.58 | 0.60 | 0.56 | 0.58 | 0.61 | 0.59 | 0.62 | 0.59 | 0.54 | 0.58 |
| | Strat α | 0.92 | 0.91 | 0.91 | 0.91 | 0.91 | 0.90 | 0.90 | 0.90 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 | 0.90 | 0.90 |

## TABLE 9-4
## RELIABILITY ANALYSIS – GRADE 5

| Content Area | Reliability | Form | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Reading | Coeff α | 0.89 | 0.89 | 0.89 | 0.90 | 0.88 | 0.88 | 0.88 | 0.89 | 0.89 | 0.89 | 0.89 | 0.87 | 0.88 | 0.89 | 0.90 | 0.89 |
| | MC α | 0.89 | 0.89 | 0.88 | 0.89 | 0.87 | 0.87 | 0.88 | 0.89 | 0.88 | 0.88 | 0.88 | 0.87 | 0.87 | 0.88 | 0.89 | 0.88 |
| | CR α | 0.59 | 0.61 | 0.60 | 0.60 | 0.65 | 0.63 | 0.64 | 0.66 | 0.57 | 0.53 | 0.61 | 0.58 | 0.59 | 0.59 | 0.66 | 0.58 |
| | Strat α | 0.90 | 0.90 | 0.89 | 0.90 | 0.89 | 0.88 | 0.89 | 0.90 | 0.89 | 0.89 | 0.89 | 0.88 | 0.89 | 0.89 | 0.90 | 0.89 |
| Mathe-matics | Coeff α | 0.91 | 0.91 | 0.91 | 0.91 | 0.90 | 0.90 | 0.90 | 0.90 | 0.91 | 0.91 | 0.90 | 0.90 | 0.90 | 0.90 | 0.92 | 0.91 |
| | MC α | 0.90 | 0.90 | 0.89 | 0.90 | 0.89 | 0.89 | 0.89 | 0.89 | 0.89 | 0.90 | 0.89 | 0.88 | 0.89 | 0.89 | 0.91 | 0.89 |
| | CR α | 0.61 | 0.61 | 0.61 | 0.62 | 0.58 | 0.63 | 0.60 | 0.54 | 0.59 | 0.60 | 0.56 | 0.57 | 0.61 | 0.59 | 0.61 | 0.57 |
| | Strat α | 0.91 | 0.91 | 0.91 | 0.91 | 0.90 | 0.91 | 0.90 | 0.90 | 0.91 | 0.91 | 0.90 | 0.90 | 0.90 | 0.91 | 0.92 | 0.91 |

## TABLE 9-5
## RELIABILITY ANALYSIS – GRADE 6

| Content Area | Reliability | Form | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Reading | Coeff α | 0.89 | 0.88 | 0.88 | 0.87 | 0.87 | 0.89 | 0.87 | 0.87 | 0.89 | 0.89 | 0.88 | 0.88 | 0.88 | 0.88 | 0.87 | 0.88 |
| | MC α | 0.89 | 0.88 | 0.88 | 0.87 | 0.87 | 0.88 | 0.87 | 0.86 | 0.88 | 0.88 | 0.87 | 0.87 | 0.87 | 0.87 | 0.86 | 0.87 |
| | CR α | 0.58 | 0.58 | 0.56 | 0.55 | 0.55 | 0.65 | 0.57 | 0.58 | 0.69 | 0.62 | 0.60 | 0.59 | 0.61 | 0.64 | 0.63 | 0.62 |
| | Strat α | 0.90 | 0.89 | 0.89 | 0.88 | 0.88 | 0.89 | 0.88 | 0.88 | 0.89 | 0.89 | 0.88 | 0.88 | 0.89 | 0.89 | 0.87 | 0.88 |
| Mathe-matics | Coeff α | 0.90 | 0.90 | 0.90 | 0.89 | 0.90 | 0.90 | 0.89 | 0.90 | 0.90 | 0.89 | 0.90 | 0.90 | 0.90 | 0.90 | 0.89 | 0.91 |
| | MC α | 0.88 | 0.89 | 0.89 | 0.88 | 0.89 | 0.89 | 0.88 | 0.89 | 0.89 | 0.88 | 0.88 | 0.89 | 0.89 | 0.88 | 0.88 | 0.90 |
| | CR α | 0.57 | 0.59 | 0.58 | 0.55 | 0.57 | 0.59 | 0.60 | 0.57 | 0.59 | 0.56 | 0.60 | 0.58 | 0.59 | 0.56 | 0.56 | 0.61 |
| | Strat α | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.89 | 0.90 | 0.91 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.89 | 0.91 |

## TABLE 9-6
## RELIABILITY ANALYSIS – GRADE 7

| Content Area | Reliability | Form | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Reading | Coeff α | 0.90 | 0.91 | 0.89 | 0.89 | 0.91 | 0.92 | 0.91 | 0.90 | 0.90 | 0.90 | 0.90 | 0.89 | 0.89 | 0.91 | 0.90 | 0.90 |
| | MC α | 0.90 | 0.90 | 0.89 | 0.89 | 0.90 | 0.92 | 0.91 | 0.90 | 0.89 | 0.90 | 0.89 | 0.89 | 0.89 | 0.91 | 0.90 | 0.90 |
| | CR α | 0.68 | 0.65 | 0.64 | 0.63 | 0.72 | 0.67 | 0.65 | 0.66 | 0.64 | 0.66 | 0.67 | 0.68 | 0.67 | 0.64 | 0.68 | 0.67 |
| | Strat α | 0.91 | 0.91 | 0.90 | 0.90 | 0.91 | 0.92 | 0.92 | 0.91 | 0.90 | 0.91 | 0.90 | 0.90 | 0.90 | 0.92 | 0.91 | 0.91 |
| Mathe-matics | Coeff α | 0.89 | 0.89 | 0.89 | 0.89 | 0.89 | 0.90 | 0.89 | 0.89 | 0.89 | 0.90 | 0.88 | 0.88 | 0.88 | 0.89 | 0.88 | 0.89 |
| | MC α | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 | 0.86 | 0.87 | 0.87 | 0.88 | 0.87 | 0.87 |
| | CR α | 0.61 | 0.55 | 0.58 | 0.56 | 0.60 | 0.60 | 0.58 | 0.58 | 0.59 | 0.59 | 0.59 | 0.51 | 0.56 | 0.60 | 0.54 | 0.58 |
| | Strat α | 0.90 | 0.89 | 0.90 | 0.89 | 0.89 | 0.90 | 0.89 | 0.89 | 0.90 | 0.90 | 0.89 | 0.88 | 0.89 | 0.90 | 0.89 | 0.89 |

## TABLE 9-7
## RELIABILITY ANALYSIS – GRADE 8

| Content Area | Reliability | Form | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Reading | Coeff α | 0.88 | 0.88 | 0.89 | 0.87 | 0.88 | 0.89 | 0.89 | 0.89 | 0.89 | 0.89 | 0.90 | 0.89 | 0.88 | 0.88 | 0.88 | 0.89 |
| | MC α | 0.88 | 0.87 | 0.89 | 0.87 | 0.87 | 0.88 | 0.88 | 0.89 | 0.88 | 0.88 | 0.89 | 0.88 | 0.87 | 0.87 | 0.87 | 0.88 |
| | CR α | 0.63 | 0.70 | 0.70 | 0.69 | 0.65 | 0.69 | 0.70 | 0.68 | 0.72 | 0.71 | 0.71 | 0.70 | 0.66 | 0.68 | 0.66 | 0.68 |
| | Strat α | 0.89 | 0.89 | 0.90 | 0.88 | 0.89 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.91 | 0.90 | 0.89 | 0.89 | 0.89 | 0.90 |
| Mathe-matics | Coeff α | 0.91 | 0.91 | 0.92 | 0.91 | 0.91 | 0.92 | 0.92 | 0.92 | 0.91 | 0.91 | 0.92 | 0.91 | 0.92 | 0.91 | 0.91 | 0.92 |
| | MC α | 0.90 | 0.90 | 0.91 | 0.90 | 0.90 | 0.91 | 0.91 | 0.91 | 0.90 | 0.90 | 0.91 | 0.90 | 0.91 | 0.90 | 0.90 | 0.91 |
| | CR α | 0.62 | 0.62 | 0.65 | 0.60 | 0.60 | 0.63 | 0.64 | 0.63 | 0.62 | 0.60 | 0.66 | 0.59 | 0.62 | 0.60 | 0.57 | 0.63 |
| | Strat α | 0.92 | 0.91 | 0.92 | 0.91 | 0.91 | 0.92 | 0.92 | 0.92 | 0.91 | 0.92 | 0.93 | 0.91 | 0.92 | 0.92 | 0.91 | 0.92 |

## TABLE 9-8
## RELIABILITY ANALYSIS – GRADE 10

| Content Area | Reliability | Form | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Reading | Coeff α | 0.90 | 0.89 | 0.89 | 0.87 | 0.89 | 0.89 | 0.90 | 0.89 | 0.89 | 0.89 | 0.89 | 0.88 | 0.90 | 0.89 | 0.89 | 0.89 |
| | MC α | 0.89 | 0.88 | 0.88 | 0.86 | 0.88 | 0.88 | 0.89 | 0.88 | 0.88 | 0.88 | 0.88 | 0.87 | 0.89 | 0.88 | 0.88 | 0.88 |
| | CR α | 0.74 | 0.72 | 0.72 | 0.71 | 0.74 | 0.74 | 0.73 | 0.72 | 0.76 | 0.74 | 0.72 | 0.74 | 0.73 | 0.69 | 0.69 | 0.72 |
| | Strat α | 0.91 | 0.90 | 0.90 | 0.88 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.89 | 0.91 | 0.90 | 0.90 | 0.89 |
| Mathe-matics | Coeff α | 0.93 | 0.93 | 0.93 | 0.92 | 0.93 | 0.92 | 0.92 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 |
| | MC α | 0.92 | 0.92 | 0.92 | 0.91 | 0.91 | 0.91 | 0.91 | 0.92 | 0.92 | 0.92 | 0.91 | 0.92 | 0.92 | 0.92 | 0.92 | 0.92 |
| | CR α | 0.72 | 0.67 | 0.70 | 0.64 | 0.66 | 0.66 | 0.65 | 0.70 | 0.68 | 0.67 | 0.67 | 0.67 | 0.69 | 0.68 | 0.66 | 0.67 |
| | Strat α | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.92 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.94 | 0.93 | 0.93 | 0.93 |

## RELIABILITY OF PERFORMANCE LEVEL CATEGORIZATION

All test scores contain measurement error; thus classifications based on test scores are also subject to measurement error. After the performance levels were specified and students were classified into those levels, empirical analyses were conducted to determine the statistical accuracy and consistency of the classifications. For the Montana CRT, students are classified into one of four performance levels: *Novice* (N), *Nearing Proficiency* (NP), *Proficient* (P), or *Advanced* (A). This section of the report explains the methodologies used to assess the reliability of classification decisions, and results are given.

## ACCURACY

Accuracy refers to the extent to which decisions based on test scores match decisions that would have been made if the scores did not contain any measurement error. Accuracy must be estimated because errorless test scores do not exist.

## CONSISTENCY

Consistency measures the extent to which classification decisions based on test scores match the decisions based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test items if two complete, parallel forms of the test are given to the same group of students. This is usually impractical, especially on lengthy tests. To overcome this issue, techniques have been developed to estimate both accuracy and consistency of classification decisions based on a single administration of a test. The technique developed by Livingston and Lewis (1995) was used for the Montana CRT because their technique can be used with both constructed-response and multiple-choice items.

## CALCULATING ACCURACY

All of the accuracy and consistency estimation techniques described below make use of the concept of "true scores" in the sense of classical test theory. A true score is the score that would be obtained on a test that had no measurement error. It is a theoretical concept that cannot be

observed, although it can be estimated. In the Livingston and Lewis method, the estimated true score distribution is used to estimate the proportion of students in each "true" performance level. After various technical adjustments (which are described in Livingston and Lewis, 1995), a $4 \times 4$ contingency table was created for each content area test and grade level. The $[i,j]$ entry of an accuracy table represents the estimated proportion of students whose true score fell into performance level $i$ and whose observed score fell into performance level $j$ on the Montana CRT. Overall accuracy, which is the proportion of students whose true and observed performance levels match one another, is the sum of the numbers on the diagonal of the accuracy table.

## CALCULATING CONSISTENCY

To estimate consistency, the true scores are used to estimate the joint distribution of classifications on two independent, parallel test forms. After statistical adjustments (see Livingston and Lewis, 1995), a new $4 \times 4$ contingency table was created for each test and grade level that shows the proportion of students who would be classified into each performance level by the two (hypothetical) parallel test forms. That is, the $[i,j]$ entry of a consistency table represents the estimated proportion of students whose observed score on the first form would fall into performance level $i$ and whose observed score on the second form would fall into performance level $j$. Overall consistency, which is the proportion of students classified into exactly the same performance level by the two forms of the test, is the sum of the numbers on the diagonal of this new contingency table.

## KAPPA

Another way to measure consistency is to use Cohen's (1960) coefficient $\kappa$ (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance. Cohen's $\kappa$ can be used to evaluate the classification consistency of a test from two parallel forms of the test. The two forms in this case were the hypothetical parallel forms used by the Livingston and Lewis method. Because $\kappa$ is corrected for chance, the values of $\kappa$ are lower than other consistency estimates.

## RESULTS OF ACCURACY, CONSISTENCY, AND KAPPA ANALYSES

Summaries of the Accuracy and Consistency analyses are provided in Tables 9-9 through 9-22. The first section of each table shows the overall accuracy and consistency indices as well as Kappa. The overall index is, as described above, the sum of the diagonal elements of the appropriate contingency table.

The second section of each table shows accuracy and consistency values conditional upon performance level. In each case, the denominator is the number of students who are associated with a given performance level. For example, the conditional accuracy value is 0.7656 for the *Proficient* category for Grade 4 Math. This indicates that, of the students whose true scores placed them in the *Proficient* category, 76.56% of them would be expected to be in the *Proficient* category if they were categorized according to their observed scores. The corresponding consistency value of .6900 indicates that 69% of students with observed scores in the *Proficient* performance level would be expected to score in *Proficient* again if a second, parallel test form were used.

For certain tests, concern may be greatest regarding decisions made about a particular threshold. For example, if a college gave credit to students who achieved an Advanced Placement test score of four or five, but not one, two, or three, one might be interested in the accuracy of the dichotomous decision, below four versus four or above. The third section of the summary tables shows information at each of the cut points. These values indicate the accuracy and consistency of the dichotomous decisions, either above or below the associated cut point. In addition, the false positive and false negative accuracy rates are also provided. These values are estimates of the proportion of students who were categorized above the cut when their true score would place them below the cut (false positive), and vice versa.

## TABLE 9-9
## ACCURACY AND CONSISTENCY -- GRADE 3 MATH

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7682 | | 0.6818 | | 0.5529 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8478 | | 0.7538 |
| | ***Nearing Proficiency*** | | 0.6180 | | 0.5017 |
| | ***Proficient*** | | 0.7746 | | 0.7051 |
| | ***Advanced*** | | 0.8337 | | 0.7365 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9425 | 0.0236 | 0.0339 | 0.9186 |
| | ***NP : P*** | 0.9113 | 0.0416 | 0.0471 | 0.8753 |
| | ***P : A*** | 0.9132 | 0.0510 | 0.0358 | 0.8785 |

## TABLE 9-10
## ACCURACY AND CONSISTENCY -- GRADE 4 MATH

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7768 | | 0.6922 | | 0.5763 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8541 | | 0.7662 |
| | ***Nearing Proficiency*** | | 0.6496 | | 0.5375 |
| | ***Proficient*** | | 0.7656 | | 0.6900 |
| | ***Advanced*** | | 0.8559 | | 0.7717 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9428 | 0.0239 | 0.0334 | 0.9191 |
| | ***NP : P*** | 0.9138 | 0.0412 | 0.0449 | 0.8791 |
| | ***P : A*** | 0.9195 | 0.0467 | 0.0338 | 0.8870 |

## TABLE 9-11
## ACCURACY AND CONSISTENCY -- GRADE 5 MATH

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7655 | | 0.6776 | | 0.5573 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8280 | | 0.7246 |
| | ***Nearing Proficiency*** | | 0.6434 | | 0.5337 |
| | ***Proficient*** | | 0.7583 | | 0.6789 |
| | ***Advanced*** | | 0.8574 | | 0.7753 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9347 | 0.0267 | 0.0386 | 0.9080 |
| | ***NP : P*** | 0.9073 | 0.0445 | 0.0482 | 0.8700 |
| | ***P : A*** | 0.9225 | 0.0446 | 0.0329 | 0.8911 |

## TABLE 9-12
## ACCURACY AND CONSISTENCY -- GRADE 6 MATH

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7600 | | 0.6717 | | 0.5502 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8068 | | 0.6954 |
| | ***Nearing Proficiency*** | | 0.6233 | | 0.5139 |
| | ***Proficient*** | | 0.7604 | | 0.6795 |
| | ***Advanced*** | | 0.8672 | | 0.7904 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9261 | 0.0303 | 0.0435 | 0.8963 |
| | ***NP : P*** | 0.9043 | 0.0465 | 0.0493 | 0.8655 |
| | ***P : A*** | 0.9282 | 0.0412 | 0.0306 | 0.8990 |

## TABLE 9-13
## ACCURACY AND CONSISTENCY -- GRADE 7 MATH

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7621 | | 0.6741 | | 0.5524 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.7992 | | 0.6874 |
| | ***Nearing Proficiency*** | | 0.6314 | | 0.5237 |
| | ***Proficient*** | | 0.7684 | | 0.6894 |
| | ***Advanced*** | | 0.8670 | | 0.7893 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9254 | 0.0310 | 0.0435 | 0.8956 |
| | ***NP : P*** | 0.9050 | 0.0466 | 0.0485 | 0.8668 |
| | ***P : A*** | 0.9304 | 0.0400 | 0.0296 | 0.9019 |

## TABLE 9-14
## ACCURACY AND CONSISTENCY -- GRADE 8 MATH

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7770 | | 0.6905 | | 0.5800 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8262 | | 0.7266 |
| | ***Nearing Proficiency*** | | 0.7034 | | 0.6041 |
| | ***Proficient*** | | 0.7510 | | 0.6651 |
| | ***Advanced*** | | 0.8685 | | 0.7940 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9346 | 0.0273 | 0.0381 | 0.9081 |
| | ***NP : P*** | 0.9128 | 0.0430 | 0.0442 | 0.8772 |
| | ***P : A*** | 0.9294 | 0.0402 | 0.0304 | 0.9007 |

## TABLE 9-15
## ACCURACY AND CONSISTENCY -- GRADE 10 MATH

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.8041 | | 0.7255 | | 0.6289 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8355 | | 0.7505 |
| | ***Nearing Proficiency*** | | 0.7540 | | 0.6663 |
| | ***Proficient*** | | 0.7766 | | 0.6947 |
| | ***Advanced*** | | 0.8850 | | 0.8213 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9372 | 0.0280 | 0.0348 | 0.9116 |
| | ***NP : P*** | 0.9251 | 0.0378 | 0.0371 | 0.8944 |
| | ***P : A*** | 0.9418 | 0.0326 | 0.0256 | 0.9180 |

## TABLE 9-16
## ACCURACY AND CONSISTENCY -- GRADE 3 READING

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.8104 | | 0.7334 | | 0.5935 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8308 | | 0.7081 |
| | ***Nearing Proficiency*** | | 0.7142 | | 0.5994 |
| | ***Proficient*** | | 0.8027 | | 0.7291 |
| | ***Advanced*** | | 0.8531 | | 0.7942 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9755 | 0.0090 | 0.0154 | 0.9651 |
| | ***NP : P*** | 0.9423 | 0.0247 | 0.0330 | 0.9186 |
| | ***P : A*** | 0.8925 | 0.0532 | 0.0543 | 0.8483 |

## TABLE 9-17
## ACCURACY AND CONSISTENCY -- GRADE 4 READING

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.8015 | | 0.7223 | | 0.5738 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8349 | | 0.7174 |
| | ***Nearing Proficiency*** | | 0.6876 | | 0.5671 |
| | ***Proficient*** | | 0.8109 | | 0.7444 |
| | ***Advanced*** | | 0.8295 | | 0.758 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9724 | 0.0104 | 0.0173 | 0.9606 |
| | ***NP : P*** | 0.9401 | 0.0257 | 0.0342 | 0.9152 |
| | ***P : A*** | 0.8888 | 0.0567 | 0.0545 | 0.8443 |

## TABLE 9-18
## ACCURACY AND CONSISTENCY -- GRADE 5 READING

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7984 | | 0.7183 | | 0.5735 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8256 | | 0.6928 |
| | ***Nearing Proficiency*** | | 0.6922 | | 0.5710 |
| | ***Proficient*** | | 0.7917 | | 0.7221 |
| | ***Advanced*** | | 0.8485 | | 0.7818 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9724 | 0.0098 | 0.0178 | 0.9605 |
| | ***NP : P*** | 0.9342 | 0.0277 | 0.0382 | 0.9068 |
| | ***P : A*** | 0.8917 | 0.0564 | 0.0519 | 0.8485 |

## TABLE 9-19
## ACCURACY AND CONSISTENCY -- GRADE 6 READING

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7778 | | 0.6920 | | 0.5360 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8270 | | 0.7012 |
| | ***Nearing Proficiency*** | | 0.6513 | | 0.5263 |
| | ***Proficient*** | | 0.7823 | | 0.7122 |
| | ***Advanced*** | | 0.8205 | | 0.7394 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9665 | 0.0123 | 0.0213 | 0.9523 |
| | ***NP : P*** | 0.9276 | 0.0306 | 0.0418 | 0.8979 |
| | ***P : A*** | 0.8833 | 0.0620 | 0.0548 | 0.8372 |

## TABLE 9-20
## ACCURACY AND CONSISTENCY -- GRADE 7 READING

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.8040 | | 0.7259 | | 0.5898 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8449 | | 0.7395 |
| | ***Nearing Proficiency*** | | 0.6880 | | 0.5708 |
| | ***Proficient*** | | 0.8115 | | 0.7458 |
| | ***Advanced*** | | 0.8384 | | 0.7682 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9682 | 0.0124 | 0.0194 | 0.9548 |
| | ***NP : P*** | 0.9378 | 0.0274 | 0.0349 | 0.9120 |
| | ***P : A*** | 0.8978 | 0.0528 | 0.0494 | 0.8568 |

## TABLE 9-21
## ACCURACY AND CONSISTENCY -- GRADE 8 READING

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7926 | | 0.7116 | | 0.5754 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8211 | | 0.6953 |
| | ***Nearing Proficiency*** | | 0.6593 | | 0.5417 |
| | ***Proficient*** | | 0.7963 | | 0.7351 |
| | ***Advanced*** | | 0.8627 | | 0.7828 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9605 | 0.0146 | 0.0249 | 0.9439 |
| | ***NP : P*** | 0.9205 | 0.0353 | 0.0442 | 0.8880 |
| | ***P : A*** | 0.9111 | 0.0512 | 0.0377 | 0.8749 |

## TABLE 9-22
## ACCURACY AND CONSISTENCY -- GRADE 10 READING

| Accuracy and Consistency of Classification Indices | | | | | |
|---|---|---|---|---|---|
| **Overall Indices** | **Accuracy** | | **Consistency** | | **Kappa (κ)** |
| | 0.7894 | | 0.7082 | | 0.5707 |
| **Indices Conditional on Level** | | | **Accuracy** | | **Consistency** |
| | ***Novice*** | | 0.8232 | | 0.7013 |
| | ***Nearing Proficiency*** | | 0.6366 | | 0.5175 |
| | ***Proficient*** | | 0.7997 | | 0.7400 |
| | ***Advanced*** | | 0.8597 | | 0.7764 |
| **Indices at Cut Points** | | **Accuracy** | | | **Consistency** |
| | | **Accuracy** | ***False Positives*** | ***False Negatives*** | |
| | ***N : NP*** | 0.9567 | 0.0162 | 0.0270 | 0.9386 |
| | ***NP : P*** | 0.9187 | 0.0362 | 0.0451 | 0.8855 |
| | ***P : A*** | 0.9132 | 0.0504 | 0.0364 | 0.8778 |

# CHAPTER 10— SCALING AND EQUATING

The purpose of equating is to ensure that scores obtained from different forms of a test are equivalent to each other. Equating may be used if multiple test forms are administered in the same year, as well as to equate one year's forms to those given in the previous year. Equating ensures that students are not given an unfair advantage or disadvantage because the test form given in one year is easier or harder than the form given in the other year. Once test scores for the forms are placed on an equivalent raw score scale, they then get translated, through the scaling process, to the score scale that is used for reporting. For the 2006 MontCAS, Phase 2 CRT, equating was performed for reading and mathematics, grades 4, 8 and 10; the remaining tests (reading and mathematics, grades 3, 5, 6, and 7) were new in 2006.

## GENERAL RULES

The following general rules are containing in the equating plan for the CRT:

- The goal is to have as many items as possible on the common form constitute the equating set.
- Items used for equating cannot be altered from their appearance in the previous form in any way.
- Whenever possible, items in the equating set should be selected so that they are within five positions of their location on the previous form.
- Passage sets selected for equating should consist of all, or most, of the items associated with the passage.
- The equating set, as a whole group of items, should mirror the characteristics of the common form in terms of content and statistics.

To determine the final set of equating items for each grade level and subject combination, a differential item functioning (DIF) approach using the delta plot method was applied. The 2006 and 2005 p-values of each multiple-choice item were transformed to the delta metric. The delta scale is an inverse normal transformation of percentage correct to a linear scale with a mean of

13 and standard deviation of 4 (Holland & Wainer, 1993). A high delta value indicates a difficult item. For constructed-response items, the average score divided by the maximum possible score, or adjusted p-value, was transformed to the delta metric. The delta values for the potential equating items were computed for each subject in each grade level.

Once all the delta values were calculated, a trend line was fit to the set of points. The perpendicular distance of each item to the regression line was then computed. Items that were not more than three standard deviations away from the regression line were used as equating items. As a result of the delta analyses, two items on the grade 8 reading test were excluded for use as equating items; all equating items were used for the remaining tests.

## IRT EQUATING

Equating for the MontCAS, Phase 2 CRT used the *anchor-test-nonequivalent-groups* design described by Petersen, Kolen, & Hoover (1989). The fixed common-item IRT procedure was used, in which the anchor items from the previous year's administration were identified during this year's calibrations, and their IRT parameters were fixed to last year's values. This method results in all person and item parameters being on the same θ scale as last year. Because of the equating model that is used for the Montana CRT, the process of equating and scaling does not change the rank ordering of students, give more weight to particular items, or change students' performance-level classifications. Note that the groups of students who took the Montana CRT in 2004-05 and 2005-06 were not equivalent. Item Response Theory (IRT) is particularly useful in equating for nonequivalent groups (Allen & Yen, 1979).

IRT uses mathematical models to define a relationship between an unobserved measure of student ability, usually referred to as theta ($\theta$), and the probability ($p$) of getting a dichotomous item correct or of getting a particular score on a polytomous item. In IRT, it is assumed that all items are independent measures of the same construct or ability (i.e., the same $\theta$). There are several IRT models commonly used to specify the relationship between $\theta$ and $p$. For the Montana CRT tests, the 1 parameter logistic (1PL) model was used for multiple-choice and short-answer items and the partial credit model was used for the constructed-response items.

For polytomous items, the generalized partial credit model can be defined as:

$$P_{jk}(\theta) = \frac{\exp \sum_{v=0}^{k} \left[ Da_j \left( \theta - b_j + d_v \right) \right]}{\sum_{c=1}^{m} \exp \sum_{v=1}^{c} \left[ Da_j \left( \theta - b_j + d_v \right) \right]}$$

where $j$ indexes the items,
$k$ indexes students,
$a$ represents item discrimination,
$b$ represents item difficulty,
$d$ represents category step parameter, and
$D$ is a normalizing constant equal to 1.701.

In the case of the Montana CRT, the $a_j$ term in the above equation is equal to 1.0 for all items. For the dichotomous items, because there are no step parameters ($d_v$) the above equation reduces to the following:

$$P_j(\theta) = \frac{\exp\left(\theta - b_j\right)}{1 + \exp\left(\theta - b_j\right)}$$

For more information on IRT and IRT models the reader is referred to Hambleton and Swaminathan (1985).

The process of determining the specific mathematical relationship between $\theta$ and $p$ is referred to as item calibration. Once items are calibrated, they are defined by a set of parameters which specify a non-linear relationship between $\theta$ and $p$. For more information about item calibration the reader is referred to Lord and Novick (1968) or Hambleton and Swaminathan (1985).

PARSCALE v3.5 (Muraki & Bock, 1999) was the software used to do the IRT analyses. The item parameter files resulting from the analyses are provided in Appendix A. Each item occupied only one block in the calibration run, and the 1.701 normalizing constant was used. A default convergence criterion of 0.001 was used, and all calibrations converged within 35 iterations.

## TRANSLATING RAW SCORES TO SCALED SCORES AND PERFORMANCE LEVELS

Montana CRT scores in each content area are reported on a scale that ranges from 200 to 300. Scaled scores supplement the Montana CRT performance-level results by providing information about the position of a student's results within a performance level. School- and district-level scaled scores are calculated by computing the average of student-level scaled scores. Students' raw scores, or total number of points, on the Montana CRT tests are translated to scaled scores using a data analysis process called **scaling**. Scaling simply converts raw points from one scale to another. In the same way that distance can be expressed in miles or kilometers, or monetary value can be expressed in terms of U.S. dollars or Canadian dollars, student scores on each Montana CRT could be expressed as raw scores (i.e., number right) or scaled scores. It is also important to notice that the raw score to scale score conversion formulae vary from CRT to CRT, analogous to how currency exchange formulae vary from country to country. For example, the scaling conversion formula for Montana's Grade 4 Reading CRT differs from that of the Grade 8 Reading CRT.

It is important to note that converting from raw scores to scaled scores does not change the students' performance-level classifications. Given the relative simplicity of raw scores, it is fair to ask why scaled scores are used in Montana CRT reports instead of raw scores. Foremost, scaled scores offer the advantage of simplifying the reporting of results across content areas, grade levels, and subsequent years. Because the standard-setting process typically results in different cut scores across content areas on a raw score basis, it is useful to transform these raw cut scores to a scale that is more easily interpretable and consistent. For the Montana CRT, a score of 225 is the cut score between the *Novice* and *Nearing Proficiency* performance levels. This is true regardless of which content area, grade, or year one may be concerned with. If one were to use raw scores, the raw cut score between *Novice* and *Nearing Proficiency* may be, for example, 35 in mathematics at grade 8, but may be 33 in mathematics at grade 10. Using scaled scores greatly simplifies the task of understanding how a student performed.

Cut points for all tests for the MontCAS, Phase 2 CRT were set at standard setting meetings held in June and July, 2006 (see Appendix C: Standard Setting Report). Standards validation occurred

in June for reading and mathematics, grades 4, 8 and 10, and in July for the remaining grades. Cut points were established on the raw score scale, and these raw score cuts were used to determine the scaling coefficients for calculating the scores used for reporting (see description below and Appendix D). Cut points were also determined on the ?-scale. For scaling in 2007 (and future years), raw score equivalents for these ?-scale cut points will be determined using the test characteristic curve (TCC), and these 2007 raw cuts will be used to calculate transformation constants to be used in 2007.

Once the 2006 raw score cut points were determined via standard setting, the next step was to calculate the transformation coefficients that would be used to place students' raw scores onto the score scale used for reporting. As previously stated, student scores on the Montana CRT are reported in integer values from 200 to 300 with three scores representing cut scores on each assessment. Two of the three cut points (*Novice*/*Nearing Proficiency* and *Nearing Proficiency*/*Proficient*) were pre-set at 225 and 250, respectively; the third cut point, between *Proficient* and *Advanced*, was allowed to vary across tests, depending on where the raw score cuts were placed. Allowing the upper cut to float results in a single conversion equation for each test, which simplifies interpretation of scaled scores and their summary statistics. Table 10-1 presents the scaled score range for each performance level in each grade/content area combination.

**TABLE 10-1**

| Grade | Content Area | SCALED SCORE RANGE FOR EACH PERFORMANCE LEVEL | | | |
|---|---|---|---|---|---|
| | | *Novice* | *Nearing Proficiency* | *Proficient* | *Advanced* |
| 3 | Reading | 200–224 | 225–249 | 250–284 | 285–300 |
| | Mathematics | 200–224 | 225–249 | 250–290 | 291–300 |
| 4 | Reading | 200–224 | 225–249 | 250–287 | 288–300 |
| | Mathematics | 200–224 | 225–249 | 250–286 | 287–300 |
| 5 | Reading | 200–224 | 225–249 | 250–286 | 287–300 |
| | Mathematics | 200–224 | 225–249 | 250–290 | 291–300 |
| 6 | Reading | 200–224 | 225–249 | 250–288 | 289–300 |
| | Mathematics | 200–224 | 225–249 | 250–290 | 291–300 |
| 7 | Reading | 200–224 | 225–249 | 250–288 | 289–300 |
| | Mathematics | 200–224 | 225–249 | 250–290 | 291–300 |
| 8 | Reading | 200–224 | 225–249 | 250–290 | 291–300 |
| | Mathematics | 200–224 | 225–249 | 250–282 | 283–300 |
| 10 | Reading | 200–224 | 225–249 | 250–289 | 290–300 |
| | Mathematics | 200–224 | 225–249 | 250–277 | 278–300 |

The scaled scores are obtained by a simple linear transformation of the raw scores using the values of 225 and 250 on the scaled score metric and the associated 2006 raw score cut points to define the transformation. The scaling coefficients were calculated using the following formulae:

$$b = 225 - m(x_1)$$

$$m = \frac{225 - 250}{x_1 - x_2}$$

where m is the slope of the line providing the relationship between the raw and scaled scores, b is the intercept, $x_1$ is the cut score on the raw score metric for the *Novice/Nearing Proficiency* cut, and $x_2$ is the cut score on the raw score metric for the *Nearing Proficiency/Proficient* cut. Scaled scores were then calculated using the following linear transformation:

$$ss = m(x) + b$$

where x represents a student's raw score. The values obtained using this formula were rounded to the nearest integer and truncated, as necessary, such that no student received a score below 200 or higher than 300. Additional information regarding raw scores, scaled scores, performance level descriptors, and content-specific descriptors may be found in Appendix D.

# CHAPTER 11—REPORTING

The CRT assessments were designed to measure student performance against Montana's Content Standards. Consistent with this purpose, results on the CRT were reported in terms of performance levels that describe student performance in relation to these established state standards. There are four performance levels: *Advanced*, *Proficient*, *Nearing Proficiency*, and *Novice* (CRT Performance Level Descriptors, Content-Specific Descriptors, Scaled Score Ranges, and Raw Scores are described in greater detail in Appendix D). Students receive a separate performance-level classification (based on total scaled score) in each content area.

School- and system-level results are reported as the number and percentage of students attaining each performance level at each grade level tested. Disaggregations of students are also reported at the school and system levels. The CRT reports are

- Student Reports;
- Class Roster & Item-Level Reports;
- School Summary Reports; and
- System Summary Reports.

"Decision Rules" were formulated in late spring 2006 by OPI and Measured Progress to identify students, during the reporting process, to be excluded from school and system-level reports. A copy of these "Decision Rules" is included in this report as Appendix F.

State summary results were provided to OPI on confidential CDs and via a secure Web site. The report formats are included in Appendix E. These reports were shipped to System Test Coordinators on or before September 15, 2006 for distribution to schools within their respective systems/districts. System Test Coordinators and teachers were also provided with copies of the *Guide to Interpreting the 2006 Criterion-Referenced Test and CRT-ALT Assessment Reports* and *iAnalyze*, to assist them in understanding the connection between the assessment and the classroom. The guide provides information about the assessment and the use of assessment results.

## Local Data Analysis and Interpretation

Using advanced Web technology, *iAnalyze* gives Montana educators and administrators the ability to filter data based on test year, grade level, and subject. Data can be sorted to isolate areas of strong or poor performance. Cross sections of data may be viewed by groupings based on demographics such as gender, Title 1 status, etc.

The confidential nature of the data therein necessitates the strict enforcement of site security. All transmissions are done over Secure Socket Layers (SSL). A system of user role definitions and permissions dictates the scope of access granted to individual users. Organizations (system or school levels) are given administrative power to grant or deny access to their data within the system, and have the ability to specify password durations, disable users, and create custom roles. Personnel using *iAnalyze* may be granted permission to view students' results at an organizational level, or only a select group as defined by the administrator. Each organization is also able to create custom data fields, and import/export functionality is provided. Predefined reports are included in the system, as is the ability to render and print additional copies.

# CHAPTER 12—VALIDITY SUMMARY

The purpose of this manual is to describe several technical aspects of the CRT in an effort to contribute to the accumulation of validity evidence to support CRT score interpretations. Because it is the interpretations of test scores that are evaluated for validity, not the test itself, this manual presents documentation to support the validity of intended interpretations (AERA et al., 1999). Each of the chapters in this manual contributes important information to the validity argument by addressing one or more of the following aspects of the CRT: test development, test alignment, test administration, scoring, equating, item analyses, reliability, scaled scores, performance levels and reporting.

The CRT assessments are based on, and aligned to, Montana's Content Standards in Reading and Mathematics. Intended inferences from the CRT results are about student achievement on Montana's reading and mathematics content standards, and these achievement inferences are meant to be useful for program and instructional improvement and as a component of school accountability.

As stated in the overview chapter, the *Standards for Educational and Psychological Testing* (AERA, et al., 1999) provides a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five general areas: test content, response processes, internal structure, relationship to other variables, and consequences of testing. Although each of these sources may speak to a different *aspect* of validity, they are not distinct *types* of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

A measure of test content validity is to determine how well the assessment tasks represent the curriculum and standards for each subject and grade level. This is informed by the item development process, including how the test blueprints and test items align to the curriculum and standards. Viewed through this lens provided by the Standards, evidence based on test content was described in Chapters 2 through 5. Item alignment with Montana content standards; item bias, sensitivity and content appropriateness review processes; adherence to the test blueprint;

use of multiple item types; use of standardized administration procedures, with accommodated options for participation; and appropriate test administration training are all components of validity evidence based on test content. As discussed earlier, all CRT test questions are aligned by Montana educators to specific Montana Content Standards, and undergo several rounds of review for content fidelity and appropriateness. Items are presented to students in multiple formats (constructed-response, short-answer and multiple-choice). Finally, tests are administered according to state-mandated standardized procedures, with allowable accommodations, and all test proctors are required to attend annual training sessions.

The scoring information in Chapter 7 describes the steps taken to train and monitor hand-scorers, as well as quality control procedures related to scanning and machine scoring. To speak to student response processes, however, additional studies would be helpful and might include an investigation of students' cognitive methods using think-aloud protocols. The OPI is currently working with its TAC to develop an approach to evaluating student responses.

Evidence based on internal structure is presented in great detail in the discussions of equating and item analyses in Chapters 8 and 9. Technical characteristics of the internal structure of the assessments are presented in terms of classical item statistics (item difficulty, item-test correlation), differential item functioning analyses, a variety of reliability coefficients, standard errors of measurement, and item response theory parameters and procedures. Each test is equated to the same grade and content test from the prior year in order to preserve the meaning of scores over time. In general, item difficulty and discrimination indices were in acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that most items were assessing consistent constructs, and students who performed well on individual items tended to perform well overall.

Evidence based on the consequences of testing is addressed in the scaled scores and reporting information in Chapters 10 and 11, as well as in the test interpretation guide, which is a separate document that is referenced in the discussion of reporting. Each of these chapters speaks to the efforts undertaken to promote accurate and clear information provided to the public regarding test scores. Scaled scores offer the advantage of simplifying the reporting of results across

content areas, grade levels, and subsequent years. Performance levels provide users with reference points for mastery at each grade level, which is another useful and simple way to interpret scores. Several different standard reports are provided to stakeholders. In addition, a data analysis tool is provided to each school system to allow educators the flexibility to customize reports for local needs. Additional evidence of the consequences of testing could be supplemented with broader investigation of the impact of testing on student learning. The OPI is currently working with its TAC to develop an approach to evaluating consequential aspects of validity.

To further support the validity argument, additional studies to provide evidence regarding the relationship of CRT results to other variables include the extent to which scores from the CRT assessments converge with other measures of similar constructs, and the extent to which they diverge from measures of different constructs. Relationships among measures of the same or similar constructs can sharpen the meaning of scores and appropriate interpretations by refining the definition of the construct.

As stated in Chapter 1, the MontCAS Phase 2 assessment program CRTs are designed to measure student acquisition of the knowledge and skills in Montana's content standards for reading and mathematics. The assessments were developed to provide information at the student, class, school, and system level. The evidence presented in this manual supports inferences of student achievement on the content represented on the Montana Content Standards for Reading and Mathematics for the purposes of program and instructional improvement and as a component of school accountability.

# SECTION IV—REFERENCES

Allen, Mary J. & Yen, Wendy M. (1979). *Introduction to Measurement Theory*. Belmont, CA: Wadsworth, Inc.

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

Bock, R. D., and E. Muraki (1999). *PARSCALE: Parameter Scaling of Rating Data* [Computer program]. Chicago, IL: Scientific Software.

Brown, F. G. (1983). *Principles of Educational and Psychological Testing* 3rd ed. Fort Worth, TX: Holt, Rinehart, and Winston.

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

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

Dorans, N. J., and P. W. Holland (1993). DIF detection and description. In P. W. Holland and H. Wainer (Eds.), *Differential item functioning* pp. 35–66. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dorans, N. J., and E. Kulick (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, *23*, 355–368.

Hambleton, R. K., and W. J. van der Linden (1997). *Handbook of Modern Item Response Theory*. New York: Springer-Verlag.

Hambleton, R. K., and H. Swaminathan (1985). *Item Response Theory: Principles and Applications*. Boston: Kluwer Academic Publishers.

Joint Committee on Testing Practices (2004). *Code of Fair Testing Practices in Education*. Washington, DC: American Psychological Association. Available for download at **http://www.apa.org/science/fairtestcode.html**.

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., and M. R. Novick (1968). *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.

Petersen, N.S., Kolen, M.J., & Hoover, H.D. (1989). *Scaling, Norming, and Equating*. In R.L. Linn (Ed.), Educational measurement (3rd ed., pp. 221-262).

# Appendix A: Item Parameter Files

## Grade 3 Math

## Grade 4 Math

## Grade 5 Math

## Grade 6 Math

## Grade 7 Math

## Grade 8 Math

## Grade 10 Math

## Grade 3 Reading

## Grade 4 Reading

## Grade 5 Reading

## Grade 6 Reading

## Grade 7 Reading

## Grade 8 Reading

## Grade 10 Reading

# APPENDIX A: ITEM PARAMETER FILES

## Grade 3 Math

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 165562 | 1 | 1 | -0.7112 | 0 | | | | |
| 165586 | 1 | 1 | -0.4142 | 0 | | | | |
| 165590 | 1 | 1 | -0.6056 | 0 | | | | |
| 165599 | 1 | 1 | -1.5045 | 0 | | | | |
| 165639 | 1 | 1 | -0.9151 | 0 | | | | |
| 165692 | 1 | 1 | 0.0217 | 0 | | | | |
| 165694 | 1 | 1 | -1.2475 | 0 | | | | |
| 170479 | 1 | 1 | 0.4720 | 0 | | | | |
| 170522 | 1 | 1 | -0.7509 | 0 | | | | |
| 175701 | 1 | 1 | -0.6122 | 0 | | | | |
| 175702 | 1 | 1 | 0.1812 | 0 | | | | |
| 175704 | 1 | 1 | -0.0495 | 0 | | | | |
| 175705 | 1 | 1 | 0.2094 | 0 | | | | |
| 175708 | 1 | 1 | -0.7341 | 0 | | | | |
| 175709 | 1 | 1 | -1.0240 | 0 | | | | |
| 175712 | 1 | 1 | -0.3691 | 0 | | | | |
| 175713 | 1 | 1 | -1.0450 | 0 | | | | |
| 175714 | 1 | 1 | -0.6732 | 0 | | | | |
| 175715 | 1 | 1 | 0.2261 | 0 | | | | |
| 175717 | 1 | 1 | -0.3855 | 0 | | | | |
| 175719 | 1 | 1 | -0.5682 | 0 | | | | |
| 175723 | 1 | 1 | -1.1018 | 0 | | | | |
| 175724 | 1 | 1 | -0.3648 | 0 | | | | |
| 175725 | 1 | 1 | -0.9423 | 0 | | | | |
| 175726 | 1 | 1 | -0.6948 | 0 | | | | |
| 175727 | 1 | 1 | -0.7278 | 0 | | | | |
| 175729 | 1 | 1 | -0.0294 | 0 | | | | |
| 175732 | 1 | 1 | -0.9987 | 0 | | | | |
| 175733 | 1 | 1 | -0.3889 | 0 | | | | |
| 175734 | 1 | 1 | -0.9551 | 0 | | | | |
| 175739 | 1 | 1 | 0.0998 | 0 | | | | |
| 175772 | 1 | 1 | -1.3259 | 0 | | | | |
| 175773 | 1 | 1 | 0.2498 | 0 | | | | |
| 175774 | 1 | 1 | -0.8064 | 0 | | | | |
| 175775 | 1 | 1 | -0.5581 | 0 | | | | |
| 175776 | 1 | 1 | -0.4965 | 0 | | | | |
| 175777 | 1 | 1 | -1.0389 | 0 | | | | |
| 175782 | 1 | 1 | -1.1538 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 175784 | 1 | 1 | -0.6583 | 0 | | | | |
| 175785 | 1 | 1 | -0.4546 | 0 | | | | |
| 175787 | 1 | 1 | -0.7832 | 0 | | | | |
| 175789 | 1 | 1 | -1.2332 | 0 | | | | |
| 175790 | 1 | 1 | -0.9529 | 0 | | | | |
| 175791 | 1 | 1 | -1.1194 | 0 | | | | |
| 175792 | 1 | 1 | -0.9212 | 0 | | | | |
| 175793 | 1 | 1 | 0.3586 | 0 | | | | |
| 175796 | 1 | 1 | -0.7385 | 0 | | | | |
| 175799 | 1 | 1 | -0.7363 | 0 | | | | |
| 175800 | 1 | 1 | -0.6747 | 0 | | | | |
| 175801 | 1 | 1 | -1.2869 | 0 | | | | |
| 175803 | 1 | 1 | -1.2349 | 0 | | | | |
| 175804 | 1 | 1 | -0.4372 | 0 | | | | |
| 244045 | 1 | 1 | -0.7763 | 0 | | | | |
| 247767 | 1 | 1 | -0.5206 | 0 | | | | |
| 247768 | 1 | 1 | -0.1237 | 0 | | | | |
| 247770 | 1 | 1 | -1.5691 | 0 | | | | |
| 247774 | 1 | 1 | -0.2152 | 0 | | | | |
| 247775 | 1 | 1 | 0.0676 | 0 | | | | |
| 175900 | 4 | 1 | -0.2177 | 0 | 0.9546 | 0.7887 | -1.118 | -0.6315 |
| 175901 | 4 | 1 | -0.5928 | 0 | 0.6667 | 0.6994 | -0.5197 | -0.8465 |

# Grade 4 Math

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 165016 | 1 | 1 | 0.1574 | 0 | | | | |
| 165022 | 1 | 1 | -0.3110 | 0 | | | | |
| 165030 | 1 | 1 | 0.1257 | 0 | | | | |
| 166204 | 1 | 1 | -1.1722 | 0 | | | | |
| 166221 | 1 | 1 | -0.9857 | 0 | | | | |
| 166233 | 1 | 1 | 0.2181 | 0 | | | | |
| 166261 | 1 | 1 | -0.4629 | 0 | | | | |
| 166358 | 1 | 1 | 0.2782 | 0 | | | | |
| 166390 | 1 | 1 | -0.1191 | 0 | | | | |
| 166391 | 1 | 1 | -0.5767 | 0 | | | | |
| 211221 | 1 | 1 | 0.0294 | 0 | | | | |
| 211227 | 1 | 1 | -0.3875 | 0 | | | | |
| 211236 | 1 | 1 | -0.1173 | 0 | | | | |
| 211240 | 1 | 1 | -0.3098 | 0 | | | | |
| 211264 | 1 | 1 | -1.0912 | 0 | | | | |
| 211271 | 1 | 1 | 0.0380 | 0 | | | | |
| 211281 | 1 | 1 | -1.0511 | 0 | | | | |
| 211288 | 1 | 1 | -0.3422 | 0 | | | | |
| 211292 | 1 | 1 | 0.0764 | 0 | | | | |
| 211301 | 1 | 1 | -0.7167 | 0 | | | | |
| 211306 | 1 | 1 | -0.1428 | 0 | | | | |
| 211411 | 1 | 1 | -0.0145 | 0 | | | | |
| 211414 | 1 | 1 | -0.5931 | 0 | | | | |
| 211416 | 1 | 1 | -1.4810 | 0 | | | | |
| 211417 | 1 | 1 | -0.1351 | 0 | | | | |
| 211419 | 1 | 1 | -0.0763 | 0 | | | | |
| 211420 | 1 | 1 | 0.5797 | 0 | | | | |
| 211422 | 1 | 1 | -0.0961 | 0 | | | | |
| 211424 | 1 | 1 | -1.3112 | 0 | | | | |
| 211426 | 1 | 1 | 0.1229 | 0 | | | | |
| 211429 | 1 | 1 | -0.5216 | 0 | | | | |
| 211430 | 1 | 1 | 0.5221 | 0 | | | | |
| 211431 | 1 | 1 | -0.5944 | 0 | | | | |
| 211433 | 1 | 1 | -0.9153 | 0 | | | | |
| 211435 | 1 | 1 | -0.3201 | 0 | | | | |
| 211437 | 1 | 1 | 0.2408 | 0 | | | | |
| 211441 | 1 | 1 | -0.2604 | 0 | | | | |
| 211444 | 1 | 1 | -0.2594 | 0 | | | | |
| 211445 | 1 | 1 | -0.0811 | 0 | | | | |
| 211448 | 1 | 1 | -1.0951 | 0 | | | | |
| 211450 | 1 | 1 | 0.6955 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 211452 | 1 | 1 | -0.5598 | 0 | | | | |
| 211453 | 1 | 1 | 0.0729 | 0 | | | | |
| 211454 | 1 | 1 | -0.7291 | 0 | | | | |
| 211456 | 1 | 1 | -0.0943 | 0 | | | | |
| 211458 | 1 | 1 | 0.3278 | 0 | | | | |
| 211460 | 1 | 1 | -0.9008 | 0 | | | | |
| 211465 | 1 | 1 | 0.4891 | 0 | | | | |
| 211466 | 1 | 1 | -0.3373 | 0 | | | | |
| 211468 | 1 | 1 | -0.1796 | 0 | | | | |
| 211469 | 1 | 1 | -0.5177 | 0 | | | | |
| 211470 | 1 | 1 | -0.8825 | 0 | | | | |
| 211473 | 1 | 1 | -0.1417 | 0 | | | | |
| 211487 | 1 | 1 | -0.0321 | 0 | | | | |
| 211496 | 1 | 1 | -0.1632 | 0 | | | | |
| 211499 | 1 | 1 | -0.8328 | 0 | | | | |
| 211502 | 1 | 1 | 0.1257 | 0 | | | | |
| 213640 | 1 | 1 | 0.1574 | 0 | | | | |
| 211346 | 4 | 1 | -0.0571 | 0 | 0.2141 | 0.3616 | -0.8273 | 0.2516 |
| 246933 | 4 | 1 | -0.1023 | 0 | 0.7640 | 0.1102 | -0.2014 | -0.6729 |

## Grade 5 Math

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 165043 | 1 | 1 | -0.6218 | 0 | | | | |
| 165053 | 1 | 1 | -0.5549 | 0 | | | | |
| 165056 | 1 | 1 | -0.9715 | 0 | | | | |
| 165065 | 1 | 1 | -0.2278 | 0 | | | | |
| 165073 | 1 | 1 | -0.6967 | 0 | | | | |
| 166298 | 1 | 1 | -0.5207 | 0 | | | | |
| 166302 | 1 | 1 | -0.8111 | 0 | | | | |
| 166438 | 1 | 1 | -0.6321 | 0 | | | | |
| 166499 | 1 | 1 | -0.1403 | 0 | | | | |
| 175143 | 1 | 1 | -0.7172 | 0 | | | | |
| 175151 | 1 | 1 | -0.0336 | 0 | | | | |
| 175155 | 1 | 1 | 0.0988 | 0 | | | | |
| 175158 | 1 | 1 | 0.0169 | 0 | | | | |
| 175159 | 1 | 1 | 0.1037 | 0 | | | | |
| 175164 | 1 | 1 | 0.2818 | 0 | | | | |
| 175171 | 1 | 1 | -0.1338 | 0 | | | | |
| 175174 | 1 | 1 | -1.2692 | 0 | | | | |
| 175176 | 1 | 1 | 0.0781 | 0 | | | | |
| 175177 | 1 | 1 | -0.4262 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 175185 | 1 | 1 | 0.1237 | 0 | | | | |
| 175186 | 1 | 1 | -0.1376 | 0 | | | | |
| 175499 | 1 | 1 | -0.1943 | 0 | | | | |
| 175500 | 1 | 1 | -0.3451 | 0 | | | | |
| 175501 | 1 | 1 | 0.1404 | 0 | | | | |
| 175502 | 1 | 1 | -0.368 | 0 | | | | |
| 175505 | 1 | 1 | -0.2752 | 0 | | | | |
| 175507 | 1 | 1 | -1.4693 | 0 | | | | |
| 175508 | 1 | 1 | -0.3297 | 0 | | | | |
| 175509 | 1 | 1 | -0.3663 | 0 | | | | |
| 175510 | 1 | 1 | -0.5198 | 0 | | | | |
| 175511 | 1 | 1 | -0.533 | 0 | | | | |
| 175515 | 1 | 1 | 0.435 | 0 | | | | |
| 175517 | 1 | 1 | -0.9634 | 0 | | | | |
| 175518 | 1 | 1 | -0.1603 | 0 | | | | |
| 175520 | 1 | 1 | 0.1609 | 0 | | | | |
| 175524 | 1 | 1 | -0.3146 | 0 | | | | |
| 175525 | 1 | 1 | 0.342 | 0 | | | | |
| 175528 | 1 | 1 | -0.2949 | 0 | | | | |
| 175530 | 1 | 1 | 0.1341 | 0 | | | | |
| 175531 | 1 | 1 | 0.1965 | 0 | | | | |
| 175532 | 1 | 1 | -0.2101 | 0 | | | | |
| 175537 | 1 | 1 | -0.2087 | 0 | | | | |
| 175683 | 1 | 1 | -0.2012 | 0 | | | | |
| 175684 | 1 | 1 | -0.4832 | 0 | | | | |
| 175685 | 1 | 1 | 0.2593 | 0 | | | | |
| 175687 | 1 | 1 | -1.0401 | 0 | | | | |
| 175689 | 1 | 1 | -0.7409 | 0 | | | | |
| 175691 | 1 | 1 | 0.0034 | 0 | | | | |
| 175692 | 1 | 1 | 0.0153 | 0 | | | | |
| 177173 | 1 | 1 | -0.3611 | 0 | | | | |
| 178371 | 1 | 1 | -0.3382 | 0 | | | | |
| 228808 | 1 | 1 | 0.3673 | 0 | | | | |
| 228809 | 1 | 1 | -0.2051 | 0 | | | | |
| 228811 | 1 | 1 | 0.5266 | 0 | | | | |
| 241846 | 1 | 1 | 0.2076 | 0 | | | | |
| 241848 | 1 | 1 | -0.1201 | 0 | | | | |
| 246604 | 1 | 1 | -0.5664 | 0 | | | | |
| 250226 | 1 | 1 | -0.6559 | 0 | | | | |
| 189314 | 4 | 1 | -0.4132 | 0 | 0.5295 | -0.1527 | 0.3179 | -0.6947 |
| 206606 | 4 | 1 | 0.5705 | 0 | 0.5635 | 0.3911 | -0.8389 | -0.1157 |

# Grade 6 Math

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 165033 | 1 | 1 | -1.5861 | 0 | | | | |
| 165035 | 1 | 1 | 0.2005 | 0 | | | | |
| 166526 | 1 | 1 | -0.55 | 0 | | | | |
| 166528 | 1 | 1 | -0.9562 | 0 | | | | |
| 166549 | 1 | 1 | -0.2854 | 0 | | | | |
| 166553 | 1 | 1 | -0.7483 | 0 | | | | |
| 166581 | 1 | 1 | -1.0276 | 0 | | | | |
| 166592 | 1 | 1 | -0.0623 | 0 | | | | |
| 166623 | 1 | 1 | -0.0745 | 0 | | | | |
| 166625 | 1 | 1 | -0.2317 | 0 | | | | |
| 177034 | 1 | 1 | -1.0387 | 0 | | | | |
| 177036 | 1 | 1 | 0.0251 | 0 | | | | |
| 177039 | 1 | 1 | -0.2572 | 0 | | | | |
| 177042 | 1 | 1 | 0.2212 | 0 | | | | |
| 177049 | 1 | 1 | -0.2397 | 0 | | | | |
| 177051 | 1 | 1 | -0.3271 | 0 | | | | |
| 177060 | 1 | 1 | 0.1432 | 0 | | | | |
| 177069 | 1 | 1 | 0.4245 | 0 | | | | |
| 177073 | 1 | 1 | -0.2058 | 0 | | | | |
| 177075 | 1 | 1 | -0.4481 | 0 | | | | |
| 177089 | 1 | 1 | 0.0429 | 0 | | | | |
| 177458 | 1 | 1 | 0.235 | 0 | | | | |
| 177459 | 1 | 1 | -0.0452 | 0 | | | | |
| 177461 | 1 | 1 | 0.42 | 0 | | | | |
| 177462 | 1 | 1 | -0.1171 | 0 | | | | |
| 177464 | 1 | 1 | -0.5855 | 0 | | | | |
| 177466 | 1 | 1 | -0.1077 | 0 | | | | |
| 177467 | 1 | 1 | 0.9442 | 0 | | | | |
| 177468 | 1 | 1 | 0.1889 | 0 | | | | |
| 177469 | 1 | 1 | 0.7565 | 0 | | | | |
| 177470 | 1 | 1 | 0.5901 | 0 | | | | |
| 177473 | 1 | 1 | -0.2196 | 0 | | | | |
| 177476 | 1 | 1 | -0.5678 | 0 | | | | |
| 177478 | 1 | 1 | 0.0321 | 0 | | | | |
| 177479 | 1 | 1 | -0.4954 | 0 | | | | |
| 177480 | 1 | 1 | 0.1555 | 0 | | | | |
| 177481 | 1 | 1 | -0.2423 | 0 | | | | |
| 177483 | 1 | 1 | -0.15 | 0 | | | | |
| 177484 | 1 | 1 | 0.0731 | 0 | | | | |
| 177486 | 1 | 1 | -0.7812 | 0 | | | | |
| 177489 | 1 | 1 | 0.0991 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 177490 | 1 | 1 | 1.1477 | 0 | | | | |
| 177492 | 1 | 1 | 0.2398 | 0 | | | | |
| 177493 | 1 | 1 | -0.2771 | 0 | | | | |
| 177495 | 1 | 1 | 0.389 | 0 | | | | |
| 177496 | 1 | 1 | 0.5308 | 0 | | | | |
| 177498 | 1 | 1 | -0.0799 | 0 | | | | |
| 177500 | 1 | 1 | -0.1343 | 0 | | | | |
| 177513 | 1 | 1 | 0.636 | 0 | | | | |
| 177533 | 1 | 1 | -0.5825 | 0 | | | | |
| 177536 | 1 | 1 | -0.5035 | 0 | | | | |
| 177537 | 1 | 1 | 0.0866 | 0 | | | | |
| 177539 | 1 | 1 | 0.2802 | 0 | | | | |
| 239792 | 1 | 1 | -0.3616 | 0 | | | | |
| 241853 | 1 | 1 | -0.6951 | 0 | | | | |
| 254145 | 1 | 1 | 1.0171 | 0 | | | | |
| 254147 | 1 | 1 | 0.6583 | 0 | | | | |
| 254148 | 1 | 1 | 1.039 | 0 | | | | |
| 254142 | 4 | 1 | 0.4146 | 0 | 0.2686 | -0.0483 | -0.1709 | -0.0493 |
| 254150 | 4 | 1 | 0.3156 | 0 | 0.7831 | 0.5604 | -0.3173 | -1.0262 |

## Grade 7 Math

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 164902 | 1 | 1 | -0.5621 | 0 | | | | |
| 164920 | 1 | 1 | 0.5982 | 0 | | | | |
| 164930 | 1 | 1 | 0.2089 | 0 | | | | |
| 164961 | 1 | 1 | -0.4322 | 0 | | | | |
| 165006 | 1 | 1 | 0.3354 | 0 | | | | |
| 165008 | 1 | 1 | -0.5592 | 0 | | | | |
| 165222 | 1 | 1 | -0.4573 | 0 | | | | |
| 165228 | 1 | 1 | -0.0712 | 0 | | | | |
| 178141 | 1 | 1 | -0.2047 | 0 | | | | |
| 178143 | 1 | 1 | -0.6280 | 0 | | | | |
| 178144 | 1 | 1 | 0.2481 | 0 | | | | |
| 178146 | 1 | 1 | 0.0001 | 0 | | | | |
| 178147 | 1 | 1 | 0.1149 | 0 | | | | |
| 178149 | 1 | 1 | -0.2162 | 0 | | | | |
| 178152 | 1 | 1 | -0.1562 | 0 | | | | |
| 178154 | 1 | 1 | -0.5705 | 0 | | | | |
| 178156 | 1 | 1 | -0.7268 | 0 | | | | |
| 178158 | 1 | 1 | -0.2134 | 0 | | | | |
| 178159 | 1 | 1 | -0.2995 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 178168 | 1 | 1 | -0.3684 | 0 | | | | |
| 178169 | 1 | 1 | 0.1987 | 0 | | | | |
| 178170 | 1 | 1 | -0.5009 | 0 | | | | |
| 178172 | 1 | 1 | -0.2218 | 0 | | | | |
| 178173 | 1 | 1 | -0.0544 | 0 | | | | |
| 178175 | 1 | 1 | 0.0977 | 0 | | | | |
| 178201 | 1 | 1 | -1.4627 | 0 | | | | |
| 178203 | 1 | 1 | -0.4675 | 0 | | | | |
| 178204 | 1 | 1 | -0.0888 | 0 | | | | |
| 178205 | 1 | 1 | -0.6808 | 0 | | | | |
| 178208 | 1 | 1 | -0.3625 | 0 | | | | |
| 178210 | 1 | 1 | 0.5437 | 0 | | | | |
| 178212 | 1 | 1 | 0.5209 | 0 | | | | |
| 178213 | 1 | 1 | -0.1597 | 0 | | | | |
| 178214 | 1 | 1 | -0.2390 | 0 | | | | |
| 178215 | 1 | 1 | 0.0496 | 0 | | | | |
| 178216 | 1 | 1 | -1.9679 | 0 | | | | |
| 178218 | 1 | 1 | -0.5077 | 0 | | | | |
| 178220 | 1 | 1 | 1.6822 | 0 | | | | |
| 178222 | 1 | 1 | 0.0758 | 0 | | | | |
| 178223 | 1 | 1 | -0.4817 | 0 | | | | |
| 178224 | 1 | 1 | 0.4018 | 0 | | | | |
| 178225 | 1 | 1 | -0.0586 | 0 | | | | |
| 178226 | 1 | 1 | 0.0884 | 0 | | | | |
| 178227 | 1 | 1 | 0.5130 | 0 | | | | |
| 178229 | 1 | 1 | -0.0956 | 0 | | | | |
| 178246 | 1 | 1 | -0.2628 | 0 | | | | |
| 178250 | 1 | 1 | -0.2639 | 0 | | | | |
| 178263 | 1 | 1 | 0.6276 | 0 | | | | |
| 239786 | 1 | 1 | -0.3724 | 0 | | | | |
| 241855 | 1 | 1 | 0.2038 | 0 | | | | |
| 241858 | 1 | 1 | 0.8696 | 0 | | | | |
| 246731 | 1 | 1 | -0.1971 | 0 | | | | |
| 249728 | 1 | 1 | 0.3053 | 0 | | | | |
| 254451 | 1 | 1 | 1.4275 | 0 | | | | |
| 254453 | 1 | 1 | -0.4864 | 0 | | | | |
| 254454 | 1 | 1 | 0.2048 | 0 | | | | |
| 254455 | 1 | 1 | 0.3925 | 0 | | | | |
| 254593 | 1 | 1 | 0.7743 | 0 | | | | |
| 254448 | 4 | 1 | 0.0649 | 0 | 0.6546 | 0.6727 | 0.0423 | -1.3696 |
| 241859 | 4 | 1 | 0.3171 | 0 | 0.2455 | -0.0528 | -0.2540 | 0.0612 |

# Grade 8 Math

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 165386 | 1 | 1 | -0.0008 | 0 | | | | |
| 165732 | 1 | 1 | -0.2496 | 0 | | | | |
| 165736 | 1 | 1 | 0.1813 | 0 | | | | |
| 165837 | 1 | 1 | -0.4260 | 0 | | | | |
| 165865 | 1 | 1 | -0.4075 | 0 | | | | |
| 165890 | 1 | 1 | 0.1422 | 0 | | | | |
| 166326 | 1 | 1 | -0.0236 | 0 | | | | |
| 212284 | 1 | 1 | 0.3909 | 0 | | | | |
| 212304 | 1 | 1 | -0.1934 | 0 | | | | |
| 212334 | 1 | 1 | 0.0007 | 0 | | | | |
| 212345 | 1 | 1 | -0.6038 | 0 | | | | |
| 212348 | 1 | 1 | -0.0134 | 0 | | | | |
| 212353 | 1 | 1 | -1.2716 | 0 | | | | |
| 212365 | 1 | 1 | -0.0276 | 0 | | | | |
| 212424 | 1 | 1 | 0.2611 | 0 | | | | |
| 212426 | 1 | 1 | -0.0558 | 0 | | | | |
| 212427 | 1 | 1 | -0.6599 | 0 | | | | |
| 212428 | 1 | 1 | -0.7561 | 0 | | | | |
| 212429 | 1 | 1 | -0.1477 | 0 | | | | |
| 212430 | 1 | 1 | 0.1877 | 0 | | | | |
| 212432 | 1 | 1 | 0.5009 | 0 | | | | |
| 212433 | 1 | 1 | -1.0082 | 0 | | | | |
| 212434 | 1 | 1 | -0.7423 | 0 | | | | |
| 212435 | 1 | 1 | -0.0643 | 0 | | | | |
| 212436 | 1 | 1 | 0.1551 | 0 | | | | |
| 212438 | 1 | 1 | 0.0581 | 0 | | | | |
| 212439 | 1 | 1 | 0.2906 | 0 | | | | |
| 212440 | 1 | 1 | 0.4452 | 0 | | | | |
| 212443 | 1 | 1 | 0.4129 | 0 | | | | |
| 212444 | 1 | 1 | 0.5086 | 0 | | | | |
| 212445 | 1 | 1 | -0.3687 | 0 | | | | |
| 212446 | 1 | 1 | -0.2355 | 0 | | | | |
| 212447 | 1 | 1 | 0.5475 | 0 | | | | |
| 212448 | 1 | 1 | -0.0296 | 0 | | | | |
| 212450 | 1 | 1 | -0.7234 | 0 | | | | |
| 212459 | 1 | 1 | -0.6131 | 0 | | | | |
| 212462 | 1 | 1 | 0.2378 | 0 | | | | |
| 213286 | 1 | 1 | -0.5575 | 0 | | | | |
| 213288 | 1 | 1 | -0.2009 | 0 | | | | |
| 213290 | 1 | 1 | -0.8041 | 0 | | | | |
| 213291 | 1 | 1 | -0.9098 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 213292 | 1 | 1 | 0.4340 | 0 | | | | |
| 213293 | 1 | 1 | -0.0546 | 0 | | | | |
| 213297 | 1 | 1 | -0.0176 | 0 | | | | |
| 213299 | 1 | 1 | -0.2049 | 0 | | | | |
| 213301 | 1 | 1 | 0.1321 | 0 | | | | |
| 213302 | 1 | 1 | -0.6712 | 0 | | | | |
| 213303 | 1 | 1 | -0.4785 | 0 | | | | |
| 213308 | 1 | 1 | 0.1783 | 0 | | | | |
| 213309 | 1 | 1 | -0.0249 | 0 | | | | |
| 213313 | 1 | 1 | -0.1358 | 0 | | | | |
| 213320 | 1 | 1 | -0.0900 | 0 | | | | |
| 213406 | 1 | 1 | 0.3671 | 0 | | | | |
| 241849 | 1 | 1 | 0.2190 | 0 | | | | |
| 241850 | 1 | 1 | -0.0719 | 0 | | | | |
| 241852 | 1 | 1 | 0.0212 | 0 | | | | |
| 246579 | 1 | 1 | 0.4472 | 0 | | | | |
| 249032 | 1 | 1 | -0.3556 | 0 | | | | |
| 212471 | 4 | 1 | -0.1590 | 0 | -0.0841 | 0.3490 | -0.4437 | 0.1787 |
| 249031 | 4 | 1 | 0.9584 | 0 | 1.1253 | 0.2054 | -0.3605 | -0.9702 |

# Grade 10 Math

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 166131 | 1 | 1 | -0.7716 | 0 | | | | |
| 166134 | 1 | 1 | -0.2145 | 0 | | | | |
| 166142 | 1 | 1 | 0.0442 | 0 | | | | |
| 166751 | 1 | 1 | 0.0165 | 0 | | | | |
| 166775 | 1 | 1 | 0.0299 | 0 | | | | |
| 166907 | 1 | 1 | 0.3915 | 0 | | | | |
| 166919 | 1 | 1 | -0.5054 | 0 | | | | |
| 166957 | 1 | 1 | 0.0017 | 0 | | | | |
| 189370 | 1 | 1 | 0.7078 | 0 | | | | |
| 212495 | 1 | 1 | -0.4596 | 0 | | | | |
| 212497 | 1 | 1 | 0.0290 | 0 | | | | |
| 212499 | 1 | 1 | -0.2048 | 0 | | | | |
| 212500 | 1 | 1 | 0.0130 | 0 | | | | |
| 212529 | 1 | 1 | -0.8275 | 0 | | | | |
| 212535 | 1 | 1 | -0.3813 | 0 | | | | |
| 212536 | 1 | 1 | -0.5829 | 0 | | | | |
| 212537 | 1 | 1 | -0.2777 | 0 | | | | |
| 212539 | 1 | 1 | -0.3037 | 0 | | | | |
| 212543 | 1 | 1 | -0.1778 | 0 | | | | |
| 212544 | 1 | 1 | 0.3711 | 0 | | | | |
| 212547 | 1 | 1 | -0.5254 | 0 | | | | |
| 212558 | 1 | 1 | -0.5126 | 0 | | | | |
| 212561 | 1 | 1 | 0.1472 | 0 | | | | |
| 212564 | 1 | 1 | 0.2416 | 0 | | | | |
| 212565 | 1 | 1 | -0.5126 | 0 | | | | |
| 212567 | 1 | 1 | 0.1881 | 0 | | | | |
| 212569 | 1 | 1 | 0.1136 | 0 | | | | |
| 212570 | 1 | 1 | 0.2426 | 0 | | | | |
| 212573 | 1 | 1 | -0.3529 | 0 | | | | |
| 212577 | 1 | 1 | 0.0397 | 0 | | | | |
| 212579 | 1 | 1 | 0.0282 | 0 | | | | |
| 212587 | 1 | 1 | -0.5237 | 0 | | | | |
| 212589 | 1 | 1 | -0.0824 | 0 | | | | |
| 212591 | 1 | 1 | 0.1542 | 0 | | | | |
| 212595 | 1 | 1 | -0.1778 | 0 | | | | |
| 212604 | 1 | 1 | -0.5517 | 0 | | | | |
| 212607 | 1 | 1 | -0.8802 | 0 | | | | |
| 212609 | 1 | 1 | 0.2805 | 0 | | | | |
| 212612 | 1 | 1 | -0.1254 | 0 | | | | |
| 212614 | 1 | 1 | -0.6115 | 0 | | | | |
| 213245 | 1 | 1 | -0.2186 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 213252 | 1 | 1 | -0.1718 | 0 | | | | |
| 213253 | 1 | 1 | -0.3880 | 0 | | | | |
| 213256 | 1 | 1 | -0.0373 | 0 | | | | |
| 213264 | 1 | 1 | 0.5916 | 0 | | | | |
| 213282 | 1 | 1 | 0.1328 | 0 | | | | |
| 239273 | 1 | 1 | -1.0087 | 0 | | | | |
| 239288 | 1 | 1 | -0.3465 | 0 | | | | |
| 239294 | 1 | 1 | 0.0596 | 0 | | | | |
| 239304 | 1 | 1 | -0.3705 | 0 | | | | |
| 239308 | 1 | 1 | -0.0192 | 0 | | | | |
| 239309 | 1 | 1 | -0.0456 | 0 | | | | |
| 239310 | 1 | 1 | 0.1072 | 0 | | | | |
| 239316 | 1 | 1 | -0.8322 | 0 | | | | |
| 242243 | 1 | 1 | 0.0415 | 0 | | | | |
| 242245 | 1 | 1 | -0.9274 | 0 | | | | |
| 242247 | 1 | 1 | -0.8154 | 0 | | | | |
| 249340 | 1 | 1 | 0.6224 | 0 | | | | |
| 249343 | 1 | 1 | -0.1896 | 0 | | | | |
| 249344 | 1 | 1 | -0.6569 | 0 | | | | |
| 249345 | 1 | 1 | -0.3922 | 0 | | | | |
| 249346 | 1 | 1 | 0.3488 | 0 | | | | |
| 249347 | 1 | 1 | 0.0820 | 0 | | | | |
| 212619 | 4 | 1 | 0.5085 | 0 | 1.6243 | -0.2093 | 0.6071 | -2.0221 |
| 212637 | 4 | 1 | 0.5949 | 0 | 0.9006 | -0.3802 | 0.2885 | -0.8089 |

# Grade 3 Reading

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 170668 | 1 | 1 | -0.4480 | 0 | | | | |
| 170707 | 1 | 1 | -0.6433 | 0 | | | | |
| 170715 | 1 | 1 | 0.2715 | 0 | | | | |
| 170720 | 1 | 1 | -0.3822 | 0 | | | | |
| 170721 | 1 | 1 | -0.7043 | 0 | | | | |
| 170725 | 1 | 1 | -0.2053 | 0 | | | | |
| 176023 | 1 | 1 | -0.4851 | 0 | | | | |
| 176047 | 1 | 1 | -1.2517 | 0 | | | | |
| 176114 | 1 | 1 | -0.6208 | 0 | | | | |
| 181346 | 1 | 1 | -0.4395 | 0 | | | | |
| 181357 | 1 | 1 | -0.7206 | 0 | | | | |
| 181361 | 1 | 1 | -0.4224 | 0 | | | | |
| 181363 | 1 | 1 | -0.1009 | 0 | | | | |
| 181366 | 1 | 1 | -0.4515 | 0 | | | | |
| 182280 | 1 | 1 | -1.4029 | 0 | | | | |
| 183799 | 1 | 1 | -1.0567 | 0 | | | | |
| 183852 | 1 | 1 | -0.2970 | 0 | | | | |
| 183860 | 1 | 1 | -1.1775 | 0 | | | | |
| 183908 | 1 | 1 | -0.5127 | 0 | | | | |
| 183919 | 1 | 1 | -1.5435 | 0 | | | | |
| 183921 | 1 | 1 | -0.4601 | 0 | | | | |
| 183924 | 1 | 1 | -0.8741 | 0 | | | | |
| 183928 | 1 | 1 | -1.0530 | 0 | | | | |
| 183933 | 1 | 1 | -0.6319 | 0 | | | | |
| 183934 | 1 | 1 | -1.0595 | 0 | | | | |
| 183935 | 1 | 1 | -1.1842 | 0 | | | | |
| 183942 | 1 | 1 | 0.1588 | 0 | | | | |
| 183957 | 1 | 1 | -0.9010 | 0 | | | | |
| 183961 | 1 | 1 | -1.2325 | 0 | | | | |
| 183964 | 1 | 1 | -0.6315 | 0 | | | | |
| 183965 | 1 | 1 | -0.4841 | 0 | | | | |
| 183966 | 1 | 1 | -0.0665 | 0 | | | | |
| 183967 | 1 | 1 | -0.7022 | 0 | | | | |
| 183968 | 1 | 1 | -0.5659 | 0 | | | | |
| 183970 | 1 | 1 | -0.3392 | 0 | | | | |
| 183974 | 1 | 1 | -0.3499 | 0 | | | | |
| 183975 | 1 | 1 | -0.5545 | 0 | | | | |
| 183977 | 1 | 1 | -0.9255 | 0 | | | | |
| 230160 | 1 | 1 | -0.8761 | 0 | | | | |
| 235933 | 1 | 1 | 0.1672 | 0 | | | | |
| 235935 | 1 | 1 | -1.3794 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 238560 | 1 | 1 | -1.0174 | 0 | | | | |
| 242308 | 1 | 1 | -0.1224 | 0 | | | | |
| 247735 | 1 | 1 | 0.4327 | 0 | | | | |
| 247739 | 1 | 1 | -0.7265 | 0 | | | | |
| 247742 | 1 | 1 | -1.3895 | 0 | | | | |
| 247748 | 1 | 1 | -0.4761 | 0 | | | | |
| 247753 | 1 | 1 | 0.2496 | 0 | | | | |
| 247755 | 1 | 1 | -0.7094 | 0 | | | | |
| 247756 | 1 | 1 | -1.3952 | 0 | | | | |
| 247758 | 1 | 1 | -0.5911 | 0 | | | | |
| 254456 | 1 | 1 | -0.6531 | 0 | | | | |
| 175999 | 4 | 1 | 0.4112 | 0 | 1.7544 | 0.3227 | -0.8442 | -1.2328 |
| 181370 | 4 | 1 | 0.1456 | 0 | 1.7136 | 0.2028 | -0.7474 | -1.1690 |

# Grade 4 Reading

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 170990 | 1 | 1 | -0.4052 | 0 | | | | |
| 170993 | 1 | 1 | 0.3126 | 0 | | | | |
| 170994 | 1 | 1 | -0.2153 | 0 | | | | |
| 170995 | 1 | 1 | 0.4899 | 0 | | | | |
| 171041 | 1 | 1 | 0.2755 | 0 | | | | |
| 171045 | 1 | 1 | -1.5015 | 0 | | | | |
| 210958 | 1 | 1 | 0.2404 | 0 | | | | |
| 210962 | 1 | 1 | -0.7969 | 0 | | | | |
| 210966 | 1 | 1 | -0.2516 | 0 | | | | |
| 210973 | 1 | 1 | -0.4940 | 0 | | | | |
| 210976 | 1 | 1 | -0.6019 | 0 | | | | |
| 210986 | 1 | 1 | 0.5609 | 0 | | | | |
| 210991 | 1 | 1 | -0.8104 | 0 | | | | |
| 210992 | 1 | 1 | 0.0705 | 0 | | | | |
| 210999 | 1 | 1 | -0.3430 | 0 | | | | |
| 211000 | 1 | 1 | -0.9773 | 0 | | | | |
| 211006 | 1 | 1 | -0.2810 | 0 | | | | |
| 211008 | 1 | 1 | -1.2804 | 0 | | | | |
| 211011 | 1 | 1 | -0.6625 | 0 | | | | |
| 211021 | 1 | 1 | -0.8092 | 0 | | | | |
| 211022 | 1 | 1 | -0.5011 | 0 | | | | |
| 211027 | 1 | 1 | 0.4089 | 0 | | | | |
| 211032 | 1 | 1 | -0.6379 | 0 | | | | |
| 211037 | 1 | 1 | -0.4355 | 0 | | | | |
| 211038 | 1 | 1 | -0.1258 | 0 | | | | |
| 211040 | 1 | 1 | -0.8351 | 0 | | | | |
| 211042 | 1 | 1 | -0.3826 | 0 | | | | |
| 211049 | 1 | 1 | -1.0177 | 0 | | | | |
| 211053 | 1 | 1 | -0.7006 | 0 | | | | |
| 211055 | 1 | 1 | -0.6790 | 0 | | | | |
| 211061 | 1 | 1 | -0.0496 | 0 | | | | |
| 211101 | 1 | 1 | -0.3728 | 0 | | | | |
| 211106 | 1 | 1 | 0.1507 | 0 | | | | |
| 211114 | 1 | 1 | -0.3469 | 0 | | | | |
| 211116 | 1 | 1 | -0.8811 | 0 | | | | |
| 211118 | 1 | 1 | 0.1108 | 0 | | | | |
| 211222 | 1 | 1 | -0.3402 | 0 | | | | |
| 211228 | 1 | 1 | -0.3952 | 0 | | | | |
| 211229 | 1 | 1 | 0.1012 | 0 | | | | |
| 211231 | 1 | 1 | -0.7978 | 0 | | | | |
| 211235 | 1 | 1 | -0.1830 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 235965 | 1 | 1 | -0.1832 | 0 | | | | |
| 235976 | 1 | 1 | -1.1089 | 0 | | | | |
| 235979 | 1 | 1 | -0.4333 | 0 | | | | |
| 242341 | 1 | 1 | -1.2677 | 0 | | | | |
| 242346 | 1 | 1 | -0.0029 | 0 | | | | |
| 242350 | 1 | 1 | 0.2251 | 0 | | | | |
| 242353 | 1 | 1 | -1.1208 | 0 | | | | |
| 246927 | 1 | 1 | -1.1133 | 0 | | | | |
| 246929 | 1 | 1 | 0.2750 | 0 | | | | |
| 246930 | 1 | 1 | -0.7315 | 0 | | | | |
| 246932 | 1 | 1 | -0.3442 | 0 | | | | |
| 210994 | 4 | 1 | 0.6069 | 0 | 1.3357 | 0.4907 | -0.7858 | -1.0406 |
| 211045 | 4 | 1 | 0.2285 | 0 | 1.1115 | 0.3030 | -0.5624 | -0.8521 |

# Grade 5 Reading

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 171065 | 1 | 1 | -0.0544 | 0 | | | | |
| 171066 | 1 | 1 | -0.9550 | 0 | | | | |
| 171068 | 1 | 1 | -1.2636 | 0 | | | | |
| 171069 | 1 | 1 | -1.1132 | 0 | | | | |
| 171087 | 1 | 1 | -0.2728 | 0 | | | | |
| 171088 | 1 | 1 | -0.7604 | 0 | | | | |
| 171091 | 1 | 1 | 0.2614 | 0 | | | | |
| 171092 | 1 | 1 | -0.2415 | 0 | | | | |
| 171094 | 1 | 1 | 0.0370 | 0 | | | | |
| 176829 | 1 | 1 | -0.6333 | 0 | | | | |
| 176842 | 1 | 1 | -0.3839 | 0 | | | | |
| 176846 | 1 | 1 | -0.3051 | 0 | | | | |
| 176849 | 1 | 1 | -0.6043 | 0 | | | | |
| 176878 | 1 | 1 | -0.6215 | 0 | | | | |
| 176879 | 1 | 1 | -0.8010 | 0 | | | | |
| 176880 | 1 | 1 | 0.3866 | 0 | | | | |
| 176883 | 1 | 1 | -0.7604 | 0 | | | | |
| 176886 | 1 | 1 | -0.4780 | 0 | | | | |
| 176890 | 1 | 1 | -1.2100 | 0 | | | | |
| 176891 | 1 | 1 | -0.5939 | 0 | | | | |
| 176892 | 1 | 1 | -0.5878 | 0 | | | | |
| 176893 | 1 | 1 | -1.2474 | 0 | | | | |
| 176894 | 1 | 1 | -0.4723 | 0 | | | | |
| 176895 | 1 | 1 | -0.3023 | 0 | | | | |
| 176903 | 1 | 1 | -0.6205 | 0 | | | | |
| 176905 | 1 | 1 | -0.1176 | 0 | | | | |
| 176906 | 1 | 1 | -0.5988 | 0 | | | | |
| 176909 | 1 | 1 | -0.3220 | 0 | | | | |
| 176911 | 1 | 1 | -0.0716 | 0 | | | | |
| 176916 | 1 | 1 | -0.6628 | 0 | | | | |
| 176917 | 1 | 1 | -0.8684 | 0 | | | | |
| 176922 | 1 | 1 | -0.0915 | 0 | | | | |
| 176928 | 1 | 1 | -0.5840 | 0 | | | | |
| 176931 | 1 | 1 | -0.6588 | 0 | | | | |
| 176932 | 1 | 1 | -0.4302 | 0 | | | | |
| 176939 | 1 | 1 | -0.1116 | 0 | | | | |
| 176941 | 1 | 1 | -0.3769 | 0 | | | | |
| 176942 | 1 | 1 | -1.4231 | 0 | | | | |
| 176944 | 1 | 1 | -1.1615 | 0 | | | | |
| 176946 | 1 | 1 | -0.9787 | 0 | | | | |
| 180998 | 1 | 1 | -0.4264 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 181334 | 1 | 1 | 0.3848 | 0 | | | | |
| 181352 | 1 | 1 | -0.0170 | 0 | | | | |
| 184553 | 1 | 1 | -0.9616 | 0 | | | | |
| 238578 | 1 | 1 | -0.8858 | 0 | | | | |
| 238581 | 1 | 1 | -0.5381 | 0 | | | | |
| 238583 | 1 | 1 | -0.1646 | 0 | | | | |
| 238584 | 1 | 1 | -0.2298 | 0 | | | | |
| 238585 | 1 | 1 | -0.7643 | 0 | | | | |
| 244291 | 1 | 1 | 0.2056 | 0 | | | | |
| 244292 | 1 | 1 | 1.1010 | 0 | | | | |
| 254076 | 1 | 1 | -1.4039 | 0 | | | | |
| 176887 | 4 | 1 | 0.2523 | 0 | 1.6875 | 0.4679 | -0.8318 | -1.3236 |
| 176910 | 4 | 1 | 0.3564 | 0 | 1.26058 | 0.5404 | -0.6581 | -1.1429 |

## Grade 6 Reading

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 171111 | 1 | 1 | -1.4415 | 0 | | | | |
| 171114 | 1 | 1 | -0.5592 | 0 | | | | |
| 171115 | 1 | 1 | 0.4097 | 0 | | | | |
| 171117 | 1 | 1 | -0.2712 | 0 | | | | |
| 171141 | 1 | 1 | -0.5372 | 0 | | | | |
| 171144 | 1 | 1 | -1.3223 | 0 | | | | |
| 171147 | 1 | 1 | -1.4711 | 0 | | | | |
| 171148 | 1 | 1 | -0.6897 | 0 | | | | |
| 176705 | 1 | 1 | 0.1548 | 0 | | | | |
| 176708 | 1 | 1 | -0.1748 | 0 | | | | |
| 176709 | 1 | 1 | -1.1066 | 0 | | | | |
| 176710 | 1 | 1 | -1.5632 | 0 | | | | |
| 176711 | 1 | 1 | 0.1946 | 0 | | | | |
| 176713 | 1 | 1 | -0.4464 | 0 | | | | |
| 176714 | 1 | 1 | -0.0228 | 0 | | | | |
| 176732 | 1 | 1 | -1.2568 | 0 | | | | |
| 176734 | 1 | 1 | -0.9196 | 0 | | | | |
| 176735 | 1 | 1 | -1.3490 | 0 | | | | |
| 176737 | 1 | 1 | -1.1229 | 0 | | | | |
| 176738 | 1 | 1 | -0.1823 | 0 | | | | |
| 176766 | 1 | 1 | -1.0156 | 0 | | | | |
| 176767 | 1 | 1 | -0.7769 | 0 | | | | |
| 176768 | 1 | 1 | -1.9934 | 0 | | | | |
| 176769 | 1 | 1 | -1.1423 | 0 | | | | |
| 176784 | 1 | 1 | -0.3765 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 176785 | 1 | 1 | -0.7824 | 0 | | | | |
| 176786 | 1 | 1 | -0.1844 | 0 | | | | |
| 176787 | 1 | 1 | -0.3132 | 0 | | | | |
| 176790 | 1 | 1 | -0.5357 | 0 | | | | |
| 176792 | 1 | 1 | -0.4606 | 0 | | | | |
| 176793 | 1 | 1 | -0.1637 | 0 | | | | |
| 176798 | 1 | 1 | -1.1941 | 0 | | | | |
| 176806 | 1 | 1 | -1.1758 | 0 | | | | |
| 176807 | 1 | 1 | -0.1501 | 0 | | | | |
| 176810 | 1 | 1 | -1.1238 | 0 | | | | |
| 176811 | 1 | 1 | -0.2579 | 0 | | | | |
| 176812 | 1 | 1 | -0.9753 | 0 | | | | |
| 176821 | 1 | 1 | -1.3745 | 0 | | | | |
| 176824 | 1 | 1 | -0.6212 | 0 | | | | |
| 176825 | 1 | 1 | -0.8200 | 0 | | | | |
| 178230 | 1 | 1 | -0.3333 | 0 | | | | |
| 181002 | 1 | 1 | -0.5708 | 0 | | | | |
| 181005 | 1 | 1 | -0.3331 | 0 | | | | |
| 184207 | 1 | 1 | -0.9491 | 0 | | | | |
| 230173 | 1 | 1 | -0.3903 | 0 | | | | |
| 237098 | 1 | 1 | -0.3414 | 0 | | | | |
| 238603 | 1 | 1 | -0.9797 | 0 | | | | |
| 242319 | 1 | 1 | -0.7451 | 0 | | | | |
| 242320 | 1 | 1 | -0.7779 | 0 | | | | |
| 242327 | 1 | 1 | -0.8343 | 0 | | | | |
| 246909 | 1 | 1 | -0.6965 | 0 | | | | |
| 253991 | 1 | 1 | 0.6198 | 0 | | | | |
| 176804 | 4 | 1 | 0.2813 | 0 | 1.7579 | 0.5049 | -0.8038 | -1.4591 |
| 253992 | 4 | 1 | 0.1217 | 0 | 1.8959 | 0.2620 | -0.8501 | -1.3078 |

# Grade 7 Reading

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 171316 | 1 | 1 | -0.8263 | 0 | | | | |
| 171319 | 1 | 1 | -1.2138 | 0 | | | | |
| 171320 | 1 | 1 | -1.0185 | 0 | | | | |
| 171321 | 1 | 1 | -0.6404 | 0 | | | | |
| 171322 | 1 | 1 | -1.2275 | 0 | | | | |
| 171345 | 1 | 1 | -0.8039 | 0 | | | | |
| 171346 | 1 | 1 | -0.5045 | 0 | | | | |
| 171347 | 1 | 1 | -1.2534 | 0 | | | | |
| 171350 | 1 | 1 | -1.1692 | 0 | | | | |
| 177220 | 1 | 1 | -0.5033 | 0 | | | | |
| 177221 | 1 | 1 | -0.7294 | 0 | | | | |
| 177224 | 1 | 1 | -0.1782 | 0 | | | | |
| 177226 | 1 | 1 | -0.0355 | 0 | | | | |
| 177526 | 1 | 1 | -1.4486 | 0 | | | | |
| 177527 | 1 | 1 | -0.3073 | 0 | | | | |
| 177528 | 1 | 1 | -1.1407 | 0 | | | | |
| 177529 | 1 | 1 | -0.3037 | 0 | | | | |
| 177545 | 1 | 1 | -0.6473 | 0 | | | | |
| 177546 | 1 | 1 | -0.3850 | 0 | | | | |
| 177558 | 1 | 1 | -1.0296 | 0 | | | | |
| 177559 | 1 | 1 | -0.4515 | 0 | | | | |
| 177560 | 1 | 1 | -0.4986 | 0 | | | | |
| 177561 | 1 | 1 | -0.1316 | 0 | | | | |
| 177567 | 1 | 1 | -0.2272 | 0 | | | | |
| 177568 | 1 | 1 | 0.0022 | 0 | | | | |
| 177570 | 1 | 1 | -0.8702 | 0 | | | | |
| 177578 | 1 | 1 | 0.1503 | 0 | | | | |
| 177581 | 1 | 1 | -0.3918 | 0 | | | | |
| 177583 | 1 | 1 | -0.6796 | 0 | | | | |
| 177586 | 1 | 1 | -0.3814 | 0 | | | | |
| 177590 | 1 | 1 | -0.5175 | 0 | | | | |
| 177592 | 1 | 1 | -0.9026 | 0 | | | | |
| 177593 | 1 | 1 | -0.3086 | 0 | | | | |
| 177594 | 1 | 1 | -0.4351 | 0 | | | | |
| 177599 | 1 | 1 | -0.7414 | 0 | | | | |
| 177602 | 1 | 1 | -0.3428 | 0 | | | | |
| 177605 | 1 | 1 | -0.5199 | 0 | | | | |
| 178291 | 1 | 1 | -0.1147 | 0 | | | | |
| 178294 | 1 | 1 | -0.1274 | 0 | | | | |
| 178295 | 1 | 1 | -0.5887 | 0 | | | | |
| 179489 | 1 | 1 | -0.2504 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 230178 | 1 | 1 | -0.5312 | 0 | | | | |
| 237093 | 1 | 1 | -1.0350 | 0 | | | | |
| 237094 | 1 | 1 | -1.3216 | 0 | | | | |
| 237095 | 1 | 1 | -1.5384 | 0 | | | | |
| 237096 | 1 | 1 | -0.0800 | 0 | | | | |
| 237097 | 1 | 1 | -0.7653 | 0 | | | | |
| 242336 | 1 | 1 | -0.5007 | 0 | | | | |
| 254375 | 1 | 1 | -1.3673 | 0 | | | | |
| 254376 | 1 | 1 | -1.1484 | 0 | | | | |
| 254377 | 1 | 1 | -1.0169 | 0 | | | | |
| 254586 | 1 | 1 | -0.0266 | 0 | | | | |
| 177574 | 4 | 1 | 0.0285 | 0 | 1.5888 | 0.4075 | -0.8113 | -1.1850 |
| 177595 | 4 | 1 | 0.1188 | 0 | 1.3512 | 0.4650 | -0.59043 | -1.2258 |

# Grade 8 Reading

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 171197 | 1 | 1 | -0.1628 | 0 | | | | |
| 171201 | 1 | 1 | -0.1652 | 0 | | | | |
| 171203 | 1 | 1 | -0.3198 | 0 | | | | |
| 171204 | 1 | 1 | -1.1414 | 0 | | | | |
| 171206 | 1 | 1 | -0.5843 | 0 | | | | |
| 171207 | 1 | 1 | 0.2280 | 0 | | | | |
| 171211 | 1 | 1 | -0.2862 | 0 | | | | |
| 171212 | 1 | 1 | -1.1199 | 0 | | | | |
| 171214 | 1 | 1 | -0.8172 | 0 | | | | |
| 211757 | 1 | 1 | -0.2262 | 0 | | | | |
| 211760 | 1 | 1 | -0.3662 | 0 | | | | |
| 211762 | 1 | 1 | 0.1951 | 0 | | | | |
| 211764 | 1 | 1 | -0.8165 | 0 | | | | |
| 211775 | 1 | 1 | -0.4347 | 0 | | | | |
| 211852 | 1 | 1 | -0.3988 | 0 | | | | |
| 211859 | 1 | 1 | -1.2589 | 0 | | | | |
| 211864 | 1 | 1 | 0.2100 | 0 | | | | |
| 211866 | 1 | 1 | -0.9641 | 0 | | | | |
| 211868 | 1 | 1 | -0.8492 | 0 | | | | |
| 211871 | 1 | 1 | -0.7126 | 0 | | | | |
| 211874 | 1 | 1 | -0.2394 | 0 | | | | |
| 211876 | 1 | 1 | -0.1422 | 0 | | | | |
| 211878 | 1 | 1 | -0.4977 | 0 | | | | |
| 211881 | 1 | 1 | -0.4825 | 0 | | | | |
| 211884 | 1 | 1 | -0.0086 | 0 | | | | |
| 211886 | 1 | 1 | -0.7530 | 0 | | | | |
| 211889 | 1 | 1 | -0.3143 | 0 | | | | |
| 211891 | 1 | 1 | -1.0287 | 0 | | | | |
| 211895 | 1 | 1 | -1.1627 | 0 | | | | |
| 211998 | 1 | 1 | 0.3415 | 0 | | | | |
| 212002 | 1 | 1 | -0.0812 | 0 | | | | |
| 212009 | 1 | 1 | -0.9865 | 0 | | | | |
| 212010 | 1 | 1 | -1.0481 | 0 | | | | |
| 212013 | 1 | 1 | -0.1765 | 0 | | | | |
| 212021 | 1 | 1 | -0.5383 | 0 | | | | |
| 212025 | 1 | 1 | -0.6264 | 0 | | | | |
| 212026 | 1 | 1 | -1.0889 | 0 | | | | |
| 212028 | 1 | 1 | 0.4659 | 0 | | | | |
| 212033 | 1 | 1 | -0.5634 | 0 | | | | |
| 212034 | 1 | 1 | -0.6305 | 0 | | | | |
| 212036 | 1 | 1 | -0.1696 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 212045 | 1 | 1 | 0.0297 | 0 | | | | |
| 212046 | 1 | 1 | -0.5009 | 0 | | | | |
| 212047 | 1 | 1 | 0.4210 | 0 | | | | |
| 212049 | 1 | 1 | -0.8141 | 0 | | | | |
| 212057 | 1 | 1 | -1.6634 | 0 | | | | |
| 212059 | 1 | 1 | -0.3385 | 0 | | | | |
| 212060 | 1 | 1 | 0.0252 | 0 | | | | |
| 212061 | 1 | 1 | 0.2255 | 0 | | | | |
| 212067 | 1 | 1 | -0.9246 | 0 | | | | |
| 239867 | 1 | 1 | -1.3029 | 0 | | | | |
| 249023 | 1 | 1 | -0.3931 | 0 | | | | |
| 211900 | 4 | 1 | -0.0036 | 0 | 1.3737 | 0.6628 | -0.6017 | -1.4349 |
| 212051 | 4 | 1 | 0.1632 | 0 | 1.0088 | 0.7531 | -0.3925 | -1.3694 |

## Grade 10 Reading

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 170894 | 1 | 1 | -0.4291 | 0 | | | | |
| 170897 | 1 | 1 | -0.0480 | 0 | | | | |
| 170899 | 1 | 1 | -0.7982 | 0 | | | | |
| 170900 | 1 | 1 | 0.1249 | 0 | | | | |
| 170901 | 1 | 1 | -0.4534 | 0 | | | | |
| 170903 | 1 | 1 | 0.1200 | 0 | | | | |
| 170925 | 1 | 1 | 0.1288 | 0 | | | | |
| 170930 | 1 | 1 | 0.2679 | 0 | | | | |
| 170934 | 1 | 1 | -0.7994 | 0 | | | | |
| 170936 | 1 | 1 | -0.9589 | 0 | | | | |
| 170937 | 1 | 1 | -0.4540 | 0 | | | | |
| 170939 | 1 | 1 | -0.9212 | 0 | | | | |
| 170941 | 1 | 1 | 1.2993 | 0 | | | | |
| 170942 | 1 | 1 | -0.9415 | 0 | | | | |
| 211538 | 1 | 1 | -0.5724 | 0 | | | | |
| 211541 | 1 | 1 | -0.6459 | 0 | | | | |
| 211542 | 1 | 1 | 0.1112 | 0 | | | | |
| 211545 | 1 | 1 | -0.8534 | 0 | | | | |
| 211546 | 1 | 1 | -0.6573 | 0 | | | | |
| 211550 | 1 | 1 | -0.0508 | 0 | | | | |
| 211554 | 1 | 1 | -0.8174 | 0 | | | | |
| 211559 | 1 | 1 | 0.2029 | 0 | | | | |
| 211564 | 1 | 1 | -0.8075 | 0 | | | | |
| 211571 | 1 | 1 | -0.3869 | 0 | | | | |
| 211617 | 1 | 1 | -0.0626 | 0 | | | | |

| ITEM | MAX | A | B | C | D1 | D2 | D3 | D4 |
|---|---|---|---|---|---|---|---|---|
| 211619 | 1 | 1 | -0.5424 | 0 | | | | |
| 211621 | 1 | 1 | -0.6096 | 0 | | | | |
| 211631 | 1 | 1 | -0.6096 | 0 | | | | |
| 211632 | 1 | 1 | -0.1599 | 0 | | | | |
| 211703 | 1 | 1 | -1.1773 | 0 | | | | |
| 211706 | 1 | 1 | -0.7473 | 0 | | | | |
| 211715 | 1 | 1 | -0.4422 | 0 | | | | |
| 211717 | 1 | 1 | -0.0488 | 0 | | | | |
| 211720 | 1 | 1 | -0.4652 | 0 | | | | |
| 211722 | 1 | 1 | -0.2482 | 0 | | | | |
| 211725 | 1 | 1 | -0.4335 | 0 | | | | |
| 211731 | 1 | 1 | 0.4362 | 0 | | | | |
| 211732 | 1 | 1 | 0.1256 | 0 | | | | |
| 211733 | 1 | 1 | -0.3279 | 0 | | | | |
| 211789 | 1 | 1 | -0.1251 | 0 | | | | |
| 211791 | 1 | 1 | -0.6406 | 0 | | | | |
| 211792 | 1 | 1 | -0.1374 | 0 | | | | |
| 211793 | 1 | 1 | -0.5150 | 0 | | | | |
| 211794 | 1 | 1 | -0.2401 | 0 | | | | |
| 211834 | 1 | 1 | -0.5993 | 0 | | | | |
| 211838 | 1 | 1 | -1.4040 | 0 | | | | |
| 211840 | 1 | 1 | -1.1479 | 0 | | | | |
| 211843 | 1 | 1 | 0.3300 | 0 | | | | |
| 211844 | 1 | 1 | -0.4366 | 0 | | | | |
| 211849 | 1 | 1 | -0.1115 | 0 | | | | |
| 211853 | 1 | 1 | -0.6288 | 0 | | | | |
| 211856 | 1 | 1 | -0.3279 | 0 | | | | |
| 211860 | 1 | 1 | 0.0073 | 0 | | | | |
| 238788 | 1 | 1 | -0.2282 | 0 | | | | |
| 238792 | 1 | 1 | -0.6729 | 0 | | | | |
| 238795 | 1 | 1 | -0.5330 | 0 | | | | |
| 238797 | 1 | 1 | -1.0953 | 0 | | | | |
| 211741 | 4 | 1 | 0.0650 | 0 | 1.2426 | 0.3785 | -0.4957 | -1.1255 |
| 212168 | 4 | 1 | -0.0905 | 0 | 0.8970 | 0.8748 | -0.5424 | -1.2295 |

# APPENDIX B: TECHNICAL ADVISORY COMMITTEE

| 2006 Technical Advisory Committee (TAC) Members | | | | |
|---|---|---|---|---|
| **First Name** | **Last Name** | **Position** | **Department** | **Organization** |
| Art | Bangert, Ph.D. | Assistant Professor | Adult and Higher Education | Montana State University |
| Susan | Brookhart, Ph.D. | President | | Brookhart Enterprises, LLC |
| Ellen | Forte, Ph.D. | President | | edCount, LLC |
| Michael | Kozlow, Ph.D. | Program Director | Assessment Program | Northwest Regional Educational Lab |
| Scott | Marion, Ph.D. | Vice-President | | Center for Assessment |
| Madalyn | Quinlan | Chief Exectutive Officer | | OPI |
| Stanley | Rabinowitz, Ph.D. | Program Director | Assessment & Standards Development Services | WestEd |
| Stephen | Sireci, Ph.D. | Professor | | University of Massachusetts Amherst |
| Judy | Snow | State Assessment Director | | OPI |
| Bud | Williams | Assistant Superintendent | | OPI |

# APPENDIX C: CRT STANDARD SETTING REPORT

# June 21-22, 2006
# July 26-27, 2006
# Helena, MT

### Overview of Standard-Setting Meetings

Standard setting for the MontCAS, Phase 2 CRT in Reading and Mathematics, Grades 3 through 8 and 10, occurred in two stages. In the first stage, which occurred on Wednesday and Thursday, June 21st and 22nd, standards were set for Grades 4, 8 and 10. In the second stage, which occurred on Wednesday and Thursday, July 26th and 27th, standards were set for Grades 3, 5, 6 and 7. For the second stage, each panel set standards for two grade levels; the groups were comprised as follows:

- Math, Grades 3 & 5
- Math, Grades 6 & 7
- Reading, Grades 3 & 5
- Reading, Grades 6 & 7

The standard-setting method implemented for both content areas and all grades was a modified version of the bookmark method. An overview of this method is described below. All panels followed the same procedures.

To help ensure consistency of procedures between panels, each panel was led through the standard-setting process by trained facilitators from Measured Progress.

## Overview of Process

This section of the report provides an overview of the standard-setting process as it was implemented in Montana. The process was divided into the following three stages, each with several constituent tasks.

❖ Tasks completed prior to the standard-setting meeting
- 2004 performance levels and Performance Level Descriptors
- Preparation of materials for panelists
- Preparation of presentation materials
- Preparation of Instructions for Facilitators Document
- Preparation of systems and materials for analysis during the meeting
- Selection of panelists
- Calculation of starting cut points

- Tasks completed during the standard-setting meeting
  - Opening Sessions: welcome and orientation (overview of process)
  - Completion of standard-setting activities
    - Reviewing assessment materials
    - Completion of item map
    - Reviewing Performance Level Descriptors created in 2004
    - Round 1 judgments
    - Tabulation of Round 1 results and presentation of data to all panel groups
    - Round 2 judgments
    - Feedback on Performance Level Descriptors
  - Modification of process for grades 3, 5, 6 and 7 (Performance Level Descriptors were created)
  - Round 2 results from standard-setting meetings
  - Evaluation
- Tasks completed after the standard-setting meeting
  - Analysis and review of panelists' feedback
  - Preparation of recommended cut scores
  - Preparation of this standard-setting report

## Tasks Completed Prior to the Standard-Setting Meeting

Creation of Performance Levels and Performance Level Descriptors

The Performance Level Descriptors provided panelists the official description of the knowledge, skills and abilities students are expected to be able to display to be classified into each performance level. These Performance Level Descriptors were presented to panelists. Panelists were given the option at the end of the standard-setting process to recommend additions and refinements to the Descriptors. The Descriptors are provided in Appendix B of this document.

Preparation of Materials for Panelists

The following materials were provided to the panelists at the standard-setting meeting:

- Meeting agenda (see Appendix A)
- Non-disclosure agreement
- Performance Level Descriptors (see Appendix B)
- Assessment booklet
- Answer key/scoring rubrics
- Ordered Item Booklet
- Item Map (see Appendix D)

- Rating forms (see Appendix E)
- Student profiles
- Evaluation Form (see Appendix F)

The agendas, Performance Level Descriptors, sample item map, sample rating form, sample student profile, and evaluation results are provided in the appendices.

Preparation of Presentation Materials

The PowerPoint presentations used in the opening sessions were prepared prior to the meeting. Copies of the PowerPoint slides are included in Appendix C of this document

Preparation of Instructions for Facilitators Document

A document, "General Instructions for Standard Setting Group Facilitators," was created for the group facilitators to refer to as they worked through the process. Copies of these instructions (one for Grades 4, 8 and 10, and one for Grades 3, 5, 6 and 7) are included in Appendix C of this document.

Preparation of Systems and Materials for Analysis During the Meeting

The programming of all analyses to be conducted during the standard-setting meeting was completed and thoroughly tested prior to the standard-setting meeting.

Selection of Panelists

Panelists were selected prior to the standard-setting meeting. The goal was to have 15 panelists for each of the panels, for a total of 150. The actual number of panelists who participated was 105 (59 in stage 1 and 46 in stage 2), distributed as follows:

- Math, Grade 4: 11
- Math, Grades 3 & 5: 13
- Math, Grades 6 & 7: 15
- Math Grade 8: 11
- Math Grade 10: 12
- Reading, Grade 4: 8
- Reading, Grades 3 & 5: 11
- Reading, Grades 6 & 7: 7
- Reading, Grade 8: 7
- Reading, Grade 10: 10

Of the 105 panelists, there were 72 teachers, 21 administrators, and 12 other (parents, librarians, counselors, etc.) All panelists were white, and 78 were female and 27 male.

Calculation of Starting Cut Points

The starting cut points for Grades 4, 8 and 10 were the cuts that were established in a standard-setting meeting in the summer of 2004. Once the stage 1 standard setting was completed for

Grades 4, 8 and 10, starting cut points for stage 2 (grades 3, 5, 6 and 7) were calculated by interpolating (or extrapolating) from the cutpoints obtained for Grades 4, 8 and 10. The process for calculating the cuts was:

1. find the percentage of students who fell below each raw score cut for grades 4, 8 and 10,
2. standardize the percent-below values using the *z*-transformation,
3. calculate a line of best fit across grades,
4. use the inverse-*z*-transformation to translate the *z*'s back into percent-below values and,
5. for grades 3, 5, 6 and 7, find the raw score associated with the observed percent-below value closest to, but not lower than, the smoothed value.

These five steps were repeated for each of the cut points. The observed percent-below values associated with the starting cuts are presented in Figures 1 and 2 in the final section of this report.

## Tasks Completed During the Standard-Setting Meeting

Day 1 of each standard-setting meeting began with a general orientation session that was attended by all panelists. The purpose of this session was to provide some background information, provide an introduction to the issues of standard setting, explain the activities that would occur during the standard-setting meeting, and go over some of the materials that would be used. At the conclusion of the opening session the floor was opened to questions about the standard-setting process.

After the large-group session, the panelists assembled into their grade/content area groups. Each group was in a separate room.

Completion of Standard-Setting Activities

As mentioned previously, panelists for Grades 3, 5, 6 and 7 set standards for two separate tests during the meeting, while panelists for Grades 4, 8 and 10 set standards for a single test. As such, the processes followed during the June and July standard-setting meetings were somewhat different, although the general steps were the same. These steps are described below.

Reviewing Assessment Materials. Each panel began by taking the test for their assigned grade level and content area. The purpose of this step was to make sure the panelists were thoroughly familiar with the assessment and what the students needed to do. Once panelists had completed the test, an answer key was distributed. At this point, panelists could discuss any issues that arose regarding items or scoring.

Completion of item map. The purpose of the next step was to ensure that panelists become very familiar with the ordered item booklet and understand the relationships among the ordered items. The ordered item booklet contained one item (or item score category) per page, and was ordered from the easiest item (or item score category) to the most difficult. The ordered item booklet was created by sorting items by their IRT-based difficulty values (b corresponding to p+ = 0.67 was used). A one-parameter logistic IRT model was used for the dichotomous items and the partial credit IRT model was used for the polytomous items. The group facilitators explained to the panelists that each four-point constructed-response item would appear four times in the ordered item booklet, once for each possible score point.

The item map listed the items in the same order as they were presented in the ordered item booklet, and had spaces for the panelists to write in the knowledge, skills, and abilities required to successfully complete each item (or get a certain score on a polytomous item); there was also a space for the panelists to write in why they felt the current ordered item was more difficult than the previous one. A sample item map is provided in Appendix D.

Because starting cuts were provided, and because the item mapping process can be very time-consuming, panelists were instructed to start approximately five ordered items prior to, and stop approximately five ordered items after, each starting cut point. However, panelists were told that the range of plus or minus five ordered items was a guideline, and that they were free to expand that range as appropriate.

Each panelist stepped through the ordered item booklet, item by item, and considered the knowledge, skills, and abilities students needed to answer each item correctly. They wrote that information onto the item map for each ordered item, as well as why the item was more difficult than the previous item. Panelists were able to refer to the Performance Level Descriptors and the definitions of the 'borderline' students they had developed earlier to help them make these determinations. After they were finished working individually, the panelists had an opportunity to discuss the item map as a group and make any necessary additions or adjustments.

Reviewing Performance Level Descriptors. Next, the panelists reviewed the Performance Level Descriptors (see Appendix B). This step of the process was very important; it was designed to ensure that the panelists thoroughly understood the knowledge, skills, and abilities that students needed to demonstrate in order to be classified as *Nearing Proficiency*, *Proficient*, and *Advanced*. The panelists began by individually reviewing the descriptors, and then they discussed them as a group, providing clarification for each level. Once they finished discussing the descriptors, the panelists developed definitions of borderline students, i.e., students who are "just able enough" to be categorized into each performance level. After the discussions were completed, bulleted lists of characteristics for

each level were generated, based on the whole group discussion, and were posted in the room for panelists to refer to throughout the bookmark process.

Round 1 Judgments. In the first round, panelists worked as a group to evaluate and, if necessary, revise the starting cut points. For this task, panelists used the Performance Level Descriptors, the borderline definitions, the completed item map, and the ordered item booklet. Beginning with the ordered item approximately five items before the starting *Nearing Proficiency* cut point, panelists considered the skills and abilities students needed to complete each ordered item and asked themselves the question, "Would at least 2 out of 3 students performing at the borderline of *Nearing Proficiency* answer this question correctly?" The panelists discussed each ordered item in turn, asking the same question and referring to the item map and the definition of the borderline *Nearing Proficiency* student. The panelists would place their bookmark at the point in the ordered item booklet where their answer to the question changed from "yes" (or predominantly "yes") to "no" (or predominantly "no"). Once the discussion was done for the *Novice* vs. *Nearing Proficiency* cut, the panelists repeated the process for the remaining two cuts. Each panelist used the Rating Form provided to record his/her ratings for each cut. A sample rating form is provided in Appendix E.

Although the panelists were working as a group, the facilitators made sure they understood that they should set the bookmark according to their individual best judgment and that they did not need to come to consensus as a group. They were encouraged to listen to the points made by their colleagues, but told that no one should feel compelled to change his or her bookmark placement.

Tabulation of Round 1 results and presentation of data to all panel groups. After Round 1 ratings were completed, Measured Progress staff calculated the room average cut points and the associated impact data. The impact data showed the percentage of students statewide who would fall into each performance level category according to the room average cut points.

Prior to beginning Round 2 judgments, all panels convened together as a whole group and were shown the Round 1 results for all grades/content areas. The rationale for this step was to give the groups an opportunity to see whether their ratings were consistent with those of the other grade groups and, if not, to discuss whether they needed to make any adjustments.

Round 2 judgments. The purpose of Round 2 was for the panelists to revisit their Round 1 placements as a whole group and to revise their ratings on the basis of that discussion. The panelists shared their rationale for their bookmark placements in terms of the knowledge and skills students must have in order to be classified as *Nearing Proficiency*, *Proficient*, or *Advanced*. The panelists were asked to pay particular attention to how their ratings compared to those of the other panelists in order to get a sense for whether they were unusually stringent or lenient.

To assist them with their discussion, panelists were provided with the room average cut points and the associated impact data. The panelists were asked to consider that information along with the input of their colleagues in deciding whether or not they should make any changes to their Round 1 ratings. The facilitators emphasized to the panelists that they should not base their decisions on the percentages, but instead on the Performance Level Descriptors and the test content. The purpose of providing the impact data was to give the panelists another check on the reasonableness of their bookmark placements.

One final resource that the panelists were given to help them decide whether they felt the cut points had been placed appropriately was the student profiles. Each row in the profile represented a typical pattern of item scores for a student who achieved a certain total raw score. The rows were presented in order from lowest to highest total raw score, and the columns were ordered first by item type, then from the easiest to the most difficult item. Panelists were asked to look at the pattern of item responses for the profiles that fell into each performance level according to the Round 1 average cut points, and determine whether they felt the item scores for each profile were consistent with the description of the level into which that profile had been placed. If the answer was no, the panelists considered whether they needed to modify their cut point placements.

After all panelists had an opportunity to share their rationale for where they placed their cut points and the room completed their discussions, the panelists then had the opportunity to change or revise their Round 1 ratings. Once again, the facilitators emphasized to the panelists that they did not need to come to consensus and that they were making their own individual bookmark placements. Each panelist once again used the Rating Form to record his/her ratings.

Feedback on Performance Level Descriptors. After completing the rating process, the panelists made suggestions for modifications to the Performance Level Descriptors based on the round 2 results of the standard-setting process (see revised Performance Level Descriptors, Appendix D, Technical Report).

Modification of Process for Grades 3, 5, 6 and 7

For Grades 4, 8 and 10, the standard-setting activities described above were completed in the order presented: reviewing the assessment materials, completing the item map, discussing the Performance Level Descriptors and creating definitions of borderline students, Round 1 ratings, sharing of Round 1 results across groups, Round 2 ratings, and providing feedback on the Performance Level Descriptors. For Grades 3, 5, 6 and 7, in which each group set standards for two tests, the order of the steps was changed slightly to make the process as efficient as possible. Specifically, each group completed Round 1 ratings for both tests, then reconvened as a large group. In the large-group meeting, the Round 1 results for all tests were presented, as well as the final results from the June

standard setting for Grades 4, 8 and 10. After the large-group session, each group completed Round 2 ratings and provided feedback on the Performance Level Descriptors for both tests.

Round 2 Results from Standard-Setting Meetings

The raw score ranges and percentage of students classified into each performance level, based on the group average cut scores from Round 2, are presented in Tables 1 through 7. In addition, the percent-below values associated with the Round 2 cuts are presented in Figures 1 and 2 in the final section of this report.

Table 1
Round 2 Cut Scores and Impact Data -- Grade 3

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 45-60 | 37.9 | 55-66 | 24.7 |
| *Proficient* | 32-44 | 41.4 | 42-54 | 44.6 |
| *Nearing Proficiency* | 21-31 | 14.6 | 34-41 | 16.6 |
| *Novice* | 0-20 | 6.1 | 0-33 | 14.2 |

Table 2
Round 2 Cut Scores and Impact Data -- Grade 4

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 47-60 | 33.5 | 55-66 | 22.9 |
| *Proficient* | 33-46 | 47.1 | 41-54 | 44.0 |
| *Nearing Proficiency* | 24-32 | 12.6 | 33-40 | 16.8 |
| *Novice* | 0-23 | 6.8 | 0-32 | 16.2 |

Table 3
Round 2 Cut Scores and Impact Data -- Grade 5

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 44-60 | 31.5 | 48-66 | 23.0 |
| *Proficient* | 30-43 | 48.5 | 35-47 | 37.0 |
| *Nearing Proficiency* | 23-29 | 11.5 | 28-34 | 17.8 |
| *Novice* | 0-22 | 8.6 | 0-27 | 22.3 |

Table 4

Round 2 Cut Scores and Impact Data -- Grade 6

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 48-60 | 23.3 | 44-66 | 20.5 |
| *Proficient* | 35-47 | 53.0 | 31-43 | 35.7 |
| *Nearing Proficiency* | 25-34 | 17.1 | 22-30 | 24.9 |
| *Novice* | 0-24 | 6.5 | 0-21 | 18.8 |

Table 5

Round 2 Cut Scores and Impact Data -- Grade 7

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 48-60 | 27.4 | 43-66 | 24.1 |
| *Proficient* | 33-47 | 50.5 | 31-42 | 34.6 |
| *Nearing Proficiency* | 24-32 | 13.3 | 20-30 | 30.7 |
| *Novice* | 0-23 | 8.8 | 0-19 | 10.7 |

Table 6

Round 2 Cut Scores and Impact Data -- Grade 8

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 47-60 | 35.2 | 47-66 | 22.8 |
| *Proficient* | 38-46 | 36.0 | 32-46 | 38.2 |
| *Nearing Proficiency* | 29-37 | 18.4 | 24-31 | 19.2 |
| *Novice* | 0-28 | 10.5 | 0-23 | 19.8 |

Table 7

Round 2 Cut Scores and Impact Data -- Grade 10

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 53-65 | 25.0 | 51-71 | 24.4 |
| *Proficient* | 39-52 | 51.3 | 40-50 | 23.7 |
| *Nearing Proficiency* | 31-38 | 13.3 | 27-39 | 28.8 |
| *Novice* | 0-30 | 10.4 | 0-26 | 23.1 |

Evaluation

At the end of the process, panelists anonymously completed an evaluation form. The results of the evaluation are presented in Appendix F.

## Tasks Completed After the Standard-Setting Meeting

Upon conclusion of the standard-setting meeting, several important tasks were completed. These tasks centered on reviewing the standard-setting meeting and addressing anomalies that may have occurred in the process or in the results.

Analysis and review of panelists' feedback

Upon completion of the evaluation forms, panelists' responses were reviewed. This review did not reveal any anomalies in the standard-setting process or indicate any reason that a particular panelist's data should not be incorporated in obtaining the final results. It appeared that all panelists understood the rating task and attended to it appropriately. Panelist responses to the evaluation items are presented in Appendix F.

Preparation of Recommended Cut Scores

Results of the two standard-setting meetings were presented to the Montana Technical Advisory Committee (TAC) on August 3. The information provided to the TAC included historical data (percent proficient or above in 2003-04 and 2004-05 for grades 4, 8 and 10), starting cuts, and Round 2 results. In addition, for Stage 1 (grades 4, 8 and 10), results obtained by averaging the starting and Round 2 cuts were presented. Finally, results were smoothed across all seven grades for each content area; these smoothed cuts were also presented to the TAC as the cuts recommended by Measured Progress for operational use. The smoothed cuts were approved by the TAC and by OPI, and were adopted for use in reporting.

The final adopted cuts are shown below in Tables 8 through 14 and Figures 1 and 2. The tables show the raw score range for each performance level, as well as the percentage of students who fall into each performance level category. The figures show the starting cut points, the Round 2 results, and the final (smoothed) cut points.

Table 8

Final (Smoothed) Cut Scores and Impact Data -- Grade 3

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 45-60 | 37.9 | 55-66 | 24.7 |
| *Proficient* | 31-44 | 43.2 | 43-54 | 41.7 |
| *Nearing Proficiency* | 21-30 | 12.8 | 35-42 | 17.9 |
| *Novice* | 0-20 | 6.1 | 0-34 | 15.7 |

Table 9

Final (Smoothed) Cut Scores and Impact Data -- Grade 4

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 47-60 | 33.5 | 54-66 | 26.0 |
| *Proficient* | 33-46 | 47.1 | 42-53 | 38.3 |
| *Nearing Proficiency* | 24-32 | 12.6 | 33-41 | 19.5 |
| *Novice* | 0-23 | 6.8 | 0-32 | 16.2 |

Table 10

Final (Smoothed) Cut Scores and Impact Data -- Grade 5

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 43-60 | 36.0 | 47-66 | 25.4 |
| *Proficient* | 30-42 | 44.0 | 34-46 | 37.4 |
| *Nearing Proficiency* | 21-29 | 13.3 | 25-33 | 21.5 |
| *Novice* | 0-20 | 6.7 | 0-24 | 15.7 |

Table 11

Round 2 Cut Scores and Impact Data -- Grade 6

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 46-60 | 33.0 | 42-66 | 25.2 |
| *Proficient* | 34-45 | 45.9 | 29-41 | 37.3 |
| *Nearing Proficiency* | 26-33 | 13.4 | 21-28 | 21.2 |
| *Novice* | 0-25 | 7.7 | 0-20 | 16.3 |

Table 12

Final (Smoothed) Cut Scores and Impact Data -- Grade 7

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 47-60 | 31.7 | 43-66 | 24.1 |
| *Proficient* | 33-46 | 46.2 | 30-42 | 37.5 |
| *Nearing Proficiency* | 24-32 | 13.3 | 22-29 | 22.9 |
| *Novice* | 0-23 | 8.8 | 0-21 | 15.5 |

Table 13

Final (Smoothed) Cut Scores and Impact Data -- Grade 8

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 48-60 | 30.6 | 46-66 | 25.2 |
| *Proficient* | 36-47 | 46.0 | 33-45 | 33.1 |
| *Nearing Proficiency* | 28-35 | 14.1 | 22-32 | 26.0 |
| *Novice* | 0-27 | 9.3 | 0-21 | 15.8 |

Table 14

Final (Smoothed) Cut Scores and Impact Data -- Grade 10

| | Reading | | Mathematics | |
|---|---|---|---|---|
| Proficiency Level | Raw Score Range | % in Level | Raw Score Range | % in Level |
| *Advanced* | 52-65 | 29.6 | 51-71 | 24.4 |
| *Proficient* | 39-51 | 46.7 | 37-50 | 30.8 |
| *Nearing Proficiency* | 31-38 | 13.3 | 24-36 | 28.2 |
| *Novice* | 0-30 | 10.4 | 0-23 | 16.6 |

**Figure 1: Montana CRT 2005-06 Standard Setting Results -- Reading**

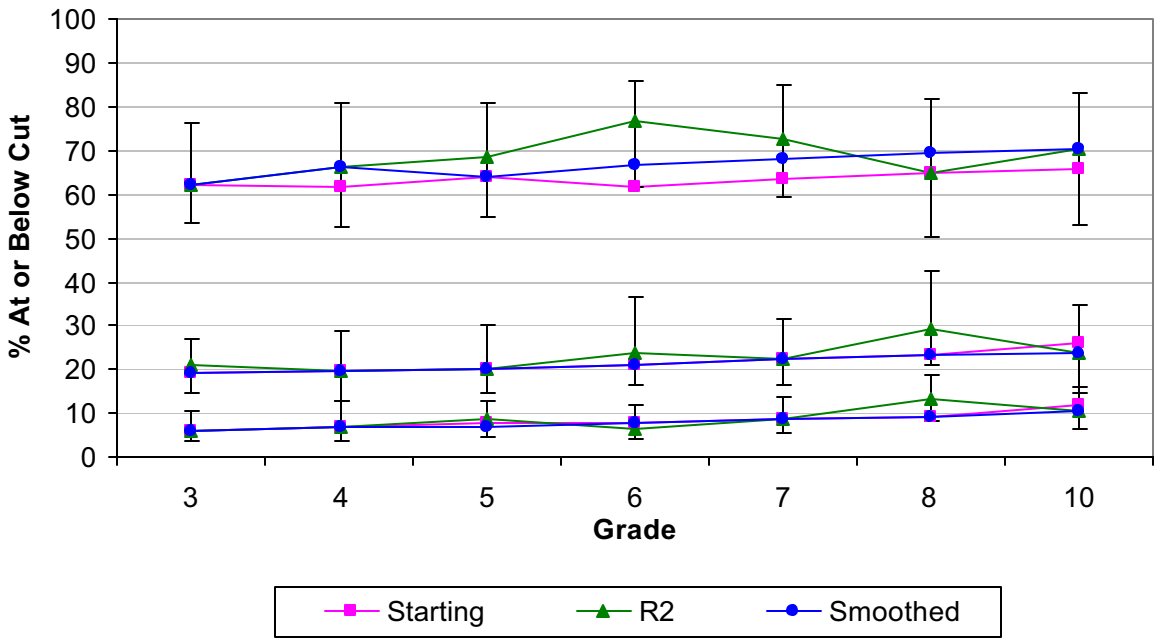
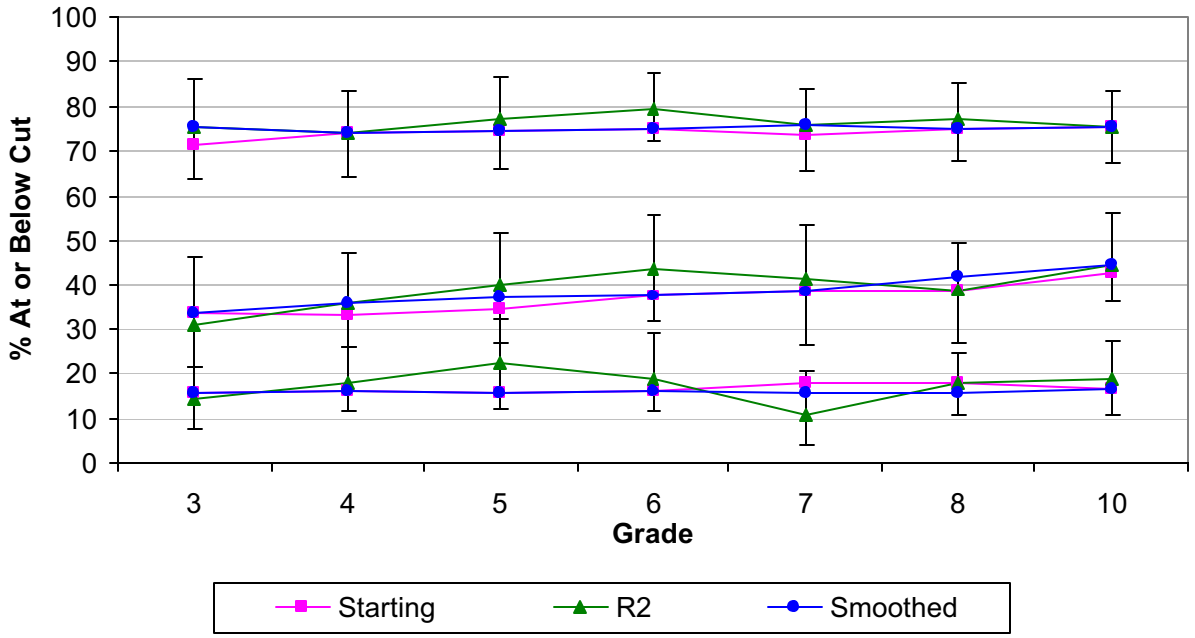



**Figure 2: Montana CRT 2005-06 Mathematics Standard Setting Results**



Preparation of This Standard-Setting Report

Following final compilation of standard-setting results for 2006, Measured Progress prepared this report, which documents the procedures and results of the 2006 standard-setting meetings in order to establish performance standards for the MontCAS, Phase 2 CRT.

# Appendix A

# Standard Setting Meeting Agendas

# MONTCAS, PHASE 2
# CRT STANDARD SETTING
# MATHEMATICS AND READING (GRADES 4, 8, AND 10)

## JUNE 21-22, 2006

## AGENDA

**WEDNESDAY, JUNE 21**

| | |
|---|---|
| 8:00 – 8:30 | Registration & Breakfast (Executive Room) |
| 8:30 – 10:30 | Introduction, Overview, and Training of Standard Setting Process |
| 10:30 – 10:45 | Break |
| 10:45 – 12:00 | Move to Grade Level/Content Area Work Rooms |
| 12:00 – 12:45 | Lunch (Executive Room) |
| 12:45 – 2:30 | Continue in Work Rooms |
| 2:30 – 2:45 | Break |
| 2:45 – 4:30 | Continue in Work Rooms |
| 4:30 | Adjourn |

**THRUSDAY, JUNE 22**

| | |
|---|---|
| 8:00 – 8:30 | Breakfast (Executive Room) |
| 8:30 – 10:30 | Move to Grade Level/Content Area Work Rooms |
| 10:30 – 10:45 | Break |
| 10:45 – 12:00 | Continue in Work Rooms |
| 12:00 – 12:45 | Lunch (Executive Room) |
| 12:45 – 2:30 | Continue in Work Rooms |
| 2:30 – 2:45 | Break |
| 2:45 – 4:30 | Continue in Work Rooms |
| 4:30 | Adjourn |

# MONTCAS, PHASE 2
# CRT STANDARD SETTING
# MATHEMATICS AND READING (GRADES 3/5 AND 6/7)

## JULY 26-27, 2006

## AGENDA

**WEDNESDAY, JULY 26**

| | |
|---|---|
| 8:00 – 8:30 | Registration & Breakfast – Capitol Room |
| 8:30 – 10:30 | Introduction, Overview, and Training of Standard Setting Process |
| 10:30 – 10:45 | Break – Second Floor Mezzanine |
| 10:45 – 12:00 | Move to Grade Level/Content Area Work Rooms |
| 12:00 – 12:45 | Lunch - Natatorium |
| 12:45 – 2:30 | Continue in Work Rooms |
| 2:30 – 2:45 | Dessert Break – Second Floor Mezzanine |
| 2:45 – 5:00 | Continue in Work Rooms |
| 5:00 | Adjourn |

**THURSDAY, JULY 27**

| | |
|---|---|
| 8:00 – 8:30 | Breakfast - Executive Room |
| 8:30 – 10:30 | Move to Grade Level/Content Area Work Rooms |
| 10:30 – 10:45 | Break – Second Floor Mezzanine |
| 10:45 – 12:00 | Continue in Work Rooms |
| 12:00 – 12:45 | Lunch - Executive Room |
| 12:45 – 2:00 | Continue in Work Rooms |
| 2:00 – 2:30 | Dessert Break – Executive Room |
| 2:30 – 3:00 | Large Group Meeting |
| 2:45 – 5:00 | Continue in Work Rooms |
| 5:00 | Adjourn |

# Appendix B

## Performance Level Descriptors

### Grade 4 Math Performance Level Descriptors

| | |
|---|---|
| Advanced | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br>• Students can recognize and understand geometric sequencing.<br>• Students can demonstrate ability to use complex problem-solving.<br>• Students can use and apply strategies and procedures to solve algebraic problems.<br>• Students can recognize and understand place value to 100,000 and beyond.<br>• Students can use and apply strategies to solve algebraic problems.<br>• Students can use formulas to measure two and three dimensional basic shapes.<br>• Students can express the probability of an event using correct vocabulary.<br>• Students multiply an divide using multiple digits (e.g. 2 digits by 2 digits)<br>• Students can create and extend patterns to real world situations. |
| Proficient | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br>• Students can add and subtract with several re-grouping steps.<br>• Students can recognize place value to one hundred thousandth place.<br>• Students can multiply three digits by one digit.<br>• Students can choose appropriate tools and techniques in applying measurement skills to everyday situations.<br>• Students can use basic vocabulary of chance (likely, more likely).<br>• Students can create a variety of patterns.<br>• Students can complete basic addition, subtractions and multiplication facts automatically.<br>• Students can use data, probability, and statistics to make consistent predictions and reasonable decisions.<br>• Students can recognize and identify geometric vocabulary (i.e. lines, points, angles, perpendiculars, etc.) |
| Nearing Proficiency | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br>• Students can select and use appropriate problem-solving strategies for simple problems.<br>• Students can present solutions with limited organization and support information.<br>• Students have a limited communication of math concepts.<br>• Students can use whole numbers to estimate and compute with limited regrouping.<br>• Students can identify place value to thousands.<br>• Students can apply basic algebraic understanding of concrete and symbolic representations. |

| | |
|---|---|
| | • Students can describe, model, and identify some geometric shapes.<br>• Students can determine measurable attributes of objects and usually select the appropriate tools for measurement.<br>• Students can interpret, organize, and recognize simple data.<br>• Students can identify and extend simple patterns.<br>• Students can recognize common fractions.<br>• Students can tell time to quarter hour. |
| Novice | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br>• Students can use only a limited number of problem-solving strategies for simple problems.<br>• Students' solutions lack organization and supporting information.<br>• Students have difficulty in communicating math concepts.<br>• Students can use manipulatives to estimate and compute whole numbers.<br>• Students can identify place value to hundreds.<br>• Students can demonstrate emerging algebraic understanding of concrete and symbolic representations.<br>• Students can describe, model, or identify a limited number of geometric shapes.<br>• Students can determine some measurable attributes of objects, but often do not select appropriate tools for measurement.<br>• Students often make inaccurate decisions based on data.<br>• Students can recognize and represent simple geometric and numerical patterns and describe the relationship (rule) in that pattern.<br>• Students can identify basic fractions pictorially or manipulatives.<br>• Students can tell time to even hour, half hour.<br>• Students can identify and count coins to $1.00. |

**Grade 4 Reading Performance Level Descriptors**

| | |
|---|---|
| Advanced | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br>• Uses advanced vocabulary<br>• Reads/interprets maps and charts<br>• Interprets and compares information<br>• Synthesizes information<br>• Critically evaluates text |
| Proficient | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br>• Understands personification, figurative language, and literary devices<br>• Distinguishes fact from opinion<br>• Makes inferences<br>• Identifies author's purpose<br>• Analyzes and organizes information<br>• Interprets and responds to text |

| | |
|---|---|
| | • Compares and contrasts<br>• Rereads to find information<br>• Understands main idea and supports with details<br>• Uses prior knowledge to make meaning of text<br>• Uses substantial reading vocabulary<br>• Reads a variety of materials<br>• Reads maps and diagrams<br>• Uses resource materials<br>• Justifies predictions |
| Nearing Proficiency | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br>• Can recall what was read<br>• Has some ability to summarize<br>• Uses basic vocabulary<br>• Begins to use context to gain understanding<br>• Demonstrates understanding of main idea<br>• Understands word parts (prefixes)<br>• Identifies supporting details<br>• Makes predictions and draws conclusions<br>• Follows directions |
| Novice | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br>• Has prior knowledge that helps student answer questions<br>• Can identify details<br>• Can make comparisons<br>• Can identify subheadings<br>• Can make simple inferences |

**Grade 8 Math Performance Level Descriptors**

| | |
|---|---|
| Advanced | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br>• Recognize & extend arithmetic & geometric patterns<br>• Solve multi-step equations<br>• Recognize 2-dimensional representations of 3-dimensional shapes<br>• Solve equations with negative exponents<br>• Knowing & applying Pythagorean Theorem<br>• Simplifying expressions with like terms<br>• Able to filter extraneous information not needed to solve the problem<br>• Solve problem & communicate strategy<br>• Solve inequalities<br>• Draw inferences; construct & evaluate based on data analysis |
| Proficient | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br>• Know order of operations (+, -, *, ÷, ( ), exponents)<br>• Calculate basic operations using all real numbers |

| | |
|---|---|
| | • Solve proportions<br>• Identify & interpret graphs<br>• Convert: fractions ? decimals; decimals ? percentages; percentages ? fractions<br>• Use fractions in real world applications<br>• Understand & manipulate geometric formulas<br>• Understand difference between and be able to calculate mean, median & mode<br>• Solve 2-step equations<br>• Draw visual combinations (using trees, tables or another strategy)<br>• Plot all real numbers on a number line<br>• Given a formula, calculate perimeter, area & volume of geometric shapes<br>• Know and be able to apply definitions of similar & congruent<br>• Solve problems with 2 steps and extend<br>• Recognize simple patterns<br>• Attempts to communicate strategies<br>• Recognize & use inequality symbols<br>• Represent geometric figures on a coordinate plane/grid<br>• Convert measurements within a system |
| Nearing Proficiency | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br>• Given a simple formula, perform variable replacement with a number<br>• Solve 1-step equations with positive numbers<br>• Know order of operations for addition, subtraction, multiplication, division & parentheses<br>• Plot points on a coordinate plane<br>• Be able to perform simple interpretations of basic types of graphs.<br>• Recognize reflections & rotations<br>• Calculate basic operations (addition, subtraction, multiplication, division) with whole numbers<br>• Understand simple probability with independent outcomes (e.g., coin flips)<br>• Plot integers on a number line |
| Novice | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems. |

**Grade 8 Reading Performance Level Descriptors**

| | |
|---|---|
| Advanced | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br>• Application of literary elements<br>• Applies a rich and varied content vocabulary<br>• Abstract comprehension<br>• Emerging analytical thinking<br>• Applies inferential thinking |

| | |
|---|---|
| | • Understands different genres<br>• Interprets figurative language |
| Proficient | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br>• Emerging understanding of literary elements<br>• Emerging content vocabulary<br>• Emerging/basic figurative comprehension<br>• Uses word structures to enhance meaning<br>• Metaphorical thinking<br>• Emerging inference skills<br>• Recognizes different genres<br>• Basic recognition of figurative language |
| Nearing Proficiency | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br>• Limited understanding of literary elements<br>• Limited content vocabulary<br>• Literal comprehension<br>• Understands basic word structures<br>• Makes some connections to prior knowledge<br>• Able to find answers when stated in text<br>• Understands difference between fiction and non-fiction |
| Novice | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br>• Minimal understanding of literary elements<br>• Minimal content vocabulary<br>• Concrete comprehension<br>• Emerging understanding of basic word structures<br>• Makes minimal connections to prior knowledge<br>• Sometimes able to find answers when stated in text<br>• Sometimes understands difference between fiction and non-fiction |

**Grade 10 Math Performance Level Descriptors**

| | |
|---|---|
| Advanced | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br>• Can analyze a problem to identify the real question<br>• Can solve problems never encountered before<br>• Can solve nontraditional presentations of a problem<br>• Can convert between abstract & concrete<br>• Can formulate a decision-making strategy |
| Proficient | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br>• Solid math vocabulary including definitions and properties<br>• Can consistently solve multi-step problems<br>• Can translate & apply language descriptions to a variety of problems |

| | |
|---|---|
| | • Can translate between multiple representations of a problem or concept<br>• Can convert written to symbolic |
| Nearing Proficiency | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br>• Recognizes and understands some concepts at basic level<br>• Can be confused by context in questions<br>• Has difficulty converting decimals/fractions<br>• Can consistently solve single-step problems |
| Novice | Students at this level demonstrate a minimal understanding of subject matter and may not solve simple problems. |

**Grade 10 Reading Performance Level Descriptors**

| | |
|---|---|
| Advanced | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br>• Extends & connects ideas<br>• Describes abstract themes & ideas<br>• Makes complex predictions<br>• Analyzes & evaluates causal relationships<br>• Formulates complex arguments with strong supporting evidence<br>• Flexibly uses a variety of strategies to interpret language, literary characteristics & overall intent<br>• Uses an enriched reading vocabulary<br>• Consistently applies complex thinking skills |
| Proficient | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br>• Makes & revises predictions, explains inferences & analyzes causal relationships<br>• Usually paraphrases accurately<br>• Formulates arguments with supporting evidence<br>• Uses a variety of strategies to interpret language, literary characteristics & overall intent<br>• Uses a substantial reading vocabulary<br>• Applies complex thinking skills<br>• Analyzes the author's uses of literary devices |
| Nearing Proficiency | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br>• Makes predictions, identifies inferences, describes causal relationships<br>• Frequently paraphrases accurately<br>• Formulates arguments with limited supporting evidence<br>• Uses a limited variety of strategies to interpret the language, literary characteristics & overall intent<br>• Uses a limited 10th grade vocabulary<br>• Limited range of reading purposes<br>• Identifies elements of an author's style |

| | |
|---|---|
| | • Occasionally applies complex thinking skills |
| Novice | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br>• Makes simple predictions & inferences<br>• Does not often grasp the meaning of causal relationships<br>• Sometimes accurately paraphrases<br>• Sometimes formulates arguments with limited supporting evidence, provides simple responses<br>• Relies primarily on a few strategies to interpret language, literary characteristics & overall intent<br>• Uses a reading vocabulary below Grade 10<br>• Limited range of reading purposes<br>• Compares & contrasts but infrequently analyzes or applies complex thinking skills |

# Appendix C

# Instructions for Group Facilitators

# GENERAL INSTRUCTIONS FOR CRT STANDARD SETTING GROUP FACILITATORS

# READING AND MATHEMATICS

## Prior to Round 1 Ratings

Introductions:

1. Welcome group, introduce yourself (name, affiliation, a little selected background information).
2. Have each participant introduce him/herself.
3. Ask participants to complete Non-Disclosure Forms. Collect forms.

## Take the Test

***Overview:*** In order to establish an understanding of the MontCAS, Phase 2 CRT test items and for panelists to gain an understanding of the experience of the students who take the test, each participant will take the test for their grade level and content area. Panelists may wish to discuss or take issue with the items in the test. Tell them we will gladly take their feedback to the OPI. However, this is the actual assessment that students took and it is the set of items on which we must set standards.

***Activities:***

1) Introduce CRT and convey/do each of the following:
    a. Tell panelists that they are about to take the actual CRT assessment.
    b. The purpose of the exercise is to help them establish a good understanding of the test items and to gain an understanding of the experience of the students who take the assessment. Let panelists know they do not need to completely answer the constructed-response questions; they can just jot a few notes.
    c. Tell panelists that only common items are scored on the CRT. Let panelists know that they will only be taking the common items. Session 2 contains a few common items but mostly field test items; therefore, Session 2 will be a short session.
2) Give each panelist a test booklet.
3) Tell panelists to try to take on the perspective of a student as they complete the test.
4) When the majority of the panelists have finished, pass out answer key.

## Fill Out Item Map

***Overview:*** The primary purpose of filling out the item map is for panelists to think about and document the knowledge, skills, and abilities students need to answer each question. Panelists should have an understanding of what makes one test item harder or easier than another. The notes panelists take here will be useful in helping them place their bookmarks and in discussions.

***Activities:***

1. Pass out the following materials:
   a. Item map
   b. Ordered item book

2. Provide an overview of the task paraphrasing the following:
   a. The primary purpose of this activity is for panelists to think about what makes one question harder or easier than another. For example, it may be that the concept tested is a difficult concept, or that the concept isn't difficult but that the particular wording of the question makes it a difficult question. Similarly, the concept may be a difficult one, but the wording of the question makes it easier.
   b. Panelists should take notes about their thoughts regarding each question. These will be useful in the rating activities and later discussions.

3. Tell panelists they will work individually at first. After they have completed the item map, they will then discuss it as a group.

4. Review the ordered item book and item map with the panelists. Explain what each is, and point out the correspondence of the ordered items between the two. Explain that the items are ordered from easiest to hardest, and that 4-pt CRs will appear once for each possible score point. There are two CRs.

5. Write the starting cut points (i.e., between which two ordered items) on the chart paper and post it on a wall visible to all panelists. Ask panelists to place bookmarks in the appropriate places in their ordered item booklet. For each cut point, the panelists will begin the item mapping process approximately five ordered items prior to the starting cut.

6. Each panelist will begin with the starting ordered item and compare it to the next ordered item. What makes the second item harder than the first? Panelists should not agonize over these decisions. It may be that the second item is only slightly harder than the first.

7. Panelists should work their way through the item map, stopping about five ordered items after the ***Novice/Nearing Proficiency*** starting cut.

8. Panelists will then do the same process for the ***Nearing Proficiency/Proficient*** and ***Proficient/Advanced*** cuts; for these cuts, they will start approximately five ordered items before each cut and end approximately five ordered items after each cut.

9. Note that panelists may feel that they need to expand the range of items they consider in one direction or the other. Five ordered items before and after the starting cuts is a guideline, but they may consider more items if necessary.

10. Once panelists have completed the item map, they should discuss them as a group. The group does not need to discuss the item maps in detail; the purpose of this step is for the panelists to discuss any particular questions or issues that arise as they are filling in the item map.

11. Based on the group discussion, the panelists should modify their own item map (make additional notes, cross things out, etc…)

## Discuss Performance Level Descriptors and Describe Characteristics of the "Borderline" Student

***Overview:*** In order to establish an understanding of the expected performance of borderline students on the test, panelists must have a clear understanding of:

1) The definition of the four achievement levels, and

2) Characteristics of students who are "just able enough" to be classified into each achievement level. These students will be referred to as borderline students, since they are right on the border between achievement levels.

The purpose of this activity is for the panelists to obtain an understanding of the Performance Level Descriptors with an emphasis on characteristics that describe students at the borderline -- both what these students can and cannot do.

This activity is critical since the ratings panelists will be making in Rounds 1 and 2 will be based on these understandings.

***Activities:***

1) Introduce the task. In this activity they will:
   a. Individually review the Performance Level Descriptors;
   b. discuss the Descriptors as a group; and
   c. generate bulleted lists of borderline ***Nearing Proficiency***, ***Proficient*** **and** ***Advanced*** students on chart paper.

2) Pass out the Performance Level Descriptors and have panelists individually review them. Panelists can make notes if they like.

3) After individually reviewing the Descriptors, have panelists discuss each one as a group, starting with *Nearing Proficiency,* and provide clarification. The goal here is for the panelists to have a collegial discussion in which to bring up/clarify any issues or questions, and to come to a common understanding of what it means to be in each achievement level. It is not unusual for panelists to disagree with the Descriptors they will see; almost certainly there will be some panelists who will want to change them.

However, the task at hand is for panelists to have a common understanding of what knowledge, skills, and abilities (KSAs) are described by each Performance Level Descriptor. Panelists will have an opportunity to provide feedback and suggestions for edits to the Descriptors after the standard setting activities are completed.

4) Once panelists have a solid understanding of the Performance Level Descriptors, have them focus their discussion on the knowledge, skills, and abilities of students who are in the *Nearing Proficiency* category, but just barely. The focus should be on those characteristics and KSAs that best describe the lowest level of performance necessary to warrant a *Nearing Proficiency* classification.

5) After discussing *Nearing Proficiency*, have the panelists discuss characteristics of the borderline *Proficient* student and then characteristics of the borderline *Advanced* student. Panelists should be made aware of the importance of the *Proficient* cut.

6) Using chart paper, generate a bulleted list of characteristics for each of the levels based on the discussion. Post these on the wall of the room.

## Round 1

***Overview of Round 1:*** The primary purpose of Round 1 is to ask the panelists to evaluate and, if necessary, revise the starting cut points. Panelists will work individually at first, and then as a group. Beginning with the starting cut between *Novice* and *Nearing Proficiency*, panelists will evaluate each item, starting approximately five ordered items before the starting cut and ending approximately five ordered items after the starting cut. (Note, again, that panelists may feel that they need to expand the range of items they consider. Five ordered items after the starting cut is a guideline, but they may consider more items if necessary.) The panelists will gauge the level of difficulty of each of the items for those students who barely meet the definition of *Nearing Proficiency*. The task that panelists are asked to do is to estimate whether a borderline *Nearing Proficiency* student would answer each question correctly. More specifically panelists should answer:

- Would *at least* 2 out of 3 students performing at the borderline answer the question correctly?

In the case of constructed-response questions, panelists should ask:

- Would *at least* 2 out of 3 students performing at the borderline get this score point *or higher*?

After the panelists have completed the individual review and the group discussion of the *N/NP* starting cut, the process is then repeated for the five or so items above and below the starting *Nearing Proficiency*/*Proficient* cut and the starting *Proficient*/*Advanced* cut.

***Activities:***

1. Panelists should have their ordered item books, item maps, and the Performance Level Descriptors. Pass out one rating form to each panelist.

2. Have panelists write round number 1 and their ID number on the rating form. The ID number is on the back of their name tags.

3. Provide an overview of Round 1, covering each of the following:
    a. Remind panelists of where the starting cuts fall in the ordered item book and that they will be starting the rating process approximately five ordered items before the starting cut for each cut point.

    b. The primary purpose of this activity is for the panelists to discuss the initial placement of each of the bookmarks and discuss whether they have been placed appropriately, or whether they feel they should be moved. Remind panelists that they should be thinking about two-thirds of the borderline students.

    c. The panelists will work individually at first, reviewing each of the ordered items around the starting cut for *Novice* vs. *Nearing Proficiency*, and making a preliminary determination about where the bookmark should be placed. Specifically, the panelists should ask themselves whether students whose performance is barely *Nearing Proficiency* have at least a two-thirds chance of correctly answering each item. Each panelist should place his/her *Novice/Nearing Proficiency* bookmark where they believe the answer of 'yes' turns to 'no.'

    d. Once the panelists have finished making their initial individual determination, they will then discuss the starting *Novice* vs. *Nearing Proficiency* cut point as a group. Panelists should be encouraged both to share their rationale for where they placed their bookmark, and to listen to the points made by their colleagues.

    e. After the discussion is complete for the *Novice/Nearing Proficiency* cut point, the panelists will repeat steps (*c)* and (*d)* for the *Nearing Profeciency/Proficient* cut point, and, finally, for the *Proficient/Advanced* cut point.

    f. Each panelist needs to base his/her judgments on his/her experience with the content, understanding of students, and the definitions of the borderline students generated previously.

    g. If panelists are struggling with placing a particular bookmark they should use their best judgment based on personal knowledge and experience.

4. Tell panelists that they will be discussing each cut point with the other panelists, but that they will be placing the bookmarks individually. **It is not necessary for the panelists to come to consensus about whether and how the cut points should be revised**.

5. Go over the rating form with panelists.
    a. Lead panelists through a step-by-step demonstration of how to fill in the rating form.
    b. Answer questions the panelists may have about the work in Round 1.
    c. Once everyone understands what they are to do in Round 1, tell them to begin.

6. Using the ordered item book, and working individually at first, the panelists begin approximately five ordered items before the starting *N/NP* cut. Once they have completed their initial, individual placement of the *N/NP* bookmark, they discuss that placement as a group.

7. Once the group discussions are completed for the first bookmark, they proceed to the *Nearing Proficiency*/*Proficient* cut, beginning approximately five ordered items prior to the starting cut. Again, they first make an individual bookmark placement, then discuss it as a group.

8. Once they have placed the second bookmark, they proceed to the *Proficient*/*Advanced* cut, again beginning approximately five ordered items prior to the starting cut, and complete the individual placement and group discussion.

9. As panelists complete the task, ask them to carefully inspect their rating forms to ensure they are filled out properly.
    a. **The round and ID number must be filled in**.
    b. **The item numbers identifying each cut score must be adjacent**.
    c. Check each panelist's rating form before you allow them to leave for a short break.
    d. When all the rating forms have been collected, the group will take a break. Immediately bring the rating forms to the R&A work room for tabulation.

## Tabulation of Round 1 Results

R&A will tabulate the results of Round 1 as quickly as possible after receipt of the rating forms; however, this data will not be shared with the group.

## Recommendations for Enhancements or Modifications to Performance Level Descriptors

Ask panelists to review the Performance Level Descriptors and the items that fall into each level according to the final recommended cut points. Working as a group, the panelists will then compile a list of recommended modifications or enhancements to the Performance Level Descriptors to reflect the specific KSAs required to successfully complete the items in each achievement level. Panelists may also recommend edits that reflect skills that are measured on the test but don't appear in the KSAs, or vice versa. Make sure panelists know that these are recommendations and that they may not all be implemented.

## Complete Evaluation Form

Upon completion of Modifications to Performance Level Descriptors, have panelists fill out the evaluation form. Emphasize that their honest feedback is important.

# Appendix D

## Sample Item Map

| Item | What does this item measure? | Why is this item more difficult than the preceding item? |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |
| 11 | | |
| 12 | | |
| 13 | | |
| 14 | | |
| 15 | | |
| 16 | | |
| 17 | | |
| 18 | | |
| 19 | | |
| 20 | | |
| 21 | | |
| 22 | | |
| 23 | | |
| 24 | | |
| 25 | | |

# Appendix E

## Sample Rating Form

**Grade** _____

**Round** _____

**ID** _____

**Table** _____

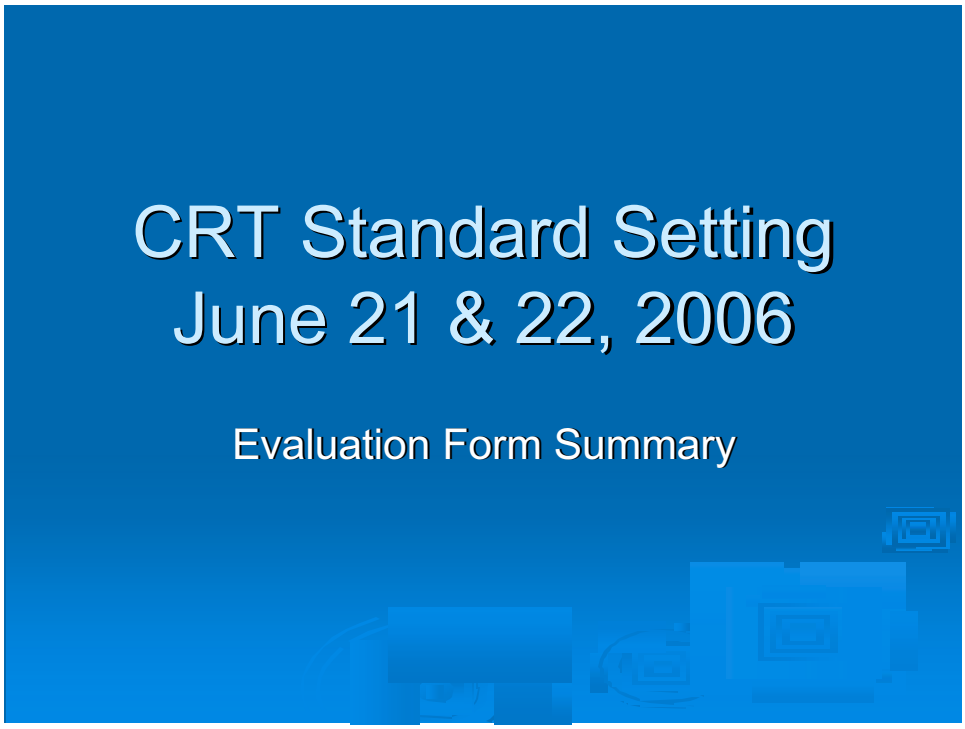
| **Novice** Ordered Item Numbers | | **Nearing Proficiency** Ordered Item Numbers | | **Proficient** Ordered Item Numbers | | **Advanced** Ordered Item Numbers | |
|---|---|---|---|---|---|---|---|
| First | Last | First | Last | First | Last | First | Last |
| 1 | | | | | | | 66 |

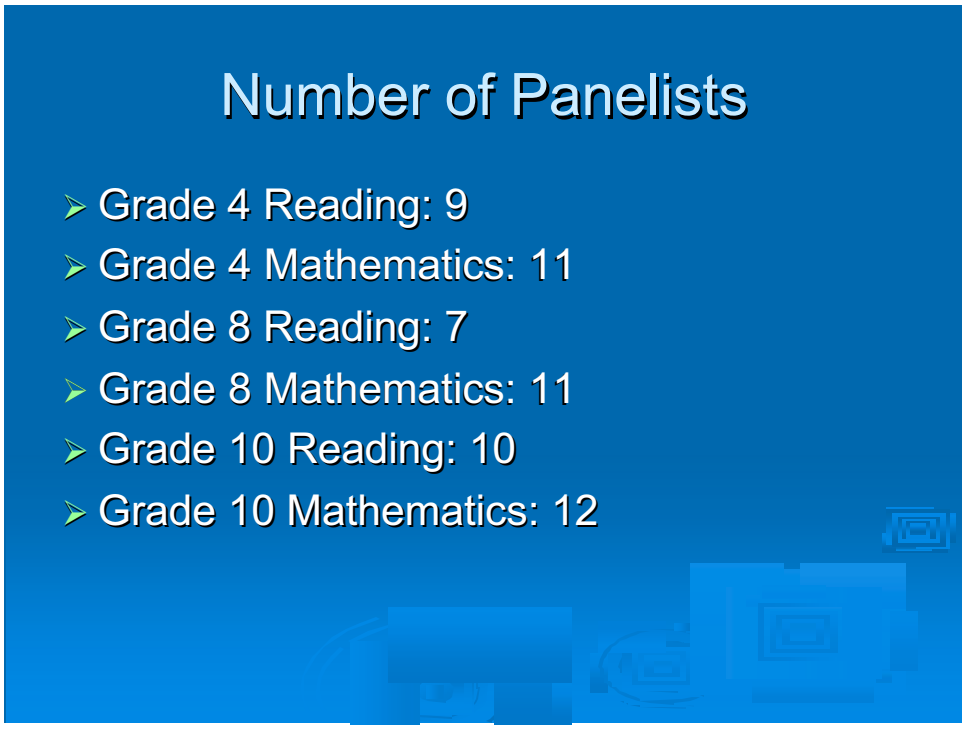
**<u>Directions</u>: Please enter the range of ordered item numbers that fall into each performance level category according to where you placed your cutpoints.**

**<u>Note</u>: The ranges must be adjacent to each other. For example: Novice 1-14, Nearing Proficiency 15-28, Proficient 29-42, Advanced 43-66.**

# Appendix F

## Results of Evaluation



# Question 1

1. Please mark the subject for which you set standards.

Reading or Mathematics

2. What was your comfort level with the standard setting process
2. What was your comfort level with the standard setting process
at the beginning of the process?

1 = Extremely Uncomfortable
3 = Somewhat Comfortable 5 = Extremely Comfortable

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | 1 | 5 | 2 | |
| Gr 4 M | | 1 | 6 | 3 | 1 |
| Gr 8 R | | 2 | 4 | | 1 |
| Gr 8 M | 2 | 4 | 4 | 1 | |
| Gr 10 R | | 1 | 6 | 2 | 1 |
| Gr 10 M | | 4 | 2 | 4 | 2 |

3. What was your comfort level with the standard setting process at the end of the process?

1 = Extremely Uncomfortable 3 = Somewhat Comfortable
5 = Extremely Comfortable

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | 2 | 6 |
| Gr 4 M | | | | 5 | 6 |
| Gr 8 R | | 1 | | 6 | |
| Gr 8 M | | | 1 | 5 | 5 |
| Gr 10 R | | | | 1 | 9 |
| Gr 10 M | | | | 3 | 9 |

4. To what extent did the training prepare you to complete the task of standard setting?

1 = Not at all 3 = Somewhat well
5 = Extremely well

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | 5 | 3 |
| Gr 4 M | | | 4 | 6 | 1 |
| Gr 8 R | | | 2 | 3 | 2 |
| Gr 8 M | | 1 | 2 | 4 | 4 |
| Gr 10 R | | | 1 | 6 | 3 |
| Gr 10 M | | | | 10 | 1 |

5. How clear were the performance level definitions?

1 = Not at all clear
3 = Somewhat clear
5 = Very clear

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | 3 | 4 | | 1 |
| Gr 4 M | 1 | 1 | 3 | 4 | 2 |
| Gr 8 R | | | 2 | 1 | 4 |
| Gr 8 M | | 2 | 2 | 4 | 3 |
| Gr 10 R | | | 1 | 7 | 2 |
| Gr 10 M | | | 1 | 6 | 5 |

6. How clear was the bookmarking task?

1 = Not at all clear
3 = Somewhat clear
5 = Very clear

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | 2 | 6 |
| Gr 4 M | | | 1 | 6 | 4 |
| Gr 8 R | | 1 | 1 | 2 | 3 |
| Gr 8 M | | | 2 | 4 | 5 |
| Gr 10 R | | | | 5 | 5 |
| Gr 10 M | | | | 7 | 5 |

7. To what extent was the length of this meeting appropriate for the task of setting performance standards?

1 = Too little time 3 = About right
5 = Too much time

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | 7 | 1 | |
| Gr 4 M | | | 6 | 5 | |
| Gr 8 R | | | | 6 | 1 |
| Gr 8 M | | | 9 | 2 | |
| Gr 10 R | | | 9 | 1 | |
| Gr 10 M | | 1 | 9 | 1 | 1 |

8. What was your level of confidence in the bookmarks you placed?

1 = Very low
5 = Very high

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | 6 | 2 |
| Gr 4 M | | | | 2 | 9 |
| Gr 8 R | | | | 6 | 1 |
| Gr 8 M | | | 3 | 3 | 5 |
| Gr 10 R | | | | 4 | 6 |
| Gr 10 M | | | 1 | 7 | 4 |

9. Do you believe the standards set by the panel are correctly placed on the exam score scale?

1 = No 3 = Unsure
5 = Yes

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | 6 | 2 |
| Gr 4 M | | | 2 | 5 | 4 |
| Gr 8 R | | | | 4 | 3 |
| Gr 8 M | | | 1 | 4 | 6 |
| Gr 10 R | | | 1 | 5 | 4 |
| Gr 10 M | | | | 7 | 5 |

10 A. How influential were the performance level descriptors in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced
5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | 1 | 5 | 2 |
| Gr 4 M | | 2 | 5 | 4 | |
| Gr 8 R | | | | 1 | 6 |
| Gr 8 M | | | 3 | 6 | 2 |
| Gr 10 R | | | 1 | 3 | 6 |
| Gr 10 M | | | 1 | 4 | 6 |

10 B. How influential were the assessment items in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced
5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | 2 | 6 |
| Gr 4 M | | 2 | 2 | 4 | 3 |
| Gr 8 R | | 1 | 2 | 3 | 1 |
| Gr 8 M | | | | 4 | 7 |
| Gr 10 R | | | | 2 | 8 |
| Gr 10 M | | | | 5 | 7 |

10 C. How influential were the other panelists' comments in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced
5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | 3 | 4 | 1 |
| Gr 4 M | | 1 | 2 | 7 | 1 |
| Gr 8 R | | 1 | | 6 | |
| Gr 8 M | | 1 | 3 | 6 | |
| Gr 10 R | | 2 | 3 | 4 | 1 |
| Gr 10 M | | | 3 | 7 | 2 |

10 D. How influential was your professional experience in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced
5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | 2 | 6 |
| Gr 4 M | | | | 6 | 5 |
| Gr 8 R | | | | 2 | 5 |
| Gr 8 M | | | 1 | 3 | 7 |
| Gr 10 R | | | 1 | 2 | 7 |
| Gr 10 M | | | | 3 | 9 |

10 E. How influential was the impact data in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced
5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | 1 | 2 | 4 | 1 |
| Gr 4 M | | 2 | 3 | 5 | 1 |
| Gr 8 R | | | 2 | 4 | 1 |
| Gr 8 M | | | 7 | 4 | |
| Gr 10 R | | | 2 | 3 | 5 |
| Gr 10 M | | | 3 | 7 | 2 |

10 F. How influential was the political climate in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced
5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | 5 | 3 | | | |
| Gr 4 M | 3 | 1 | 5 | 2 | |
| Gr 8 R | 4 | 1 | 2 | | |
| Gr 8 M | 3 | 1 | 7 | | |
| Gr 10 R | 3 | 3 | 3 | 1 | |
| Gr 10 M | | 5 | 5 | 1 | 1 |

11. Did you find this standard setting session to be professionally rewarding?

1 = No, not at all 3 = Somewhat
5 = Yes, extremely

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | | 8 |
| Gr 4 M | | 1 | 1 | 1 | 8 |
| Gr 8 R | | | | 3 | 4 |
| Gr 8 M | | | | 5 | 6 |
| Gr 10 R | | | | 1 | 9 |
| Gr 10 M | | | | 3 | 9 |

12. How would you characterize the organization of the standard setting session activities?

1 = Disorganized 3 = Somewhat organized
5 = Extremely organized

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 4 R | | | | 3 | 5 |
| Gr 4 M | | | | 7 | 4 |
| Gr 8 R | | | 1 | 2 | 4 |
| Gr 8 M | | | | 3 | 8 |
| Gr 10 R | | | | 2 | 8 |
| Gr 10 M | | | | 1 | 11 |

**Grade 4 Reading**

- Better describe performance level descriptor. Don't flow or gel very well. Not clear.
  Involve more people in each level for both math and reading. 8 people seems very small to impact statewide testing.
- I learned a ton about Montcass and Standard Setting. Our group and facilitator worked extremely well with each other. I feel we did an excellent job as a group and our discussions were valuable and thought provoking – thanks so much!! I would like to see the results of round 2, but understand why.
- I would like to see the results of Round 2 – would that discussion following Round 1 make a difference?
  I felt the workshop was very beneficial as a classroom instructor. Being a part of the "process" is so valuable.
- This was a great process to be a part of. It is wonderful to know teachers are an important aspect to this process.
- A follow up letter in June to confirm hotel registration would've been helpful. (Fortunately I called the hotel to check and was told they needed a credit card to hold my room past 4pm.)
  Perhaps a question/answer session (with questions submitted in writing beforehand) on the whole AYP/NCLB/MONTCAS/ITBS process would be helpful in dispelling rumors, etc.
- I appreciate how well our facilitator "facilitated." She was very nonjudgmental and neutral. She let us talk, diverse, stray off, (somewhat), while still keeping us task-oriented and on schedule. It was an educational and rewarding experience.
- It was very beneficial to me as an educator. I thought the process was overall quite effective and organized. I think it is very important to involve working, active educators in this whole process!

- This is a good process. It involves classroom professionals in a meaningful way. It was done very professionally and should be commended.

**Grade 8 Reading**

- Next year I will definitely feel more comfortable. Facilitator was very helpful without expressing own feelings or influencing.
  Group discussion was great, although at times avid!
  In all, a good experience.
- Need to have experts of the content area to visit each group several times during the conference. We saw ours at the very end – would have helped tremendously.
  Also – it was a positive note to know that MT will listen to us a little bit on all of our work.
  I had a great learning experience – Thank you!
- This was my first time and felt the process was organized. The climate was very comfortable that everyone felt comfortable to share thoughts – both positive and negative. I am very glad I did this and believe I gained professionally.
- Best of three – provide data on the cuts for second cut mark placements for review and reflection. Measured Progress people were very helpful.
- It is important to have practicing teachers do this process. Others are inclined to bring non-specific attitudes to the process.
  The facilitator (Julie-Ann) was excellent!
- Thank you for allowing me to participate in this experience. At the risk of sounding not open to others' ideas and opinions, I was terribly annoyed with the group I was in. Our facilitator was excellent – she was patient in letting us discuss and come to our own conclusions. The group members (not all of them) did not stay on track, were poor listeners, talked while others were talking, and were extremely arrogant in their opinions. What is the old adage? The worst students are teachers. I will work on being more patient if I am to be included in other standard setting procedures.

**Grade 10 Mathematics**

- Ginger is awesome. She is professional and fair. She was great at keeping everyone on task in a very tactful manner. Great experience.
- It would be immensely helpful (but impossible to do) if every teacher of mathematics could do this task. It has helped me realize the content lacking in our state.
- A little concerned with the politics (which I know is an issue).
  Thanks for the experience.
  I loved being able to visit with other 10th grade math teachers!!
- I think we need to be more clear about what each question is <u>supposed</u> to be assessing.
  Climate control in rooms is important.
- I learned a lot through the whole process.
- Well organized. Thanks.
- The state standards/Progress Through Standards should be handed out to panelists to allow us to focus on what "should" happen. We tend to focus on our district's curriculum and students. Also, Montana NEEDS to revisit its standards to reflect "grade 10" and not "upon graduation." OPI needs to set this as a priority.
- This was a professionally rewarding experience.

- Ginger did an excellent job of keeping us on task. I think the first day was a little much – not a lot of time to digest some of the information. Felt rushed thru the bookmarking in Round 1. Liked that there was no rush and lots of time to do the bookmarking for round 2.
- I appreciated the wide variety of educators in the panel. I felt much more confident in the results knowing that all perspectives were represented.

**Grade 10 Reading**

- Very good work on all parts in this process. The set bookmarks streamlined the process and kept it in the right direction. Outstanding group which was very diverse but came to good accurate consensus and could agree to disagree, much to the thanks of a good moderator.
- The Performance Level Descriptors very helpful when placing bookmarks. I thought the cross section of individuals kept refocusing all to student skills/abilities at each level.
- Could OPI provide workshop for our English and math teachers to explain our Montana standards, Grade 10 Reading Performance Level Descriptors, the test, how bookmarks determined (what they need to have on their curriculum)
- Performance Predictors need work; Impact scores important; Good to have to psychometricians there; Borderline discussions useful;
  ?Student profiles – ? were they realistic or so leveled out?
  Some people saw impact charts as not clear (options – go up to %100 – bar charts? different visuals
- This was an interesting experience that I'm glad I had the opportunity to contribute to.
- It would have been helpful, at the beginning if a demonstration of how the item map should have been used i.e. looking at question #1. What does this item measure? (give examples) It seemed that after the break things went faster and I was more confident as to what I was doing. I felt that at the beginning I was confused and spent too much time analyzing each question.
- CRT performance descriptors need to be included on parent sheet – the salmon sheet would be more helpful and the white OPI sheet definitions are not necessary.
- Thank you for an informative session. I wish more info could be brought back to our districts – I understand the non-disclosure, but it is so helpful to share our own deepened understanding of the CRT with others.
- I found this to be very productive. I wish several individuals at our school would pay more attention to the testing environment and stress the importance of consistency for all students.
- I believe that more information would have been helpful. For example, knowing the %, state-wide, of students who scored proficient or advanced on past tests would have been helpful. I can understand the concern about not influencing our decision with data, however there was a feeling that data was being held back or hidden. Overall, I found the experience to be professionally rewarding – well worth my time. Thank you.

**Grade 4 Mathematics**

- There is a certain element of futility. Someone has already done it; someone (decision making) will change it. Just how important is our work and input? Is this expensive token teacher input? I hope not, but it seems that way.

- Donna facilitated quite well! She kept us on track and encouraged us to do the right thing! I really enjoyed this experience and learned a lot more about the CRT and teaching!
  What a great experience! Thank you
- I believe this whole process is very valuable and I would recommend all teachers be a part of this at some time in their career. I also find taking the "test" very useful. Thanks!
- Dona was well-organized and kept our group on task in order to use time wisely. The group I worked with were thoughtful in recommending ideas in the standard setting.
- I enjoyed this time and found it to be rewarding. This process has expanded my knowledge and understanding of CRT Standard. This type of session would be beneficial to all teachers/educators in Montana.
- I enjoyed being able to take part in this process. I understand the testing process better now and how some items were selected. I was able to reflect on what a question was asking and how they vary in difficulty. Our facilitator was very accommodating and knowledgeable.
- This is a very interesting process. I think every educator should participate at some time in their career.
- It would be helpful for more teachers to learn the process and understand how the benchmarks are set. I can see benefits from having more clearly defined Nearly Proficient, Proficient, and Advanced students.
  The state giving local control of how curriculum is taught showed up and is supported by this process. The pacing of individual teaching in schools across the state was varied, but overall students were exposed to all topics tested.
  Our facilitator was fantastic!
- This helped me understand the process of how items are evaluated for use on the state tests. It was interesting. I would have liked more information before coming to Helena about what we would be doing here. I had no idea about what I would be doing. It was fun to hear how districts from around the state handle the MONT Cas tests. Our facilitator made the sessions more relaxed and fun.
- I would have liked to have some materials to look over prior to arriving at this sessions. As a visual learner having something to view prior to or during the sessions would help. Otherwise, I enjoyed the experience.
- I would suggest the facilitator sit at the center of the table or walk around. I felt that only half the group was fully engaged in the conversation due to the proximity of the facilitator. The accommodation (meals, lodging) were outstanding – Thank you!

**Grade 8 Mathematics**

- Better explanation of beginning tasks.
- Great overall process once I experienced it. Individual commitment to bookmarks (cut points) is needed to allow (require) individual buy in. Group discussion following was great.
  Might be good for participants to bring a text to check grade level content when questions arise.
  Dan was good to get answers to questions we posed that he was unsure of.
- I was in a group that had some <u>very</u> strong personalities who didn't really allow for any disagreement. We discussed a lot but certain members did not allow for any other suggestions if they didn't jive with their thoughts. I learned a lot about the process, but didn't contribute much due to I like my head where it is.

- It was wonderful to be encouraged to give input. I was unsure how the "cut" scores affected to overall outcome, but was happy to give my input to such an important issue.
- Facilitator (Dan) clarified at the beginning that no one person should not dominate the conversation. This was helpful for me to listen more – talk less. It helped that everyone knew they were a part of the conversation.
- I wish I had known to bring my standards and benchmarks for reference.
  This opportunity has certainly been enriching! Thank you.
- Overall, things were great. At the beginning we were cautioned to avoid NCLB, AYP, political, discussions. An opportunity to discuss these issues (other than at meals) would be beneficial. Not only to discuss the issue, but to ask relevant questions of the experts.
- As a first time participant, I needed a more explicit orientation to the process. This could be achieved through a written clear explanation sent with the application or agenda form mailed prior to the workshop.
  This process will make a positive impact on my teaching. Thank you for involving educators in standard setting.
- A chart explaining the steps in the process would help people to remain task oriented.

# CRT Standard Setting
# July 26 & 27, 2006

Evaluation Form Summary

## Number of Panelists

- Grade 3/5 Reading: 11
- Grade 3/5 Mathematics: 13
- Grade 6/7 Reading: 10
- Grade 6/7 Mathematics: 15

## Question 1

1. Please mark the subject for which you set standards.

Reading or Mathematics

2. What was your comfort level with the standard setting process at the beginning of the process?

1 = Extremely Uncomfortable
3 = Somewhat Comfortable 5 = Extremely Comfortable

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | 1 | 1 | 7 | 2 | |
| Gr 3/5 Mathematics | | 5 | 5 | 1 | 2 |
| Gr 6/7 Reading | 1 | | 4 | 2 | |
| Gr 6/7 Mathematics | | 3 | 6 | 4 | 2 |

3. What was your comfort level with the standard setting process at the end of the process?

1 = Extremely Uncomfortable 3 = Somewhat Comfortable
5 = Extremely Comfortable

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | | 6 | 5 |
| Gr 3/5 Mathematics | | | | 8 | 5 |
| Gr 6/7 Reading | | | | 4 | 3 |
| Gr 6/7 Mathematics | | 1 | 3 | 7 | 4 |

4. To what extent did the training prepare you to complete the task of standard setting?

1 = Not at all 3 = Somewhat well
5 = Extremely well

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | 1 | 8 | 2 |
| Gr 3/5 Mathematics | | 1 | 4 | 4 | 4 |
| Gr 6/7 Reading | | | 2 | 3 | 2 |
| Gr 6/7 Mathematics | | 2 | 7 | 6 | |

5. How clear were the performance level definitions?

1 = Not at all clear
3 = Somewhat clear
5 = Very clear

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | 1 | 4 | 4 | 2 |
| Gr 3/5 Mathematics | | 1 | 3 | 8 | 1 |
| Gr 6/7 Reading | | 1 | | 5 | 1 |
| Gr 6/7 Mathematics | | 4 | 6 | 5 | |

6. How clear was the bookmarking task?

1 = Not at all clear
3 = Somewhat clear
5 = Very clear

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | | 6 | 5 |
| Gr 3/5 Mathematics | | | | 5 | 8 |
| Gr 6/7 Reading | | | | 3 | 4 |
| Gr 6/7 Mathematics | | | 4 | 6 | 5 |

7. To what extent was the length of this meeting appropriate for the task of setting performance standards?

1 = Too little time 3 = About right
5 = Too much time

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | 5 | 5 | 1 |
| Gr 3/5 Mathematics | | | 11 | 2 | |
| Gr 6/7 Reading | | | 5 | 2 | |
| Gr 6/7 Mathematics | | | 8 | 5 | |

8. What was your level of confidence in the bookmarks you placed?

1 = Very low
5 = Very high

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | 1 | 9 | 1 |
| Gr 3/5 Mathematics | | | 1 | 8 | 4 |
| Gr 6/7 Reading | | | 1 | 4 | 2 |
| Gr 6/7 Mathematics | | | 3 | 8 | 4 |

9. Do you believe the standards set by the panel are correctly placed on the exam score scale?

1 = No 3 = Unsure
5 = Yes

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | | 7 | 4 |
| Gr 3/5 Mathematics | | | 1 | 7 | 5 |
| Gr 6/7 Reading | | | | 4 | 3 |
| Gr 6/7 Mathematics | | 1 | 4 | 4 | 6 |

10 A. How influential were the performance level descriptors in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced
5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | 3 | 3 | 5 |
| Gr 3/5 Mathematics | | 2 | 4 | 7 | |
| Gr 6/7 Reading | | | 3 | 2 | 2 |
| Gr 6/7 Mathematics | | | 3 | 6 | 6 |

10 B. How influential were the assessment items in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced 5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | | 3 | 8 |
| Gr 3/5 Mathematics | | | 1 | 4 | 8 |
| Gr 6/7 Reading | | | | 6 | 1 |
| Gr 6/7 Mathematics | | | 1 | 6 | 8 |

10 C. How influential were the other panelists' comments in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced 5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | 6 | 1 | 4 |
| Gr 3/5 Mathematics | | | 4 | 7 | 2 |
| Gr 6/7 Reading | | 1 | 2 | 4 | |
| Gr 6/7 Mathematics | | | 8 | 5 | 2 |

10 D. How influential was your professional experience in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced 5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | | 5 | 6 |
| Gr 3/5 Mathematics | | | | 3 | 10 |
| Gr 6/7 Reading | | | | 5 | 2 |
| Gr 6/7 Mathematics | | | 3 | 6 | 6 |

10 E. How influential was the impact data in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced 5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | 3 | 8 | |
| Gr 3/5 Mathematics | | 1 | 3 | 8 | 1 |
| Gr 6/7 Reading | | | 2 | 5 | |
| Gr 6/7 Mathematics | | | 2 | 12 | 1 |

10 F. How influential was the political climate in determining the standards you set?

1 = Not at all influenced 3 = Somewhat influenced
5 = Very influenced

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | 7 | 2 | 1 | 1 |
| Gr 3/5 Mathematics | 2 | 4 | 5 | 2 | |
| Gr 6/7 Reading | | 3 | 1 | 1 | 2 |
| Gr 6/7 Mathematics | 5 | | 5 | 5 | |

11. Did you find this standard setting session to be professionally rewarding?

1 = No, not at all 3 = Somewhat
5 = Yes, extremely

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | | 3 | 8 |
| Gr 3/5 Mathematics | | | | 5 | 8 |
| Gr 6/7 Reading | | | | 1 | 6 |
| Gr 6/7 Mathematics | | | 1 | 6 | 8 |

12. How would you characterize the organization of the standard setting session activities?

1 = Disorganized 3 = Somewhat organized
5 = Extremely organized

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gr 3/5 Reading | | | | 2 | 9 |
| Gr 3/5 Mathematics | | | 1 | 3 | 9 |
| Gr 6/7 Reading | | | | 1 | 6 |
| Gr 6/7 Mathematics | | | 3 | 4 | 8 |

**Grade 3/5 Mathematics**

- Ginger did a fantastic job reviewing and explaining the process! This is a great experience.
- Ginger did a great job as our facilitator! This was an interesting learning process.
- Making sure there are teacher above and below the grade level being set would be helpful. Ex. for third grade, make sure there are teachers from 2nd and fourth grade.
  There also seems to be more teachers from the smaller districts and not enough from the larger school districts.
- Ginger did a great job of keeping us on task. Liz and Donna were helpful answering our questions (Abdullah did not come in to answer). All 4 of you were a great help at breaks and lunches my questions were clearly answered. Thank you for your patience and help.
- It was well organized and I felt I went through it more comfortable than I expected.
- I think Ginger did a very good job. She kept the group on task and focused on the process.
  Room temperature was a big problem. Working in a warm environment was very difficult.
- I believe it would be beneficial to split the grade levels when working on projects such as this.
- Our facilitator was well organized with the materials and the flow of the activities. I did feel she tended to focus on one person each day as the "expert" based on the grade level she taught. One participant in particular did more than her fair share of talking.
- I thought this process was outstanding. Every teacher would benefit from this process -- more understanding / better teaching / better over-all (statewide) results.
- The process could be somewhat smoother if expectations were a little clearer about who we should think about, all of Montana, not just their own classes.
- The process was very interesting and clearly facilitated.

- The people from Measured Progress were very informative and good at their job. I did find frustrating that the teachers were using different curriculum objectives because our state standards are so general.
- Process was great – no problem. I do have concerns regarding consistency of language between grade levels.
  Grade 5 test is truly quite difficult for low level students.

**Grade 6/7 Mathematics**

- I felt a little more guidance and possibly some examples of what performance level descriptions should look like in the beginning would have been more helpful. The second day we seemed to figure things out a little better.
- Keep previous drafts out of the conversation. Otherwise, well done, I think.
- Interesting / Shed light on expectations of what each test was designed to test, however the consistency with test questions in relation to grade level became somewhat confusing, therefore leading to frustration. The teachers I worked with were knowledgeable and informative!
- After our group developed guidelines for defining NP, P, and Advanced, the process became crystal clear – 2nd day! The first day out comments were very vague and not done well! Thank you Judy Snow for your comments about the beginning teacher and what they would need to see to bring our students to proficiency!
- It was great to be a part of this process. It is frustrating to take the data, the reality of my students, the MT standards and NCLB and try to make them all jive. It was great to work with educators from around the state.
  We had a hard time getting focused but did get there. It might be easiest for those facilitating math to suggest breaking the subject into the categories and asking what students would need in each of the categories.
- I felt a bit frustrated because we had to deal with the test as it was. I feel the problems we had in book-marking were really a test problem. The test not be indicative of what is taught or what students are able to do.
- More time was needed to understand the data and graphs. I didn't exactly feel rushed, but needed to digest the material a bit longer.
- This process is a lot of hard work.
- We struggled a lot at creating performance level descriptors due to lack of group focus. Once we redirected our approach, we were more successful.
- The part where we set performance indicators was difficult because we weren't allowed to look at the state standards. There is a lack of alignment that is disturbing.
- The actual setting of the performance level descriptors in our small group at novice, nearing proficient, proficient, and advanced should have been explained more clearly – also our OBJECTIVE for setting them should have been clearly explained at the onset – NOT at the end! Our facilitator was not as specific or thorough at the onset – it would have been much easier had he been more thorough to begin with. Thanks!
- This session was great. Having attended the standard setting in June, and being lost, I wasn't sure I wanted to attend this one. I'm glad I did – I learned a lot and this has helped me understand the

testing and standards process much better! Great session leader – kept us on task and kept things moving!

- Set performance level descriptor before you look at the questions.
- Explain the Raw Scores, percentages, etc. at the beginning of the session before meeting in separate rooms. It was somewhat confusing of what was expected.

**Grade 3/5 Reading**

- I found the process very confusing at first (and sometimes during) but the facilitator and other educators were helpful. This was a great group to work with.
- Participation in the standards setting process has helped me to become more aware of how decisions are made with regards to testing. It has helped me to become more aware of what our school needs to be doing.
- In the beginning, it was difficult. By the afternoon of the first day, I was comfortable w/ the process. Our group was very compatible, which helped. Our manager was good about providing guidance, but making us find our own way, which I thought was important.
- The two days flew by. I felt enlightened and left feeling like I had really learned something new!
- It was set up well. We struggled with the performance level descriptors, but that was a pre-existing problem.
- Great group of panelists. Helpful facilitator was prepared and knowledgeable. Pacing was appropriate and task was clear.
- A very educational and rewarding experience that I'd love to go through again!
- Many thanks. The leaders and organizers were very professional and helpful. Excellent leadership qualities.
- I think overall it was a good process and I learned more about the process.
- For our state, the performance level descriptors need to be review/modified and clarified. This first step would enable the standard setting process to go more smoothly and efficiently. We had a good project leader and were lucky our group was on the same page when it came to defining the descriptors. Thanks for the experience.
- I feel honored to be a part of this process and look forward to participating in further sessions. The wide variety of experience levels and types of the committee members contributed to insightful information/sharing.

**Grade 6/7 Reading**

- Michael was extremely patient, knowledgeable, and competent. He kept our group on task, which was not always easy to do.
  I am extremely impressed with what I understood about how Measured Progress works as a company. They gave me a renewed sense of the importance of my profession as an educator.
- These days always are beneficial to me! The group ideas are invaluable… Different personalities and different teaching experiences are so informational

Michael did a great job! He kept us on task… We were all agreeable for the most part. Thank you for the experience and the mints!!

- I felt the facilitators were very good to work with. They allowed discussion, humor and questions but kept us on task. I felt our session was productive! It was even fun and informative!
- I am not sure taking the time to identify standards for the "borderline" students was worth it and helped that much. We have all been teaching long enough that we already understand who these students are in our classes.
- This process is a valuable tool to help discover a more efficient way to determine student growth and achievement. The ability to learn from other educators across the state creates an environment of support that will continue to build student success. Thanks for the opportunity to work in a professional environment!
- Our facilitator, Mike, did a great job at keeping us on task and focused. During the times when we would get bogged down, he would remind us of the importance of the content.

  The only part I would change would be the "borderline" posters. I think it was too much and could have been used during discussion only.

  Thanks for everything.
- This was very helpful to me as a teacher in understanding testing. My involvement helps me understand my student's test results. I enjoyed the working relationship with professionals across the state.

# APPENDIX D: CRT PERFORMANCE LEVEL DESCRIPTORS, SCALED SCORES, AND RAW SCORES

## CRT Performance Level Descriptors (General)

| | |
|---|---|
| **Advanced** | This level denotes superior performance. |
| **Proficient** | This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter. |
| **Nearing Proficiency** | This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark. |
| **Novice** | This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark. |

## Content-Specific Performance Level Descriptors (final versions)

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 3 Reading Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br><br>Using grade level text or above, the student is able to:<br>• Use advanced reading vocabulary<br>• Understand simple figurative language and literary elements<br>• Distinguish fact from opinion<br>• Identify author's purpose<br>• Retell information read<br>• Interpret and respond to text<br>• Compare and contrast<br>• Reread to find information<br>• Understand main idea and support with details<br>• Use prior knowledge to make meaning of text<br>• Read a variety of materials<br>• Use resource materials<br>• Describe reading successes and set reading goals |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br><br>Using grade level text, the student is able to:<br>• Use appropriate reading vocabulary<br>• Understand simple figurative language and literary elements<br>• Distinguish fact from opinion<br>• Identify author's purpose<br>• Retell information read<br>• Interpret and respond to text<br>• Compare and contrast<br>• Reread to find information<br>• Understand main idea and support with details<br>• Use prior knowledge to make meaning of text<br>• Read a variety of materials<br>• Use resource materials<br>• Describe reading successes and set reading goals |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br><br>Using grade level text or near-grade level text, the student is able to:<br>• Sometimes use vocabulary approaching grade level<br>• Make obvious predictions<br>• Identify main idea and one supporting detail<br>• Decode some unknown words<br>• Occasionally understand simple figurative language and literary elements<br>• Work inconsistently when working independently<br>• Sometimes describe reading successes |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems. Using near-grade level text, the student is able to:<br>• Use below grade level vocabulary<br>• Make obvious predictions<br>• Sometimes identify main idea and one supporting detail<br>• Decode some unknown words<br>• Occasionally recognize literary elements in works of literature<br>• Work inconsistently when working independently<br>• Rarely describe reading successes |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 3 Math Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter.<br><br>• Select and use multiple problem solving strategies to solve two-step problems involving the four operations and clearly communicate strategies<br>• Read, identify, and interpret place value of numbers to 9,999 and apply in problem solving situations<br>• Add and subtract numbers with multiple regroupings and estimate addition and subtract problems involving three-digit numbers<br>• Multiply and divide whole numbers with basic facts and interpret remainders in simple division situations<br>• Add and subtract simple fractions with common denominators<br>• Solve algebraic problems involving equations, number patterns, geometric patterns, and change and clearly communicate strategies<br>• Use properties and extensive vocabulary to describe and identify two- and three-dimensional figures<br>• Solve complex geometric problems involving coordinate systems, symmetry, transformations, visual and spatial reasoning and clearly communicate strategies<br>• Select and apply appropriate formulas, units, and tools to use in everyday measurement situations and perform simple conversions<br>• Collect, display, read, and interpret data in complex graphs and judge the probability of simple events |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter.<br><br>• Select and use problem-solving strategies to solve two-step problems involving the four operations and clearly communicate strategies<br>• Read, identify, and interpret place value of numbers to 9,999<br>• Add and subtract numbers with multiple regroupings and estimate addition and subtract problems involving two-digit numbers<br>• Multiply and divide whole numbers with basic facts<br>• Identify and write simple fractions<br>• Solve algebraic problems involving equations, number patterns, geometric patterns, and change<br>• Use properties and limited vocabulary to describe and identify two- and three-dimensional figures<br>• Solve geometric problems involving coordinate systems, symmetry, transformations, and visual and spatial reasoning<br>• Select and apply appropriate units and tools to use in everyday measurement situations and perform simple conversions<br>• Collect, display, read, and interpret data and judge the probability of simple events |

| | |
|---|---|
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter<br><br>• Select and use a problem-solving strategy to solve one-step problems involving the four operations and communicates strategies with limited organization or support information<br>• Read and identify place value of numbers to 9,999<br>• Add and subtract numbers with limited regrouping<br>• Multiply and divide whole numbers using models<br>• Identify and write simple fractions with models<br>• Solve algebraic problems involving simple equations, number patterns, and geometric patterns<br>• Name two- and three-dimensional figures and recognize the results of subdividing and combining shapes<br>• Recognize symmetric figures, transformations, shapes from different perspectives, use simple vocabulary to describe direction and position on a grid, and solve geometric problems involving visual and spatial reasoning<br>• Select and apply appropriate units and tools to use in everyday measurement situations<br>• Display, read, and interpret data |
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter.<br><br>• Solve simple one-step problems involving the four operations with models and communicates strategies without organization or support information<br>• Read and identify place value of numbers using models<br>• Add and subtract numbers involving basic facts<br>• Relate multiplication and division to pictorial models of the operations<br>• Recognize and explore meaning of fractions<br>• Solve problems involving simple number and geometric patterns<br>• Recognize and name basic two- and three- dimensional figures and recognize figures that have the same size and shape<br>Recognize symmetric figures and use simple vocabulary to describe direction and position on a grid<br>• Select appropriate units and tools to use in everyday measurement situations<br>• Display, read, and interpret data in simple graphs |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 4 Reading Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br><br>Using grade level text or above, the student is able to:<br>• Use advanced reading vocabulary<br>• Understand main idea and support with details<br>• Use prior knowledge to make meaning of text<br>• Read a variety of materials<br>• Understand personification, figurative language, and literary devices<br>• Distinguish fact from opinion<br>• Make inferences<br>• Identify author's purpose<br>• Read and interpret maps and charts<br>• Interpret and respond to text<br>• Analyze, organize, and synthesize information<br>• Critically evaluate text<br>• Read maps and diagrams<br>• Use resource materials<br>• Describe reading successes and set reading goals |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br><br>Using grade level text, the student is able to:<br>• Use appropriate reading vocabulary<br>• Understand personification, figurative language, and literary devices<br>• Distinguish fact from opinion<br>• Make inferences<br>• Identify author's purpose<br>• Analyze and organize information<br>• Interpret and respond to text<br>• Compare and contrast<br>• Reread to find information<br>• Understand main idea and support with details<br>• Use prior knowledge to make meaning of text<br>• Read a variety of materials<br>• Read maps and diagrams<br>• Use resource materials<br>• Justify predictions<br>• Describe reading successes and set reading goals |

| | |
|---|---|
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br><br>Using grade level text or near-grade level text, the student is able to:<br>• Sometimes use vocabulary approaching grade level<br>• Recall what was read<br>• Demonstrate some ability to summarize<br>• Begin to use context to gain understanding<br>• Understand main idea<br>• Identify supporting details<br>• Understand word parts (prefixes)<br>• Make predictions and draw conclusions<br>• Follow directions<br>• Sometimes describe reading successes and set reading goals |
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br><br>Using near-grade level text, the student is able to:<br>• Use below grade level vocabulary<br>• Use prior knowledge to answer questions<br>• Sometimes summarize main idea<br>• Identify some details<br>• Make comparisons<br>• Identify subheadings<br>• Make simple inferences<br>• Rarely describe reading successes and set reading goals |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 4 Math Performance Level Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter.<br><br>• Select and use multiple problem solving strategies to solve complex problems involving the four operations and clearly communicate strategies<br>• Read, identify, and interpret place value of numbers to 1 million and apply in problem solving situations<br>• Solve multi-step addition and subtraction problems involving whole numbers and decimals with multiple regroupings<br>• Solve multi-step multiplication problems with multi-digit numbers with multiple regrouping<br>• Divide by two-digit divisor and interpret remainder<br>• Solve problems involving addition and subtraction of simple fractions with common denominators and recognize equivalent fractions<br>• Use and apply strategies and procedures to solve complex multi-step algebraic problems involving equations, number patterns, geometric patterns, and change and clearly describe the relationship<br>• Use properties and extensive vocabulary to describe and identify two- and three-dimensional figures and the relationships among them<br>• Solve complex geometric problems involving points on coordinate grids, symmetry, transformations, visual and spatial reasoning and clearly communicate strategies<br>• Apply tools, procedures, and formulas of measurement to solve complex problems<br>• Collect, organize, display, read, and interpret complex data and use data in problem solving situations and find all possible outcomes of an experiment |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter<br><br>• Select and use problem-solving strategies to solve multi-step problems involving the four operations and clearly communicate strategies<br>• Read, identify, and interpret place value of numbers to 1 million<br>• Solve addition and subtraction problems involving whole numbers and decimals with multiple regroupings<br>• Solve multiplication problems with multi-digit numbers with multiple regrouping<br>• Divide by one-digit divisor and interpret remainder<br>• Add and subtract simple fractions with common denominators<br>• Use and apply strategies and procedures to solve multi-step algebraic problems involving equations, number patterns, geometric patterns, and change and clearly describe the relationship<br>• Use properties and vocabulary to describe and identify two- and three-dimensional figures and the relationships among them<br>• Solve geometric problems involving points on coordinate grids, symmetry, transformations, visual and spatial reasoning and clearly communicate strategies<br>• Apply tools, procedures, and formulas of measurement to solve problems<br>• Collect, organize, display, read, and interpret data and use data in problem solving situations and judge the probability of a simple event as impossible, unlikely, likely, or certain and determine which outcomes are most or least likely |

| | |
|---|---|
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter.<br><br>• Select and use problem-solving strategies to solve two-step problems involving the four operations and communicates strategies with limited organization or support information<br>• Read, identify, and interpret place value of numbers to 100,000<br>• Solve addition and subtraction problems with whole numbers and decimals with limited regrouping<br>• Multiply three-digit numbers by one digit numbers with multiple regrouping<br>• Divide by one-digit divisor<br>• Add and subtract simple fractions with common denominators with models<br>• Use and apply strategies and procedures to solve algebraic problems involving equations, number patterns, geometric patterns, and change<br>• Use properties and limited vocabulary to describe and identify two- and three-dimensional figures<br>• Solve geometric problems involving, symmetry, transformations, visual and spatial reasoning and describe direction and position using the cardinal directions<br>• Select and apply appropriate units, tools, and simple formulas to use in everyday measurement situations<br>• Collect, organize, display, read, and interpret data and judge the probability of a simple event as impossible, unlikely, likely, or certain and determine which outcomes are most or least likely |
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter.<br><br>• Solve two-step problems involving the four operations and communicates strategies without organization or support information<br>• Read and identify place value of numbers using models<br>• Solve addition and subtraction problems with limited regrouping<br>• Multiply numbers with limited or no regrouping<br>• Divide numbers with basic facts<br>• Solve simple problems involving basic fractions<br>• Use and apply strategies and procedures to solve simple algebraic problems involving equations, number patterns, geometric patterns, and change<br>• Recognize and name two- and three-dimensional figures<br>• Solve geometric problems involving symmetry, visual and spatial reasoning, and use simple vocabulary to describe direction and position on a grid<br>• Select and apply appropriate units and tools to use in everyday measurement situations<br>• Collect, display, read, and interpret data and judge the probability of a simple event as impossible, unlikely, likely, or certain and determine which outcomes are most or least likely |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 5 Reading Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br><br>Using grade level text or above, the student is able to:<br>• Use advanced reading vocabulary<br>• Understand literary elements and devices<br>• Distinguish fact from opinion<br>• Identify inferred and stated main ideas<br>• Identify author's purpose<br>• Analyze and organize information<br>• Interpret and respond to text in a variety of ways<br>• Compare and contrast information from variety of sources<br>• Reread to find information<br>• Make connections with prior knowledge<br>• Read a variety of materials<br>• Interpret maps and diagrams<br>• Use resource materials<br>• Justify predictions<br>• Set reading goals and describe reading progress |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br><br>Using grade level text, the student is able to:<br>• Use appropriate reading vocabulary<br>• Understand literary elements and devices<br>• Distinguish fact from opinion<br>• Begin to identify inferred and stated main ideas<br>• Identify author's purpose<br>• Analyze and organize information<br>• Interpret and respond to text<br>• Compare and contrast information from variety of sources<br>• Reread to find information<br>• Make connections with prior knowledge<br>• Read a variety of materials<br>• Read maps and diagrams<br>• Use resource materials<br>• Justify predictions<br>• Set reading goals and describe reading progress |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br><br>Using grade level text or near-grade level text, the student is able to:<br>• Sometimes use vocabulary approaching grade level<br>• Make obvious predictions<br>• Identify main idea and one supporting detail<br>• Decode some unknown words<br>• Sometimes recognize literary elements in works of literature<br>• Sometimes set reading goals and describe reading progress |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br><br>Using near-grade level text, the student is able to:<br>• Use below grade level vocabulary<br>• Use prior knowledge to answer questions<br>• Sometimes summarize main idea<br>• Identify some details<br>• Make comparisons<br>• Make simple inferences<br>• Occasionally recognize literary elements in works of literature<br>• Rarely describe reading successes and set reading goals |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 5 Math Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of the following rigorous subject matter.<br><br>• Employ sophisticated or efficient problem solving strategies to solve multi-step or complex problems involving the four operations and elegantly communicate strategies using sophisticated mathematical language, symbols, and/or visual representation<br>• Employ sophisticated or efficient problem solving strategies to solve multi-step multiplication and division problems involving numbers with money<br>• Employ sophisticated or efficient problem solving strategies to solve multi-step addition and subtraction problems involving numbers with decimals and fractions<br>• Use and apply strategies and procedures to solve single-step algebraic problems<br>• Use properties and vocabulary to describe and identify two- and three-dimensional figures<br>• Locate position on a coordinate plane after a transformation<br>• Convert measures within metric and standard systems<br>• Solve geometric problems using formulas<br>• Collect, organize, display, read, and interpret complex data and use data in problem solving situations<br>• Predict the outcome of events |
| **Proficient** | Students at this level demonstrate a solid understanding of the following challenging subject matter<br>.<br>• Employ appropriate problem-solving strategies to solve single-step problems and some multi-step problems involving the four operations and clearly communicate strategies using appropriate mathematical language, symbols, and/or visual representation<br>• Employ appropriate problem-solving strategies to solve single-step multiplication and division problems involving numbers with money<br>• Employ appropriate problem-solving strategies to solve simple and some complex addition and subtraction problems involving numbers with decimals and fractions<br>• Identify and use rules for numeric and geometric patterns<br>• Use and apply strategies and procedures to solve simple algebraic problems<br>• Use properties and vocabulary to describe and identify two- and three-dimensional figures<br>• Identify transformations<br>• Locate position on a coordinate plane<br>• Solve simple geometric problems using formulas<br>• Organize, display, read, and interpret data and use data in problem solving situations<br>• Find all possible outcomes of an experiment |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of the following subject matter<br><br>• Employ appropriate problem-solving strategies to single-step and/or straight computation problems of the four operations and communicate strategies using minimal mathematical language, symbols, and/or visual representation<br>• Employ appropriate problem-solving strategies to solve simple addition and subtraction problems involving numbers with decimals and fractions<br>• Use and apply some strategies and procedures to solve simple algebraic problems<br>• Use properties and vocabulary to describe and identify two- and some three-dimensional figures<br>• Organize, display, read, and interpret data and use data in simple problem solving situations |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of the following subject matter.<br><br>• Employ appropriate problem-solving strategies to single-step and/or straight computation problems for some of the four operations and communicate strategies using limited mathematical language, symbols, and/or visual representation<br>• Read and identify numbers into the millions<br>• Employ appropriate problem-solving strategies to solve single-step and/or straight computation multiplication problems involving numbers with money<br>• Employ some appropriate problem-solving strategies to solve simple addition and subtraction problems involving numbers with decimals<br>• Identify numeric and geometric patterns<br>• Use and apply some strategies to solve simple algebraic problems<br>• Use some properties and vocabulary to describe and identify some two- and three-dimensional figures<br>• Identify some transformations<br>• Select appropriate units, tools, and techniques in applying measurement skills to everyday situations<br>• Display, read, and interpret data<br>• Find possible outcomes of a simple experiment |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 6 Reading Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br><br>Using grade level text or above, the student is able to:<br>• Use a rich and varied reading and listening vocabulary<br>• Apply a variety of strategies (e.g., decode unknown words, use comprehension strategies) when reading literature and content area material<br>• Consistently apply complex thinking skills – identify a variety of purposes for reading, select appropriate reading material to meet a variety of purposes, compare and contrast information and begin to draw conclusions, make simple connections, make predictions, interpret stated and inferred main and subordinate ideas, identify important supporting details, and respond using a variety of modes to reading material,<br>• Compare and begin to analyze a variety of literary elements/devices<br>• Recognizes how authors use literary devices for various purposes<br>• Set, monitor progress toward, and meet reading goals |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br><br>Using grade level text, the student is able to:<br>• Use a broad vocabulary at grade level<br>• Generally make predictions and connections<br>• Begin to identify inferred and stated main ideas<br>• Identify several purposes for reading and often select appropriate reading material to meet a variety of purposes<br>• Occasionally recognize an author's point of view<br>• Compare and contrast information from variety of sources<br>• Usually apply, articulate, and self-monitor decoding and comprehension strategies when reading literature and content area material<br>• Identify and compare some literary devices<br>• Set, monitor progress toward, and often meet reading goals |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br><br>Using grade level text or near-grade level text, the student is able to:<br>• Sometimes use vocabulary approaching grade level<br>• Actively make obvious and simple predictions and connections between new material and prior knowledge<br>• Sometimes summarize main idea<br>• Usually decode unknown words and apply a few strategies when reading<br>• Identify and compare a few literary elements and devices<br>• Generally apply, articulate, and self-monitor decoding and comprehension strategies; and sometimes evaluates reading progress<br>• Sometimes set and sometimes meet reading goals<br>• Identify several purposes for reading and sometimes select appropriate material<br>• Occasionally recognize an author's point of view and identify a few literary devices |
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br><br>Using near-grade level text, the student is able to:<br>• Use below grade level vocabulary<br>• Sometimes make obvious and simple predictions and connections between new material and prior knowledge<br>• Sometimes summarize main idea<br>• Sometimes decode unknown words and apply a few strategies when reading<br>• Occasionally identify purposes for reading and sometimes select appropriate material<br>• Sometimes identify literary elements and devices<br>• Rarely apply, articulate, or self-monitor decoding and comprehension strategies; rarely evaluate reading progress<br>• Rarely set or meet reading goals |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 6 Math Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced**<br>• | Students at this level demonstrate a comprehensive and in-depth understanding of the following rigorous subject matter<br><br>• Employ sophisticated or efficient problem solving strategies to solve multi-step or complex problems involving the four operations of whole numbers, decimals, and fractions and clearly and elegantly communicate strategies using sophisticated mathematical language, symbols, and/or visual representation<br>• Employ sophisticated or efficient problem-solving strategies to solve problems involving common multiples, percents, negative numbers and ratios<br>• Use and apply strategies and procedures to solve multi-step algebraic problems<br>• Locate position on a coordinate plane after a transformation<br>• Solve complex measurement problems using formulas and conversions within the metric and standard system<br>• Collect, organize, and display data using a variety of graphic representations. Read and interpret complex data representations using statistics<br>• Select and use multiple problem solving strategies to determine all outcomes of simple and compound events and clearly communicate strategies |
| **Proficient** | Students at this level demonstrate a solid understanding of the following challenging subject matte<br><br>• Employ appropriate problem solving strategies to solve single and multi-step problems involving the four operations of whole numbers, decimals, and fractions and clearly communicate strategies using appropriate mathematical language, symbols, and/or visual representation<br>• Employ appropriate problem-solving strategies to solve problems involving common multiples, percents, negative numbers and ratios<br>• Identify and use rules for numeric and geometric patterns<br>• Use and apply strategies and procedures to solve single-step algebraic problems<br>• Solve geometric problems using appropriate properties and formulas<br>• Locate position on a coordinate plane<br>• Solve measurement problems using formulas and conversions within the metric and standard system<br>• Collect, organize, display data using a variety of graphic representations. Read and interpret data representations using statistics<br>• Determine all outcomes of simple events and clearly communicate strategies |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of the following subject matter<br><br>• Employ appropriate problem solving strategies to solve single-step and/or straight computation problems involving the four operations of whole numbers, decimals, and fractions and may communicate strategies using limited mathematical language, symbols, and/or visual representation<br>• Employ appropriate problem-solving strategies to solve problems involving percents, negative numbers and simple ratios<br>• Use and apply strategies and procedures to solve simple algebraic problems<br>• Solve some geometric problems using appropriate properties and formulas<br>• Solve simple measurement problems using formulas and simple conversions within the metric and standard system<br>• Collect, organize, display data using some graphic representations. Read and interpret data representations using some statistics<br>• Determine the outcome of a simple event |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of the following subject matter<br><br>• Employ appropriate problem solving strategies to solve single-step and/or straight computation problems involving the four operations of whole numbers, and simple decimal and fractions fraction problems and may communicate strategies using limited mathematical language, symbols, and/or visual representation<br>• Read and identify numbers into the billions<br>• Employ some appropriate problem-solving strategies to solve simple problems involving percents and negative numbers<br>• Identify numeric and geometric patterns<br>• Use and apply some strategies to solve simple algebraic problems<br>• Solve some simple geometric problems using appropriate properties and formulas<br>• Identify some transformations<br>• Locate position on the first quadrant of a coordinate plane<br>• Solve simple measurement problems sometimes using formulas and simple conversions within the standard system<br>• Collect, organize, display, read, and interpret simple data representations<br>• Judge the probability of a simple event as impossible, very likely, unlikely, or certain |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 7 Reading Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br><br>Using grade level text or above, the student is able to:<br>• Use a substantial reading and listening vocabulary<br>• Apply a variety of strategies (e.g., decode unknown words, use comprehension strategies) when reading literature and content area material<br>• Consistently apply complex thinking skills – define purposes for reading, select appropriate material to meet reading purposes, recognize an author's point of view and purpose, make connections, make predictions, make complex connections between new material and prior knowledge, interpret stated and inferred main and subordinate ideas, identify important supporting details, and respond using a variety of modes to reading material, use information from a variety of print and non-print sources to support an argument<br>• Begin to analyze and evaluate a variety of literary elements and devices<br>• Set, monitor progress toward, and meet reading goals |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br><br>Using grade level text, the student is able to:<br>• Use reading and listening vocabulary appropriate to grade level<br>• Usually define purposes for reading and select appropriate material to meet reading purpose<br>• Sometimes make predictions and connections between new material and prior knowledge<br>• Begin to infer stated and inferred main ideas, identifies important supporting details<br>• Begin to interpret at grade level<br>• Usually recognize author's point of view<br>• Compare, contrast, and integrate information from several print and non-print sources<br>• Decode unknown words and apply several strategies when reading literature and content area material<br>• Compare literary elements and devices<br>• Set, monitor progress towards, and often meet reading goals |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br><br>Using grade level text or near-grade level text, the student is able to:<br>• Use a limited reading and listening vocabulary approaching grade level<br>• Sometimes select appropriate material to meet reading purposes, and sometimes define purposes for reading<br>• Most of time make obvious predictions and simple connections between new material and prior knowledge<br>• Summarize and begin to interpret stated main ideas and important supporting details<br>• Sometimes recognize purpose an author's point of view and purpose<br>• Compare, contrast, and integrate information from print and non-print source<br>• Generally apply, articulate, and self-monitor decoding and a few comprehension strategies when reading literature and content area material<br>• Identify and compare some literary elements and devices<br>• Often set and sometimes meet reading goals, sometimes evaluate reading progress |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br><br>Using near-grade level text, the student is able to:<br>• Use a limited reading and listening vocabulary below grade level<br>• Occasionally select appropriate material to meet reading purposes, and occasionally define purposes for reading<br>• Sometimes make obvious predictions and simple connections between new material and prior knowledge<br>• Sometimes summarize and begin to interpret stated main ideas and important supporting details<br>• Occasionally recognize an author's point of view and purpose<br>• Sometimes compare, contrast, and integrate information from print and non-print source<br>• Rarely apply, articulate, and self-monitor decoding and comprehension strategies when reading literature and content area material<br>• Rarely identify and compare literary elements and devices<br>• Rarely set or sometimes meet reading goals |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 7 Math Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter<br><br>• Identify the least common multiple of a pair of natural numbers<br>• Evaluate a numerical expression with multiple operations and fractions<br>• Solve a multi-step problem involving rate, time and distance<br>• Extend a repeating number sequence more than 3 terms beyond those given<br>• Identify the linear expression that generalizes a given arithmetic sequence<br>• Determine whether a pair of rectangles are congruent or similar<br>• Locate vertices of a rectangle on the coordinate plane<br>• Estimate the volume of a cylinder given its dimensions in nonstandard units<br>• Compare the relative volumes of rectangular prisms<br>• Interpret a box-and-whisker plot<br>• Identify a strategy to collect data that is most representative of a group<br>• Calculate a compound probability from a table of data |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter<br><br>• Use a proportion to solve a problem involving fractions, decimals or percents<br>• Order rational numbers in different forms<br>• Identify the solution to a linear equation<br>• Write an equation with two variables to describe a real-world situation<br>• Identify the image of a polygon after a transformation<br>• Identify a pair of transformations applied to a figure that produces a given image<br>• Estimate equal units in different systems of measure<br>• Calculate the area of a quadrilateral<br>• Use the relationships among the radius, diameter and circumference of a circle to solve a problem |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter<br><br>• Write numbers in words or standard form to trillions<br>• Express ratios as percents<br>• Perform multiple operations on decimal numbers to solve a problem<br>• Identify a linear graph that describes an arithmetic sequence given a rule in words<br>• Identify the coordinates of the *x*-intercept of a line graphed on the coordinate plane<br>• Determine the change in a dependent variable of a single operation function given a change in the independent variable<br>• Identify a single transformation applied to a figure that produces a given image<br>• Apply the Triangle Sum Theorem to find a missing angle measure<br>• Determine the measures of central tendency of a set of data displayed in a line plot<br>• Interpret a circle graph |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter<br><br>• Evaluate a numerical expression involving exponents<br>• Apply a single operation to decimal numbers to solve a problem<br>• Describe in words a rule given a arithmetic sequence<br>• Evaluate a linear algebraic expression for a given value<br>• Associate a net with a sketch of its prism<br>• Identify the side view of a three dimensional figure drawn in perspective<br>• Estimate the lengths of figures in a scale drawing<br>• Create a bar graph for a given set of data<br>• Interpret a bar graph<br>• Make a prediction given the probability of an outcome |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 8 Reading Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br><br>Using grade level text or above, the student is able to:<br>• * Apply a rich and varied content vocabulary<br>• * Consistently apply complex thinking skills – connect ideas, make predictions, explain causal relationships<br>• * Demonstrate emerging analytical thinking<br>• * Apply literary elements<br>• Apply inferential thinking<br>• Understand different genres<br>• Interpret figurative language<br>• Set, monitor progress toward, and meet reading goals |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br><br>Using grade level text, the student is able to:<br>• * Use emerging content vocabulary<br>• * Apply complex thinking skills – connect ideas, make predictions, explain causal relationships, use metaphorical thinking and emerging inference skills<br>• * Emerging understanding of literary elements and emerging/basic figurative comprehension<br>• * Use word structures to enhance meaning<br>• Recognize different genres<br>• Basic recognition of figurative language<br>• Set, monitor progress toward, and meet reading goals |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br><br>Using grade level text or near-grade level text, the student is able to:<br>• * Use limited content vocabulary<br>• * Apply some thinking skills – demonstrate literal comprehension, make some connections to prior knowledge, find answers when stated in text<br>• * Limited understanding of literary elements<br>• * Understand basic word structures<br>• Understand difference between fiction and non-fiction<br>• Often set and sometimes meet reading goals, sometimes evaluate reading progress |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br><br>Using near-grade level text, the student is able to:<br>• * Minimal content vocabulary<br>• * Minimal understanding of literary elements<br>• * Concrete comprehension<br>• * Emerging understanding of basic word structures<br>• Make minimal connections to prior knowledge<br>• Sometimes find answers when stated in text<br>• Sometimes understand difference between fiction and non-fiction<br>• Rarely set or meet reading goals |

* top 4 in each category are <u>most</u> important

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 8 Math Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter<br><br>• Write and compare numbers in scientific notation<br>• Evaluate an algebraic expression involving exponents and more than one variable for given values<br>• Identify and interpret the slope of a linear function from a graph, equation, or written description<br>• Sort triangles by sides and angles<br>• Determine the coordinates of the image of a vertex of a polygon after a rotation<br>• Apply properties of similar triangles to solve a problem<br>• Calculate the area of a quadrilateral<br>• Calculate the volume of a solid composed of triangular and rectangular prisms<br>• Apply the Pythagorean Theorem to find the length of a segment on the coordinate plane<br>• Interpret a histogram or bar graph |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter<br><br>• Calculate and compare unit costs<br>• Use proportions and percents to solve a problem<br>• Write an equation with two variables to describe a real-world situation<br>• Apply properties of the real numbers to manipulate formulas and simplify expressions<br>• Identify the equation of a nonlinear function from a table<br>• Identify the graph of a function that best represents a described real-world situation<br>• Solve a two-step linear equation<br>• Identify the coordinates of the image of a vertex of a polygon after a translation or reflection<br>• Determine whether points on a coordinate plane can be vertices of a parallelogram<br>• Identify the net of a cube<br>• Estimate equal units in different systems of measure<br>• Identify a scatterplot given a description of the variables being compared<br>• Use data in a table or scatter plot to make a prediction<br>• Interpret a line graph |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter<br><br>• Express a ratio as a fraction or decimal<br>• Identify a proportion that can be used to relate quantities in a real-world situation<br>• Identify an inequality or equation with one variable that describes a real-world situation<br>• Extend an arithmetic sequence at least three terms<br>• Interpret the meaning of the $y$-intercept of the graph of a linear function in a real-world context<br>• Identify a single transformation applied to a figure that produces a given image<br>• Calculate the area of a triangle or quadrilateral on or off the coordinate plane<br>• Identify a strategy to collect data that is most representative of a group<br>• Calculate the simple probability of a single event |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter<br><br>• Perform multiple operations on decimal numbers to solve a problem<br>• Evaluate numerical expressions involving square roots of perfect squares<br>• Identify the *y*-intercept of a linear function given the graph of the line<br>• Order rational numbers in decimal form<br>• Describe in words a rule for a given a geometric sequence<br>• Identify a linear graph that represents an arithmetic sequence given a rule in words<br>• Identify the side view of a three dimensional figure drawn in perspective |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 10 Reading Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter and provide sophisticated solutions to complex problems.<br><br>Using grade level text or above, the student is able to:<br>• Apply a rich and varied content vocabulary<br>• Consistently apply complex thinking skills – extend and connect ideas, make complex predictions, analyze and evaluate causal relationships, describe abstract themes & ideas, formulate complex arguments with strong supporting evidence<br>• Apply literary elements<br>• Apply inferential thinking<br>• Understand different genres<br>• Flexibly use a variety of strategies to interpret language, literary characteristics & overall intent<br>• Set, monitor progress toward, and meet reading goals |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.<br><br>Using grade level text, the student is able to:<br>• Use a substantial reading and content vocabulary<br>• Apply complex thinking skills – make & revise predictions, explains inferences, analyze causal relationships, formulate arguments with supporting evidence<br>• Usually paraphrase accurately<br>• Use a variety of strategies to interpret language, literary characteristics & overall intent<br>• Analyze the author's uses of literary devices<br>• Set, monitor progress toward, and meet reading goals |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.<br><br>Using grade level text or near-grade level text, the student is able to:<br>• Use a limited 10th grade vocabulary<br>• Occasionally apply complex thinking skills – make predictions, identify inferences, describe causal relationships, frequently paraphrase accurately, formulate arguments with limited supporting evidence, identify elements of an author's style<br>• Use a limited variety of strategies to interpret the language, literary characteristics & overall intent<br>• Use a limited range of reading purposes<br>• Often set and sometimes meets reading goals, sometimes evaluate reading progress |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.<br><br>Using near-grade level text, the student is able to:<br>• Use a reading vocabulary below Grade 10<br>• Make simple predictions & inferences<br>• Not often grasp the meaning of causal relationships<br>• Sometimes paraphrase accurately<br>• Sometimes formulate arguments with limited supporting evidence and provide simple responses<br>• Rely primarily on a few strategies to interpret language, literary characteristics & overall intent<br>• Understand a limited range of reading purposes<br>• Compare & contrast but infrequently analyze or apply complex thinking skills<br>• Rarely set or meet reading goals |

**OFFICE OF PUBLIC INSTRUCTION**
**PO BOX 202501**
**HELENA MT 59620-2501**
**www.opi.mt.gov**
**(406) 444-3095**
**(888) 231-9393**
**(406) 444-0169 (TTY)**

**Linda McCulloch**
**Superintendent**

# Grade 10 Math Performance/Achievement Descriptors

| | |
|---|---|
| **Advanced** | Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matte<br><br>• Write a linear equation with slope other than one or zero given a table of values, graph, or description in words<br>• Solve an equation in one variable that requires more than two steps<br>• Write an equation involving trigonometric ratios to solve a real-world problem using<br>• Identify the relevant theorem that justifies the congruence of two given triangles<br>• Sort quadrilaterals on and off the coordinate plane by properties involving angles, sides, or diagonals<br>• Identify a geometric shape that provides a counter-example to a given statement<br>• Apply the Pythagorean Theorem to solve a problem that requires multiple steps<br>• Calculate the area of a composite figure<br>• Determine the number of unique combinations given a set of objects<br>• Calculate the probability of a desired outcome given the probabilities of all other possible outcomes<br>• Display data in a circle graph |
| **Proficient** | Students at this level demonstrate a solid understanding of challenging subject matter<br><br>• Order rational numbers written as fractions, mixed numbers and decimals<br>• Describe the effect of operations on arbitrary real numbers<br>• Determine and interpret the slope of a linear function from a graph<br>• Generalize a linear sequence of numbers with an algebraic expression<br>• Describe the characteristics of smaller figures used to construct a three dimensional figure<br>• Use relationships of angle and segments in a figure to determine similarity of polygons<br>• Apply the Pythagorean Theorem to determine the length of leg of a right triangle<br>• Convert among derived units to solve a problem<br>• Apply the distance formula to problems involving the coordinate grid<br>• Use probability to make predictions<br>• Identify the appropriate display of a given set of data<br>• Calculate the median of a set of data displayed in a frequency table |
| **Nearing Proficiency** | Students at this level demonstrate a partial understanding of subject matter<br><br>• Evaluate a numerical expression with multiple operations on fractions<br>• Use proportions or percents to solve a problem<br>• Determine whether a given number is rational<br>• Write and compare numbers in scientific notation<br>• Write an inequality or equation with two variables to describe a real-world situation<br>• Evaluate an algebraic expression for a given value<br>• Determine whether a graphed function is linear or nonlinear<br>• Associate a line graphed on the coordinate plane with its equation<br>• Determine the coordinates of the image of a vertex of a polygon after a transformation<br>• Identify the relationships among angles formed by parallel lines and a transversal<br>• Apply the Pythagorean Theorem to determine the length of the hypotenuse of a right triangle<br>• Compare the relative volumes of rectangular prisms<br>• Identify a positive or negative correlation between two variables in a scatter plot |

| | |
|---|---|
| **Novice** | Students at this level demonstrate a minimal understanding of subject matte<br><br>• Evaluate numerical expression with multiple operations on whole numbers<br>• Identify a proportion that can be used to relate quantities in a real-world situation<br>• Identify an inequality or equation with one variable that describes a real-world situation<br>• Read a graph of a function on the coordinate grid to determine intervals of increasing and decreasing<br>• Identify the shape of the cross section of a three dimensional figure with a drawing<br>• Determine whether two variables have a correlation given a scatter plot<br>• Interpret a circle, line, or bar graph |

# CRT Scaled Score Ranges for Performance Levels

## Grade 3

| | Reading | Mathematics |
|---|---|---|
| Advanced | 285-300 | 291-300 |
| Proficient | 250-284 | 250-290 |
| Nearing Proficiency | 225-249 | 225-249 |
| Novice | 200-224 | 200-224 |

## Grade 4

| | Reading | Mathematics |
|---|---|---|
| Advanced | 288-300 | 287-300 |
| Proficient | 250-287 | 250-286 |
| Nearing Proficiency | 225-249 | 225-249 |
| Novice | 200-224 | 200-224 |

## Grade 5

| | Reading | Mathematics |
|---|---|---|
| Advanced | 287-300 | 291-300 |
| Proficient | 250-286 | 250-290 |
| Nearing Proficiency | 225-249 | 225-249 |
| Novice | 200-224 | 200-224 |

## Grade 6

| | Reading | Mathematics |
|---|---|---|
| Advanced | 289-300 | 291-300 |
| Proficient | 250-288 | 250-290 |
| Nearing Proficiency | 225-249 | 225-249 |
| Novice | 200-224 | 200-224 |

## Grade 7

| | Reading | Mathematics |
|---|---|---|
| Advanced | 289-300 | 291-300 |
| Proficient | 250-288 | 250-290 |
| Nearing Proficiency | 225-249 | 225-249 |
| Novice | 200-224 | 200-224 |

## Grade 8

| | Reading | Mathematics |
|---|---|---|
| Advanced | 291-300 | 283-300 |
| Proficient | 250-290 | 250-282 |
| Nearing Proficiency | 225-249 | 225-249 |
| Novice | 200-224 | 200-224 |

## Grade 10

| | Reading | Mathematics |
|---|---|---|
| Advanced | 290-300 | 278-300 |
| Proficient | 250-289 | 250-277 |
| Nearing Proficiency | 225-249 | 225-249 |
| Novice | 200-224 | 200-224 |

# Raw Score Range and Percent of Students in Each Performance Level

### GRADE 3 READING

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 45-60 | 37.9 |
| Proficient | 31-44 | 43.2 |
| Nearing Proficiency | 21-30 | 12.8 |
| Novice | 0-20 | 6.1 |

### GRADE 4 READING

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 47-60 | 33.5 |
| Proficient | 33-46 | 47.1 |
| Nearing Proficiency | 24-32 | 12.6 |
| Novice | 0-23 | 6.8 |

### GRADE 5 READING

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 43-60 | 36.0 |
| Proficient | 30-42 | 44.0 |
| Nearing Proficiency | 21-29 | 13.3 |
| Novice | 0-20 | 6.7 |

## GRADE 6 READING

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 46-60 | 33.0 |
| Proficient | 34-45 | 45.9 |
| Nearing Proficiency | 26-33 | 13.4 |
| Novice | 0-25 | 7.7 |

## GRADE 7 READING

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 47-60 | 31.7 |
| Proficient | 33-46 | 46.2 |
| Nearing Proficiency | 24-32 | 13.3 |
| Novice | 0-23 | 8.8 |

## GRADE 8 READING

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 48-60 | 30.6 |
| Proficient | 36-47 | 46.0 |
| Nearing Proficiency | 28-35 | 14.1 |
| Novice | 0-27 | 9.3 |

## GRADE 10 READING

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 52-65 | 29.6 |
| Proficient | 39-51 | 46.7 |
| Nearing Proficiency | 31-38 | 13.3 |
| Novice | 0-30 | 10.4 |

## GRADE 3 MATH

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 55-66 | 24.7 |
| Proficient | 43-54 | 41.7 |
| Nearing Proficiency | 35-42 | 17.9 |
| Novice | 0-34 | 15.7 |

## GRADE 4 MATH

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 54-66 | 26.0 |
| Proficient | 42-53 | 38.3 |
| Nearing Proficiency | 33-41 | 19.5 |
| Novice | 0-32 | 16.2 |

## GRADE 5 MATH

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 47-66 | 25.4 |
| Proficient | 34-46 | 37.4 |
| Nearing Proficiency | 25-33 | 21.5 |
| Novice | 0-24 | 15.7 |

## GRADE 6 MATH

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 42-66 | 25.2 |
| Proficient | 29-41 | 37.3 |
| Nearing Proficiency | 21-28 | 21.2 |
| Novice | 0-20 | 16.3 |

## GRADE 7 MATH

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 43-66 | 24.1 |
| Proficient | 30-42 | 37.5 |
| Nearing Proficiency | 22-29 | 22.9 |
| Novice | 0-21 | 15.5 |

## GRADE 8 MATH

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 46-66 | 25.2 |
| Proficient | 33-45 | 33.1 |
| Nearing Proficiency | 22-32 | 26.0 |
| Novice | 0-21 | 15.8 |

## GRADE 10 MATH

| Proficiency Level | Raw Score Range | % in Level |
|---|---|---|
| Advanced | 51-71 | 24.4 |
| Proficient | 37-50 | 30.8 |
| Nearing Proficiency | 24-36 | 28.2 |
| Novice | 0-23 | 16.6 |

# APPENDIX E: REPORT SHELLS

## Student Report

## Class Roster & Item-Level Report

## School Summary Report

## System Summary Report

# CRT Performance Level Descriptors

The Performance Level Descriptors below describe students' knowledge, skills, and abilities in a content area. These descriptions provide a picture or profile of student achievement at the four performance levels: **Advanced**, **Proficient**, **Nearing Proficiency**, and **Novice**. Grade and content performance level descriptors may be found on OPI's web site at http://www.opi.mt.gov/assessment/index.html

**Advanced**
This level denotes superior performance.
**Proficient**
This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.
**Nearing Proficiency**
This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.
**Novice**
This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

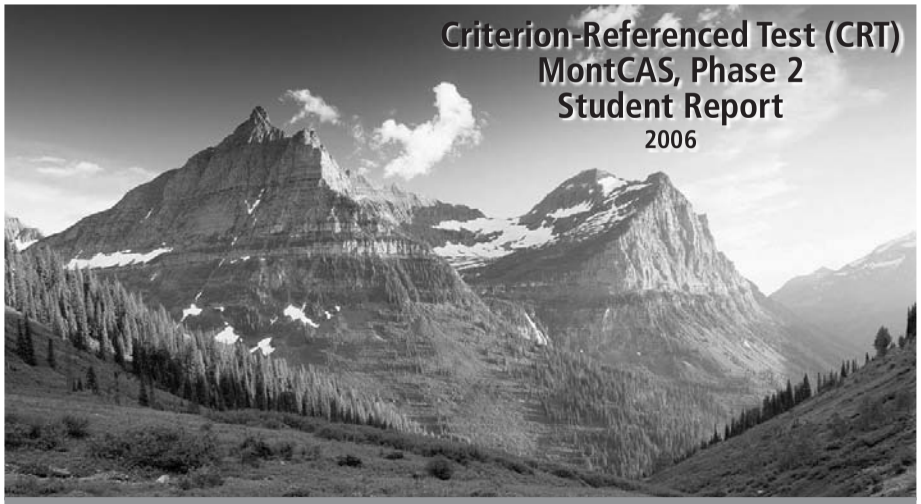
| | Score Ranges | |
|---|---|---|
| | **Reading** | **Math** |
| **Advanced** | (285-300) | (291-300) |
| **Proficient** | (250-284) | (250-290) |
| **Nearing Proficiency** | (225-249) | (225-249) |
| **Novice** | (200-224) | (200-224) |

**For more information regarding student assessments in Montana, check out the Office of Public Instruction's Parents Page at http://www.opi.mt.gov/parents.**

**OPI Contact**
**Judy Snow, State Assessment Director**
**406-444-3656**
**jsnow@mt.gov**

OPI
**Linda McCulloch, Superintendent**
**Montana Office of Public Instruction**
**PO Box 202501**
**Helena, Montana 59620-2501**
**http://www.opi.mt.gov**





# Criterion-Referenced Test (CRT)
## MontCAS, Phase 2
## Student Report
**2006**

**Student Name:**
**School:**
**System:**
**Grade:** 03

Dear Parents/Guardians:

This report contains the results of the Montana Comprehensive Assessment System Criterion-Referenced Test (CRT) that your child took in March. The major purpose of the CRT is to provide schools with solid information to evaluate and improve curriculum and instruction to help all students meet Montana's reading and mathematics standards. This report provides important information about your child's performance on the assessment, along with state results.

The CRT contains multiple-choice and short-answer questions. The test measures a student's knowledge of subject matter identified in the Montana State Standards for Reading and Mathematics. Your child's results in reading and mathematics are reported in one of four performance levels. These performance levels are defined on the back cover of this report.

It is important to remember that the CRT is just one measure of your child's academic progress. Your local school staff can provide further information about your child's performance in school. The CRT, which is required by the No Child Left Behind Act, is part of an ongoing statewide educational improvement process. Working together, we can ensure that Montana's children continue to receive a high-quality education.

Sincerely,

Linda McCulloch
Montana Superintendent of Public Instruction

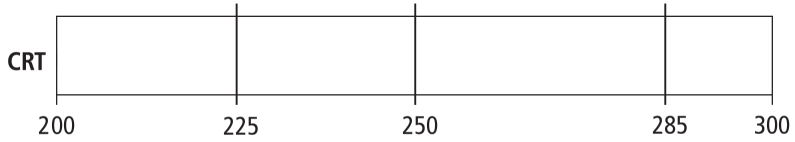
## Scaled Scores on the CRT

The criterion-referenced test (CRT) is designed to measure student performance against the learning goals described in the Montana Content Standards (http://www.opi.state.mt.us/standards/index.html). Consistent with this purpose, results on the CRT are reported according to performance levels that describe student performance in relation to the established state standards. There are four performance levels: **Advanced**, **Proficient**, **Nearing Proficiency**, and **Novice**. Your child's performance levels in reading and mathematics are based on a total scaled score in each content area. Scaled scores in each content area range from 200 to 300. Your child's performance levels, based on the scaled scores, are shown in the bar graphs below.

## Scaled Scores

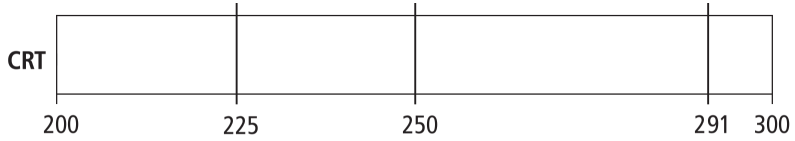
**STUDENT RESULTS FOR READING**

**Performance Level:**
**Student Scaled Score:**




**STUDENT RESULTS FOR MATHEMATICS**

**Performance Level:**
**Student Scaled Score:**




## Scores on Montana Content Standards

In addition to performance levels, CRT results are reported for Montana Content Standards in reading and mathematics. Unlike scaled scores which provide a total performance level score, Montana Content Standard Scores provide more specific information about your child's achievement on the CRT. The chart on the following page shows your child's performance in each area of study within subject areas (Montana Content Standards for reading and math). These results can be used to show your child's relative strengths or weaknesses.

Contact your student's school for more information about the following symbols:
† Student did not complete the assessment.
§ Student took non-standard accommodation.



| Scores on Montana Standards | Points Possible | Student Percentage / State Percentage | Percentage of Points Earned (0 25 50 75 100) |
|---|---|---|---|
| **Reading Standards** | | | |
| 1. Students construct meaning as they comprehend, interpret, and respond to what they read. | 21 | | |
| 2. Students apply a range of skills and strategies to read. | 19 | | |
| 3. Students set goals, monitor, and evaluate their reading progress. | This standard is not measurable in a statewide assessment. | | |
| 4. Students select, read, and respond to print and nonprint material for a variety of purposes. | 9 | | |
| 5. Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audiences. | 11 | | |
| **Math Standards** | | | |
| 1. Problem Solving | 9 | | |
| 2. Numbers and Operations | 12 | | |
| 3. Algebra | 6 | | |
| 4. Geometry | 11 | | |
| 5. Measurement | 9 | | |
| 6. Data Analysis, Statistics, and Probability | 12 | | |
| 7. Patterns, Relations, and Functions | 7 | | |

Student percentage of points earned

State percentage of points earned

**MontCAS, Phase 2 CRT**

# Reading
## Roster & Item-Level Report
**Confidential**

Spring 2006

Class:
School:
System:

Grade: 03
Page: 1 of 2

| Name | Item Number | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 46 | Scaled Score | Perf. Level |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Standard | 2 | 2 | 2 | 1 | 5 | 1 | 2 | 4 | 4 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 4 | 2 | 4 | 1 | 5 | 5 | 4 | 1 | 1 | 5 | 1 | 5 | 2 | 5 | 1 | 5 | 2 | | |
| | Correct Response | B | D | C | A | A | B | D | C | A | C | D | B | C | C | D | A | B | A | B | D | A | | C | B | B | D | A | C | A | D | A | C | B | | |
| | Total Possible Points | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

**MontCAS, Phase 2 CRT**

# Reading
## Roster & Item-Level Report
**Confidential**

Class:
School:
System:

Spring 2006

Grade: 03
Page: 2 of 2

| Name | Item Number | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | | | | | | | | | | | | | Scaled Score | Perf. Level |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Standard | 4 | 2 | 4 | 1 | 1 | 2 | 1 | 4 | 2 | 2 | 5 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 4 | 1 | 2 | | | | | | | | | | | | | | |
| | Correct Response | B | C | C | C | C | A | A | C | D | C | B | D | A | B | A | C | C | D | D | B | | | | | | | | | | | | | | | |
| | Total Possible Points | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | |

**MontCAS, Phase 2 CRT**

# Mathematics
## Roster & Item-Level Report
**Confidential**

Class:
School:
System:

Spring 2006

Grade: 03
Page: 1 of 2

| Name | Item Number | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 35 | 36 | 37 | Scaled Score | Perf. Level |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Standard | 2 | 4 | 7 | 4 | 2 | 4 | 7 | 5 | 6 | 3 | 7 | 5 | 2 | 5 | 2 | 6 | 4 | 1 | 3 | 2 | 7 | 6 | 5 | 4 | 6 | 4 | 6 | 2 | 5 | 4 | 5 | 1 | 3 | | |
| | Correct Response | A | B | B | D | B | D | A | C | C | D | A | D | C | D | B | D | C | B | A | D | C | C | A | A | | D | C | D | B | B | C | C | A | | |
| | Total Possible Points | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

**MontCAS, Phase 2 CRT**

# Mathematics
## Roster & Item-Level Report
**Confidential**

Class:
School:
System:

Spring 2006

Grade: 03
Page: 2 of 2

| Name | Item Number | 38 | 39 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | | | | | | | Scaled Score | Perf. Level |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Standard | 7 | 1 | 4 | 5 | 2 | 6 | 7 | 3 | 6 | 4 | 5 | 3 | 1 | 4 | 6 | 1 | 7 | 4 | 6 | 3 | 2 | 5 | 2 | 2 | 2 | 2 | 1 | | | | | | | | |
| | Correct Response | B | A | C | A | A | D | D | C | A | B | A | C | A | B | B | D | A | D | B | D | D | C | C | | | | | | | | | | | | |
| | Total Possible Points | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

# MontCAS, Phase 2
# CRT

System:
Grade: 03
Spring 2006

**Reading**

## System Summary Report

### I. Distribution of Scores

| Perf. Level | Scores | System | | | State | | |
|---|---|---|---|---|---|---|---|
| | | Number | % of Students | % of Students in Cat. | Number | % of Students | % of Students in Cat. |
| Advanced | 298-300 | | | | | | |
| | 295-297 | | | | | | |
| | 291-294 | | | | | | |
| | 288-290 | | | | | | |
| | 285-287 | | | | | | |
| Proficient | 278-284 | | | | | | |
| | 271-277 | | | | | | |
| | 264-270 | | | | | | |
| | 257-263 | | | | | | |
| | 250-256 | | | | | | |
| Nearing Proficiency | 245-249 | | | | | | |
| | 240-244 | | | | | | |
| | 235-239 | | | | | | |
| | 230-234 | | | | | | |
| | 225-229 | | | | | | |
| Novice | 220-224 | | | | | | |
| | 215-219 | | | | | | |
| | 210-214 | | | | | | |
| | 205-209 | | | | | | |
| | 200-204 | | | | | | |

### II. Subtest Results

| Reading | | Possible Points | Average Points Earned | |
|---|---|---|---|---|
| | | | System | State |
| Total Points | | 60 | | |
| Standards | 1. Students construct meaning as they comprehend, interpret, and respond to what they read | 21 | | |
| | 2. Students apply a range of skills and strategies to read | 19 | | |
| | 3. Students set goals, monitor, and evaluate their reading progress | This standard is not measurable in a statewide assessment. | | |
| | 4. Students select, read, and respond to print and nonprint material for a variety of purposes | 9 | | |
| | 5. Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audiences | 11 | | |

**CRT Performance Level Descriptors**

**Advanced (285-300)**
This level denotes superior performance.

**Proficient (250-284)**
This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

**Nearing Proficiency (225-249)**
This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

**Novice (200-224)**
This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

# MontCAS, Phase 2 CRT

**Confidential**

**Reading** **System Summary Report**

System:
Grade: 03
Spring 2006

## III. Results for Subgroups of Students

| Reporting Category | System | | | | | State | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Number | % in N | % in NP | % in P | % in A | Number | % in N | % in NP | % in P | % in A |
| **All Students** | | | | | | | | | | |
| Gender | | | | | | | | | | |
| Male | | | | | | | | | | |
| Female | | | | | | | | | | |
| Ethnicity | | | | | | | | | | |
| American Indian or Alaska Native | | | | | | | | | | |
| Asian | | | | | | | | | | |
| Hispanic | | | | | | | | | | |
| Black or African American | | | | | | | | | | |
| Native Hawaiian or Other Pacific Islander | | | | | | | | | | |
| White | | | | | | | | | | |
| Significant Cognitive Disability | | | | | | | | | | |
| Special Education | | | | | | | | | | |
| Students with a 504 Plan | | | | | | | | | | |
| Title I (optional) | | | | | | | | | | |
| Tested with Standard Accommodation | | | | | | | | | | |
| Tested with Non-Standard Accommodation | | | | | | | | | | |
| Alternate Assessment | | If a student in your system or school took the CRT-Alternate, please refer to Table III on the CRT-Alternate System or School Summary Report | | | | | | | | |
| Migrant | | | | | | | | | | |
| Gifted/Talented | | | | | | | | | | |
| LEP/ELL | | | | | | | | | | |
| Former LEP Student | | | | | | | | | | |
| LEP Student Enrolled for First Time in a U.S. School | | Performance levels are not reported for 1st year LEP students | | | | | | | | |
| Free/Reduced Lunch | | | | | | | | | | |
| Special Education Disability(ies): | | | | | | | | | | |
| Autism | | | | | | | | | | |
| Cognitive Delay | | | | | | | | | | |
| Deaf-Blindness Impairment | | | | | | | | | | |
| Deafness | | | | | | | | | | |
| Emotional Disturbance | | | | | | | | | | |
| Hearing Impairment | | | | | | | | | | |
| Learning Disability | | | | | | | | | | |
| Other Health Impairment | | | | | | | | | | |
| Orthopedic Impairment | | | | | | | | | | |
| Speech/Language | | | | | | | | | | |
| Traumatic Brain Injury | | | | | | | | | | |
| Visual Impairment | | | | | | | | | | |

*Less than ten (10) students were assessed

Shells

# MontCAS, Phase 2
# CRT

System:
Grade: 03
Spring 2006

**Mathematics**

## System Summary Report

### I. Distribution of Scores

| Perf. Level | Scores | System | | | State | | |
|---|---|---|---|---|---|---|---|
| | | Number | % of Students | % of Students in Cat. | Number | % of Students | % of Students in Cat. |
| Advanced | 299-300 | | | | | | |
| | 297-298 | | | | | | |
| | 295-296 | | | | | | |
| | 293-294 | | | | | | |
| | 291-292 | | | | | | |
| Proficient | 283-290 | | | | | | |
| | 275-282 | | | | | | |
| | 266-274 | | | | | | |
| | 258-265 | | | | | | |
| | 250-257 | | | | | | |
| Nearing Proficiency | 245-249 | | | | | | |
| | 240-244 | | | | | | |
| | 235-239 | | | | | | |
| | 230-234 | | | | | | |
| | 225-229 | | | | | | |
| Novice | 220-224 | | | | | | |
| | 215-219 | | | | | | |
| | 210-214 | | | | | | |
| | 205-209 | | | | | | |
| | 200-204 | | | | | | |

### II. Subtest Results

| Mathematics | | Possible Points | Average Points Earned | |
|---|---|---|---|---|
| | | | System | State |
| Total Points | | 66 | | |
| Standards | 1. Problem Solving | 9 | | |
| | 2. Numbers and Operations | 12 | | |
| | 3. Algebra | 6 | | |
| | 4. Geometry | 11 | | |
| | 5. Measurement | 9 | | |
| | 6. Data Analysis, Statistics, and Probability | 12 | | |
| | 7. Patterns, Relations, and Functions | 7 | | |

**CRT Performance Level Descriptors**

**Advanced (291-300)**
This level denotes superior performance.

**Proficient (250-290)**
This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

**Nearing Proficiency (225-249)**
This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

**Novice (200-224)**
This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

# MontCAS, Phase 2 CRT

**Confidential**

**Mathematics** **System Summary Report**

System:
Grade: 03
Spring 2006

## III. Results for Subgroups of Students

| Reporting Category | System | | | | | State | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Number | % in N | % in NP | % in P | % in A | Number | % in N | % in NP | % in P | % in A |
| **All Students** | | | | | | | | | | |
| Gender | | | | | | | | | | |
| Male | | | | | | | | | | |
| Female | | | | | | | | | | |
| Ethnicity | | | | | | | | | | |
| American Indian or Alaska Native | | | | | | | | | | |
| Asian | | | | | | | | | | |
| Hispanic | | | | | | | | | | |
| Black or African American | | | | | | | | | | |
| Native Hawaiian or Other Pacific Islander | | | | | | | | | | |
| White | | | | | | | | | | |
| Significant Cognitive Disability | | | | | | | | | | |
| Special Education | | | | | | | | | | |
| Students with a 504 Plan | | | | | | | | | | |
| Title I (optional) | | | | | | | | | | |
| Tested with Standard Accommodation | | | | | | | | | | |
| Tested with Non-Standard Accommodation | | | | | | | | | | |
| Alternate Assessment | | If a student in your system or school took the CRT-Alternate, please refer to Table III on the CRT-Alternate System or School Summary Report | | | | | | | | |
| Migrant | | | | | | | | | | |
| Gifted/Talented | | | | | | | | | | |
| LEP/ELL | | | | | | | | | | |
| Former LEP Student | | | | | | | | | | |
| LEP Student Enrolled for First Time in a U.S. School | | Performance levels are not reported for 1st year LEP students | | | | | | | | |
| Free/Reduced Lunch | | | | | | | | | | |
| Special Education Disability(ies): | | | | | | | | | | |
| Autism | | | | | | | | | | |
| Cognitive Delay | | | | | | | | | | |
| Deaf-Blindness Impairment | | | | | | | | | | |
| Deafness | | | | | | | | | | |
| Emotional Disturbance | | | | | | | | | | |
| Hearing Impairment | | | | | | | | | | |
| Learning Disability | | | | | | | | | | |
| Other Health Impairment | | | | | | | | | | |
| Orthopedic Impairment | | | | | | | | | | |
| Speech/Language | | | | | | | | | | |
| Traumatic Brain Injury | | | | | | | | | | |
| Visual Impairment | | | | | | | | | | |

*Less than ten (10) students were assessed

Shells

# MontCAS, Phase 2
# CRT

School:
System:
Grade: 03
Spring 2006

**Reading**

## School Summary Report

### I. Distribution of Scores

| Perf. Level | Scores | School | | | System | | | State | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | N | % of Students | % of Students in Cat. | N | % of Students | % of Students in Cat. | N | % of Students | % of Students in Cat. |
| Advanced | 298-300 | | | | | | | | | |
| | 295-297 | | | | | | | | | |
| | 291-294 | | | | | | | | | |
| | 288-290 | | | | | | | | | |
| | 285-287 | | | | | | | | | |
| Proficient | 278-284 | | | | | | | | | |
| | 271-277 | | | | | | | | | |
| | 264-270 | | | | | | | | | |
| | 257-263 | | | | | | | | | |
| | 250-256 | | | | | | | | | |
| Nearing Proficiency | 245-249 | | | | | | | | | |
| | 240-244 | | | | | | | | | |
| | 235-239 | | | | | | | | | |
| | 230-234 | | | | | | | | | |
| | 225-229 | | | | | | | | | |
| Novice | 220-224 | | | | | | | | | |
| | 215-219 | | | | | | | | | |
| | 210-214 | | | | | | | | | |
| | 205-209 | | | | | | | | | |
| | 200-204 | | | | | | | | | |

### II. Subtest Results

| Reading | | Possible Points | Average Points Earned | | |
|---|---|---|---|---|---|
| | | | School | System | State |
| Total Points | | 60 | | | |
| Standards | 1. Students construct meaning as they comprehend, interpret, and respond to what they read | 21 | | | |
| | 2. Students apply a range of skills and strategies to read | 19 | | | |
| | 3. Students set goals, monitor, and evaluate their reading progress | This standard is not measurable in a statewide assessment. | | | |
| | 4. Students select, read, and respond to print and nonprint material for a variety of purposes | 9 | | | |
| | 5. Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audiences | 11 | | | |

**CRT Performance Level Descriptors**

**Advanced (285-300)**
This level denotes superior performance.

**Proficient (250-284)**
This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

**Nearing Proficiency (225-249)**
This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

**Novice (200-224)**
This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.



ShellsShells

# MontCAS, Phase 2 CRT

**Confidential**

**Reading**

**School Summary Report**

School:
System:
Grade: 03
Spring 2006

## III. Results for Subgroups of Students

| Reporting Category | School | | | | | System | | | | | State | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | % in N | % in NP | % in P | % in A | Number | % in N | % in NP | % in P | % in A | Number | % in N | % in NP | % in P | % in A |
| **All Students** | | | | | | | | | | | | | | | |
| Gender | | | | | | | | | | | | | | | |
| Male | | | | | | | | | | | | | | | |
| Female | | | | | | | | | | | | | | | |
| Ethnicity | | | | | | | | | | | | | | | |
| American Indian or Alaska Native | | | | | | | | | | | | | | | |
| Asian | | | | | | | | | | | | | | | |
| Hispanic | | | | | | | | | | | | | | | |
| Black or African American | | | | | | | | | | | | | | | |
| Native Hawaiian or Other Pacific Islander | | | | | | | | | | | | | | | |
| White | | | | | | | | | | | | | | | |
| Significant Cognitive Disability | | | | | | | | | | | | | | | |
| Special Education | | | | | | | | | | | | | | | |
| Students with a 504 Plan | | | | | | | | | | | | | | | |
| Title I (optional) | | | | | | | | | | | | | | | |
| Tested with Standard Accommodation | | | | | | | | | | | | | | | |
| Tested with Non-Standard Accommodation | | | | | | | | | | | | | | | |
| Alternate Assessment | | If a student in your system or school took the CRT-Alternate, please refer to Table III on the CRT-Alternate System or School Summary Report | | | | | | | | | | | | | |
| Migrant | | | | | | | | | | | | | | | |
| Gifted/Talented | | | | | | | | | | | | | | | |
| LEP/ELL | | | | | | | | | | | | | | | |
| Former LEP Student | | | | | | | | | | | | | | | |
| LEP Student Enrolled for First Time in a U.S. School | | Performance levels are not reported for 1st year LEP students | | | | | | | | | | | | | |
| Free/Reduced Lunch | | | | | | | | | | | | | | | |
| Special Education Disability(ies): | | | | | | | | | | | | | | | |
| Autism | | | | | | | | | | | | | | | |
| Cognitive Delay | | | | | | | | | | | | | | | |
| Deaf-Blindness Impairment | | | | | | | | | | | | | | | |
| Deafness | | | | | | | | | | | | | | | |
| Emotional Disturbance | | | | | | | | | | | | | | | |
| Hearing Impairment | | | | | | | | | | | | | | | |
| Learning Disability | | | | | | | | | | | | | | | |
| Other Health Impairment | | | | | | | | | | | | | | | |
| Orthopedic Impairment | | | | | | | | | | | | | | | |
| Speech/Language | | | | | | | | | | | | | | | |
| Traumatic Brain Injury | | | | | | | | | | | | | | | |
| Visual Impairment | | | | | | | | | | | | | | | |

*Less than ten (10) students were assessed

# MontCAS, Phase 2
# CRT

School:
System:
Grade: 03
Spring 2006

**Mathematics**

## School Summary Report

### I. Distribution of Scores

| Perf. Level | Scores | School | | | System | | | State | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | N | % of Students | % of Students in Cat. | N | % of Students | % of Students in Cat. | N | % of Students | % of Students in Cat. |
| Advanced | 299-300 | | | | | | | | | |
| | 297-298 | | | | | | | | | |
| | 295-296 | | | | | | | | | |
| | 293-294 | | | | | | | | | |
| | 291-292 | | | | | | | | | |
| Proficient | 283-290 | | | | | | | | | |
| | 275-282 | | | | | | | | | |
| | 266-274 | | | | | | | | | |
| | 258-265 | | | | | | | | | |
| | 250-257 | | | | | | | | | |
| Nearing Proficiency | 245-249 | | | | | | | | | |
| | 240-244 | | | | | | | | | |
| | 235-239 | | | | | | | | | |
| | 230-234 | | | | | | | | | |
| | 225-229 | | | | | | | | | |
| Novice | 220-224 | | | | | | | | | |
| | 215-219 | | | | | | | | | |
| | 210-214 | | | | | | | | | |
| | 205-209 | | | | | | | | | |
| | 200-204 | | | | | | | | | |

### II. Subtest Results

| Mathematics | | Possible Points | Average Points Earned | | |
|---|---|---|---|---|---|
| | | | School | System | State |
| Total Points | | 66 | | | |
| Standards | 1. Problem Solving | 9 | | | |
| | 2. Numbers and Operations | 12 | | | |
| | 3. Algebra | 6 | | | |
| | 4. Geometry | 11 | | | |
| | 5. Measurement | 9 | | | |
| | 6. Data Analysis, Statistics, and Probability | 12 | | | |
| | 7. Patterns, Relations, and Functions | 7 | | | |

**CRT Performance Level Descriptors**

**Advanced (291-300)**
This level denotes superior performance.

**Proficient (250-290)**
This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

**Nearing Proficiency (225-249)**
This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

**Novice (200-224)**
This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.



ShellsShells

# MontCAS, Phase 2 CRT

**Confidential**

**Mathematics**

**School Summary Report**

School:
System:
Grade: 03
Spring 2006

## III. Results for Subgroups of Students

| Reporting Category | School | | | | | System | | | | | State | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | % in N | % in NP | % in P | % in A | Number | % in N | % in NP | % in P | % in A | Number | % in N | % in NP | % in P | % in A |
| **All Students** | | | | | | | | | | | | | | | |
| Gender | | | | | | | | | | | | | | | |
| Male | | | | | | | | | | | | | | | |
| Female | | | | | | | | | | | | | | | |
| Ethnicity | | | | | | | | | | | | | | | |
| American Indian or Alaska Native | | | | | | | | | | | | | | | |
| Asian | | | | | | | | | | | | | | | |
| Hispanic | | | | | | | | | | | | | | | |
| Black or African American | | | | | | | | | | | | | | | |
| Native Hawaiian or Other Pacific Islander | | | | | | | | | | | | | | | |
| White | | | | | | | | | | | | | | | |
| Significant Cognitive Disability | | | | | | | | | | | | | | | |
| Special Education | | | | | | | | | | | | | | | |
| Students with a 504 Plan | | | | | | | | | | | | | | | |
| Title I (optional) | | | | | | | | | | | | | | | |
| Tested with Standard Accommodation | | | | | | | | | | | | | | | |
| Tested with Non-Standard Accommodation | | | | | | | | | | | | | | | |
| Alternate Assessment | | If a student in your system or school took the CRT-Alternate, please refer to Table III on the CRT-Alternate System or School Summary Report | | | | | | | | | | | | | |
| Migrant | | | | | | | | | | | | | | | |
| Gifted/Talented | | | | | | | | | | | | | | | |
| LEP/ELL | | | | | | | | | | | | | | | |
| Former LEP Student | | | | | | | | | | | | | | | |
| LEP Student Enrolled for First Time in a U.S. School | | Performance levels are not reported for 1st year LEP students | | | | | | | | | | | | | |
| Free/Reduced Lunch | | | | | | | | | | | | | | | |
| Special Education Disability(ies): | | | | | | | | | | | | | | | |
| Autism | | | | | | | | | | | | | | | |
| Cognitive Delay | | | | | | | | | | | | | | | |
| Deaf-Blindness Impairment | | | | | | | | | | | | | | | |
| Deafness | | | | | | | | | | | | | | | |
| Emotional Disturbance | | | | | | | | | | | | | | | |
| Hearing Impairment | | | | | | | | | | | | | | | |
| Learning Disability | | | | | | | | | | | | | | | |
| Other Health Impairment | | | | | | | | | | | | | | | |
| Orthopedic Impairment | | | | | | | | | | | | | | | |
| Speech/Language | | | | | | | | | | | | | | | |
| Traumatic Brain Injury | | | | | | | | | | | | | | | |
| Visual Impairment | | | | | | | | | | | | | | | |

*Less than ten (10) students were assessed

# Appendix F: Reporting Decision Rules

**CRT Spring 05-06**
**Decision Rules**

| Participation | | | Relationship w/ Data File Layouts | Impact on Analyses | Impact on Student report | Impact on School/System/State reports | Impact on Student Roster and I-Analyze | Impact on student level data Excel files for System CD's | Impact on student level data Excel files for State CD |
|---|---|---|---|---|---|---|---|---|---|
| Number of Students ("N") | 1 | Number of students included in state aggregation | NA | N=total number of students with 2 or more responses minus students tested at a private accredited school (PRAS) minus students tested in a non-accredited Title I private school (PRNAT1) minus foreign exchange (FXS) students minus students not enrolled (SNE) minus student enrolled part-time (PSNE) minus students tested at a private non-accredited school (PRNAS) minus LEP student enrolled first time in U.S. school minus students who took a content area using a nonstandard accommodation. | | | | | |
| No class header provided | 2 | No class indicators provided | Tfname=' ' and Tlname=' ' | Class aggregations calculated are actually school level. | No impact | No impact | Report produced | No impact | No impact |
| Number of Students for Reporting | 3 | Schools (Systems) has **less than 10 included** students in both content areas | NA | | No impact | School/system report Produced. Page 2: For each category numbers will be suppressed if number of included students less than ten. The N-size is always reported. Footnote *'Less than 10 students were assessed" | No Impact | No Impact | No Impact |
| Student Names Not Provided | 4 | No student barcode label and no name bubbled on answer sheet | Lname, Fname | No Impact on analyses.<br><br>Student included in DP report to systems.<br>Student counted in N. | No Impact<br>Student name is "Name Not Provided" | Student included based on inclusion rules stated in this document. | Student included based on inclusion rules stated in this document. | Student included based on inclusion rules stated in this document. | Student included based on inclusion rules stated in this document. |

| Participation | | | Relationship w/ Data File Layouts | Impact on Analyses | Impact on Student report | Impact on School/System/State reports | Impact on Student Roster and I-Analyze | Impact on student level data Excel files for System CD's | Impact on student level data Excel files for State CD |
|---|---|---|---|---|---|---|---|---|---|
| Form Number Not Coded | 5 | DP codes as Form 1, only common items scored | Form | Student counted in N | No Impact | Student Included | No Impact | No Impact | No impact. |
| Tested but Fewer than 2 of the answers marked | 6 | Student answered fewer than 2 of the common MC questions | All common items | Student not counted in N; student excluded from item analysis | Score given with a footnote (†): "Student did not complete the assessment" | Student not Included | Score given with a footnote (†): "Student did not complete the assessment" | Student Included | No Impact |
| Tested with Standard Accom-modations | 7 | Student requires an accommodation(s) by content area | Any REASA1-REASA29 bubbled and MATSA1-MATSA29 | If one or more standard accommodations (#1-29) are coded, student is counted as Tested with Standard Accommodation(s) | No Impact | Counted as Tested with Standard Accommodation(s) | No Impact | No Impact | No Impact |
| Tested with Non-standard Accom-modations | 8 | Student requires a non-standard accommodation(s) by content area | Any REANSA30-REANSA33 bubbled and MATNSA30-MATNSA33 | If one or more non-standard accommodations (#30-33) are coded, student is be counted as Tested with Non-standard Accommodation(s), counted as NOT participated, and will not be included in aggregations in content area(s) where non-standard accommodations were coded. | Student report will indicate raw score with an (§) and a footnote stating that the student took a non-standard accommodation. The scaled score is the scaled score associated with the earned raw score | Student will NOT be included in count ("n") and will count as not participated in the content area student took a nonstandard accommodation except on the reporting categories page; they will be counted in the nonstandard accommodations category for the appropriate content area. | Student record will indicate raw score with an (§) stating that the student took a non-standard accommodation. | Student included in system CD. | Student included in state CD. |

| Participation | | | Relationship w/ Data File Layouts | Impact on Analyses | Impact on Student report | Impact on School/System/State reports | Impact on Student Roster and I-Analyze | Impact on student level data Excel files for System CD's | Impact on student level data Excel files for State CD |
|---|---|---|---|---|---|---|---|---|---|
| Program Information | 9 | Student is identified as participating in an identified program. | SE='1' or Plan504='1' or Migrant='1' or GT='1' or LEP='1' or Lowincome ='1' or Title1Mat or Title1Rea='1' or Disability='1' or FormerLEP='1' | If one or more Program Information codes are bubbled, student is counted as a program participant **First year LEP students are not included in counts for LEP students.** | No Impact | Reported on school & system Reporting Category reports. All numbers except the N-size are suppressed if N-size less than 10. Footnote *'Less than 10 students were assessed.' | No Impact | No Impact | No Impact |
| Special Education- not optional. Can have more than one bubbled for a student | 10 | Student is has an identified disability under IDEA -97. | AU='1',CD='1',DB='1',DE='1', ED='1',HI='1',LD='1',OI='1',OH='1',SL='1',TB='1',VI='1' | Student is counted in their respective disability group on page 2 of summary reports. | No Impact | Student is counted in their respective disability group on page 2 of summary reports. All numbers except the N-size are suppressed if N-size less than 10.Footnote * 'Less than 10 students were assessed.' | No Impact | No Impact | No Impact |
| First year LEP student And not native American | 11 | Student is identified as being a first year LEP | LEPFirst='1' and ethnic not ='I' | Student is excluded from all aggregations for both content areas. | Student receives report. Student does not receive scaled score or performance level for reading. Performance Level='LEP' on report for reading. Student receives earned score in Math. Student receives earned score if he/she attempted the test. | Student is excluded from aggregations**. Student is counted in First year LEP student enrolled first time in U.S. school.** | Student is included. Student's scaled score is blank. Student's performance level ='LEP' for Reading. If student took the Reading test the responses are shown. The student included in Math with earned scores and responses shown. PL="LEP" for both content areas. Student not included in aggregations for both content areas. | Student is included Receives earned score if he/she attempted the test. | Student is included. Receives earned score if he/she attempted the test. |

| Participation | | | Relationship w/ Data File Layouts | Impact on Analyses | Impact on Student report | Impact on School/System/State reports | Impact on Student Roster and I-Analyze | Impact on student level data Excel files for System CD's | Impact on student level data Excel files for State CD |
|---|---|---|---|---|---|---|---|---|---|
| Foreign Exchange Student (FXS) | 12 | Student is identified as a foreign exchange student | FXS='1' | Student is not included in any school/system/state aggregations. | Student receives report. | Not included on Reports | Not Included on Reports | Students are not included on System CD | Included on State CD; identified as FXS |
| Student Not Enrolled (SNE) (Homeschooled) | 13 | Student is identified as not enrolled in an accredited public school. | SNE='1' | Student is not included in any school/system/state aggregations | Student receives report. | Not Included on Reports | Not Included on Reports | Students are not included on System CD | Not included on State CD |
| Private Accredited School (PRAS) | 14 | Student is identified as testing at a private accredited school | PRAS='1' | Student is not included in any state aggregations | Student receives report. | School report produced. System report produced. They are their own system. | Report produced | Students are included on System CD | Included on State CD; identified as PRAS |
| Private Non-accredited Title I School (PRNAT1) | 15 | Student is identified as testing in a non-accredited Title I school | PRNAT1='1' | Student is not included in any state aggregations | Student receives report. | School report produced. System report produced. They are their own system. | Report produced | Students are included on System CD | Included on State CD; identified as PRNAT1 |
| Private Non-Accredited School (PRNAS) | 16 | Student is identified as testing in a non-accredited school. | PRNAS='1' | Student is not included in state aggregations | Student receives report. | School report produced. System report produced. They are their own system. | Report produced. | Students are included on System CD. | Included on state CD; identified as PRNAS |
| Student enrolled part-time (<180 hours) (PSNE) | 17 | Student is identified as enrolled part-time | PSNE='1' | Student is not included in any school/system/state aggregations | Student receives report. | Student not included | Not included on reports. | Students are not included in system CD. | Student not included on state CD. |

| Participation | | | Relationship w/ Data File Layouts | Impact on Analyses | Impact on Student report | Impact on School/System/State reports | Impact on Student Roster and I-Analyze | Impact on student level data Excel files for System CD's | Impact on student level data Excel files for State CD |
|---|---|---|---|---|---|---|---|---|---|
| Participation Information (NSAY & NDAY) | 21 | Student participated in CRT but has not been a student in school or district for entire academic year. | NSAY='1' , NDAY='1' | Student is included in participation. If student is marked as NSAY only then student is not included in school aggregations. If student is marked as NDAY then student is not included in either school or district aggregations. | No impact. | If student is marked as NSAY only then student is not included in school data. If student is marked as NDAY then student is not included in school or district data. | If student is NSAY or NDAY student is included on roster with footnote(¥) "Not in school and/or district full academic year." Student excluded from school (if NSAY or NDAY) and/or district (if NDAY) aggregations. | No Impact | No Impact |
| Alternate Assessment Student | 22 | Student participated through alternate assessment this year | Alt='1' | Student is excluded from CRT aggregations. | Student receives CRT-Alternate student report. | Student included in CRT-Alternate aggregations. Student included in count of alternate students on page 2 of CRT summary reports. Only N-size is reported. The rest of the line is covered with a watermark. | Student not included in I-Analyze. Student included (unless otherwise excluded based on CRT-Alternate decision rules) on CRT-Alternate Roster | Student included in CRT-Alternate system CDs. Student included on CRT CDs with alt flagged. | Student included in CRT-Alternate state CD. Student not included in CRT. |
| Student has bubbled SNE and PSNE | 23 | Student has bubbled both not enrolled and part-time | SNE='1' and PSNE='1' | Student is not included in any school/system/state aggregations | Student receives report. | Student is not included | Student is not included | Student is not included | Student is not included |
| Student is SNE and enrolled part-time(PSNE) in a private school | 24 | Student has bubbled not enrolled and part-time and a private school | SNE='1' ,PSNE='1', PRNAT1='1', PRNAS='1', PRAS='1' | Student is not included in any school/system/state aggregations | Student receives report. | Student is not included | Student is not included | Student is not included | Student is not included. |

**Additional Rules:**

**1. Only common items are used to calculate scores.**

**Grade 3 = 391**
**Grade 4 = 403**
**Grade 5 = 400**
**Grade 6 = 361**
**Grade 7 = 298**
**Grade 8 = 299**
**Grade 10 = 187**

**Systems: 351**

**Scores:**

**Reading Subtest:** **Raw score is number of correct responses to common items. Total possible is:**
**Grades 3-8: 60 score points**
**Grade 10: 65 score points**

**Math Subtest:** **Raw score is number of correct responses to common items. Total possible is:**
**Grades 3-8: 66 score points**
**Grade 10: 71 score points**

**<u>Addenda:</u>**

1) **Students identified with a testing irregularity will be treated in the following manner:**
a) **They will have a parent/guardian report produced with their earned score (after blanking out responses to the items which the testing irregularity occurred) and performance level displayed. Footnote will indicate that there was a testing irregularity.**
b) **They will be included in aggregations (class, school, system, or state) .**
c) **On the class roster for the content area in which the testing irregularity occurred will be listed with 'IR' appearing in the pl column. The responses to the items which they cheated on will be blanked out. The scaled score column will be blank.**
2) **If alt is not bubbled and significant cognitive disability is not bubbled and student took both CRT and CRT-Alternate for both content areas, the Alt scores will be invalidated.**
3) **If a student took the CRT-Alternate in one content area and CRT in the other content area then the student will be considered alt and will be a non-participant for the content area in which he did the CRT.**