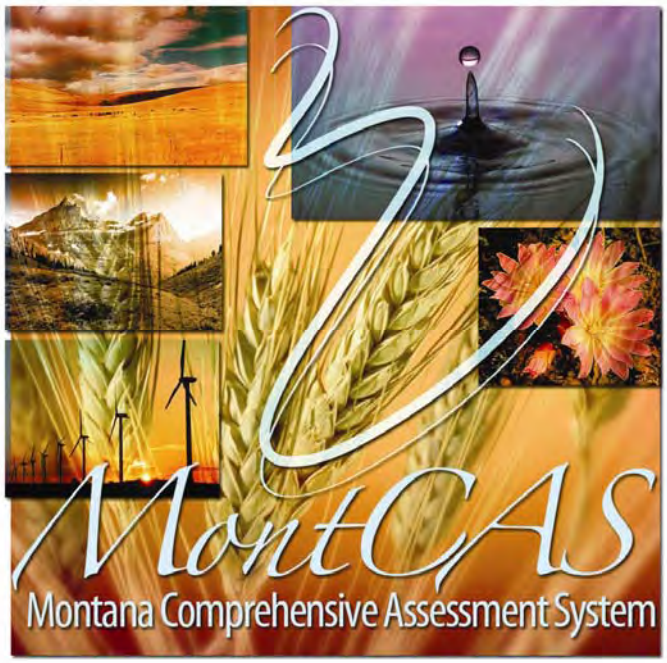


MontCAS
Criterion-Referenced Test
Alternate Assessment
(CRT-Alternate)

2008–09

Technical Report

TABLE OF CONTENTS

SECTION I: ASSESSMENT DEVELOPMENT **1**

CHAPTER 1. BACKGROUND AND OVERVIEW 1
1.1 Purpose of This Report *1*
1.2 Purpose of the CRT-Alternate *1*
1.3 Development of the Reading, Mathematics, and Science Expanded Benchmarks *2*
1.4 Test Scheduling *4*
1.5 Organization of This Report *5*
CHAPTER 2. ASSESSMENT PARTICIPATION 7
2.1 Participants *7*
2.2 Participation Guidelines *7*
CHAPTER 3. OVERVIEW OF TEST DESIGN 9
3.1 CRT-Alternate *9*
3.2 Assessment Type *9*
3.3 Test Design *10*
3.3.2 CRT-Alternate Items 11
3.3.3 Evidence and Evidence Templates 12
3.3.4 Test Administration Survey 12
3.4 Scaffolding as Scoring *14*
CHAPTER 4. TEST DEVELOPMENT PROCESS 15
4.1 Item and Activity Development *15*
4.2 CRT-Alternate Item Development Process Overview *15*
4.3 Item/Activity Editing *18*
CHAPTER 5. ASSESSMENT DESIGN BLUEPRINTS 19
5.1 Reading Assessment Blueprint *19*
5.2 Mathematics Assessment Blueprint *20*
5.3 Science Assessment Blueprint *21*

SECTION II: TEST ADMINISTRATION **23**

CHAPTER 6. TEST ADMINISTRATION 23
6.1 Responsibility for Administration *23*
6.2 Procedures *23*
6.3 Training *24*

SECTION III: DEVELOPMENT AND REPORTING OF SCORES **27**

CHAPTER 7. SCORING 27
7.1 Scoring the Assessment *27*
7.2 Using Scaffolding to Gather Student Performance Information *27*
7.3 Scoring Rubric *29*
7.4 Scoring Rules *29*
7.5 Scanning Procedures and Quality Control *30*
7.5.1 Gatekeeping 30
7.5.2 Guillotining 31
7.5.3 Scanning Procedures 31
7.5.4 Machine Scored Items 31
7.5.5 Quality Control 32
7.6 Electronic Data Files *32*
CHAPTER 8. ITEM ANALYSES 33
8.1 Difficulty Indices (p-value) *33*
8.2 Item-Test Correlations (Discrimination) *34*
8.3 Summary of Item Analysis Results *34*
CHAPTER 9. SCALING 37
9.1 Translating Raw Scores to Scaled Scores and Performance Levels *37*
CHAPTER 10. RELIABILITY 41
10.1 A Note on Scorer Interrater Reliability *41*
10.2 Other Reliability Results *41*
10.3 Reliability of Performance-Level Categorization *43*

10.3.1 Accuracy, Consistency, and Kappa 43
10.3.2 Results of Accuracy, Consistency, and Kappa Analyses 44
CHAPTER 11. REPORTING 47
CHAPTER 12. VALIDITY SUMMARY 49

SECTION IV: REFERENCES 51

APPENDICES 53

Appendix A Montana-Alt Standard Setting Report
Appendix B Performance Level Descriptors—Raw & Scaled Scores
Appendix C Sample Tasklet
Appendix D Interrater Reliability Report
Appendix E Technical Advisory Committee Members
Appendix F CRT Alt Released Items
Appendix G Item Level Classical Statistics Results
Appendix H Decision Accuracy and Consistency Results
Appendix I Analysis and Reporting Decision Rules
Appendix J Report Shells

SECTION I: ASSESSMENT DEVELOPMENT

Chapter 1. BACKGROUND AND OVERVIEW

1.1 Purpose of This Report

The purpose of this report is to document the technical aspects of the 2009 MontCAS Criterion-Referenced Test-Alternate Assessment (CRT-Alternate). In the spring of 2009, students in grades 3–8 and 10 participated in the administration of the CRT-Alternate in both reading and mathematics. Students in grades 4, 8, and 10 were also assessed in science. Due to new development, a standard setting meeting for reading and mathematics (grades 4, 8, and 10) was held in May 2009 using data from the spring 2009 administration (see Appendix A for standard setting report). This report provides information about the technical quality of those assessments, including a description of the processes used to develop, administer, and score the tests and to analyze results.

Historically, the intended audience of a technical report has been experts in psychometrics and educational research. This edition of the CRT-Alternate technical report is intended to be more accessible and useful to educators and other stakeholders by providing rich descriptions of general categories of information. In making some of the information more accessible, we have purposefully preserved the depth of technical information provided in our past 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 be familiar with advanced topics in measurement and statistics.

1.2 Purpose of the CRT-Alternate

The Individuals with Disabilities Education Act (IDEA) requires students with disabilities be included in each state’s system of accountability and that students with disabilities have access to the general curriculum. The No Child Left Behind Act (NCLB) speaks to the inclusion of all children in a state’s accountability system by requiring states to report student achievement for all students, as well as for groups of students on a disaggregated basis. These federal laws reflect an ongoing concern about equity: all students should be academically challenged and taught to high standards. It is also necessary that all students be involved in the educational accountability system.

To ensure the participation of all students in the state’s accountability system, Montana has developed the CRT-Alternate. The CRT-Alternate is a point-in-time, direct measure of a student’s performance based on alternate achievement standards aligned with Montana’s Content Standards and Expanded Benchmarks. Only those IDEA-eligible students with the most significant cognitive disabilities are expected to participate in the CRT-Alternate.

1.3 Development of the Reading, Mathematics, and Science Expanded Benchmarks

Expanded benchmarks were developed for students with significant cognitive disabilities who are not working at the same level as their age-level counterparts. The benchmarks correspond to the standards for (a) end of grade 4, (b) end of grade 8, and (c) upon graduation—end of grade 12. Expansion is toward foundational skills and is keyed to grade-span rather than grade-level expectations due to the wide diversity of students in this population.

The expanded benchmarks were developed using Montana's Content Standards and Benchmarks for reading, mathematics, and science. Measured Progress's curriculum and special education specialists developed a draft of the expanded benchmarks. The Montana Office of Public Instruction (OPI), beta test teachers, the advisory committee, and the development and revision workshop participants all provided input and recommendations for changes to the original draft. Measured Progress revised the expanded benchmarks using these recommendations, and the document was further revised to include grade-span expectations per new federal legislation. This document was then used as the basis for developing the assessment performance indicators. Table 1-1 shows how the document is organized and gives an example for each content area. The full Montana Content Standards and Expanded Benchmarks for the content areas are not included in this report because of their length. They are located on the OPI Web site at www.opi.state.mt.us and the Measured Progress Web site at www.measuredprogress.org.

Table 1-1. 2008–09 Montana CRT-ALT: Breakdown of Standards and Expanded Benchmarks

Term/Description	***Example***		
Content Area	**Reading**	**Mathematics**	**Science**
Standard Learning outcome expected for all students throughout all grades	Standard 2: Students apply a range of skills and strategies to read.	Standard 2: Students demonstrate understanding of and ability to use Numbers and Operations.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
Essence of the Standard A statement of the standard separating the essential components	Interpret print and non-print information.	Number concepts, concepts of operations, computing and estimating.	Matter exists in a variety of forms. All physical interactions involve changes in energy. Therefore, knowledge of matter and energy is essential to interpreting, explaining, predicting, and influencing change in our world.
Benchmark **Grade Level Expectation (GLE)** Expectation for typical students described for each grade level	2.6, Grade 8: Students will develop vocabulary through the use of context clues, analysis of word parts, auditory clues, and reference sources (e.g., dictionary, thesaurus, and glossary).	2.2, Grade 4: Students will use the number system by counting, grouping, and applying place value concepts.	2.2, Grade 4: Examine, describe, compare, and classify objects in terms of common physical properties.
Expanded Benchmark Benchmark skill or concept expanded from the typical GLE to a basic level	2.6.2: Student will use words/pictures/symbols/objects to communicate.	2.2.1: Student will demonstrate an understanding of whole numbers.	2.2.2: Student will compare the common physical properties of two objects.
Performance Indicator Expanded benchmark expressed in a measurable and observable statement of a specific performance	2.6.2.1: Student will identify a word/picture/symbol/object used to name a familiar place.	2.2.1.2: Student will demonstrate the concept of one (e.g., “Hit the switch one time”; “Give me one”).	2.2.2.1 Student will identify the similarities and differences in the size of two objects or substances.
Prompt The script for the directions the test administrator delivers to the student, calling for the specific behavior	Item 4: “Show me the word/picture/symbol/object that means ‘library.’”	Item 4: “These are counters. We are going to use these in our activity. Show me one counter.”	Item 2: “This box has a hole in it. Which object is small enough to fit through this hole?”

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 and Expanded Benchmarks for their grade span. The underlying principle of the assessment is that all students should be taught using Montana's Content Standards and Expanded Benchmarks in reading, mathematics, and science. The tests are intended to measure how a student is performing in relation to those content standards. Results should be used to inform future instruction in the Montana content standards.

The 2008–09 administration of the CRT-Alternate was the sixth year of implementation. After the first year, extensive revisions were made based on feedback from teachers who administered the assessment. Alternate assessments, ranging from checklists to portfolios and performance-based tests, have been in place nationally since 2000 due to federal requirements. We are still learning appropriate ways to address reliability and validity for alternate assessments. To address the reliability of the CRT-Alternate, Cronbach's α, accuracy and consistency of performance-level categorization, and kappa analyses were performed. These analyses are summarized in Chapter 10. Each chapter in this report contributes important information to the validity argument by addressing one or more of the following aspects of the CRT-Alternate:

- test development
- test alignment
- test administration
- scoring item analyses
- reliability
- scaling
- performance levels
- reporting

These aspects, as well as other information on validity, are addressed in Chapter 12.

1.4 Test Scheduling

The CRT-Alternate was administered during the spring: reading and mathematics were administered in grades 3–8 and 10, and science in grades 4, 8, and 10, during a six-week window (February 10–March 25, 2009). Schools were able to schedule testing sessions at any time during this period. This window, longer than that for the CRT, allowed teachers administering the CRT-Alternate extra time to prepare and adapt test materials needed for administering the assessment.

The CRT-Alternate is an untimed assessment. Teachers administering the assessments were instructed to watch for indications that students might need a break. Recommendations for breaks are inserted throughout each grade-specific test booklet. Teachers could choose to stop at the breaks or at other points in the assessment.

1.5 Organization of This Report

The organization of this report is based on the conceptual flow of an assessment's life span. It begins with the initial test specification and addresses all the intermediate steps that lead to final score reporting.

Section I consists of five chapters:

- Background and Overview
- Assessment Participation
- Overview of Test Design
- Test Development Process
- Assessment Design Blueprints

Section II consists of a single chapter:

- Test Administration

Section III consists of six chapters:

- Scoring
- Item Analysis
- Scaling
- Reliability
- Reporting
- Validity Summary

Section IV contains references and is followed by the appendices.

Chapter 2. ASSESSMENT PARTICIPATION

2.1 Participants

Because the general CRT provides full access to the vast majority of students, only about 100 students per grade are expected to participate in the CRT-Alternate. Table 2-1 displays the number of students who participated in the CRT-Alternate by grade and content area in 2008–09.

Table 2-1. 2008–09 Montana CRT-ALT: Counts of Participating Students by Grade and Content Area

Grade	*Content Area*	*N*
3	Mathematics	94
	Reading	92
4	Mathematics	104
	Reading	104
	Science	104
5	Mathematics	98
	Reading	97
6	Mathematics	109
	Reading	109
7	Mathematics	72
	Reading	72
8	Mathematics	103
	Reading	103
	Science	103
10	Mathematics	130
	Reading	128
	Science	129

In accordance with 34 CFR 200.13 Adequate Yearly Progress in general, there is a 1% cap applied to the number of proficient and advanced scores based on the alternate assessment that may be included in AYP calculations at both the state and district levels.

2.2 Participation Guidelines

How a student with disabilities will participate in the state's accountability system is decided by the student's Individualized Education Program (IEP) team. When considering whether students with disabilities should participate in the CRT-Alternate, the IEP team should address each of the questions shown in Figure 2-2.

Table 2-2. 2008–09 Montana CRT-ALT: Participation Guidelines

Participation Guidelines: For each of the statements below, answer YES or NO		
Does the student have an active IEP and receive services under the Individuals with Disabilities Education Act (IDEA)?	YES	NO
Do the student's demonstrated cognitive abilities and adaptive behavior require substantial adjustments to the general curriculum?	YES	NO
Do the student's learning objectives and expected outcomes focus on functional application of skills, as illustrated in the student's IEP's annual goals and short-term objectives?	YES	NO
Does the student require direct and extensive instruction to acquire, maintain, generalize, and transfer new skills?	YES	NO

If the IEP team determines the answer is "no" to <u>any</u> of the above questions, the student must participate in the general CRT. If all answers are "yes," the student is eligible to take the alternate assessment and is considered to have a significant cognitive disability. IEP teams are informed that the decision to have a student participate in the CRT-Alternate may not be based on excessive or extended absence; disability category; social, cultural, or economic factors; the amount of time receiving special education services; or academic achievement significantly lower than his or her same-age peers.

Chapter 3. OVERVIEW OF TEST DESIGN

3.1 CRT-Alternate

CRT-Alternate test items are directly linked to Montana's Content Standards and Expanded Benchmarks. (See section 1.3 for more information about the expanded benchmarks.) The content standards are the basis for the reporting categories developed for each content area and are used to help guide the development of test items. An item may address part, all, or several of the benchmarks within a standard or standards.

3.2 Assessment Type

Due to separate development cycles through the life span of the assessment program, the CRT-Alternate format varied slightly depending on the grade and content area assessed until this year. The original format of the CRT-Alternate consisted of one task activity per content area with 22–35 items. The original format, with one task activity (e.g., activity based around baking cake) narrowed the student's opportunity for success if the student was averse to that topic. Through feedback from the field, it was determined that a variety of activities within each content area would be more appropriate for this population. Furthermore, a variety of activities within a content area provides students more opportunities to demonstrate their knowledge and skills.

Designing the test around a series of short activities, or "tasklets," allows the teacher and student to break the administration into smaller time segments with less concern about disruption in continuity. With the recent redevelopment of grades 4, 8, and 10 in reading and mathematics, all content areas and grades now use the tasklet model. This consistency across every grade and content area provides ease and fluidity for test administration. Each content area tested is composed of five tasklets that consist of five to six questions each where teachers are given a script, written directions, and scaffolding levels. (See section 3.4 for more information on scaffolding.)

The tasklets are developed from the expanded benchmarks, follow the scaffolding rubric, and are designed to show a student's performance in relation to the Montana reading, mathematics, and science standards and benchmarks. Students are encouraged to engage in the tasklet and show performance on the items through appropriate prompting by the test administrator. The teacher who administers the tasklet scores the student on each item through observation using a five-point scoring rubric. Every student takes the same form of the test. Test items are kept secure, but the performance indicators, which come from the Montana reading, mathematics, and science Content Standards and Expanded Benchmarks, are released every year on the OPI and Measured Progress Web sites. The 2009 released performance items are located in Appendix F.

3.3 Test Design

Table 3.1 outlines the design of the CRT-Alternate and its related components. The first page of each tasklet provides a useful guide for test administrators by listing the following information:

- content standards and expanded benchmarks
- a brief explanation of the suggested tasklet
- parameters of the tasklet
- materials provided and other materials that are needed

Each tasklet contains one introductory item, as well as a suggested break at the end of the tasklet. Passages are provided on the second page of reading tasklets, as well as in the Materials Kit. The Materials Kit contains associated test materials needed to administer the assessment, such as student response cards, passages in storybook format, and specially adapted materials that provide symbol-text pairings for students who require a higher level of support. In order to collect evidence within each content area of the CRT-Alternate, the test administrator must complete two forms for specified test items. Specific scoring rules have been developed for the assessment, for which students are required to attempt every tasklet.

Table 3-1. 2008–09 Montana CRT-ALT: Test Design

Format	Tasklet—five short activities of five or six items each per content area Total of 25–28 items
Introductory Items	First item in each tasklet Designed to gain student's attention, introduce the activity, and show materials to be used Scored at levels 4 or 0 of the rubric
Breaks	Breaks between tasklets
Reading Passage	Page 2 of each reading tasklet
Student Evidence	1–2 tasklets in each content area require student evidence Two forms need to be filled out for each item that requires evidence
Scoring Rule	Student must try every tasklet Halt the administration of a tasklet only if the student scores a 0 for three consecutive items after administering the tasklet during two different test sessions
Material Kits	Tabs in the Materials Kits are labeled by content area and tasklet number

3.3.2 CRT-Alternate Items

Each item of the CRT-Alternate consists of the following:

- materials needed to administer the item
- communication support strategies the teacher may use to administer the item
- setup instructions and script for the teacher to follow
- scaffolding script for the suggested test activity
- the correct student response
- the performance indicator (The performance indicator—a description of what the question is measuring—is derived from the Montana Content Standards and Expanded Benchmarks.)

Figure 3-1 describes the information presented in each column of every test item in the CRT-Alternate. A sample item is provided in Figure 3-2.

Figure 3-1. 2008–09 Montana CRT-ALT: Information Presented in Test Items

Materials for the Activity	*Activity Teacher will:*	*Student Work Student will:*	*Performance Indicators Use Scoring Guide Transfer scores to student response booklet*
The materials that are needed for each item and suggested student communication supports and strategies that may be helpful for some students are described in this column. Most materials can be found in the Material Kits, but teachers need to supply some materials.	This column contains information about how to display tasklet materials and prepare the student for the question. A script for the teacher appears in bold and italicized print and suggests language that can be used to present the item. Information on how to scaffold levels 3, 2, and 1 of the rubric for items that are scored at levels 4 through 0 is also provided in this column.	The correct student response and/or an explanation of how the student should be responding are provided in this column.	The performance indicator that is assessed by each item is identified in this column. The performance indicators come from the Montana Content Standards and Expanded Benchmarks.

Figure 3-2. 2008–09 Montana CRT-ALT: Grade 3 Mathematics Sample Item

Materials for the Activity	Activity Teacher will:	Student Work Student will:	Performance Indicators Use Scoring Guide Transfer scores to student response booklet
2. • 1 large square • 1 large triangle • 1 large circle • 1 large rectangle Communication support strategies: • Student may look at/point to task materials to express a choice. • Request may be rephrased to require a yes/no response (e.g., "Is this the CIRCLE?"). • Student may tell teacher to "stop" at desired response as teacher sequentially points to each of the 4 choices.	2. Place all the shapes in random order on the work space. ***"Show me the circle."*** Scaffold: Level 3: Remove an incorrect response. Repeat task request. Level 2: Remove another incorrect response. Repeat task request. Level 1: "This is the circle." Assist the student as needed to identify the circle.	2. Identify a circle.	2. Identifies (names) shapes as circles, squares, triangles, rectangles, and ovals. 4 3 2 1 0 Performance Indicator: 4.1.1.6 Expanded Benchmark: 4.1.1

(For a complete sample tasklet see Appendix C.)

3.3.3 Evidence and Evidence Templates

Evidence on how the student performed in each content area must be collected during the course of the assessment. Templates are provided in the CRT-Alternate test booklet for all evidence that is required. A magnifying glass icon in the "Student Work, Student will" column of the test booklet indicates when evidence must be collected. One form is used to document the way in which the student responded to the item; a second form captures the student's final response. The Evidence Template Teacher Recording Sheet provides a format to document the student's entire sequence of responses to the test item. As the test item is presented to the student, the test administrator documents the modality used by the student to communicate a response, as well as the accuracy of the response at each step of the scaffolding process. Recording ends when the student demonstrates a correct response, with or without scaffolding. An Evidence Template is used to document the student's FINAL response for the test item for which evidence is being collected. By reviewing the information contained on these two forms, it is possible to visualize the student's complete response to the test item. The evidence must be submitted along with the used test booklet.

3.3.4 Test Administration Survey

The last page of the test booklet contains a list of questions regarding preparation and administration for the teacher to answer after the administration of the reading, mathematics, and science tasklets. Question 11 asks the teacher to report how much time he or she spent preparing for the assessment. Question 12 asks the

teacher to report how much time was spent administering the assessment to the student. According to this year's embedded survey, there was no significant difference from last year in the amount of time used by teachers to both prepare and administer the assessment. The lowest average preparation and administration times in both mathematics and science were reported in Grade 10. Tables 3-2 and 3-3 summarize survey responses to questions 11 and 12.

Table 3-2. 2008–09 Montana CRT-ALT: Survey Responses—Question 11 Setup Time/Planning

Grade	*Content Area*	*Average # of Hours*
3	Reading	1.19
	Mathematics	1.20
4	Reading	1.13
	Mathematics	1.11
	Science	1.06
5	Reading	1.05
	Mathematics	1.05
6	Reading	.92
	Mathematics	.90
7	Reading	.95
	Mathematics	.86
8	Reading	1.12
	Mathematics	1.04
	Science	.97
10	Reading	.89
	Mathematics	.84
	Science	.77

Table 3-3. 2008–09 Montana CRT-ALT: Survey Responses—Question 12 Time Spent Administering Assessment

Grade	*Content Area*	*Average # of Hours*
3	Reading	1.33
	Mathematics	1.32
4	Reading	1.27
	Mathematics	1.31
	Science	1.19
5	Reading	1.24
	Mathematics	1.28
6	Reading	1.32
	Mathematics	1.22
7	Reading	1.30
	Mathematics	1.30
8	Reading	1.14
	Mathematics	1.12
	Science	1.05
10	Reading	1.21
	Mathematics	1.06
	Science	1.06

3.4 Scaffolding as Scoring

As Gail McGregor of the University of Montana–Missoula notes in her paper titled, "Examining the Interrator Reliability of Montana's CRT-Alternative" (Appendix D), "Administration of the CRT-Alt incorporates a response prompting methodology known as the 'system of least prompts' (Wolery, Ault & Doyle, 1992). This is a well-established strategy that has been found to be effective as a teaching procedure for students with severe disabilities across a wide range of applications (Doyle, Wolery, Ault & Gast, 1988)." The system of least prompts, or scaffolding, requires the teacher (or test administrator) to administer each test item beginning at the highest level of independence. The student is asked the question and allowed sufficient time to produce the answer. If the student produces the answer, the teacher records the student's score for that question at the highest level. If the student answers incorrectly, the test administrator asks the question again, this time using the second-highest level of independence for that particular question.

The levels of independence are standardized and scripted within the test. The second-highest level of independence usually amounts to removing one or two choices from the set of possible answers. If the student provides the correct answer, the test administrator will record the score at the second-highest level of independence. If the student cannot provide the correct answer, the test administrator moves on to the next-highest level of independence, and so on, until the student is guided (hand-over-hand) to the correct answer and the student's score for that particular item is recorded at the lowest level of independence. More information regarding the research base of this method and a discussion regarding the selection of this method can be found in Appendix D.

Chapter 4. TEST DEVELOPMENT PROCESS

4.1 Item and Activity Development

The CRT-Alternate was developed as a collaborative project between Measured Progress and OPI divisions of Assessment, Special Education, and Educational Opportunity and Equity.

An advisory committee, representing the perspectives of parents, teachers, administrators, and faculty in higher education, provided input during the development of this assessment. In addition, educator work groups were formed at several points in the development and revision processes. Reading, mathematics, and science item development work groups were composed of general and special education educators, as well as school administrators. These educators and administrators developed tasklets that are the basis of the performance tasks for this assessment. A third group of special education teachers and administrators participated in the beta testing of this assessment, providing valuable feedback about the test design.

OPI and Measured Progress were responsible for organizing and facilitating committees to review reading passages and items for bias and sensitivity. OPI reviewed the feedback and approved appropriate changes to the items and reading passages. Table 4-1 outlines the total number of items developed in each grade and content area.

Table 4-1. 2008–09 Montana CRT-ALT: Total Numbers of Items Developed by Grade and Content Area

Grade	*Reading*	*Mathematics*	*Science*
3	25	25	
4	25	25	26
5	25	25	
6	25	25	
7	25	25	
8	25	25	26
10	25	25	28

4.2 CRT-Alternate Item Development Process Overview

Four separate development process cycles comprise the current CRT-Alternate. The separate development cycles for reading, mathematics, and science occurred as follows: (1) Reading and mathematics, grades 4, 8, and 10 were developed in the original task activity model between August 2003 and October 2004 (an overview of the test development process for these grades is outlined in the technical report for 2005); (2) Development for reading and mathematics, grades 3, 5, 6, and 7 in the new tasklet model took place between March 2005 and January 2006 (outlined in the technical report for 2006); (3) The science assessment in grades 4, 8, and 10 was developed in the new tasklet model between April 2006 and February 2008 (an overview of the test development process is outlined in the technical report for 2008).

The most recent development cycle began in March 2008 for the redesign of grades 4, 8, and 10, in both reading and mathematics. These grades were redeveloped in order to transition from the former task activity model to the tasklet model. The reading and mathematics test-development process for grades 4, 8, and 10 began with a review of the original blueprints for the grades and content areas. Using the original benchmarks for end of grades 4, 8, and 12, staff from Measured Progress in collaboration with OPI created a test blueprint for each grade and content area for the new tasklet model. The original task-activity design consisted of 22–35 items per content area, while the tasklet model consists of a total of 25 items. The new blueprints were created in order to reflect the same amount of emphasis on standards and benchmarks as in the original blueprints.

The blueprints indicated which benchmarks should be tested for each grade and content area. Once the blueprints were approved by the state, staff from Measured Progress further identified which expanded benchmarks and performance indicators were to be used as the base for selecting tasklet topics and creating reading passages and test items. The state was involved in every step of the process in order to provide feedback and/or give approval. In April 2008, a bias and item review, in conjunction with an item development workshop, was held consisting of various stakeholders including special education teachers, general education teachers, and school/system administrators. During the bias and item review, committee members identified draft passages and items that were potentially unsuitable for the assessment in terms of cultural, socioeconomic, religious, age-appropriateness, and accessibility concerns. For example, one of the drafted passages was based on the movie *Harry Potter*. This selection was ultimately deemed inappropriate due to its controversial wizardry content. Committee members reviewed draft items and developed new items for both reading and mathematics. Measured Progress and OPI compiled and reviewed the feedback from both meetings. OPI made final decisions on which passages and items should be replaced, and which items developed by committee members should be incorporated into the assessment.

After the editorial-and-approval phase, the tasklets were beta tested by Montana educators and their students. Beta test feedback included concerns regarding the consistency of graphics and the feasibility of educators being able to supply real life objects in place of the provided materials when needed. Beta test educators also suggested that some of the mathematics materials be modified so student responses would not be distinguishable by color. This modification to test materials was made to ensure equal access to students who have varying levels of vision. OPI and Measured Progress revised the reading and mathematics tests based on feedback from the field. The development steps are described in Table 4-2.

Table 4-2. 2008–09 Montana CRT-ALT: Development Process Overview

Development Step	*Step Details*
Original blueprint review	Measured Progress curriculum and special education specialists reviewed original blueprints and the covered benchmarks. OPI reviewed and approved the new blueprints.
Blueprint redesign	Measured Progress redesigned the blueprints with the same amount of emphasis on standards and benchmarks as the original blueprints. Blueprints were approved by the OPI.
Passage/tasklet ideas and item creation	Measured Progress curriculum and special education specialists used the blueprint to further select the expanded benchmarks and performance indicators that fall within the benchmarks. Measured Progress developed passages/tasklet ideas and test items based on the expanded benchmarks and performance indicators. OPI was given the passages/topics to approve as a final draft. OPI made its initial approval of drafts.
Bias and item review workshops	Measured Progress and the OPI: provided bias and sensitivity and item development training to Montana committee members; facilitated the bias and sensitivity review, as well as the item review and development, with Montana committee members; and incorporated committee member feedback into the drafted passages/topics and items.
Passage/topic selection and development	Measured Progress used the draft passages/tasklet ideas and items in combination with the newly developed tasklet ideas and items developed by committee members to create final passages/tasklet ideas and items for the state. OPI was given the passages/tasklet ideas and items for approval. OPI made approvals.
Editorial review of items	All items were reviewed by members of the Measured Progress publications staff to ensure: clarity and unambiguousness of items; correct grammar, punctuation, usage, and spelling; technical quality with respect to stems, options, and scoring guides; and compliance with OPI sensitivity standards and style guidelines.
Beta test	Approximately 23 students participated in the beta test. Beta test teachers tested a student on both content areas and sent feedback to Measured Progress on the assessment items and activity. Beta test participants gave additional feedback in a conference call. OPI and Measured Progress reviewed all grades and content areas and provided feedback via a form and conference call.
Revisions after beta test	Using the feedback from the beta test teachers, the OPI and Measured Progress revised the assessment.

4.3 Item/Activity Editing

Editors reviewed and edited the tasklets and items to ensure uniform style (based on the *Chicago Manual of Style*) and adherence to sound testing principles. These principles included the stipulation that the tasklets and items

- contained correct grammar, punctuation, usage, and spelling;
- were written in a clear, concise style;
- measured the performance indicator;
- had appropriate materials;
- contained unambiguous explanations for teachers as to what was required of the student;
- 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 scaffolding script for teachers; and
- were free of potentially insensitive content.

Items should assess only knowledge or skills that are identified as part of the domain being tested and should not assess 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.

Chapter 5. ASSESSMENT DESIGN BLUEPRINTS

5.1 Reading Assessment Blueprint

As indicated earlier, the framework for reading was based on Montana's reading Content Standards and Expanded Benchmarks, which identify the following five content standards that apply specifically to reading and reading comprehension:

- **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. (This standard is not measurable in a statewide assessment.)
- **Reading Standard 4:** Students select, read, and respond to print and nonprint 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 test blueprint for the CRT-Alternate was designed to mirror the same level of emphasis on concepts across all grades that are represented in the general CRT. The CRT-Alternate design reflects how students with significant cognitive disabilities are working on similar concepts and skills as students in general education classrooms who participate in the CRT, but that have been expanded to the foundational level. Table 5-1 shows the standards measured at each grade level. For a complete list of performance indicators for all reading, mathematics, and science test items (and the correlating standards assessed through each item), see Appendix F.

Table 5-1. 2008–09 Montana CRT-ALT: Distribution of Reading Standards Measured at Each Grade

	Standard 1	*Standard 2*	*Standard 3*	*Standard 4*	*Standard 5*
Grade 3	13	8	*	4	0
Grade 4	9	12	*	3	1
Grade 5	13	8	*	4	0
Grade 6	13	7	*	1	4
Grade 7	13	7	*	1	4
Grade 8	11	10	*	3	1
Grade 10	14	6	*	3	2

Note: Standards 1 and 2 for reading are measured at every grade level, and the other standards are measured evenly across grade spans (elementary 3–5, middle 6–8, and high school 10).
*Standard 3 is not measurable in a statewide assessment.

5.2 Mathematics Assessment Blueprint

The mathematics framework was based on Montana's mathematics Content Standards and Expanded Benchmarks, which identify seven content standards, as shown below:

- 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

The mathematics test blueprint for the CRT-Alternate was designed to mirror the same level of emphasis on concepts across all grades that are represented in the general CRT. The CRT-Alternate design reflects how students with significant cognitive disabilities are working on similar concepts and skills as students in general education classrooms who participate in the CRT, but that have been expanded to the foundational

level. Table 5-2 shows the standards measured at each grade level. For a complete list of performance indicators for all reading, mathematics, and science test items (and the correlating standards assessed through each item), see Appendix F.

Table 5-2. 2008–09 Montana CRT-ALT: Distribution of Mathematics Standards Measured at Each Grade

	Standard 1	*Standard 2*	*Standard 3*	*Standard 4*	*Standard 5*	*Standard 6*	*Standard 7*
Grade 3	8	10	0	10	0	0	5
Grade 4	5	8	0	0	0	8	4
Grade 5	9	10	5	0	10	0	0
Grade 6	6	10	0	5	5	0	5
Grade 7	9	10	10	0	0	5	0
Grade 8	5	4	4	0	4	8	0
Grade 10	2	10	4	4	0	0	5

Note: Standards 1 and 2 for mathematics are measured at every grade level, and the other standards are measured evenly across grade spans (elementary 3–5, middle 6–8, and high school 10).

5.3 Science Assessment Blueprint

The science framework was based on Montana's science Content Standards and Expanded Benchmarks, which identify six content standards, as shown below:

- Science Standard 1: Scientific Investigations
- Science Standard 2: Physical Science
- Science Standard 3: Life Science
- Science Standard 4: Earth/Space Science
- Science Standard 5: Impact on Society
- Science Standard 6: Historical Development

The science test blueprint for the CRT-Alternate was designed to mirror the same level of emphasis on concepts across all grades that are represented in the general CRT. The CRT-Alternate design reflects how

students with significant cognitive disabilities are working on similar concepts and skills as students in general education classrooms who participate in the CRT, but that have been expanded to the foundational level. Table 5-3 shows the standards measured at each grade level. For a complete list of performance indicators for all reading, mathematics, and science test items (and the correlating standards assessed through each item), see Appendix F.

Table 5-3. 2008–09 Montana CRT-ALT: Distribution of Science Standards Measured at Each Grade

	Standard 1	*Standard 2*	*Standard 3*	*Standard 4*	*Standard 5*	*Standard 6*
Grade 4	1	8	5	9	2*	1*
Grade 8	3	5	8	10	0*	0*
Grade 10	5	11	5	9	1*	0*

*Standards 5 and 6 subscores are not reported.

SECTION II: TEST ADMINISTRATION

Chapter 6. TEST ADMINISTRATION

6.1 Responsibility for Administration

The CRT-Alternate is administered by a special education teacher or another certified individual who has worked extensively with the student and is trained in the assessment procedures. Because this is an on-demand performance assessment, the administrator is also the scorer. This becomes a consideration with regard to reliability, where values tend to be inflated due to administrator effects. This is discussed further in Chapter 10—Reliability.

The test administrator may find it helpful to ask another person in the school to assist with the administration. The additional persons who assist in administration may include, but are not limited to, the following:

- parent
- general education teacher
- paraprofessional
- special service provider (speech/language therapist, psychologist, occupational or physical therapist, etc.)
- school counselor
- principal
- other educational professional

6.2 Procedures

A training CD with an audio PowerPoint presentation was sent to teachers who would be administering the CRT-Alternate. Test administrators were instructed to follow the steps below to prepare for the assessment:

- View training CD and participate in question/answer sessions.
- Receive the secure CRT-Alternate Test Booklet from the test coordinator.
- Receive hard copy of the test materials, CD with test materials, and training CD. Teachers may have needed to further adapt materials to meet the needs of students taking the assessment. Guidelines and examples for adapting materials were given in the "Materials" section of the test booklet and on pages 28–30 of the *CRT-Alternate Administration Manual*.
- Download the *CRT-Alternate Administration Manual* and scoring rubric from the OPI or Measured Progress Web site.

- Read the *CRT-Alternate Administration Manual* to become familiar with the administration and scoring directions.
- Read the *CRT-Alternate Test Booklet* to become familiar with the tasklets and performance indicators.
- Consider how the student will access and respond to the test and determine the adaptations and supports the student will need.
- Check to ensure all materials and resources needed are available to complete the tasklets. For example, the grade 8 mathematics tasklet asks the student to use a ruler to find the length of a street on a provided map. The test administrator needs to locate the ruler the student is most familiar with in order to administer the test item.
- Provide the assistive technologies the student needs to access the materials and respond to the test items.
- Schedule the assessment administration session for a time and place that are optimal for student effort and focus.

6.3 Training

System and school test coordinators were instructed to read the *Test Coordinator's Manual* before testing and become familiar with the instructions provided in the *CRT-Alternate Administration Manual*. The *Test Coordinator's Manual* and the *CRT-Alternate Administration 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 manuals 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. Test administrators received copies of the *Test Coordinator's Manual*, the *CRT-Alternate Administration Manual*, and the test-administrator training CD. Training materials and the PowerPoint presentations from the training CD were posted on the OPI Web site. Below is a summary of the information presented in the training CD:

- Important Dates
- CRT-Alternate Overview
- Week 1 of Testing
- Eligibility for the CRT-Alternate
- Who Should Administer the CRT-Alternate
- Materials Needed for the Presentation and for Testing
- About the Tests
- Test Booklet Organization

- Assessment Format
- Introductory Item
- Test Administration Strategies
- Scaffolding
- Scoring & Scoring Rules
- Dealing with Resistance
- Student Evidence
- Test Materials
- Student Response Booklet (SRB)
- Student Barcode Labels
- Returning Student Materials
- Final Administration Hints
- Questions and Answers

To answer any questions not addressed in the training, contact information for OPI, Measured Progress, and the University of Montana–Missoula were provided to teachers, test administrators, and test coordinators. The contact information was provided on the training CD, in the manual, and on the memo sent out with the test materials.

SECTION III: DEVELOPMENT AND REPORTING OF SCORES

Chapter 7. SCORING

7.1 Scoring the Assessment

The CRT-Alternate is administered to a student one-on-one, possibly with the help of another administrator. The teacher scores every item as it is administered using the rubric and a process called scaffolding.

7.2 Using Scaffolding to Gather Student Performance Information

Scaffolding is a process of providing the student with the support needed to respond to the questions in the test. It is similar to support during daily instruction, in which many strategies are used frequently to ensure that students experience success. For example, if a student is unable to make a correct choice from a display of four pictures, the teacher reduces the complexity by removing one of the choices. Scaffolding serves this same function and is provided so that students will experience success in completing the test items. An important result of scaffolding is that it helps students demonstrate their knowledge and skills. These skills can be described and measured, resulting in an accurate picture of what students can do.

The scoring system in the CRT-Alternate allows for increasing amounts of scaffolding, which is provided only when the student does not respond at all or responds incorrectly. This approach is sometimes described as a "least to most" prompt hierarchy (see Chapter 3 for a description of the scaffolding-as-scoring paradigm).

Each tasklet begins with an item that introduces the subject and materials that will be used in the test activity. These items are scored as either a 4 (student responds accurately and with no assistance) or a 0 (student does not respond or actively resists). Items scored this way (at a level 4 or 0) may also be found further into the tasklet when new materials are being introduced.

After the introductory items are scored, each subsequent item within the tasklet is scored on a five-point descending scale from 4 through 0, where 4 represents a correct, independent response; 1, a correct response that has been completely guided by the teacher; and 0, when the student does not respond or actively resists participation in the test activity. (The scoring rubric is presented later in this section.)

The scores from all items, including the introductory items and the subsequent items within each tasklet, are added together to produce a raw score (i.e., total score) for the test. The raw score is then scaled and a performance level assigned for the content area (see Chapter 9 for details on scaling).

A script is provided for scaffolding each of the test items. It describes the prompts to scaffold the student to a level 3, level 2, and level 1. It may be used verbatim or modified by the teacher to meet the needs of the student. For each test item, level 1 prompting is full support from the teacher, guiding the student to the

correct response. Depending on the student and the test item, this may involve physically guiding the student to the correct response or some other form of support that ensures that the student responds correctly.

It is critical that the test administrator deliver each item in a way that allows the student the opportunity to score at level 4. That is, it is first assumed that the student can respond independently to each item, even if that is not the usual instructional practice. The following are directions given to test administrators in order to standardize scaffolding procedures across the state:

- Follow the guidelines to observe the student demonstrating the performance required and allow adequate wait time for the student to process the information and respond without assistance. Do not repeat the question multiple times.
- If the student does not respond or responds incorrectly, scaffold the student to level 3—"Student responds accurately when teacher clarifies, highlights important information, or reduces the range of options to three." Again, give the student adequate wait time.
- If the student does not respond or responds incorrectly, scaffold to level 2—"Student responds accurately when teacher provides basic yes/no questions or forced choices between two options."
- If the student still does not respond with the desired behavior, scaffold to level 1—"Student is guided to correct response by teacher (e.g., modeling the correct response or providing full physical assistance)."
- If the student resists participation for an item, the test administrator will indicate a 0—"Student does not respond or actively resists."

Scaffolding, in other words, is the process for determining the amount of information the student needs to reach the correct response. If the student can respond independently (level 4), no further information is needed by the student. If the student does not respond accurately or independently, more information is given about the item (per a script in the *CRT-Alternate Test Booklet*) and/or the choices are reduced (level 3). This funneling toward the correct response continues (per script) as the student needs more assistance, by providing specific information about the item and/or a forced choice between two options (level 2) and finally by guiding the student to the correct response (level 1). In this way, the student is not expected to either "get it" or "not get it" as in most on-demand assessments. The CRT-Alternate considers the level of assistance that students require to demonstrate their knowledge and skills and thus provides more precise information about student performance and achievement. This system is designed to be sensitive to small increments of change in student performance, an important consideration in describing the learning outcomes of students with severe disabilities.

This process must be used systematically with *every* item identified for scoring within each tasklet. The intent is to give the student every opportunity to perform independently on each item. Scaffolding examples are provided in the *CRT-Alternate Administration Manual*.

The consistent use of required levels of assistance during administration/scoring will increase item intercorrelations and overall test reliability. (The effects of scaffolding and the scoring scheme are further discussed in Chapter 10—Reliability.)

7.3 Scoring Rubric

Each tasklet begins with an introductory item. Only the rubric levels of 4 and 0 are used to score these items. All five levels of the rubric are used to score remaining items. Figure 7-1 shows the scoring rubric with all five levels. Test administrators are encouraged to have the rubric available as a reference when giving the test.

Table 7-1. 2008–09 Montana CRT-ALT: Scoring Rubric

Montana Alternate Assessment Scoring Guide *Performance (independence and accuracy)* *Used to score every item during the structured observation test activity.*				
4	**3**	**2**	**1**	**0**
Student responds accurately and with no assistance.	Student responds accurately when teacher clarifies, highlights important information, or reduces the range of options to three.	Student responds accurately when teacher provides basic yes/no questions or forced choices between two options.	Student is guided to correct response by teacher (e.g., modeling the correct response or providing full physical assistance).	Student does not respond or actively resists.

7.4 Scoring Rules

Instructions and examples provided to test administrators in both the *CRT-Alternate Administration Manual* and training CD illustrate the following rules for scoring:

- Begin with the introductory items and score 4 or 0.
- Use the full scale of 4, 3, 2, 1, and 0 to score the remaining items within each tasklet. Start with level 4 and work systematically through the scaffolding system for every performance indicator as necessary, based on the student's response.
- Allow for appropriate wait time as you scaffold through each level of the scoring rubric.
- Do not repeat questions or directions numerous times.
- Visual, verbal, gestural, and physical cues are allowed in each level except 4.
- Record only one score for each item.
- Score 0 only if the student does not respond or actively resists.
- Halt the administration if the student is showing a pattern of resisting, is becoming fatigued, or is not participating in any way, and resume testing at another time.

- Score every item in a tasklet until the student scores at level 0 for three consecutive items. Stop the administration of the assessment at this point. On the following assessment session, re-administer the final three items on which the student scored a 0. If the student receives a level 0 on these three consecutive items again, halt the administration of the tasklet—leaving the remaining items in the tasklet blank—and move on to the next tasklet.

Test administrators were reminded that the student must start all five tasklets in each content area, and if the student scores at level 0 for three consecutive items, the teacher must attempt to re-administer the tasklet.

7.5 Scanning Procedures and Quality Control

This section of the report outlines the scanning procedures and quality control processes for all returned CRT-Alternate student response booklets. Once the 2008–09 test 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. For all student response booklets, this was the last step in the processing loop in which the documents themselves were handled.

7.5.1 Gatekeeping

Gatekeeping is the first step in the scanning process where the association of Scan Box and bundles of student response booklets from Login are validated before the box continues on to the guillotine station. This validation confirms that the proper Scan Boxes and student response booklet bundles are associated and aids in booklet loss prevention.

- Each box transferred from Login to Gatekeeping has a scannable label applied to it that includes specific contract, content area, and batch number and is associated with the Login Headers that were placed in the box during the Login phase.
- All bundles of student response booklets are removed from the box and the header of each bundle is scanned; if any discrepancy between the headers scanned in this process and the headers assigned to the box in Login are discovered, the box is rejected and returned to Login to be corrected.
- If no discrepancies are discovered, the bundles are replaced in the Scan Box, and the box is flagged in the system as having been gatekept. A box with missing or additional headers cannot be marked as gatekept.
- A yellow index sheet (Box Header) is generated, listing all header information for the box and is placed in the top of the box.

The box is then transferred to the guillotine station.

7.5.2 Guillotining

- Bundles of student response booklets are removed from the box and placed into a blue holding bin. (Blue holding bins are used to keep bundles or student response booklets together while they are not in a box.)
- One bundle is handled at a time.
- Student response booklets are unbundled and their spines cut off.
- The cut pages are immediately rebundled and returned to the Scan Box.

The guillotine operator records the box ID in the guillotine log as having been guillotined and transfers the box to Scanning.

7.5.3 Scanning Procedures

- The scanning operator scans the box label, marking that the box has been transferred to scanning. This scan also tells the scanning program which contract, content area, and grade is being prepared for scanning.
- All bundles are removed from the scanning box and placed into a blue temporary holding bin.
- One bundle is handled at a time.
- Each bundle is individually jogged (placed on a vibrating tray to separate and align pages).
- Each bundle is then placed in the scanner with the Login Header on top and the actual scanning begins.
- The lithocode number is checked at the time of scanning, confirming that the student response booklets being scanned are the correct grade, that the form number is within range, and that the correct number of pages are present for each grade, content area, and form number. Lithocode numbers are unique. This step also prevents booklets with any missing pages from being scanned; any such booklets are hand-edited.

Completed scanned boxes are placed on carts, re-palletized, and placed into short-term storage before being placed in the warehouse.

7.5.4 Machine Scored Items

- The image set generated from scanning is overlaid with an electronic template.
- Bubbled data is read and written to a database.
- Void Answer Documents, multiple marks, and incomplete scans are detected in the data at this time and identified in Data Processing.

The data from the bubbled database is then merged in the data analysis process after being transferred to the Research and Analysis Department (formerly MDA).

7.5.5 Quality Control

- Header QC – compares the count of headers at Gatekeeping, Scanning, and Extraction for each box. If there are any discrepancies in this check, the missing headers are traced back through the process, located, and processed.
- Booklet QC – confirms that the count of booklets scanned matches the count of booklets logged in for each header. Disagreements in these counts are resolved by a Login recount and, if necessary, rescan.
- Extraction QC – confirms that all booklets logged in and scanned have been extracted. If any booklets were not extracted, the image is checked to determine the cause and corrected.
- Multiple Response QC – confirms that any record extracted to a production database that has five or more asterisks (Double Marks) will be manually verified.
- Length Check QC – each data string has a designated number of responses. Before any data is exported to the data processing group, each record in the database is checked to make sure it has the correct string length consistent with the scanning specifications assembled for that contract.
- Spot Check QC – random booklets are selected from various batches during production. Each booklet selected is manually verified, bubble by bubble, to ensure that all hardware and software is functioning properly.
- Duplicate Record QC – before data is exported to the data processing group, any duplicate records have to be verified and resolved. These booklets are pulled and sent through the bull-pen process.
- Label Verification QC – before data is exported to the data processing group, each student ID is compared with a student label file. Any label that does not link back to the student label file is flagged for KFI (Key from Image). This process allows our employees to hand-enter any student labels that did not read correctly through the software.

7.6 Electronic Data Files

Once the data is entered and the scanning logs and other paperwork completed, the test booklets themselves are put into storage (where they are kept for at least 180 days beyond the close of the fiscal year). When it is determined that the electronic files resulting from scanning are complete and accurate, the files are duplicated electronically and made available for many other processing options.

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* and the *Code of Fair Testing Practices in Education* include standards for identifying quality items. While the specific statistical criteria identified in these publications were developed primarily for general—not alternate—assessment, the principles and some of the techniques apply within the alternate assessment framework as well.

Both qualitative and quantitative analyses were conducted to ensure that CRT-Alternate items met these standards. Qualitative analyses are described in earlier sections of this report; this section focuses on the quantitative evaluations. The statistical evaluations discussed are difficulty indices and item-test correlations. The item analyses presented here are based on the statewide administration of the CRT-Alternate in spring 2009.

8.1 Difficulty Indices (*p*-value)

All tasks 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 score for the item. CRT-Alternate items are scored polytomously, such that a student can achieve a score of 0, 1, 2, 3, or 4 for an item. By computing the difficulty index as the average proportion of points achieved, the items are placed on a scale that ranges from 0.0 to 1.0. 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 have either a very high or very low difficulty index are considered to be potentially problematic, because they are either so difficult that few students get them right or so easy that nearly all students get them right. In either case, such items should be reviewed for appropriateness for inclusion on the assessment. If an assessment were composed entirely of very easy or very hard items, all students would receive nearly the same scores, and the assessment would not be able to differentiate high-ability students from low-ability students.

It is worth mentioning that using a norm-referenced criterion such as *p*-values to evaluate test items is somewhat contradictory to the purpose of a criterion-referenced assessment like the CRT-Alternate, which has the goal not of differentiating among students but of providing evidence on student progress relative to a standard. Thus, the generally accepted criteria regarding classical item statistics are only cautiously applicable to the CRT-Alternate. Difficulty indices (i.e., item level classical stats) for each item are provided in Appendix G.

8.2 Item-Test Correlations (Discrimination)

A desirable feature of an item is that the higher-ability students perform better on the item than do 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, this 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. The discrimination index used to evaluate CRT-Alternate tasks was the Pearson product-moment correlation. The theoretical range of this statistic is –1.0 to 1.0.

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. For the CRT-Alternate, the test total score, excluding the item being evaluated, was used as the criterion score.

8.3 Summary of Item Analysis Results

A summary of the item difficulty and item discrimination statistics for each grade/content area combination is presented in Table 8-2. The mean difficulty values shown in the table indicate that, overall, students performed well on the items on the CRT-Alternate. In interpreting these values, it is important to note that item scores lower than 2 are fairly rare on the CRT-Alternate, and a score of 0 is awarded only if the student refuses to respond. These aspects of the item score scale should be considered when evaluating the difficulty values presented in Table 8-2. In contrast to alternate assessments, the difficulty values for assessments designed for the general population tend to be in the 0.4 to 0.7 range for the majority of items. Because the nature and purpose of alternate assessments are different from those of general assessments, and because very few guidelines exist as to criteria for interpreting these values for alternate assessments, the values presented in Table 8-2 should not be interpreted to mean that the students performed better on the CRT-Alternate than the students who took general assessments did on those tests.

Also shown in Table 8-2 are the mean discrimination values. A couple of factors should be considered when interpreting these values. First, all items on the CRT-Alternate are polytomously scored. In general, polytomous items will tend to have higher discrimination values than dichotomous items (e.g., multiple-choice items) because the former are less affected by a restriction of range. Second, the CRT-Alternate item score scale awards points based on the extent to which students require assistance to complete the tasklet. Because students who require assistance with one task are more likely to require assistance with other tasklets, discrimination values will be higher for items scored in this way.

As with the item difficulty values, because the nature and use of the CRT-Alternate are different from those of a general assessment such as the general CRT, and because very few guidelines exist as to criteria for interpreting these values for alternate assessments, the statistics presented in Table 8-2 should be interpreted with caution.

Table 8-1. 2008–09 Montana CRT-ALT:
Item Difficulty and Discrimination Statistics

Grade	*Content Area*	*Number of Items*	*Difficulty*		*Discrimination*	
			Mean	*SD*	*Mean*	*SD*
3	Mathematics	25	0.85	0.10	0.70	0.10
	Reading	25	0.83	0.09	0.63	0.12
4	Mathematics	25	0.81	0.11	0.66	0.12
	Reading	25	0.86	0.09	0.62	0.22
	Science	26	0.87	0.08	0.70	0.12
5	Mathematics	25	0.79	0.08	0.78	0.09
	Reading	25	0.78	0.09	0.71	0.11
6	Mathematics	25	0.86	0.08	0.71	0.09
	Reading	25	0.86	0.07	0.67	0.10
7	Mathematics	25	0.79	0.14	0.70	0.20
	Reading	25	0.83	0.11	0.69	0.12
8	Mathematics	25	0.78	0.12	0.58	0.19
	Reading	25	0.82	0.10	0.61	0.23
	Science	26	0.85	0.11	0.63	0.12
10	Mathematics	25	0.86	0.08	0.70	0.14
	Reading	25	0.84	0.08	0.70	0.10
	Science	28	0.89	0.06	0.70	0.13

Chapter 9. SCALING

9.1 Translating Raw Scores to Scaled Scores and Performance Levels

CRT-Alternate scores in each content area are reported on a scale that ranges from 200 to 300. Scaled scores supplement the CRT-Alternate 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 CRT-Alternate 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 the same temperature can be expressed on either the Fahrenheit or Celsius scale and the same distance can be expressed either in miles or kilometers, student scores on the CRT-Alternate tests can be expressed as raw scores or scaled scores.

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 CRT-Alternate 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 consistent scale. For example, a score of 225 on the CRT-Alternate is fixed as the cut score between the *Novice* and *Nearing Proficiency* performance levels.[1] This is true regardless of content area, grade, or year. If one were to use raw scores, the raw cut score between *Novice* and *Nearing Proficiency* may be, for example, 57 in mathematics at grade 8, but 66 in mathematics at grade 10, or 60 in reading at grade 8. Using scaled scores provides consistency for understanding student performance across content areas and grade levels. Raw score cutpoints for the CRT-Alternate in reading and mathematics, in grades 3–8 and 10, were established via standard setting in July 2006. (Details of the standard setting were included as an appendix in the 2006–07 CRT-Alternate technical report.) In June 2008, OPI and Measured Progress convened panels of Montana educators to participate in a standard setting process for the new science assessment in grades 4, 8, and 10 (Details of the standard setting were included as an appendix in the 2007–08 CRT-Alternate technical report.). Panels were reconvened in May 2009 in order to determine new raw score cutpoints at each performance level for reading and mathematics in grades 4, 8, and 10 due to the redevelopment of the assessments into the tasklet model (see Appendix A for the 2008–09 standard setting report).

Once raw score cutpoints are established, transformation coefficients based on them are calculated in order to place students' raw scores onto the score scale used for reporting. Student scores on the CRT-Alternate are reported in integer values from 200 to 300, with three scores representing cut scores on each

[1] Note that the cut score between *Nearing Proficiency* and *Proficient* is also fixed, at 250. The cut between *Proficient* and *Advanced* varies by grade level and content.

assessment. Two of the three cutpoints (*Novice*/*Nearing Proficiency* and *Nearing Proficiency*/*Proficient*) are pre-set at 225 and 250, respectively, in all grades/content areas. The third cutpoint, between *Proficient* and *Advanced*, is allowed to vary across tests, depending on where the raw score cuts are placed. Allowing the upper cut to float results in a single conversion equation for each test; this simplifies interpretation of scaled scores and their summary statistics. Table 9-1 presents the scaled score range for each performance level in each grade/content area combination.

Table 9-1. 2008–09 Montana CRT-ALT: Scaled Score Ranges

Grade	*Content Area*	*Scaled Score Range for each Performance Level*			
		Novice	*Nearing Proficiency*	*Proficient*	*Advanced*
3	Mathematics	200–224	225–249	250–268	269–300
	Reading	200–224	225–249	250–264	265–300
4	Mathematics	200–224	225–249	250–268	269–300
	Reading	200–224	225–249	250–266	267–300
	Science	200–224	225–249	250–273	274–300
5	Mathematics	200–224	225–249	250–296	297–300
	Reading	200–224	225–249	250–262	263–300
6	Mathematics	200–224	225–249	250–257	258–300
	Reading	200–224	225–249	250–274	275–300
7	Mathematics	200–224	225–249	250–274	275–300
	Reading	200–224	225–249	250–276	277–300
8	Mathematics	200–224	225–249	250–277	278–300
	Reading	200–224	225–249	250–274	275–300
	Science	200–224	225–249	250–270	271–300
10	Mathematics	200–224	225–249	250–260	261–300
	Reading	200–224	225--249	250–282	283–300
	Science	200–224	225–249	250–268	269–300

The scaled scores are obtained by a simple linear transformation of the raw scores using the fixed scaled score values noted above (225 and 250) and the associated 2008–09 raw score cutpoints.

The scaling coefficients were calculated using the following formula for the slope (m) of scaled scores as a function of raw scores.

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

Where:
x_1 is the raw cut score for the *Novice*/*Nearing Proficiency* cut, and
x_2 is the raw cut score for the *Nearing Proficiency*/*Proficient* cut

In other words, the slope is the ratio between the scaled score and raw score differences at the fixed cutpoints.

The intercept (b) of the function is found either by

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

or

$$b = 250 - m(x_2)$$

and represents the resultant scaled score if, at the rate of the slope, the raw score fell from one of the cutpoints to zero.

Scaled scores were then calculated using the resulting linear function:

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

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.

Chapter 10. RELIABILITY

10.1 A Note on Scorer Interrater Reliability

Because the scoring of student performance on the CRT-Alternate relies so heavily on human judgment, interrater reliability may be the form of reliability of most concern in evaluating the meaning of results. OPI designed and administered a study to review interrater reliability on the CRT-Alternate for the 2006–07 administration. Although the study was not performed again this year, the test itself has not changed; therefore, the implications from the interrater reliability study are still relevant. For one component of the study, a group of five highly qualified administrators independently observed and scored seven test administrations (a total of thirty-five students). The scoring was double-blind, meaning that the independent observers/scorers did not communicate their scores to the official test administrator of record or vice versa. For a second component, per Technical Advisory Committee (TAC) (see Appendix E) recommendation, a highly qualified administrator conducted a "read-behind" of thirty evidence templates and recording sheets from among the independently observed administrations. For both analyses the two instances were compared for accuracy. Finally, following another recommendation of the TAC, OPI developed a survey to query the level of training each administrator had received prior to testing.

The double-blind, read-behind, and survey results can be found in the paper titled "Examining the Interrater Reliability of Montana's CRT-Alternate" (Appendix D).

10.2 Other Reliability Results

For paper-and-pencil assessments administered to the general population, such as the general CRT, reliability is commonly evaluated in terms of the way items function together and complement one another. Such analyses may also be carried out on an alternate assessment such as the CRT-Alternate, with the following caveats: its items are quite different from those found on the general assessment, and item scores may be confounded with administrator/scorer affects.

Over and above the confounds inherent in alternate assessments, some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. Items that function well together produce assessments that have less measurement error (i.e., the error is small on average). Such assessments are described as "reliable."

There are a number of ways to estimate an assessment's reliability. One approach is to split all test items into two groups and then correlate students' scores on the two half-tests. This is known as a split-half estimate of reliability. If the two half-test scores correlate highly, the items on them are likely measuring very similar knowledge or skills. It suggests that measurement error will be minimal.

The split-half method requires psychometricians to select items that contribute to each half-test score. This decision may have an impact on the resulting correlation, since each different possible split of the test

halves will result in a different correlation. Another problem with the split-half method of calculating reliability is that it underestimates reliability, because test length is cut in half. All else being equal, a shorter test is less reliable than a longer test. Cronbach (1951) provided a statistic, alpha (α), which avoids these concerns of the split-half method by comparing individual item variances to total test variance. Cronbach's α was used to assess the reliability of the 2008–09 CRT Alternate:

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

Where:
i indexes the item,
n is the number of items,
$\sigma^2_{(Y_i)}$ represents individual item variance, and
σ_x^2 represents the total test variance

Table 10-1 presents Cronbach's α coefficient for each content area (reading, mathematics, and science) and each grade level. The values in Table 10-1 are all greater than or equal to 0.95, indicating that these tests have a high level of reliability. Note, however, that these high values do not necessarily indicate that the CRT-Alternate is "better" than general assessments, which tend to have reliabilities ranging from around 0.80 to around 0.95. There are several factors that may contribute to these high values. First, because the CRT-Alternate is individually administered, the reliability values are likely to be inflated due to administrator effects. In other words, the item scores awarded by the administrator may be influenced by his or her overall sense of the student's level of ability or proficiency, which may result in item scores that are more homogeneous than they would be if they were based strictly on the student's performance on each item. Second, the reliabilities are artificially inflated due to the fact that items are "bundled" together within activities. Items that are bundled together will be more highly correlated, which will increase test reliability. Finally, the use of level of assistance required in the item scoring guide (as described above) will also increase item intercorrelations and overall test reliability.

Table 10-1. 2008-09 Montana CRT-ALT: Common Item Raw Score Descriptive Statistics, Reliability, and SEM by Grade and Content Area

Grade	*Content Area*	*N*	*Possible Score*	*Min Score*	*Max Score*	*Mean Score*	*Score SD*	*Reliability (α)*	*SEM*
3	Mathematics	84	100	0	100	83.881	20.419	0.96	4.084
	Reading	83	100	0	100	81.373	20.618	0.95	4.610
4	Mathematics	100	100	0	100	79.290	21.315	0.96	4.263
	Reading	95	100	0	100	84.337	18.726	0.96	3.745
	Science	104	104	0	104	88.048	22.063	0.97	3.821
5	Mathematics	95	100	0	100	76.695	28.015	0.98	3.962
	Reading	93	100	0	99	75.903	25.683	0.97	4.448
6	Mathematics	106	100	0	100	85.726	20.570	0.97	3.563
	Reading	103	100	1	100	85.272	18.747	0.96	3.749
7	Mathematics	69	100	8	100	78.261	21.968	0.96	4.394
	Reading	69	100	8	100	82.043	20.191	0.96	4.038
8	Mathematics	98	100	16	100	77.163	18.676	0.94	4.575
	Reading	96	100	20	99	82.156	17.666	0.94	4.327
	Science	100	104	0	104	87.010	18.665	0.95	4.174
10	Mathematics	126	100	0	100	84.635	20.655	0.96	4.131
	Reading	128	100	0	100	82.609	21.974	0.97	3.806
	Science	128	112	0	112	97.984	22.361	0.97	3.873

10.3 Reliability of Performance-Level Categorization

Classification of students into performance levels based on their test scores is also subject to measurement error. After students were classified into the CRT-Alternate performance levels (*Novice* [N], *Nearing Proficiency* [NP], *Proficient* [P], and *Advanced* [A]), empirical analyses were conducted to determine the statistical accuracy and consistency of the classifications. Following is a brief explanation of the methodologies used to assess the reliability of classification decisions, after which results are presented.

10.3.1 Accuracy, Consistency, and Kappa

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 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 CRT-Alternate because it is a flexible approach that is appropriate for tests that are composed entirely of polytomous items.

All the accuracy and consistency estimation techniques described here 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 (described in Livingston and Lewis, 1995), a 4×4 contingency table was created for each content area and grade level. The [*i*,*j*] entry of an accuracy table represents the estimated proportion of students whose true scores fell into performance level *i* and whose observed scores fell into performance level *j* on the CRT-Alternate. 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.

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 4×4 contingency table was created for each content area 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.

Another way to measure consistency is to use Cohen's (1960) coefficient κ (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance. Cohen's κ 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 κ is corrected for chance, the values of κ are lower than other consistency estimates.

10.3.2 Results of Accuracy, Consistency, and Kappa Analyses

The accuracy and consistency analyses described above are tabulated in Appendix H. The first section of each table shows the overall accuracy and consistency indices, as well as κ. The overall index, as described above, is the sum of the diagonal elements of the appropriate contingency table, and κ, as described above, is a version of the overall consistency value that has been corrected for chance. Note that, as expected, the values of κ are lower than the overall consistency estimates.

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 0.8498 for the *Proficient* level for grade 4 mathematics.

This figure indicates that among the students whose true scores placed them in the *Proficient* level, 84.98% of them would be expected to be placed in *Proficient* if they were categorized according to their observed scores. The corresponding consistency value of 0.8214 indicates that 82.14% 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, for purposes of accountability, there is generally greatest interest in distinguishing between students who are *Proficient* or *Advanced* and those who have not yet reached the *Proficient* threshold. The third section of the summary tables shows information at each of the cutpoints. These values indicate the accuracy and consistency of the dichotomous decisions, either above or below the associated cutpoint. 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 scores would place them below the cut (false positive), and vice versa.

Table 10-2 summarizes most of the results of accuracy and consistency at a glance. As with other types of reliability, it is inappropriate when analyzing the decision accuracy and consistency of a given test to compare results between grades and content areas.

Table 10-2. 2008–09 Montana CRT-ALT: Summary of Decision Accuracy (and Consistency) Results

Grade	*Content Area*	*Overall*	*Conditional on Level*				*At Cutpoint*		
			N	*NP*	*P*	*A*	*N:NP*	*NP:P*	*P:A*
3	Mathematics	0.80 (0.75)	0.90 (0.87)	0.66 (0.56)	0.66 (0.59)	0.93 (0.84)	0.96 (0.94)	0.94 (0.92)	0.91 (0.88)
4	Mathematics	0.84 (0.78)	0.89 (0.85)	0.76 (0.68)	0.74 (0.68)	0.94 (0.88)	0.97 (0.95)	0.95 (0.92)	0.93 (0.91)
5	Mathematics	0.86 (0.82)	0.94 (0.93)	0.54 (0.42)	0.77 (0.73)	0.94 (0.87)	0.97 (0.95)	0.96 (0.95)	0.93 (0.91)
6	Mathematics	0.84 (0.80)	0.90 (0.87)	0.83 (0.78)	0.65 (0.58)	0.94 (0.88)	0.98 (0.97)	0.95 (0.93)	0.91 (0.89)
7	Mathematics	0.86 (0.80)	0.86 (0.81)	0.81 (0.74)	0.83 (0.81)	0.93 (0.84)	0.98 (0.97)	0.95 (0.93)	0.93 (0.90)
8	Mathematics	0.83 (0.76)	0.84 (0.78)	0.73 (0.64)	0.78 (0.73)	0.93 (0.86)	0.97 (0.95)	0.94 (0.92)	0.92 (0.89)
10	Mathematics	0.86 (0.82)	0.87 (0.83)	0.80 (0.74)	0.67 (0.59)	0.96 (0.92)	0.98 (0.97)	0.95 (0.94)	0.93 (0.91)
3	Reading	0.84 (0.79)	0.84 (0.76)	0.82 (0.77)	0.78 (0.74)	0.93 (0.86)	0.98 (0.97)	0.95 (0.93)	0.91 (0.89)
4	Reading	0.86 (0.82)	0.87 (0.81)	0.81 (0.75)	0.79 (0.74)	0.95 (0.89)	0.98 (0.97)	0.96 (0.94)	0.93 (0.90)
5	Reading	0.87 (0.82)	0.90 (0.87)	0.80 (0.73)	0.69 (0.59)	0.96 (0.92)	0.97 (0.96)	0.95 (0.94)	0.94 (0.92)
6	Reading	0.89 (0.84)	0.85 (0.79)	0.79 (0.72)	0.82 (0.78)	0.95 (0.91)	0.99 (0.98)	0.97 (0.95)	0.93 (0.91)
7	Reading	0.90 (0.86)	0.83 (0.74)	0.81 (0.75)	0.83 (0.79)	0.96 (0.92)	0.99 (0.99)	0.97 (0.96)	0.94 (0.91)
8	Reading	0.86 (0.81)	0.81 (0.73)	0.73 (0.64)	0.75 (0.67)	0.95 (0.91)	0.98 (0.97)	0.96 (0.94)	0.93 (0.90)
10	Reading	0.88 (0.84)	0.90 (0.87)	0.69 (0.59)	0.80 (0.74)	0.96 (0.93)	0.98 (0.97)	0.97 (0.95)	0.94 (0.92)
4	Science	0.89 (0.85)	0.90 (0.87)	0.73 (0.63)	0.76 (0.69)	0.97 (0.93)	0.98 (0.97)	0.97 (0.95)	0.94 (0.92)
8	Science	0.86 (0.81)	0.83 (0.76)	0.80 (0.73)	0.80 (0.76)	0.94 (0.89)	0.99 (0.98)	0.96 (0.94)	0.92 (0.89)
10	Science	0.87 (0.83)	0.90 (0.88)	0.69 (0.59)	0.71 (0.65)	0.96 (0.92)	0.98 (0.97)	0.96 (0.95)	0.93 (0.91)

N = novice; NP = nearing proficiency; P = proficient; A = advanced

Chapter 11. REPORTING

The CRT-Alternate assessment was designed to measure student performance against Montana's Content Standards and Expanded Benchmarks. Consistent with this purpose, results from the CRT-Alternate were reported in terms of performance levels that describe student performance in relation to the established state standards. There are four performance levels: *Advanced*, *Proficient*, *Nearing Proficiency*, and *Novice*. (CRT-Alternate performance level descriptors and the performance level cuts on both the raw and scaled-score scales are presented in Appendix B.) Students receive a separate performance-level classification 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 by student subgroups are also reported at the school and system levels. The CRT-Alternate reports are:

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

To establish protocols for handling data discrepancies and data clean-up processes, OPI and Measured Progress collaborated to formulate decision rules in late spring 2009. A copy of these decision rules is included as Appendix I.

State summary results were provided to OPI via a secure Web site. The report formats are included in Appendix J. All reports were made available to system and school administrators via Montana's online reporting system, Montana Analysis and Reporting System (MARS). Student reports were shipped to system test coordinators in September 2009 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 2009 Criterion-Referenced Test and CRT-Alternate Assessment Reports* 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.

Chapter 12. VALIDITY SUMMARY

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

The CRT-Alternate assessments are based on, and aligned to, Montana's Content Standards and Expanded Benchmarks in reading, mathematics and science. The CRT-Alternate results are intended to provide inferences about student achievement on Montana's reading, mathematics and science Content Standards and Expanded Benchmarks, and these achievement inferences are meant to be useful for program and instructional improvement and as a component of school accountability.

The *Standards for Educational and Psychological Testing* (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 content area 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 content standards, evidence based on test content was extensively described in chapters 3 through 5. Item alignment with Montana content standards; item bias, sensitivity, and content appropriateness review processes; adherence to the test blueprint; use of standardized administration procedures; and appropriate test administration training are all components of validity evidence based on test content. As discussed earlier, all CRT-Alternate test questions are aligned by Montana educators to specific Montana content standards and undergo several rounds of review for content fidelity and appropriateness. Finally, tests are administered according to state-mandated standardized procedures, and all test administrators are required to review the training CD.

The scoring information in chapter 7 describes the steps taken to train the teachers administering the assessment on scoring procedures, as well as quality control procedures related to scanning. In order to obtain additional validity evidence, it would be helpful to conduct a study in which a percentage of teachers administering the assessment would be videotaped to confirm validity of administration and scoring.

Evidence based on internal structure is presented in the discussions of item analyses and reliability in chapters 8 and 10. Technical characteristics of the internal structure of the assessments are presented in terms of classical item statistics (item difficulty, item-test correlation) and reliability coefficients. In general, indices were within the ranges expected.

To further support the validity argument, additional studies to provide evidence regarding the relationship of CRT-Alternate results to other variables might include the extent to which scores from the CRT-Alternate 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.

The evidence presented in this report supports inferences of student achievement on the content represented in the Montana content standards for reading, mathematics, and science for the purposes of program and instructional improvement and as a component of school accountability.

SECTION IV: REFERENCES

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

Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). Fort Worth: Holt, Rinehart and Winston.

Chicago Manual of Style (15th ed.). (2003). Chicago: University of Chicago Press.

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.

Joint Committee on Testing Practices (1988). *Code of fair testing practices in education*. Washington, D.C.: National Council on Measurement in Education.

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

Washington, DC: American Psychological Association. Available for download at **http://www.apa.org/science/fairtestcode.html**.

APPENDICES

Appendix A—Standard Setting Report

MontCAS Criterion-Referenced Test—Alternate Assessment

2008-09
STANDARD-SETTING REPORT
READING AND MATH: GRADES 4, 8, AND 10
MAY 19 & 20, 2009
HELENA, MONTANA

TABLE OF CONTENTS

1. STANDARD-SETTING PROCESS 1

2. TASKS COMPLETED PRIOR TO THE STANDARD-SETTING MEETING 3

2.1 CREATION OF PERFORMANCE LEVEL DESCRIPTORS (PLDS) 3
2.2 PREPARATION OF MATERIALS FOR PANELISTS 3
2.3 PREPARATION OF PRESENTATION MATERIALS 3
2.4 PREPARATION OF INSTRUCTIONS FOR FACILITATORS DOCUMENTS 4
2.5 PREPARATION OF SYSTEMS AND MATERIALS FOR ANALYSIS DURING THE MEETING 4
2.6 SELECTION OF PANELISTS 4

3. TASKS COMPLETED DURING THE STANDARD-SETTING MEETING 5

3.1 ORIENTATION 5
3.2 STANDARD-SETTING PROCESS 5
3.2.1 Discuss Performance Level Descriptors 6
3.2.2 Practice Round 6
3.2.3 Training Evaluation 7
3.2.4 Round 1/Intial Judgments 7
3.2.5 Tabulation of Round 1 Results for Grade 10 8
3.2.6 Round 2 Judgments 8
3.2.7 Tabulation of Round 2 Results 9
3.2.8 Final Judgments 10
3.2.9 Recommendations for Modifications to PLDs 15
3.2.10 Complete the Evaluation 15

4. TASKS COMPLETED AFTER THE STANDARD-SETTING MEETING 17

4.1 ANALYSIS AND REVIEW OF PANELISTS' FEEDBACK 17
4.2 PREPARATION OF RECOMMENDED CUT SCORES 17
4.3 PREPARATION OF STANDARD-SETTING REPORT 17

APPENDICES 19

APPENDIX A— PERFORMANCE LEVEL DESCRIPTORS 21
APPENDIX B— MEETING AGENDA 29
APPENDIX C— SCORING FLOWCHART 33
APPENDIX D— SAMPLE OF VISUAL ITEM MAP 37
APPENDIX E— SAMPLE STUDENT PROFILE/RATING SHEET 41
APPENDIX F— EVALUATION FORM 45
APPENDIX G— OPENING SESSION POWERPOINT PRESENTATIONS 51
APPENDIX H— FACILITATOR SCRIPT 63
APPENDIX I— PANELIST AFFILIATIONS 75
APPENDIX J— PANELIST DESCRIPTOR RECOMMENDATIONS 79
APPENDIX K— EVALUATION RESULTS 89

1. STANDARD-SETTING PROCESS

Standard-setting activities for the Montana Criterion-Referenced Test-Alternate Assessment (CRT-Alternate) in reading and math occurred May 19th and 20th, 2009. At the standard-setting meeting, cut-points were recommended for the alternate assessment in grades four, eight, and ten using the data from the spring 2009 administration. This report documents the procedures and results of the standard-setting meeting.

Each panel consisted of six to seven participants. Each panel completed the standard-setting process for one grade level, reading on the first day and math on the second day. The Modified Body of Work standard-setting method was implemented across all grades and contents. To help ensure consistency of procedures between panels, all participants attended a large-group training session at the beginning of the meeting. In addition, each panel was led through the standard-setting process by a trained facilitator from Measured Progress.

This report is organized into three major sections, describing tasks completed prior to, during, and following the standard-setting meeting.

2. TASKS COMPLETED PRIOR TO THE STANDARD-SETTING MEETING

2.1 Creation of Performance Level Descriptors (PLDs)

The PLDs presented to panelists provided the official description of the set of knowledge, skills, and abilities that students are expected to display in order to be classified into each performance level. These descriptors were created prior to the standard-setting meeting by staff of the Office of Public Instruction (OPI). The descriptors are provided in Appendix A of this report.

2.2 Preparation of Materials for Panelists

The following materials were assembled for presentation to the panelists at the standard-setting meeting:

- Meeting Agenda
- PLDs
- Ordered CRT-Alternate Test Booklet
- Auxiliary Assessment materials
- Scoring Flowchart
- Administration Manual
- Visual Item Map
- Student Profiles/Rating sheets
- Evaluation forms

The meeting agenda, scoring flowchart, sample visual item map, sample student profiles/rating sheet, and evaluation form are provided in Appendices B through F of this report, respectively.

2.3 Preparation of Presentation Materials

The PowerPoint presentations used in the opening session were prepared prior to the meeting. Two sets of PowerPoint slides are included as Appendix G of this document: the first set provides an overview of the CRT-Alternate, the criteria for participation in the assessment, and an explanation of the administration and scoring procedures. The second provides an overview of the issues of standard setting, specifics about the standard-setting process, and an overview of the activities the panelists would be completing during the standard-setting meeting.

2.4 Preparation of Instructions for Facilitators Documents

A document was created for the group facilitators to refer to while working through the process. The document for both reading and math is provided in Appendix H.

2.5 Preparation of Systems and Materials for Analysis During the Meeting

The computational programming to carry out all analyses during the standard-setting meeting was completed and thoroughly tested prior to the standard-setting meeting. The program designed to calculate cuts and impact data was written using SAS statistical software.

2.6 Selection of Panelists

Panelists were recruited and selected to reflect as diverse of a population as possible. Measured Progress and Montana's OPI staff worked together to recruit panelists, with OPI's final approval over participant selection.

The goal of the panelist recruitment was to assemble panels of approximately 10 participants. Ideally, each panel was to include a minimum of three special education teachers experienced in working with students with significant disabilities, three subject area content teachers, and two school administrators, higher education personnel, and/or stakeholders from interest groups related to significant disabilities. An additional goal was for the panels to reflect a balance of gender, race/ethnicity, and geographic location. Finally, panelists were selected who were familiar either with the grade level subject matter or the special education population for which they would be setting standards. The numbers of panelists who participated in the standard setting was six or seven per group, as shown in Table 1 below. A list of the panelists' affiliations and their roles can be found in Appendix I.

Table 2-1. Numbers of Participants by Group

Panel	*Number of Panelists*
Reading and Math - Grade 4	7
Reading and Math - Grade 8	6
Reading and Math - Grade 10	7
Total	20

3. TASKS COMPLETED DURING THE STANDARD-SETTING MEETING

3.1 Orientation

The standard-setting meeting began with a general orientation session that was attended by all panelists. The purpose of the orientation was to ensure that all panelists heard the same message about the need for and goals of standard setting and about their part in the process. The orientation consisted of three parts. First, OPI welcomed the panelists and thanked them for participating, provided some context about the CRT-Alternate, the need for setting standards, and some general information about their role in the process. Next, a Measured Progress Special Education Program Manager provided an overview of the CRT-Alternate, including its participation criteria, and administration and scoring procedures. Finally, a Measured Progress psychometrician gave an overview of standard setting, an introduction to the standard-setting method that was being used in Montana, and provided an overview of the activities that the standard-setting panelists would be completing.

Once the general orientation was complete, each panel reconvened into breakout rooms, where the panelists received more detailed training and completed the standard-setting activities.

3.2 Standard-Setting Process

For grades 4 and 8, the standard-setting process followed a standards validation model which included two rounds of individual recommendations following extensive group discussion. The starting cut points were calculated for each cut score by extrapolating (within a content area) from the grade 3, 5, 6, and 7 cut scores using the following five step process:

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

Although starting cuts were initially calculated for grade 10, estimation required that the grade 7 cut points be extrapolated to grade 10. The proximity of available data coupled with the small numbers of students in each grade, very few of whom were located in the lowest two performance categories, resulted in a negative regression line that placed none of the grade 10 students in level 2 for reading. Consequently, starting cut points were not used in grade 10. Instead,

the grade 10 standard-setting process followed that of a full standard setting which included three rounds; in the first round, panelists recommended cut-points individually without discussion. Then, in Rounds 2 and 3, they recommended cut-points individually following extensive group discussion.

For purposes of simplicity, and due to the similarity between the Round 1 validation procedures and the Round 2 standard setting procedures, the Round 1 results from grades 4 and 8 will be presented alongside the Round 2 results from grade 10. Round 1 for the grades 4 and 8 validation will be considered the initial, individual ratings, which were not entered for analyses.

3.2.1 Discuss Performance Level Descriptors

The first step in the process, once the panelists convened into their grade groups, was to discuss the Performance Level Descriptors (PLDs). This important step of the process was designed to ensure that panelists thoroughly understood the needed knowledge, skills, and abilities for profiles to be classified as *Novice*, *Nearing Proficiency*, *Proficient*, and *Advanced*. Panelists began by reviewing the descriptors individually and then discussed them as a group, clarifying each level and coming to consensus as to the definitions of each. Bulleted lists of characteristics for each level were generated based on the group discussion and posted in the room for panelists to refer to during all of the small group activities.

3.2.2 Practice Round

Next, the panelists completed a practice round of ratings. The purpose of the practice round was to familiarize the panelists with all of the materials they would be using as part of the standard-setting process and to walk them through the process of rating student profiles. In addition to the PLDs, panelists were given the following materials:

- Ordered CRT-Alternate Test Booklet – a copy of the CRT-Alternate items, presented in order from the easiest to the hardest, based on each item's p-value.
- Auxiliary Assessment materials: response cards, manipulatives, storybook format reading passage, etc.
- Scoring Flowchart
- Administration Manual
- Practice student profiles/rating form – the student profiles/rating sheets show typical patterns of item scores for students scoring at particular total scores; for the practice round, three profiles were included on the form. The profiles consist of a column for each item, again presented in order of difficulty; each row of the profile represents a typical student at a given total score The profiles for each grade were created based on data from the 2009 operational administration by

selecting all students at a particular total score, finding the average score for that subgroup for each item, then selecting a profile with a pattern of item scores that resembled the item averages as closely as possible. The student profiles also included two blank columns where panelists entered their rating for each profile. The profiles used in grades 4 and 8 had an additional column which reflected the initial categorization of the profiles based on the starting cut points. A copy of the practice rating form can be found in Appendix E.

The facilitator reviewed all of the materials and how panelists would use them in making their ratings. Then the facilitator reviewed the first profile with the panelists, pointing out the score on each item, then drawing the panelists' attention to the items in the Ordered Test Booklet as well as to the Scoring Flowchart. The facilitator reviewed the relationship between the particular skills required to successfully complete that item, and how the item performance corresponded to the definitions of the performance levels. The second and third profiles were reviewed with panelists in the same manner. Panelists were asked to rate each profile and then the facilitator led a discussion with the panelists to understand how they had rated each profile, and asked them to share their reasoning and justification with the group.

3.2.3 Training Evaluation

At the end of the practice round, panelists completed the training evaluation form. This section was designed as a check for understanding, to see how confident the panelists felt in their ability to complete the rating process. A copy of the evaluation is included in Appendix F; and results are summarized in Appendix K.

3.2.4 Round 1/Intial Judgments

In the first round, panelists worked individually with the PLDs, the Round 1 Profiles/Rating sheet, the Ordered Test Booklet, Scoring Flowchart, Administration Manual, and Visual Item Map. The profile sheet consisted of approximately 40 profiles, with scores ranging from the minimum observed score to the maximum possible score (i.e., approximately every second score point). As in the practice round, the profiles consist of a column for each item, presented in order of difficulty; each row of the profile represents a typical student at a given total score. The profiles for each grade were created based on data from the 2009 operational administration by selecting all students at a particular total score, finding the average score for that subgroup for each item, then selecting a profile with a pattern of item scores that resembled the item averages as closely as possible. The student profiles also included three blank columns where panelists entered their rating for each

profile during each round. In addition, the profiles used in grades 4 and 8 had an additional column which reflected the initial categorization of the profiles based on the starting cut points. A copy of a sample rating form can be found in Appendix E.

For each profile, the panelists considered the skills and abilities demonstrated by a student who had that particular pattern of scores, and decided which performance level was the best match for each profile. The panelists worked their way through the profiles, making a rating for each one, and recorded their ratings. While the profiles were presented in order of total score, panelists were not required to rate them in strictly increasing order. Instead, panelists were encouraged to take a holistic look at the *pattern* of scores, and the items the scores were associated with, rather than making a judgment based primarily on the total raw score.

3.2.5 Tabulation of Round 1 Results for Grade 10

In grade 10, after all panelists had completed their individual ratings, Measured Progress staff calculated the average cut-points for the group based on the Round 1 ratings. Cuts were calculated using SAS statistical software by first determining each panelist's individual cuts using logistic regression, then averaging across panelists to get the overall cuts. A psychometrician shared the location of the overall cuts with the group to assist them in their group discussion and Round 2 ratings. The Round 1 results are outlined in Table 3-1.

Table 3-1. Round One Results

Content	*Grade*	*Performance Level*	*Average Cut*	*Standard Error*	*Raw Score*		*Percent of Students*
					Min	Max	
Reading	10	Novice	NA	NA	0	52	8.5
		Nearing Proficiency	52.4	2.2	53	71	12.4
		Proficient	71.9	1.6	72	94	43.4
		Advanced	94.7	3.6	95	100	35.7
Mathematics	10	Novice	NA	NA	0	48	8.7
		Nearing Proficiency	48.1	2.7	49	74	9.4
		Proficient	74.4	1.4	75	90	25.2
		Advanced	90.9	1.2	91	100	56.7

3.2.6 Round 2 Judgments

Prior to beginning the group discussion, and using a show of hands, the facilitator recorded how many panelists placed each profile into each performance level on chart paper. Starting with the first profile for which there was disagreement as to how it should be categorized, or in grades 4

and 8 disagreements with the classifications based on the starting cut points, the panelists began discussing the categorization of the profiles according to their initial ratings. Panelists were encouraged both to share their own point of view as well as to listen to the thoughts of their colleagues. Facilitators made sure the panelists knew that the purpose of the discussion was not to come to consensus: at every point throughout the standard-setting process, panelists were asked to provide their own individual best judgment. Once the discussions were complete, the panelists filled in their profiles/rating sheet.

3.2.7 Tabulation of Round 2 Results

After all panelists had completed their individual ratings, Measured Progress staff calculated the average cut-points for the group based on their most recent ratings. Cuts were calculated using SAS statistical software by first determining each panelist's individual cuts using logistic regression, then averaging across panelists to get the overall cuts. In addition, impact data were calculated, consisting of the percentage of students who would fall into each performance level based on the group average ratings. A psychometrician shared this information with the group to assist them in their group discussion and subsequent ratings. The results are outlined in Table 3-2.

Table 3-2. Round Two Results

Content	*Grade*	*Performance Level*	*Average Cut*	*Standard Error*	*Raw Score*		*Percent of Students*
					Min	Max	
Reading	4	Novice	NA	NA	0	49	8.6
		Nearing Proficiency	49.7	1.1	50	76	15.2
		Proficient	76.3	0.9	77	94	43.8
		Advanced	94.5	0.4	95	100	32.4
	8	Novice	NA	NA	0	46	8.0
		Nearing Proficiency	47.0	0.0	47	61	5.0
		Proficient	61.5	0.5	62	82	25.0
		Advanced	82.8	0.3	83	100	62.0
	10	Novice	NA	NA	0	52	8.5
		Nearing Proficiency	52.1	2.0	53	71	12.4
		Proficient	71.3	1.2	72	91	27.1
		Advanced	91.3	0.9	92	100	51.9
Mathematics	4	Novice	NA	NA	0	57	15.2
		Nearing Proficiency	57.7	1.9	58	76	13.3
		Proficient	76.0	1.2	77	92	49.5
		Advanced	92.6	0.7	93	100	21.9
	8	Novice	NA	NA	0	50	12.0
		Nearing Proficiency	50.5	0.0	51	68	14.0
		Proficient	68.5	0.0	69	87	37.0
		Advanced	87.9	0.4	88	100	37.0
	10	Novice	NA	NA	0	50	8.7
		Nearing Proficiency	51.0	0.0	51	78	15.0
		Proficient	78.4	0.4	79	92	23.6
		Advanced	92.1	0.4	93	100	52.8

3.2.8 Final Judgments

Once the panelists completed their ratings, the facilitator once again asked for a show of hands and tallied the number of panelists who categorized each profile into each performance level on chart paper. As in previous rounds, starting with the first profile for which there was disagreement as to its categorization, the panelists discussed their rationale for how they rated each profile. Again, the purpose of the discussion was for the panelists to benefit from the points of view of their colleagues, not to come to consensus about the ratings.

Panelists were also asked to include the impact data as part of their discussion. In presenting the impact data, the psychometrician explained to the panelists that its purpose was to provide a "reasonableness check," and that they should resist letting it influence their decisions in isolation. Instead, if any of the percentages seemed too high or too low, they were told to return to the assessment and to the Performance Level Descriptors, and consider whether they needed to make

adjustments to their last round of ratings.

Once the discussions had been completed, the panelists recorded their ratings sheet and the rating sheets were submitted for data analysis. The results of the panelists' final round of ratings are outlined in Table 3-3.

Table 3-3. Final Results

Content	*Grade*	*Performance Level*	*Average Cut*	*Standard Error*	*Raw Score*		*Percent of Students*
					Min	Max	
Reading	4	Novice	NA	NA	0	49	8.6
		Nearing Proficiency	49.7	1.1	50	76	15.2
		Proficient	76.6	0.8	77	94	43.8
		Advanced	94.5	0.2	95	100	32.4
	8	Novice	NA	NA	0	46	8.0
		Nearing Proficiency	47.0	0.0	47	65	7.0
		Proficient	65.5	0.0	66	83	26.0
		Advanced	84.0	0.7	85	100	59.0
	10	Novice	NA	NA	0	54	8.5
		Nearing Proficiency	54.9	1.2	55	70	11.6
		Proficient	70.5	0.0	71	91	27.9
		Advanced	91.7	0.7	92	100	51.9
Mathematics	4	Novice	NA	NA	0	57	15.2
		Nearing Proficiency	57.1	1.9	58	77	14.3
		Proficient	77.1	0.9	78	92	48.6
		Advanced	92.6	0.7	93	100	21.9
	8	Novice	NA	NA	0	50	12.0
		Nearing Proficiency	50.5	0.0	51	68	14.0
		Proficient	68.5	0.0	69	88	38.0
		Advanced	88.3	0.2	89	100	36.0
	10	Novice	NA	NA	0	50	8.7
		Nearing Proficiency	51.0	0.0	51	79	16.5
		Proficient	79.4	0.3	80	92	22.0
		Advanced	92.1	0.3	93	100	52.8

A graphical display of the percent of students in each performance level for reading and math across grades is also provided in Figures 3-1 and 3-2, respectively.

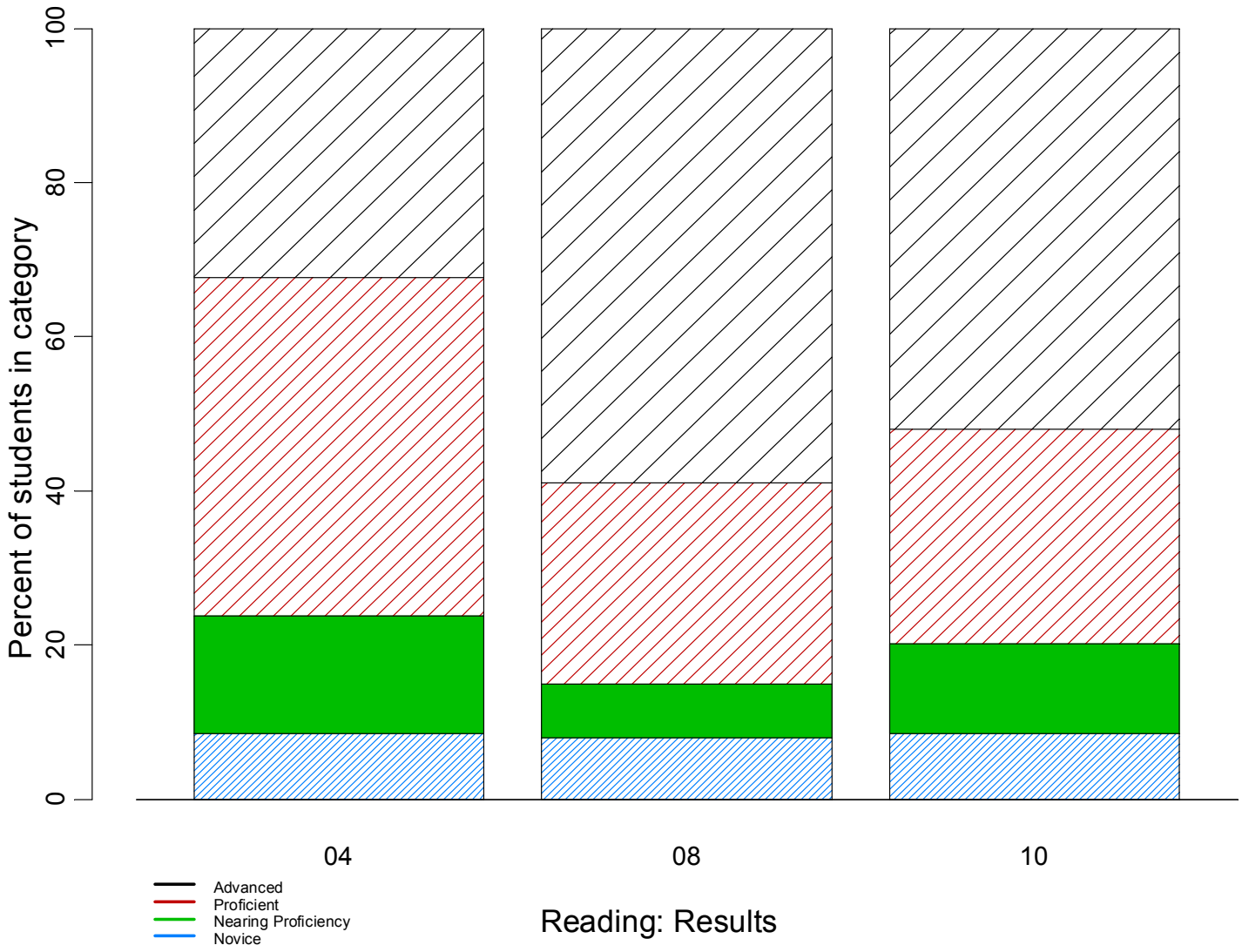


Figure 3-1. The percent of students falling at each performance level in reading

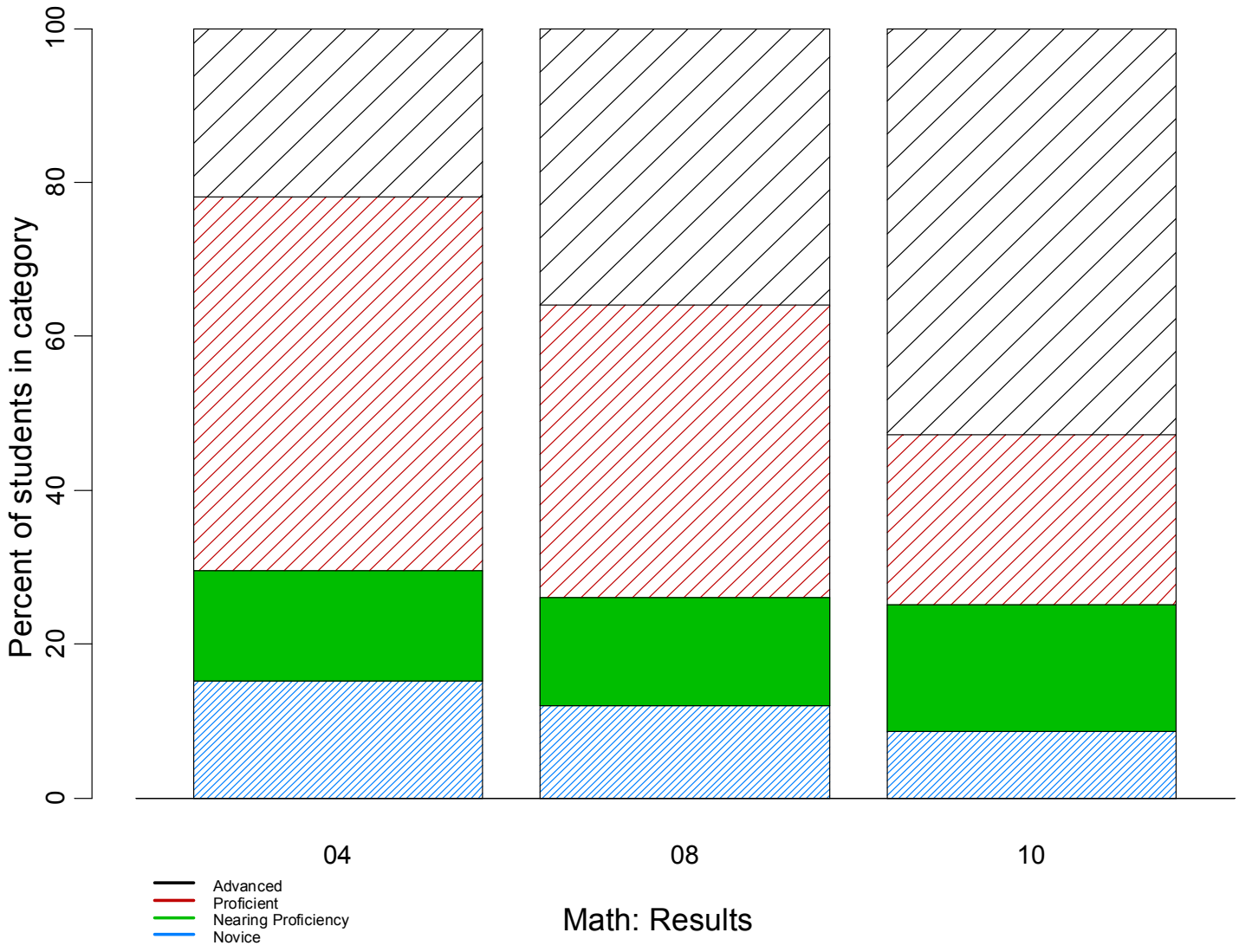


Figure 3-2. The percent of students falling at each performance level in math

3.2.9 Recommendations for Modifications to PLDs

After completing the final round of ratings, the panelists were given an opportunity to provide feedback on the Performance Level Descriptors. Panelists were asked to focus on providing language that is clearer and more teacher- and parent-friendly. Panelists were informed that the suggestions they made were just recommendations and that they may or may not be implemented by OPI. The descriptor recommendations provided by the panelists are included in Appendix J.

3.2.10 Complete the Evaluation

As the last step in the standard-setting process, panelists in all three groups anonymously completed an evaluation form. A copy of the evaluation is presented as Appendix F, and the results of the evaluations are presented as Appendix K.

4. 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 results of the standard-setting meeting and addressing anomalies that may have occurred in the process or in the results, presenting the results to the Technical Advisory Committee (TAC), and making any final revisions or adjustments.

During the standard setting meeting, it was discovered that for the first two rounds of ratings in grade 10 reading, a panelist had been using the math rating sheet, instead of the reading rating sheet. The panelist was given the correct rating sheet for the third round and the results for each round were examined with and without this panelist's ratings. Although, the panelist's ratings impacted the Round 1 and 2 results, they did not impact the final results. Consequently, the panelist was removed from the Round 1 and 2 results, but included in the final results. The results reported in Tables 3-1 through 3-4 correspond to this final decision.

4.1 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 included when the final cut-points were calculated. It appeared that all panelists understood the rating task and attended to it appropriately.

4.2 Preparation of Recommended Cut Scores

The results of the standard setting were presented to the Montana TAC on May 27th. The TAC recommended that the results from the final round of ratings be used as the official cut points for all three grades.

4.3 Preparation of Standard-Setting Report

Following final compilation of standard-setting results, Measured Progress prepared this report, which documents the procedures and results of the May 2009 standard-setting meeting in order to establish performance standards for the CRT-Alternate Assessment in reading and math, grades 4, 8 and 10.

APPENDICES

Appendix B—Performance Level Descriptors

Alternate Performance Level Descriptors for Grade 4 Reading	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • follows 3-step or more directions • communicates knowledge using expanded vocabulary • communicates a complete thought related to topic or concept • correctly answers who, what, and where questions • is able to generalize information from one setting to another • recognizes and articulates the main idea • relates and uses relevant knowledge to make connections
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • follows two-step directions • communicates knowledge of basic vocabulary and familiar words • demonstrates written words have meaning • explores pictures, symbols, and objects • answers yes and no questions • identifies beginning main idea • uses literacy materials appropriately • contributes/elaborates on responses
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • follows one-step directions consistently • understands when response is needed • needs multiple re-direction to the test material to respond to a specific item • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • begins to respond to literacy with varied prompts • uses prior knowledge to demonstrate knowledge of basic vocabulary • begins to communicate with a purpose
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a reading activity • responds to own name • attempts to communicate • attends for short periods of time to the teacher, materials, and test items • attends to pictures, symbols, objects when presented • begins/attempts to participate with support

Alternate Performance Level Descriptors for Grade 4 Mathematics	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • creates and extends a repeating pattern using objects, shapes, designs, or numbers • uses methods and tools to solve a problem involving patterns, relations, or functions • carries out a strategy to solve problems involving patterns, relations, or functions • determines which of two numbers is closer to the quantity in a given set • understands and uses comparison words (more, less, some, none) • explains reasoning about probability items
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • understands the concept of 1 and 1:1 correspondence • sorts objects into sets • understands comparison words (more, less, some, none) • extends or supplies a missing element in a repeating pattern by attribute or number • sets up a graph (i.e. labels axes) • understand words that indicate operations in word problems • demonstrates a basic understanding of math skills, concepts and vocabulary
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • demonstrates an understanding that numbers, as opposed to letters, are used to express quantity, order, or size/amount • counts with another person • recognizes quantities • identifies basic shapes (i.e. circles, squares, triangles, and rectangles) • matches two- dimensional physical shapes to pictures of the shapes in different orientations • communicates numbers correctly
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a math activity • attends to materials being displayed • attends to another person making patterns and to a person describing patterns • attends to a person demonstrating with concrete materials • attends to objects or pictures of two- and three- dimensional geometric shapes and the relationships among them • attends to another person estimating an amount of a given set

Alternate Performance Level Descriptors for Grade 8 Reading	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • connects prior knowledge to make meaning of text • identifies main idea and various supporting details • understands story lessons/author's purpose • locates title and other information from a variety of documents/sources • recognizes vowel letter-sound • uses reading strategies to gain information (i.e. rereading, use of key words, use of features of text) • reads and comprehends a paragraph
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • has basic word recognition • tracks while reading • identifies words from sentences • identifies a word/picture/symbol for content communication • identifies title and basic parts of a reading selection • identifies main idea of a story and some supporting facts/details • identifies purposes of various texts (i.e. dictionary, map) • has a firm grasp of sound/symbol association
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • recognizes that letters have names and is aware of letter sounds • recognizes difference between letters and other symbols (i.e. numerals) • identifies letters by name/sign • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • identifies a word/picture/object of familiar places and people • responds mostly through basic yes/no questions • understands story beginning and end • understands basic main idea (answer with one picture/short response)
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a reading activity • attends to materials being displayed • demonstrates readiness by following one-step directions or with teacher modeling/prompting • responds to name, words, pictures and symbols • directs attention and responds to external stimuli when requested (i.e. turns head in direction, nods head, operates switch, points to, etc.) • interacts with stimuli (i.e. teacher, words, pictures, and symbols)

Alternate Performance Level Descriptors for Grade 8 Mathematics	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • measures to the inch • compares and calculates measurements, time, and monetary amounts • understands concept of fractions • understands Algebra concepts • labels sets of data and components of a graph (i.e. label axis) • creates graph and explains conclusions drawn from graph • applies beginning connections between concrete and symbolic representations, operations, measurement, graphing and problem solving strategies
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • reads/makes simple measurements • uses comparison words (more, less, some, none) correctly • understands numbers can represent monetary amounts, measurement, and time • demonstrates basic problem solving skills • fills in data on a graph • identifies basic information from a graph • makes a statement about data • demonstrates beginning connections between concrete and symbolic representations, operation (+/-), measurement and graphing
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • identifies and/or recognizes a map and measuring tools • demonstrates solid number concept for 1:1 • can count single digits • can add/subtract single digits • can distinguish between two different elements/variables • basic understanding of graphs and data • communicates understanding of beginning connections between concrete and symbolic representations
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a math activity • attends to materials being displayed • attends to another person reviewing a map with prompting • attends to another person reviewing a graph with prompting • engages with instructor with prompts • recognizes numbers (symbol or rote recitation)

Alternate Performance Level Descriptors for Grade 10 Reading	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • identifies main idea and supporting details from various reading selections • identifies appropriate resources for gaining specific information • draws conclusions from a variety of texts (i.e. poem, fiction) • communicates meaning of new and unfamiliar vocabulary • communicates a complete thought related to topic or concept • uses word-recognition skills, context clues, and prior knowledge to understand text • rereads to gain understanding
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • has basic reading and comprehension skills • understands difference between various literacy materials • uses prior knowledge to understand text • communicates an opinion • identifies main ideas and some supporting details/facts • is beginning to identify appropriate resources for gaining specific information • identifies words/pictures/symbols and objects that are new and unfamiliar
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • begins to access to prior knowledge • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • able to match and identify familiar words/pictures/ symbols/objects • understands story beginning and end • understands basic main idea (answer with one picture/short response) • communicates an opinion • identifies resources
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a reading activity • attends to materials being displayed • responds to name, words, pictures and symbols • demonstrates readiness by following one-step directions or with teacher modeling/prompting • directs attention and responds to external stimuli when requested (i.e. turns head in direction, nods head, operates switch, points to, etc.) • interacts with stimuli (i.e. teacher, words, pictures, and symbols)

Alternate Performance Level Descriptors for Grade 10 Mathematics	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • generalizes very basic information • completes two to three-step processes of addition and subtraction • completes basic division problem • applies beginning connections between concrete and symbolic representations by using a chart/table to draw conclusions • creates graph/tables and explains conclusions drawn from graph • understands and communicates relationship between variables • solves problems using bills and their values • follows navigational directions and recalls shapes and locations
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • completes and/or extends basic patterns of data • sorts items into sets by multiple defining characteristics • demonstrates beginning connections between concrete and symbolic representations • identifies basic information from a graph/chart • makes a statement about data • understands and matches bills and their values • recognizes and identifies two-dimensional shapes • chooses correct procedures to solve simple number problems
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • recognizes properties of limited (square/circle) two-dimensional shapes • recognizes distinct categories • recognizes basic patterns of data • sorts items into sets by one defining characteristic • understands quantity • can count single digits • can add/subtract single digits • communicates understanding of beginning connections between concrete and symbolic representations
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a math activity • attends to materials being displayed • attends to another person reviewing a graph/chart • attends to another person reviewing signs and labels • shows limited understanding of quantity when given two choices • engages with instructor with prompts • recognizes numbers (symbol or rote recitation)

Appendix C—MEETING AGENDA

CRT-ALTERNATE ASSESSMENT STANDARD SETTING AGENDA

THE BEST WESTERN HELENA GREAT NORTHERN HOTEL, HELENA, MT

MAY 19 & 20, 2009

4.3.2

TUESDAY, MAY 19

8:00 – 8:30 am	Registration & Continental Breakfast
8:30 – 10:00 am	Introduction and Overview of Standard Setting Process
10:00 – 10:15 am	Break
10:15 – 12:00 pm	Groups go to Breakout Rooms*
12:00 – 1:00 pm	Lunch
1:00 – 2:30 pm	Continue in Breakout Rooms*
2:30 – 2:45 pm	Break
2:45 – 4:00 pm	Continue in Breakout Rooms*
4:00 pm	Adjourn

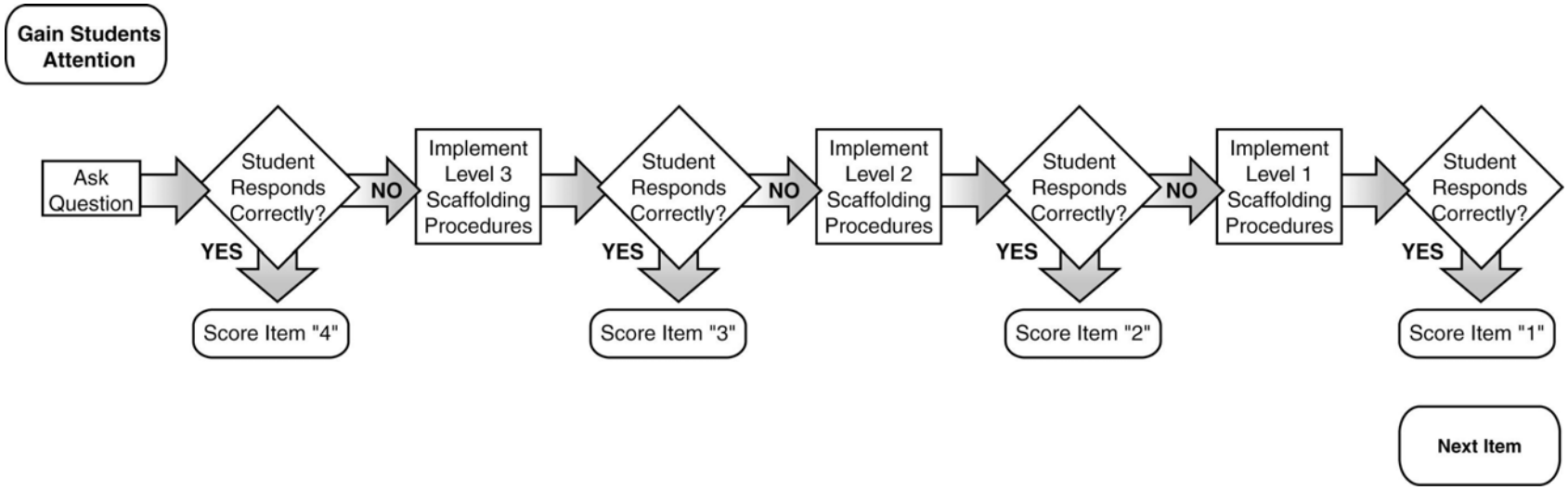
WEDNESDAY, MAY 20

8:00 – 8:30 am	Continental Breakfast
8:30 – 10:00 am	Groups go to Breakout Rooms *
10:00 – 10:15 am	Break
10:15 – 12:00 pm	Continue in Breakout Rooms*
12:00 – 1:00 pm	Lunch
1:00 – 2:30 pm	Continue in Breakout Rooms*
2:30 – 2:45 pm	Break
2:45 – 4:00 pm	Continue in Breakout Rooms*
4:00 pm	Adjourn

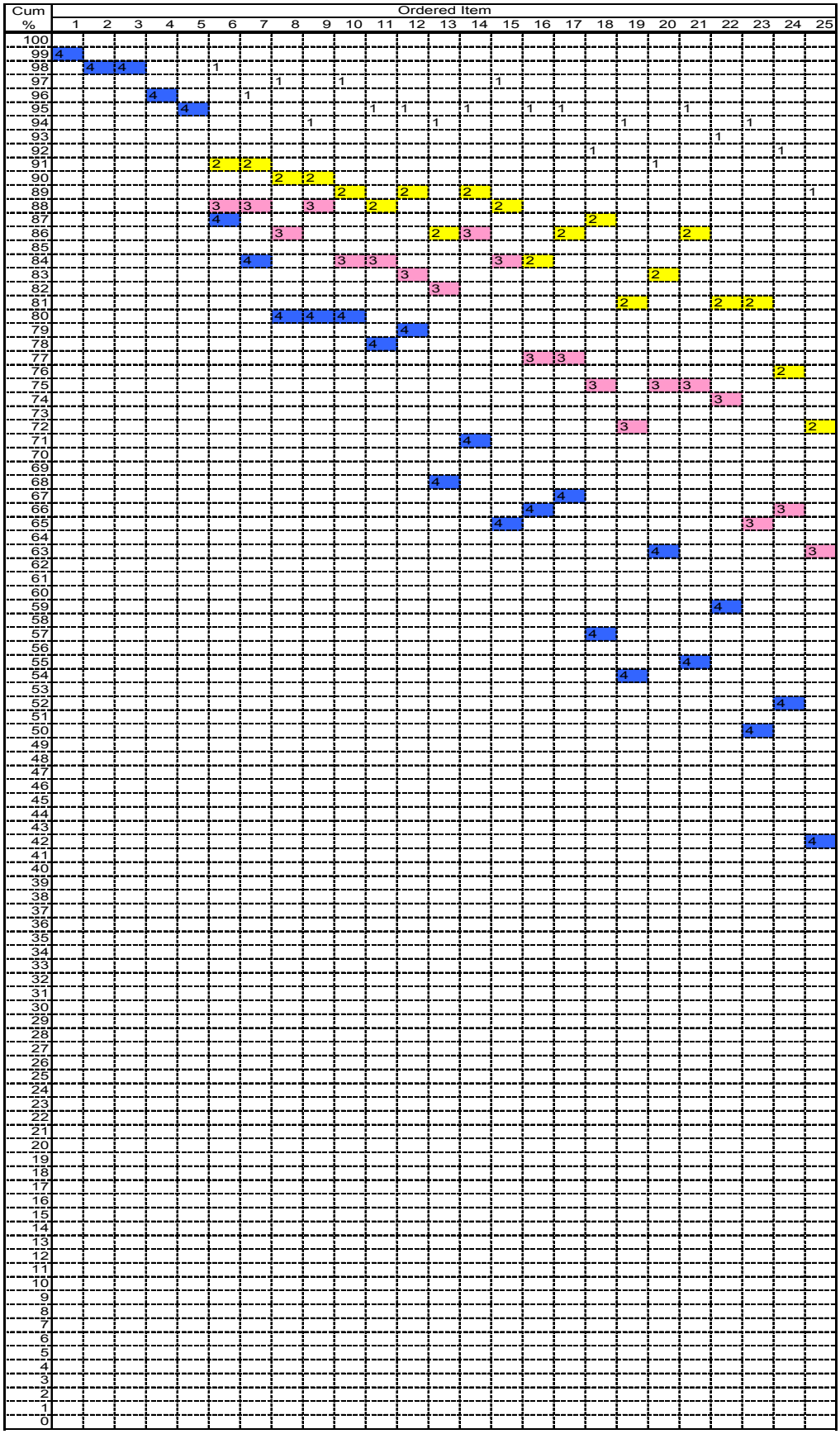
*Breakout rooms will take additional breaks as needed

Appendix D—Scoring Flowchart

SCORING FLOWCHART



Appendix E—Sample of Visual Item Map


Cum %
Ordered 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
100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

Appendix F—Sample Student Profile/Rating Sheet

Practice

Rater I.D._____

Rating (1 = Novice ,2 = Nearing Proficiency,3 = Proficient, 4 = Advanced)

Profile No.	Total Score	Starting Cuts	Initial	Round 1	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
1	**12**	1			4	0	0	0	0	1	2	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
2	**66**	2			4	4	4	4	4	3	4	3	4	2	2	2	4	2	2	1	1	1	1	4	2	2	2	2	2
3	**76**	3			4	4	4	4	4	4	4	4	3	4	4	4	3	4	3	4	2	3	1	1	1	1	2	3	1

Math Grade 8

Rater I.D._____

Rating (1 = Novice ,2 = Nearing Proficiency,3 = Proficient, 4 = Advanced)

Profile No.	Total Score	Starting Cuts	Initial	Round 1	Round 2	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
1	12	1				4	4	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	16	1				0	4	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	3	0	2	0
3	20	1				4	4	4	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	38	1				4	4	4	4	0	3	1	1	0	0	0	1	4	2	0	0	1	2	0	1	1	2	2	1	0
5	40	1				4	4	4	4	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	41	1				4	4	4	4	4	2	2	1	1	1	1	2	1	1	1	1	1	2	0	1	1	1	0	1	0
7	42	1				4	4	4	4	4	1	1	1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	1
8	43	1				4	4	4	4	4	1	2	1	1	1	1	1	1	2	1	1	1	2	1	1	1	1	1	1	1
9	46	2				4	4	4	4	4	2	4	1	1	1	1	1	2	1	1	1	1	2	1	1	1	1	1	1	1
10	49	2				4	4	4	4	4	2	2	1	4	1	1	2	1	2	1	2	1	1	2	1	1	1	1	1	1
11	52	2				4	4	4	4	4	3	2	3	1	1	3	1	1	2	1	1	1	2	1	1	1	3	1	1	2
12	54	2				4	4	4	4	4	4	4	2	1	1	1	2	1	1	1	1	1	2	3	1	1	1	3	1	2
13	58	2				4	4	4	4	4	3	4	3	2	1	3	1	1	1	4	1	1	1	1	4	1	3	1	1	1
14	60	2				4	4	4	0	4	3	4	1	4	4	0	4	4	1	0	2	1	4	1	3	3	2	0	1	2
15	62	2				4	4	4	4	4	3	4	2	3	2	4	2	1	2	1	1	1	4	2	3	2	2	1	1	1
16	63	2				4	4	4	4	4	4	4	3	4	3	1	1	1	1	1	3	4	1	3	1	2	2	1	2	1
17	67	2				4	4	4	4	4	4	3	3	3	1	4	3	1	2	2	3	3	1	4	3	0	2	2	1	2
18	68	2				4	4	4	4	4	2	4	4	2	4	2	1	2	3	1	1	1	4	1	2	2	1	4	3	4
19	69	3				4	4	4	4	4	4	4	4	3	1	3	3	3	3	1	4	0	4	1	3	1	3	3	0	1
20	70	3				4	4	4	4	4	3	1	1	4	2	4	2	2	3	4	2	2	3	4	3	1	2	4	2	1
21	72	3				4	4	4	4	4	4	4	2	4	3	1	3	2	3	4	2	2	2	3	4	2	2	1	1	3
22	73	3				4	4	4	4	4	4	4	4	2	4	2	3	2	3	1	1	4	2	3	1	1	2	4	2	4
23	74	3				4	4	4	4	4	4	2	4	3	2	4	2	3	2	4	1	1	1	3	3	2	4	4	1	4
24	75	3				4	4	4	4	4	4	3	2	3	4	4	3	3	3	2	1	3	3	1	2	4	3	1	4	2
25	78	3				4	4	4	4	4	1	4	3	4	4	4	4	2	3	3	3	4	4	3	2	4	2	1	2	1
26	80	3				4	4	4	4	4	3	4	4	2	4	2	3	3	3	4	2	3	2	2	4	3	3	3	3	3
27	81	3				4	4	4	4	4	4	4	4	3	4	3	4	3	4	4	3	4	3	1	3	1	4	1	3	1
28	83	3				4	4	4	4	4	4	4	3	3	4	4	3	3	1	3	4	1	3	2	3	4	4	3	3	4
29	85	3				4	4	4	4	4	4	4	4	4	3	4	4	4	4	2	4	4	3	2	3	3	4	2	1	2
30	87	4				4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	3	2	3	2	2	2	4	3	3
31	88	4				4	4	4	4	4	4	4	4	4	4	4	4	4	4	1	4	4	3	1	3	3	4	4	1	4
32	89	4				4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	3	3	4	4	3	3	2	4	3	1
33	91	4				4	4	4	4	4	4	4	3	4	4	4	4	4	3	4	4	4	4	4	3	4	1	3	3	3
34	92	4				4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	3	4	2	4	2	2
35	93	4				4	4	4	4	4	4	4	4	4	4	2	4	4	4	2	4	4	4	4	4	4	2	4	4	3
36	94	4				4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	3	3	4	4	4	4
37	96	4				4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	4	3	4	4	3	4	4	4	4	4
38	97	4				4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	4	3	4	4	4	3
39	98	4				4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	3
40	99	4				4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4

Appendix G—EVALUATION FORM

Content Area: _____
Grade: _____

Standard Setting Training Evaluation

The purpose of this evaluation form is to obtain your feedback about the training you have received. Please complete the information below. **Do not put your name on the form.** We want your feedback to be anonymous.

Please mark the appropriate box for each statement.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
I understand the goals of the standard setting meeting.	☐	☐	☐	☐	☐
I understand the procedures we are using to set standards.	☐	☐	☐	☐	☐
I understand how to use the standard setting materials.	☐	☐	☐	☐	☐
I understand the differences between the performance levels.	☐	☐	☐	☐	☐
I understand how to make the cut score judgment.	☐	☐	☐	☐	☐
I know what tasks to expect for the remainder of the meeting.	☐	☐	☐	☐	☐
I am confident in my understanding of the standard setting task.	☐	☐	☐	☐	☐

Please indicate any areas in which you would like more information before you continue.

Please indicate any questions you may have about the remainder of the standard setting meeting.

Standard Setting Final Evaluation

Please complete the information below. Your feedback will provide a basis for evaluating the training, methods, and materials. **Do not put your name on the form.** We want your feedback to be anonymous.

Gender: Male ☐ Female ☐

Race/ethnicity: White ☐ Black ☐ Hispanic ☐ Asian ☐ Pacific Islander ☐

American Indian ☐

Years of experience in education: 0-5 ☐ 5-10 ☐ 10-15 ☐ More than 15 ☐

Area of Expertise (Check all that apply):

- Students with Disabilities ☐
- Students with Limited English Proficiency ☐
- Economically Disadvantaged Students ☐
- Gifted and Talented Students ☐
- General Education ☐

Please mark the appropriate box for each statement.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
I understood the goals of the standard setting meeting.	☐	☐	☐	☐	☐
I understood the procedures we used to set standards.	☐	☐	☐	☐	☐
The facilitator helped me understand the process.	☐	☐	☐	☐	☐
The materials contained the information needed to set standards.	☐	☐	☐	☐	☐
I understood how to use the materials provided.	☐	☐	☐	☐	☐
The performance level descriptors were clear.	☐	☐	☐	☐	☐
I understood how to make the cut score judgments.	☐	☐	☐	☐	☐
I understood how to use the feedback provided after each round.	☐	☐	☐	☐	☐
I understood how to use the impact data.	☐	☐	☐	☐	☐
I understood how the cut scores were calculated.	☐	☐	☐	☐	☐
The facilitator was able to get answers to my questions.	☐	☐	☐	☐	☐
Sufficient time was allotted for training on the standard setting tasks.	☐	☐	☐	☐	☐
Sufficient time was allotted to complete the standard setting tasks.	☐	☐	☐	☐	☐
The facilitator helped the standard setting process run smoothly.	☐	☐	☐	☐	☐

Content Area: _____
Grade: _____
Please rate the usefulness of each of the following:

	Not at all useful				Extremely useful
The opening session.	☐	☐	☐	☐	☐
The small group activities.	☐	☐	☐	☐	☐
Becoming familiar with the assessment.	☐	☐	☐	☐	☐
Articulating the differences between the performance levels.	☐	☐	☐	☐	☐
Discussions with other participants.	☐	☐	☐	☐	☐
Providing additional details to the performance level descriptors.	☐	☐	☐	☐	☐

Please rate the influence of the following when setting standards.

	Not at all influential				Extremely influential
The performance level descriptors.	☐	☐	☐	☐	☐
My expectations of students.	☐	☐	☐	☐	☐
The difficulty of the test materials.	☐	☐	☐	☐	☐
The student responses.	☐	☐	☐	☐	☐
My experience in the field.	☐	☐	☐	☐	☐
Discussions with other participants.	☐	☐	☐	☐	☐
Cut scores of other participants.	☐	☐	☐	☐	☐
Impact data.	☐	☐	☐	☐	☐

Content Area: _____
Grade: _____

Do you believe the final recommended cut score for each of the performance levels is too low, about right, or too high?

	Too Low	About Right	Too High
Math			
Advanced/Proficient	☐	☐	☐
Proficient/Nearing Proficiency	☐	☐	☐
Nearing Proficiency/Novice	☐	☐	☐
Reading			
Advanced/Proficient	☐	☐	☐
Proficient/Nearing Proficiency	☐	☐	☐
Nearing Proficiency/Novice	☐	☐	☐

Please provide any additional comments about the standard setting process or suggestions as to how the training and process could be improved.

Appendix H—Opening Session PowerPoint Presentations

MontCAS CRT-Alternate Assessment

Orientation to Standard Setting
May 19, 2009

May 2009

Participation Guidelines

1. Does the student have an active IEP and receive services under the Individuals with Disabilities Education Act (IDEA)?
2. Do the student's demonstrated cognitive abilities and adaptive behavior require substantial adjustments to the general curriculum?
3. Do the student's learning objectives and expected outcomes focus on functional application of skills, as illustrated in the student's IEP's annual goals and short-term objectives?
4. Does the student require direct and extensive instruction to acquire, maintain, generalize and transfer new skills?

May 2009

What is the CRT-Alternate Assessment?

- The CRT-Alternate is a performance-based test that is aligned with Montana's Content Standards and Expanded Benchmarks and measures student performance based on alternate achievement standards
- The CRT-Alternate Assessment was designed for students who are unable to participate in the regular CRT, even with accommodations. Only IDEA-eligible students with significant cognitive disabilities are eligible to participate in the CRT-Alternate

May 2009

Expanded Benchmarks

- Expanded from end of grades 4, 8, and 12 to foundational skills
- Are not grade level specific, due to the wide diversity of students in this population
- Used to develop the assessment performance indicators

May 2009

Decisions about Participation

- CRT-Alternate is intended for students with significant cognitive disabilities
- Participation decisions made and documented by student IEP teams
- Guidance document provides questions to assist determining which test is appropriate for a student.
- A full range of accommodations were possible with the CRT

May 2009

The Framework

- Starts with content area, then works down to the standard, which indicates what all students should know, understand, and be able to do, then funnels down to performance indicator
- Performance indicators serve as the actual test items within the assessment

➢Content
- ➢ Standard
  - ➢Benchmark
    - ➢Expanded Benchmark
      - ➢Performance Indicator

May 2009

Framework Example

Content → **Reading**

Standard → **Standard 2 -** Students apply a range of skills and strategies to read.

Benchmark → **2.6: -** Develop vocabulary through the use of context clues, analysis of word parts, auditory clues, and reference sources, and expand and refine vocabulary related to specific academic areas, culture, and technology.

Expanded Benchmark → **2.6.1 -** Develop vocabulary.

Performance Indicator → **2.6.1.1 -** Demonstrate understanding of a new word based on context of a reading selection.

May 2009

Materials/Handouts

- Agenda
- Re-ordered Test Booklet
- Actual Test Materials
- Scoring Guide & Flow Chart
- Visual Item Map
- Rating Sheet
- Performance Level Descriptors
- Evaluation
- Reimbursement Forms
- Non-disclosure Forms

May 2009

Tasklet Format

Each content area consists of:

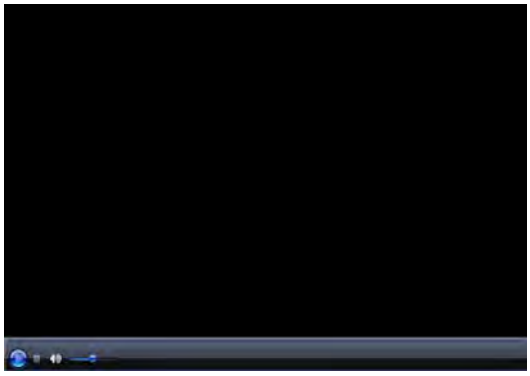
- Five distinct tasklets (short activities)
- Five items per tasklet
- Each tasklet starts with an introductory item that orients the student to the tasklet materials and topic

May 2009

Re-ordered Test Booklet

- Contains Tasklets with:
  - Materials
  - Script/Scaffold
  - Student work/student will
  - Performance Indicators/Scoring Guide
- Items are re-ordered from least to most difficult overall
  - The easiest item will be the one that most students completed accurately and independently
  - The hardest item will be the one that most students needed the greatest amount of support to complete accurately
- Shows the original item number in the test booklet

May 2009



May 2009

Test Booklet Item Example

Materials	Activity Steps Teacher will:	Student Work Student will:	Performance Indicators Use Scoring Guide TRANSFER SCORES TO STUDENT RESPONSE BOOKLET
2. Picture cards used in previous item: • island • lake • mountain • valley **Communication support strategies:** • Student may look at/point to task materials to express a choice. • Request may be rephrased to require a yes/no response (e.g. *"Is this an island?"*). • Student may tell teacher to "stop" at desired response as teacher sequentially points to each of the choices.	2. Keep same four picture cards on the work space. ***"Which one is an island?"*** Encourage student to pick up and look through the pictures. Scaffold: Level 3: Remove an incorrect response. Say *"An island is a piece of land."* Repeat task request. Level 2: Remove another incorrect response. Say *"An island is surrounded by water."* Repeat task request. Level 1: Say *"This is an island."* Assist the student as needed to identify the island.	2. Indicate an island.	2. Identifies an island. ○ ○ ○ ○ ○ 4 3 2 1 0 Performance Indicator: 4.1.1.8 Expanded Benchmark: 4.1.1

May 2009

CRT-Alternate Scoring Guide

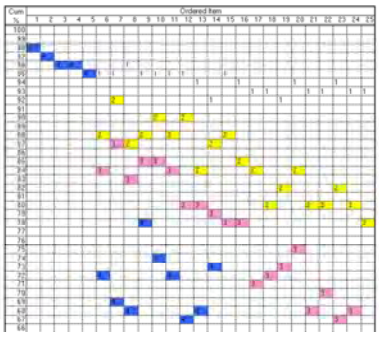
Montana Alternate Assessment Scoring Guide

Performance (independence and accuracy)
Used to score every item during the structured observation test activity.

4	3	2	1	0
Student responds accurately and with no assistance.	Student responds accurately when teacher clarifies, highlights important information or reduces the range of options to three.	Student responds accurately when teacher provides basic yes/no questions or forced choices between two options.	Student is guided to correct response by teacher (e.g., modeling the correct response or providing full physical assistance).	Student does not respond or actively resists.

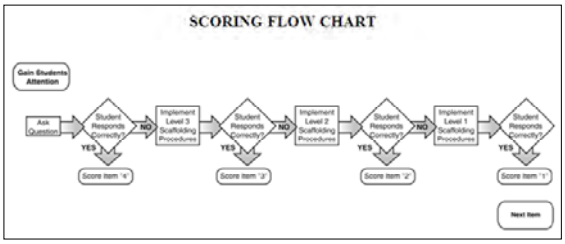
May 2009

Example Visual Item Map



May 2009

Flow Chart



May 2009

Rating Sheet

- Student Profiles listed in order
- Rating column for each round
- Assign each profile a number
  - Advanced (4)
  - Proficient (3)
  - Nearing Proficiency (2)
  - Novice (1)

***Please remember to put your ID number on the top of the form.**

May 2009

Visual Item Map

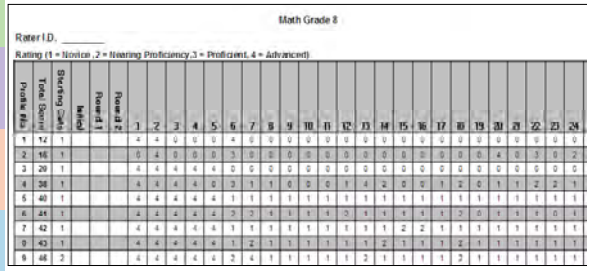
Visual Item Maps are intended to provide a visual representation of how students performed on each test item, and the item's relative level of difficulty

Columns - level of difficulty for each performance indicator (item), from easiest to hardest

Rows - percent of students who received at least that score point on the indicator

May 2009

Example Rating Sheet



Math Grade 8

Rater I.D. _____

Rating (1 = Novice ,2 = Nearing Proficiency,3 = Proficient, 4 = Advanced)

Profile No	Total Stories	Starting Cells	Retest	Round # 1	Round # 2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	12	1				4	4	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	16	1				0	4	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	3	0	2
3	20	1				4	4	4	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	36	1				4	4	4	4	0	3	1	1	0	0	0	1	4	2	0	0	1	2	0	1	1	2	2	1
5	40	1				4	4	4	4	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	44	1				4	4	4	4	4	3	3	1	1	1	3	3	1	1	1	1	1	3	0	1	1	1	0	1
7	42	1				4	4	4	4	4	1	1	1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1
8	43	1				4	4	4	4	4	1	2	1	1	1	1	1	1	2	1	1	1	2	1	1	1	1	1	1
9	46	2				4	4	4	4	4	2	4	1	1	1	1	1	2	1	1	1	1	2	1	1	1	1	1	1

May 2009

Performance Level Descriptors

Advanced (4):
The student at the Advanced level **accurately and independently** demonstrates the ability to carry out **comprehensive** content-specific performance indicators.

Proficient (3):
The student at the Proficient level, **given limited prompting**, demonstrates the ability to respond accurately in performing a **wide variety** of content specific performance indicators.

Nearing Proficiency (2):
The student at the Nearing Proficiency level, **given moderate prompting**, demonstrates the ability to respond accurately in performing a **narrow set** of content-specific performance indicators.

Novice (1):
The student at the Novice level, **given physical assistance and/or modeling**, is **supported to participate** in content specific performance indicators.

May 2009

Breakout Sessions

- Grade 4 – Depot Room
- Grade 8 – Trestle/Signal Room
- Grade 10 – Oriental C

*Lunch will be served in Oriental A/B

May 2009

Evaluations

Content Area: _____
Grade: _____

Standard Setting Training Evaluation

The purpose of this evaluation form is to obtain your feedback about the training you have received. Please complete the information below. **Do not put your name on the form.** We want your feedback to be anonymous.

Please mark the appropriate box for each statement.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
I understand the goals of the standard setting meeting.	☐	☐	☐	☐	☐
I understand the procedures we are using to set standards.	☐	☐	☐	☐	☐
I understand how to use the standard setting materials.	☐	☐	☐	☐	☐
I understand the differences between the performance levels.	☐	☐	☐	☐	☐
I understand how to make the cut score judgment.	☐	☐	☐	☐	☐
I know what tasks to expect for the remainder of the meeting.	☐	☐	☐	☐	☐
I am confident in my understanding of the standard setting task.	☐	☐	☐	☐	☐

Please indicate any areas in which you would like more information before you continue.

May 2009

Questions?

May 2009

Reimbursement Forms

- Honorarium or Substitute Reimbursement
- Mileage/Travel Reimbursement
- Allow 4-6 weeks process time

May 2009

MontCAS CRT- Alternate Assessment Setting Performance Standards

Reading and Math, Grades 4, 8 & 10

Purpose of Standard Setting Meeting

Your job is to make recommendations about the following cuts:

Novice

Nearing Proficiency ← Cut Score

Proficient ← Cut Score

Advanced ← Cut Score

Today's Training

In today's session we will cover:

- An overview of standard setting
- Details of the Body of Work process as it will be implemented for the CRT-Alternate
- Your role in this process

Note

This session is intended to be an overview.

Your room facilitator will give you more details and will guide you through the process step by step.

What is Standard Setting

Set of activities that result in the determination of threshold or cut scores on an assessment.

How much is enough?

What knowledge, skills and abilities need to be demonstrated on the assessment for a student to be classified as Novice?

Nearing Proficiency?

Proficient?

Advanced?

Why are you here?

Provide data to validate cut scores for

Math: Grades 4, 8 & 10

Reading: Grades 4, 8 & 10.

We are trying to answer the question:

- What does Novice performance look like on each test?
- What does Nearing Proficiency performance look like on each test?
- What does Proficient performance look like on each test?
- What does Advanced performance look like on each test?

Standard Setting vs. Standards Validation

The process of originally establishing cut scores is referred to as *Standard Setting*.

Panelists are not given initial cut points.

The process of examining proposed cut scores is referred to as *Standards Validation.*

Starting cut points are provided at the beginning of the process.

Multi-Step Process

Data collection phase
- Your job for the next two days

Policy/Decision making phase
- Final Decisions are put in place
- Acceptance/Rejection or modification of data component
  - State Department
  - Technical Advisory Committee
  - Legislature

Body of Work Method

A standard setting procedure that allows panelists to use samples of actual student work to make their determinations.

Is especially useful for assessments that consist primarily or entirely of constructed-response items.

Has been used successfully for setting standards on Alternate Assessment (and was used for setting standards on the CRT-Alternate in the past).

Has resulted in defensible cut points.

Data Collection Phase

There are many standard setting methods:
- Angoff
- Body of Work
- Bookmark
- Contrasting Groups
- Direct Consensus
- Item-Descriptor Matching
- Nedelsky

Details for Standard Setting using the Body of Work Procedure

How do we decide which method to use?

Prior usage/history

Recommendation/requirement by some policy making authority

Type of assessment

How does it work?

The Body of Work Procedure

A standard setting procedure that uses samples of student work, in this case, score profiles representing typical patterns of item scores for students at given total score points.

Your job is to classify each profile into the performance level in which you feel it belongs.

What is the Body of Work procedure?

You are to classify each profiles into one of 4 levels by comparing:

- Performance level descriptors
- Knowledge skills and abilities measured by the items
- How the students scored on those items (profiles)

Novice
Nearing Proficiency
Proficiency
Advanced

Steps for Body of Work Method: Training

Become Familiar with the Performance Levels

Review and revise the performance level descriptor of each performance level

Read the descriptors individually.

Discuss the descriptors with each other.

Review and revise a bulleted list of the knowledge, skills and abilities associated with each level.

You must come to a group consensus about the knowledge, skills and abilities that define student performance at each performance level.

Before you start classifying profiles….

You need to:

Become familiar with the assessment

Become familiar with the performance level descriptors

Identify the knowledge, skills and abilities necessary to be classified in each level.

Become familiar with the profiles

Understand the knowledge, skills and abilities necessary to obtain each item score.

This will happen in your grade-level break out rooms.

Practice Round

You will be given a set of practice profiles and will complete a Practice Round of ratings as a group with your facilitator.

Steps for Body of Work Method: Training

Once in the breakout rooms, you will:

Review all standard setting materials

- Profile/rating form
- Ordered test booklet and auxiliary materials
- Scoring rubric flowchart
- Administration manual
- Visual item map

Become familiar with the assessment

Go through the ordered test booklet item by item.

Discuss the the knowledge, skills and abilities being measured by the test items.

Check for Understanding

Your facilitator will check in with you for understanding and answer any questions you may have during and after the Practice Round.

You will then complete the training evaluation form.

Profiles

You will be basing your decisions on sets of student profiles.

Profiles cover the range of possible scores and are presented in order from lowest to highest total score.

Each profile shows a typical pattern of item scores for students who received a given total score.

The items are ordered from the easiest item to the most difficult item.

Note

You may disagree about the order of the student profiles; that's OK

--you may feel the complexity of an item should give it more weight in the process.

Your task is to categorize the student profiles as you see fit, whether your ratings agree with the order or not

Steps for Body of Work Method

Thinking of the student who demonstrates the knowledge, skills and abilities for each level, you will classify each profile into the level in which you feel it belongs:

Novice
Nearing Proficiency
Proficient
Advanced

Note

It is **never** necessary for panelists to come to consensus as to how the profiles should be categorized.

You may change your mind as a result of the discussions, or you may not.

You should be open-minded when listening to your colleagues' rationales for their ratings.

However; we want your **individual best judgment** in each round of rating.

Steps for Body of Work Method

Round 1:
- Panelists individually review the profiles
  - Profiles will be pre-categorized into performance levels.
- Group discussion of starting cuts
- Panelists make their first set of ratings
  - You will either validate the categorizations or recommend changes

Round 2:
- Results of Round 1 will be presented
  - Cut score locations & Impact
- Discussion of Round 1 ratings
- Panelists make their final set of ratings

Repeat Process for Second Grade/Content

Review the assessment
- Item by item review of the knowledge skills and abilities

Review the performance level descriptors
- Create a bulleted list of the knowledge, skills and abilities

Round 1 Ratings
- Individual review, discussion, ratings

Round 2 Ratings
- Feedback, discussion, ratings

Performance Level Descriptor Feedback

If time allows, you will have an opportunity to provide feedback on the Performance Level Descriptors.

We need your help to make the language clearer and more teacher- and parent-friendly.

It is important to note that while your comments will be carefully considered, they may not all be implemented.

You will then complete the final evaluation form.

Good Luck!

And that's it….

Please make sure to ask your facilitators any questions you may have about the Body of Work Procedure.

Appendix I—FACILITATOR SCRIPT

GENERAL INSTRUCTIONS FOR CONTENT/GRADE GROUP FACILITATORS
MONTCAS CRT-ALTERNATE STANDARD SETTING
READING & MATH, GRADES 4, 8, & 10

Overview

Because each group will be setting standards for two different content areas, the standard-setting activities described below will be repeated twice. Once the two rounds of ratings have been completed for the first content area, the group will begin with the "Review Assessment Materials" step for the second content area. At the end of the process, if time allows, the panelists will be given an opportunity to provide feedback on the Performance Level Descriptors and will be asked to complete the final evaluation.

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

Review Assessment Materials

Overview: Some of the panelists administered the assessment to students, while others did not. In order to ensure that all panelists have an understanding of the knowledge and skills assessed, thoroughly review the ordered assessment with the group, walking through each item and pointing out the scaffolding script.

1) Review the ordered assessment
2) Review ancillary test materials
3) Review the administration manual
4) Review the scoring guide and flow chart
5) Review the Visual Item Map

a. Ordered CRT-Alternate Test Booklet -- A photocopy of the items from the spring 2009 operational administration for the appropriate grade level will be provided to each participant. The items will be ordered by difficulty using the p-value for each item. Have the panelists read through the test booklet, reminding them that the booklet has been re-ordered to reflect the easiest to the most difficult item. Direct panelists to review scaffolding script for each item.
b. Ancillary Test Materials -- Three sets of ancillary materials, used for test administration, will be provided for each grade. These materials consist of reading selections and responses cards for each test item. Direct panelists to review the associated materials for each test item.

c. Administration Manual -- This manual lists and explains the following: participation guidelines, administration procedures, scoring/scaffolding directions, accommodations and assistive technology, and contact information. Two to three manuals per table will be provided as a resource for participants. Remind panelists that this is available on the table to refer to.
d. Scoring Guide and Flow Chart -- The scoring guide and flowchart show the direct relationship between scoring and scaffolding. The scoring system in the CRT-Alternate is built on increasing amounts of scaffolding, provided only when the student does not respond or responds incorrectly. Each tasklet begins with items that introduce the subject and materials that will be used throughout the tasklet. These items are scored as either a 4 (student responds accurately and with no assistance) or a 0 (student does not respond or actively resists). Each subsequent item within the tasklet is scored on a five-point scale 4–0, with "4" representing a correct, independent response and "1" representing a correct response that has been completely guided by the teacher. A score of "0" is used when the student does not respond, or actively resists participation. Scaffolding is based on the amount of information the student needs to reach the correct response. If the student can respond independently, no further information is needed. If the student does not respond accurately or independently, more information is given about the item and the choices are reduced for each level of scaffolding.
e. Visual Item Map (VIM) -- the visual item map provides a graphical summary of the pattern of student scores on the items, and can be helpful to panelists in understanding the relationships among the items. Each column on the VIM represents one item, presented in order from easiest to hardest, and the left-most column shows percentages, from 100 to 0. For each item, each possible score point (1, 2, 3, and 4) appears in the row corresponding to the percentage of students who got that score point or higher on that item. Therefore, for easier items, all 4 score points will be clustered in the upper part of the table, while, for harder items, at least some of the score points will appear closer to the bottom. NOTE: the purpose of the VIM is solely to help panelists understand the relationship among the items; if a panelist does not find it helpful, he/she is not required to use it.

Discuss Performance Level Descriptions

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

1) the definition of the four performance levels, and
2) the key knowledge, skills, and abilities that distinguish students in adjacent performance levels.

The purpose of this activity is for the panelists to come to consensus about what characterizes students in each of the four performance levels and to develop a list of characteristics that captures the knowledge, skills, and abilities attained at each level. The list should contain both what these students can do independently and what they can do with a level of support that still shows they possess the skills necessary at each level. This activity is critical since the ratings panelists will be making in Rounds 1 and 2 will be based on these understandings.

It is important to understand that the draft Performance Level Descriptors and the list of characteristics are to be used as a starting point only and that they will be reviewed again at the end of the entire process and any recommended adjustments will be recorded for the Office of Public Instruction (OPI).

Activities:

1. Introduce task. In this activity they will:
    a. Individually review the Performance Level Descriptors;
    b. discuss Descriptors as a group; and
    c. make any modification or revisions, upon group consensus, to bulleted lists in order to define students in each achievement level category.

2. Have panelists individually review all of the Performance Level Descriptors provided in their folders. The bulleted lists are intended to serve as a starting point for discussion. Panelists can make notes if they like. The goal here is for the panelists to come to a common understanding of what it means to be in each performance level. It is not unusual for panelists to disagree with the descriptions 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 are described by each Performance Level Descriptor.

3. After individually reviewing the Descriptors, have the panelists discuss each one as a group, starting with *Novice*, and provide clarification. Encourage panelists to make any modifications or revisions to the list in order to achieve group consensus. The purpose of this is to have a collegial discussion in which to bring up/clarify any issues or questions that any individual may have and to reach consensus on an understanding of the descriptor.

4. During the discussion of each performance level, write the bulleted lists with any modifications or revisions as agreed upon by the group that reflect the knowledge, skills and abilities that best describe students in that level, on chart paper. The panelists want to answer the question: What knowledge, skills and abilities must a student demonstrate in order to be classified in the *Novice* category? Or, put another way: What are the most important knowledge, skills and abilities that distinguish a *Novice* student from a student in the *Nearing Proficiency* category? They will then repeat this process for the *Proficient* and *Advanced* categories. Panelists will have an opportunity to provide final feedback and suggestions for edits to the Descriptors after the standard setting activities are completed.

Practice Round

Overview of Practice Round: The primary purpose of the Practice Round is to have the facilitator walk the panelists through all of the documents by completing the process with three practice profiles. They will discuss the student profiles and make their determination as to which level each should be classified into. In this round, panelists will begin by reviewing the profiles with the facilitator, and discussing them as a group.

Activities:

1. Orient panelists to the set of profiles. Point out that the profiles are presented in order, from lowest scoring to highest scoring.
2. The panelists will begin by reviewing the profiles with the facilitator. As they are reviewing the profiles, they should consider the knowledge, skills, and abilities demonstrated in each profile and how they relate to the performance level definitions. The facilitator should review the Scoring Rubric and explain how the coding on the profile relates to the scoring level of each item. The facilitator should demonstrate how the panelists will need to refer to the Ordered Test Booklets to see how the profiles relate to the knowledge, skills and abilities required by the items. The purpose of this step is to thoroughly familiarize panelists with the materials and with the rating process as well as to allow panelists to get an initial sense of how they feel the profiles should be categorized. Steps for walking through the practice profiles: (facilitator should refer to the Practice Talking Points document)
    f. For the first profile walk through it as a group, read through the scores for each item, lead them to each of the items in the test booklet and the auxiliary materials that go with each item. This profile will be selected to be clearly in the first performance level.
    g. For the second profile walk through it as a group, read through the scores for each item, lead them to each of the items in the test booklet and the auxiliary materials that go with each item. This profile will be selected to be one that could be considered either in the first or second performance level. Lead them through discussion of their rationale. Remind panelists that they are looking for the profile to be **just over the line** between the first and second performance level-they are not looking for a solid second level performance.
    h. For the third profile have the panelists rate it on his or her own and then lead them through a discussion of the rationale. This profile will be selected to be one that could be considered either in the second or third performance level.
3. Panelists may want to take notes as they work if there are particular points they would like to discuss with their colleagues.
4. Make sure panelists know to enter their ratings on the rating form.
5. Go over the rating form with panelists:
    a. Have panelists write their ID number on the rating form. The ID number is on their name tags.
    b. Lead panelists through a step-by-step demonstration of how to fill in the rating form.
6. Once panelists have completed their ratings for the practice profiles, check in with them to see make sure they understand the process and to see if they have any questions. Have panelists fill out the training evaluation form before proceeding to Round 1. Before you start the Round 1 activities, scan the completed evaluations to see if there are any problems or concerns that need to be addressed before proceeding. Return the completed evaluations to the data analysis room at the next convenient opportunity.

Round 1 Ratings

The primary purpose of Round 1 is to ask the panelists to discuss and rate the student profiles and make their determination as to which performance level, each should be classified into. The outcome from this activity is for the panelists to determine the cut points between novice and nearing proficiency; nearing proficiency and proficient; and proficient and advanced. Panelists will refer to

the lists of skills and abilities developed earlier as they consider their placements.

Overview of Round 1: Panelists will thoroughly review the reordered assessment, ancillary materials, and scoring rubric and flowchart. Panelists should be directed towards the visual item map, which they may use to help them understand the relationship among the indicators, and the rating form, which summarizes the student profiles and includes columns where the panelists will record their Initial, Round 1, and 2 ratings.

Activities:

1. Make sure panelists have the following materials:
    i. Ordered Test Booklet
    j. Ancillary Test Materials
    k. Performance Level Descriptors
    l. Scoring Guide and Flow Chart
    m. Visual Item Map
    n. Student Profiles/Rating Sheet

2. Go over the profiles/rating form with the panelists:
    o. Have panelists write their ID number on the rating form. The ID number is on their name tags.
    - Have panelists review the student profiles. Explain that the student profiles represent how the average student at each selected total score point performed on each of the indicators.
    - Point out to the panelists the first column of classifications and explain that those classifications represent the starting cuts. Explain to the panelists that their task is to either validate the placement of each profile or to make recommendations for changing the categorizations (and, therefore, the cuts).

3. Make initial classifications
    i. For each student profile determine whether a student displaying the represented knowledge, skills, and abilities belongs in the *Novice (1)*, *Nearing Proficiency (2)*, *Proficient (3)*, or *Advanced (4)* performance level.
    ii. They will start with the profile with the lowest score and in turn, work their way through all the profiles assigning students based on their performance on the assessment to one of the 4 performance levels. As panelists work, let them know they can change their designations as they work. Also let them know that the *Novice* and *Advanced* levels may be the easiest to determine.
    iii. As they are reviewing the profiles, the panelists should keep in mind the Performance Level Descriptors. They should consider the knowledge, skills, and abilities demonstrated in each profile and how they relate to the definitions of the Performance Level Descriptors. The purpose of this step is for panelists to make an initial determination of how they feel the profiles should be categorized.
    iv. In completing the rating sheet, panelists should use the following designations:
        - 1 – *Novice*
        - 2 – *Nearing Proficiency*
        - 3 – *Proficient*
        - 4 – *Advanced*

Panelists will write the appropriate number for each profile on the sheet, in the Initial column making sure there is a rating entered for each and every profile. Make sure panelists know that, even though the profiles are ordered from lowest to highest score, their ratings do not need to be in strictly ascending order.

4. Once panelists complete the individual review, using a show of hands, indicate on a piece of chart paper how many panelists assigned each profile to each performance level category. Beginning with the first profile for which there is disagreement as to how it should be categorized, the panelists should begin discussing the categorization of the profiles according to their initial ratings.
    a) Panelists only need to discuss those profiles for which there is disagreement as to how they should be categorized **or** disagreement with the categorizations based on the starting cuts.
    b) Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
    c) If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
    d) On the basis of the discussions, panelists should make adjustments to their ratings, as appropriate.
    e) The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating with which they disagree.
5. Encourage the panelists to use the discussion to assess how stringent or lenient a judge they are. If a panelist is categorizing profiles consistently higher or lower than the group, he or she may have a different understanding of the Performance Level Descriptors than the rest of the group. **It is acceptable for panelists to disagree, but that disagreement should be based on a common understanding of the Performance Level Descriptors.**

6. After the discussions are completed, the panelists will fill in their categorizations in the "Round 1" column of the profile/rating sheet.

7. As panelists complete the task, ask them to carefully inspect their rating forms to ensure they are filled out properly.
    a) The ID number must be filled in.
    b) Each profile must be assigned to one and only one performance level.
    c) Reiterate that although the profiles are presented in order from lowest- to highest-scoring, the panelists' category assignments do not need to be in strictly increasing order.

8. Facilitators should bring all the completed rating forms together to R&A for tabulation in the data analysis room. Results will be shared with the group as soon as they are available.

Tabulation of Round 1 Results

Tabulation of Round 1 results will be completed as quickly as possible after receipt of the rating forms.

Round 2 Ratings

Overview of Round 2: The primary purpose of Round 2 is to ask the panelists to discuss their ratings in the context of the ratings made by other members of the group. During Round 2, the panelists will discuss the Round 1 categorizations of the profiles. Panelists will be given the room average cut point placements, based on the results of Round 1, as well as impact data indicating the percentage of students statewide who would fall into each performance level category based on the Round 1 ratings. Focusing on the profiles that are near the cut points, the panelists will discuss why they categorized each profile as they did, making sure that all different points of view are included in the discussion.

Activities:

1. Make sure panelists have the following materials:
    - Ordered Test Booklet
    - Ancillary Test Materials
    - Performance Level Descriptors
    - Scoring Guide and Flow Chart
    - Visual Item Map
    - Student Profiles/Rating Sheet
    - Round 1 results (will be displayed on chart paper)

2. A psychometrician will review the Results of Round 1 information with the panelists:
    - The group average cut scores
    - The percentage of students in each performance level based on the group average cut scores

3. The facilitator will again lead the discussion for Round 2.
    - Using a show of hands, indicate on chart paper how many panelists assigned each profile to each performance level indicator.
    - Panelists should be given a few minutes to review the results. Encourage the panelists to use this information to assess how stringent or lenient a judge they are. If a panelist is consistently higher or lower than the group they may have a different understanding of the performance level definitions. It is O.K. for panelists to disagree, but that disagreement should be based on a common understanding of the performance level definitions.
    - The facilitator will ask the panelists to review the student profiles in the areas of disagreement and lead a discussion of those profiles, starting with the one with the lowest score, and focusing on the placement of the cut points and what those placements mean in terms of the abilities and skills of students at each performance level.
    - Each panelist should have a rationale for their placement.
    - Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
    - Panelists should discuss whether the percentage of students classified in each performance level “feels right”. They should address the question: Does it make sense to

the panelists to have XX% of the students in the *Advanced* level and YY% in the *Novice* level?

- In light of the additional information presented, if the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they should adjust their ratings to incorporate that information.

4. Following the discussion, each panelist will review his or her placement of the cut points on the rating sheet. Panelists may change any or all of their placements in light of the group discussion, or they may choose to leave them where they initially placed them**. It is not necessary to reach consensus during the standard setting process.** This set of ratings constitutes Round 2 of the standard setting process.
   - When making revised ratings, panelists should not feel compelled to change their ratings. They will make their Round 2 ratings individually, as they did in Round 1.

5. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced to making a rating they disagree with.

6. As each panelist completes the task, collect the rating form from each. When you collect the rating forms carefully inspect them to ensure they are filled out properly
   a. The ID number must be filled in.
   b. Each student profile must have a single rating.

Complete Standard-Setting Activities for Second Test

Once Round 2 ratings are completed for the first test and the panelists have completed the evaluation, the panelists will repeat the process for the second content area for their grade level, starting with the "Review Assessment Materials" step and going through all two rounds of ratings. For the Practice Round there will be no practice profiles provided, however, this time should be used to read and review all of the materials specific to the new grade level.

Finalizing Recommendations for Performance Descriptors

1. Have panelists revisit the Performance Level Descriptors and make any necessary adjustments or revisions, based on where they placed the cut points.

2. Panelist may
   - Clarify
   - Add more information
   - Add content specific detail
   - etc

3. Have panelists record changes as bullet points. Panelists do not have to agree on exact language.

Complete the Evaluation Form

After completing all standard setting activities for both tests, have panelists fill out the final evaluation form. Emphasize that their honest feedback is important.

Appendix J—PANELIST AFFILIATIONS

Last Name	*First Name*	*School/ District*	*Title*	*Grade*
Armstrong	Millie	Manhattan High School / Manhattan	School Psychologist	10
Augustine	Erin	Castle Middle Rock School / Billings	Special Educator	8
Castle	Robin	Fort Shaw Elementary / Sun River Valley	School Counselor	4
Clinch	Rose	Clancy Elementary / Clancy	Special Educator	4
Dale	Cheri	Washington School / Billings	Special Educator	4
Dehne	Rose	Charlo Elementary / Charlo	Special Educator	4
Ellis	Cassie	Hamilton Middle School / Hamilton	Special Educator	8
Feddes	Meredith	White Sulphur High School / White Sulphur	Special Educator	10
Gilboy	Kathy	East Valley Middle School / Helena	Special Educator	8
Gregory	Nina	Poplar 7-8 / Poplar	Special Educator	8
Jaquith	Martha	Victor School / Victor	Special Educator	4
Long	Leslie	Fergus High School / Lewistown	Special Educator	10
Lyndes	Mary	Franklin School / Missoula	General Educator	4
McGaugh	Rhonda	/East Middle School / Great Falls	Special Educator	8
Moore	Kevin	Whitehall 7-8 / Whitehall	Special Educator	10
Muir	Patty	Laurel Middle School /Laurel	Special Educator	4
Paeth	Susan	Sentinel High School / Missoula County	Special Educator	10
Schmeling	Corky	Hardin Middle School / Hardin	Intensive Support Special Educator	10
Sorenson	Barbara	Whitefish High School / Whitefish	General Educator	10
Thompson-Bailey	Shani	East Middle School / Butte	Special Educator	8

Appendix K—Panelist Descriptor Recommendations

Alternate Performance Level Descriptors for Grade 4 Reading	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • follows 3-step or more directions • may read independently • communicates knowledge using expanded vocabulary • communicates a complete thought related to topic or concept • correctly answers who, what, when and where questions • is able to generalize information from one setting to another • recognizes and articulates the main idea • relates and uses relevant knowledge to make connections
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • follows two-step directions • communicates knowledge of basic vocabulary and familiar words • demonstrates written words have meaning • explores pictures, symbols, and objects • answers yes and no questions • identifies beginning main idea • uses literacy materials appropriately • contributes/elaborates on responses
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • follows one-step directions consistently • understands when response is needed • needs multiple re-direction to the test material to respond to a specific item • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • begins to respond to literacy with varied prompts • uses prior knowledge to demonstrate knowledge of basic vocabulary • begins to communicate with a purpose
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • may anticipate a reading activity • responds to own name • attempts to communicate • attends for short periods of time to the teacher, materials, and test items • attends to pictures, symbols, objects when presented • begins/attempts to participate with support

Alternate Performance Level Descriptors for Grade 4 Mathematics	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • ~~creates and~~ extends a repeating pattern using objects, shapes, designs, or numbers • uses methods and tools to solve a problem involving patterns, relations, or functions • sets up a graph (i.e. labels axes) • carries out a strategy to solve problems involving patterns, relations, or functions • determines which of two numbers is closer to the quantity in a given set • understands and uses comparison words (more, less, some, none) • ~~explains~~ demonstrates reasoning about probability items • understand words that indicate operations in word problems
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • understands the concept of 1:1 correspondence • sorts objects into sets • understands comparison words (more, less, some, none) • extends or supplies a missing element in a repeating pattern by attribute or number • reads a simple graph ~~(i.e. label axes)~~ • ~~understand words that indicate operations in word problems~~ • demonstrates a basic understanding of math skills, concepts and vocabulary
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • understands the concept of "1" • may recognize a simple pattern • demonstrates an understanding that numbers, as opposed to letters, are used to express quantity, order, or size/amount • counts with another person • may recognize quantities • identifies basic shapes (i.e. circles, squares, triangles, and rectangles) and the relationships among them • matches two- dimensional physical shapes to pictures of the shapes in different orientations • may communicate some numbers correctly
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • may anticipate a math activity • attends to materials being displayed • attends to another person making patterns and to a person describing patterns • attends to a person demonstrating with concrete materials • attends to objects or pictures of two- and three- dimensional geometric shapes • attends to another person estimating an amount of a given set

Alternate Performance Level Descriptors for Grade 8 Reading	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • connects prior knowledge to make meaning of text • identifies main idea and various supporting details • understands story lessons~~/author's purpose~~ • locates title and other information from a variety of documents/sources • recognizes vowel letter-sound • uses reading and/or listening strategies when needed to gain information (i.e. rereading, use of key words, use of features of text) • ~~reads and~~ comprehends a simple paragraph
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • has basic word recognition • tracks while reading or being read to • identifies basic words and recognizes some words in different contexts • identifies a word/picture/symbol for content communication • identifies title and basic parts (beginning, middle, and end) of a reading selection • identifies main idea of a story and some supporting facts/details • identifies purposes of various texts (i.e. dictionary, map) • ~~has a firm grasp of sound/symbol association~~
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • requires a high level of rephrasing • shows an understanding of the beginning and end of a story by giving attention to the reader or the text • recognizes that letters have names and is aware of letter sounds • recognizes difference between letters and other symbols (i.e. numerals) • identifies letters by name/sign • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • identifies a word/picture/object of familiar places and people • responds mostly through basic yes/no questions • ~~understands story beginning and end~~ • understands basic main idea (answer with one picture/short response)
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a reading activity • attends to materials being displayed • demonstrates readiness by following one-step directions or with teacher modeling/prompting • responds to name, words, pictures and symbols • directs attention and responds to external stimuli when requested (i.e. turns head in direction, nods head, operates switch, points to, etc.) • interacts with stimuli (i.e. teacher, words, pictures, and symbols)

Alternate Performance Level Descriptors for Grade 8 Mathematics	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • measures to the inch • ~~compares and calculates measurements, time, and monetary amounts~~ • ~~understands concept of fractions~~ measures the distance between 2 points on a map • ~~understands Algebra concepts~~ • completes missing components in basic number sentence • recognizes and understands all operational symbols (+, -, =), measurement symbols (in. cm. etc), monetary symbols ($), and time • uses all comparison words (more, less, some, none, most, least) correctly • understands ordinal numbers beyond 3rd • selects the correct label ~~sets of data and components of~~ for a graph (i.e. label axis) • ~~creates graph and~~ explains conclusions drawn from graph • applies beginning connections between concrete and symbolic representations, operations, measurement, graphing and problem solving strategies
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • reads~~/makes~~ simple measurements • can subtract single digits • understands first 3 ordinal numbers (1st, 2nd, 3rd) • uses some comparison words (more, less, some, none, most, least) correctly • understands numbers can represent monetary amounts, measurement, and time • demonstrates basic problem solving skills • fills in data, as directed to create a representation, on a bar graph • recognizes and understands most operational symbols (+, -, =), measurement symbols (in. cm. etc), monetary symbols ($), and time • identify places on a map • ~~identifies basic information~~ answers questions about a bar graph • makes a statement about data • demonstrates knowledge of basic number sentences • ~~demonstrates beginning connections between concrete and symbolic representations, operation (+/-), measurement and graphing~~
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • identifies and/or recognizes a map and measuring tools • demonstrates solid number concept for 1:1 correspondence (consistently touch counts) • can count single digits • can add~~/subtract~~ single digits • recognizes and understands some operational symbols (+, -, =), measurement symbols (in. cm. etc), monetary symbols ($) • ~~can distinguish between two different elements/variables~~ • basic understanding of bar graphs and data • can make general statements about a bar graph • ~~communicates understanding of beginning connections between concrete and symbolic representations~~

Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a math activity • attends to materials being displayed • attends to another person reviewing a map with prompting • attends to another person reviewing a graph with prompting • engages with instructor with prompts • recognizes numbers (symbol or rote recitation)

Alternate Performance Level Descriptors for Grade 10 Reading	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • identifies main idea and supporting details from various reading selections • identifies appropriate resources for gaining specific information • draws conclusions from a variety of texts (i.e. poem, fiction) • communicates meaning of new and unfamiliar vocabulary • communicates a complete thought related to topic or concept • uses word-recognition skills, context clues, and prior knowledge to understand text • rereads to gain understanding
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • has basic reading and comprehension skills • understands difference between various literacy materials • ~~uses~~ begins to access prior knowledge to understand text • communicates ~~an opinion~~ a basic thought on topic • identifies main ideas and some supporting details/facts • is beginning to identify appropriate resources for gaining specific information • identifies words/pictures/symbols and objects that are new and unfamiliar
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • ~~begins to access to prior knowledge~~ • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • able to match and identify familiar words/pictures/ symbols/objects • ~~understands story beginning and end~~ • ~~understands~~ identifies basic main idea (answers with one picture/short response) • communicates an opinion • identifies ~~resources~~ familiar literary resources (i.e. newspaper, CDs, Internet, oral histories)
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • ~~anticipates a reading activity~~ • attends to materials being displayed • responds to name, words, pictures and symbols • demonstrates readiness by following one-step directions or with teacher modeling/prompting • directs attention and responds to external stimuli when requested (i.e. turns head in direction, nods head, operates switch, points to, etc.) • interacts with stimuli (i.e. teacher, words, pictures, and symbols)

Alternate Performance Level Descriptors for Grade 10 Mathematics	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • generalizes very basic information • completes two to three-step processes of addition and subtraction • completes basic division and multiplication problem • applies beginning connections between concrete and symbolic representations by using a chart/table to draw conclusions • creates graph/tables and explains conclusions drawn from graph • understands and communicates relationship between variables • solves problems using bills and their values • follows navigational directions and recalls shapes and locations
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • completes and/or extends basic patterns of data • sorts items into sets by multiple defining characteristics • demonstrates ~~beginning~~ basic connections between concrete and symbolic representations • identifies basic information from a graph/chart • ~~makes a statement about data~~ • ~~understands and~~ matches bills and their values • recognizes and identifies two-dimensional shapes • chooses correct procedures to solve simple number problems • adds and subtracts 2-digit numbers
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • recognizes properties of limited (square/circle) two-dimensional shapes • recognizes distinct categories • recognizes basic patterns of data • sorts items into sets by one defining characteristic • understands quantity • can count single digits • can add/subtract single digits • communicates ~~understanding of~~ beginning connections between concrete and symbolic representations
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • ~~anticipates a math activity~~ • attends to materials being displayed • ~~attends to another person reviewing a graph/chart~~ • ~~attends to another person reviewing signs and labels~~ • shows limited understanding of quantity when given two choices • ~~engages with instructor with prompts~~ • recognizes numbers (symbol or rote recitation)

Alternate Performance Level Descriptors for Grade 10 Mathematics	
Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • generalizes very basic information • completes two to three-step processes of addition and subtraction • completes basic division and multiplication problem • applies beginning connections between concrete and symbolic representations by using a chart/table to draw conclusions • creates graph/tables and explains conclusions drawn from graph • understands and communicates relationship between variables • solves problems using bills and their values • follows navigational directions and recalls shapes and locations
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content specific performance indicators. • arrives at correct answer with limited prompting • completes and/or extends basic patterns of data • sorts items into sets by multiple defining characteristics • demonstrates ~~beginning~~ basic connections between concrete and symbolic representations • identifies basic information from a graph/chart • ~~makes a statement about data~~ • ~~understands and~~ matches bills and their values • recognizes and identifies two-dimensional shapes • chooses correct procedures to solve simple number problems • adds and subtracts 2-digit numbers
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • recognizes properties of limited (square/circle) two-dimensional shapes • recognizes distinct categories • recognizes basic patterns of data • sorts items into sets by one defining characteristic • understands quantity • can count single digits • can add/subtract single digits • communicates ~~understanding of~~ beginning connections between concrete and symbolic representations
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • ~~anticipates a math activity~~ • attends to materials being displayed • ~~attends to another person reviewing a graph/chart~~ • ~~attends to another person reviewing signs and labels~~ • shows limited understanding of quantity when given two choices • ~~engages with instructor with prompts~~ • recognizes numbers (symbol or rote recitation)

Appendix L—EVALUATION RESULTS

Grade 4

Training Evaluation

	N	*Mean*	*% SD*	*% D*	*% U*	*% A*	*% SA*
I understand the goals of the standard setting meeting.	7	4.14	0	0	14	57	29
I understand the procedures we are using to set standards.	7	4.29	0	0	14	43	43
I understand how to use the standard setting materials.	7	4.29	0	0	14	43	43
I understand the differences between the performance levels.	7	4.29	0	0	0	71	29
I understand how to make the cut score judgment.	7	4	0	0	14	71	14
I know what tasks to expect for the remainder of the meeting.	7	4.43	0	0	0	57	43
I am confident in my understanding of the standard setting task.	7	3.86	0	0	29	57	14

Please indicate any areas in which you would like more information before you continue.

Please indicate any questions you may have about the remainder of the standard setting meeting.

Final Evaluation

Panelist Demographics	*N = 7*
Gender:	
Male	0
Female	7
Race/Ethnicity:	
White	7
Black	0
Hispanic	0
Asian	0
Pacific Islander	0
American Indian	0
Years of Educational Experience:	
0-5	0
5-10	0
10-15	0
More than 15	7
Professional Experience (check all that apply):	
Students with Disabilities	6
Students with Limited English Proficiency	0
Economically Disadvantaged Students	3
Gifted and Talented Students	2
General Education	5

	N	Mean	% SD	% D	% U	% A	% SA
I understood the goals of the standard setting meeting.	6	4.83	0	0	0	17	83
I understood the procedures we used to set standards.	6	5	0	0	0	0	100
The facilitator helped me understand the process.	6	5	0	0	0	0	100
The materials contained the information needed to set standards.	6	5	0	0	0	0	100
I understood how to use the materials provided.	6	5	0	0	0	0	100
The performance level descriptors were clear.	6	4.67	0	0	0	33	67
I understood how to make the cut score judgments.	6	4.5	0	0	0	50	50
I understood how to use the feedback provided after each round.	6	4.67	0	0	0	33	67
I understood how to use the impact data.	6	4.67	0	0	0	33	67
I understood how the cut scores were calculated.	6	4.83	0	0	0	17	83
The facilitator was able to get answers to my questions.	6	4.83	0	0	0	17	83
Sufficient time was allotted for training on the standard setting tasks.	6	4.83	0	0	0	17	83
Sufficient time was allotted to complete the standard setting tasks.	6	4.83	0	0	0	17	83
The facilitator helped the standard setting process run smoothly.	6	5	0	0	0	0	100

Usefulness of	*N*	*Mean*	*%Low*				*%High*
The opening session.	7	4	0	0	29	43	29
The small group activities.	7	4.57	0	0	0	43	57
Becoming familiar with the assessment.	7	4.71	0	0	0	29	71
Articulating the differences between the performance levels.	7	4.71	0	0	0	29	71
Discussions with other participants.	7	4.71	0	0	0	29	71
Providing additional details to the performance level descriptors.	7	4.71	0	0	0	29	71

Influence of	*N*	*Mean*	*%Low*				*%High*
The performance level descriptors.	7	4.86	0	0	0	14	86
My expectations of students.	7	4.43	0	0	0	57	43
The difficulty of the test materials.	7	4.86	0	0	0	14	86
The student responses.	7	4.57	0	0	0	43	57
My experience in the field.	7	4.86	0	0	0	14	86
Discussions with other participants.	7	5	0	0	0	0	100
Cut scores of other participants.	7	4.43	0	0	0	57	43
Impact data.	7	4.29	0	0	14	43	43

Recommended Cut Scores	*N*	*% Too Low*	*% About Right*	*% Too High*
	Mathematics			
Advanced/Proficient	7	0	100	0
Proficient/Nearing Proficiency	7	0	100	0
Nearing Proficiency/Novice	7	0	100	0
	Reading			
Advanced/Proficient	7	0	100	0
Proficient/Nearing Proficiency	7	0	100	0
Nearing Proficiency/Novice	7	0	100	0

Please provide any additional comments about the standard setting process or suggestions as to how the training and process could be improved.

Grade 8

Training Evaluation

	N	*Mean*	*% SD*	*% D*	*% U*	*% A*	*% SA*
I understand the goals of the standard setting meeting.	6	4.83	0	0	0	17	83
I understand the procedures we are using to set standards.	6	4.83	0	0	0	17	83
I understand how to use the standard setting materials.	6	4.5	0	0	0	50	50
I understand the differences between the performance levels.	6	4.33	0	0	0	67	33
I understand how to make the cut score judgment.	6	4.33	0	0	0	67	33
I know what tasks to expect for the remainder of the meeting.	6	4.5	0	0	0	50	50
I am confident in my understanding of the standard setting task.	6	4.33	0	0	0	67	33

Please indicate any areas in which you would like more information before you continue.

Please indicate any questions you may have about the remainder of the standard setting meeting.

Final Evaluation

Panelist Demographics	*N = 6*
Gender:	
Male	0
Female	6
Race/Ethnicity:	
White	6
Black	0
Hispanic	0
Asian	0
Pacific Islander	0
American Indian	0
Years of Educational Experience:	
0-5	0
5-10	2
10-15	1
More than 15	3
Professional Experience (check all that apply):	
Students with Disabilities	6
Students with Limited English Proficiency	1
Economically Disadvantaged Students	2
Gifted and Talented Students	1
General Education	3

	N	*Mean*	*% SD*	*% D*	*% U*	*% A*	*% SA*
I understood the goals of the standard setting meeting.	6	4.83	0	0	0	17	83
I understood the procedures we used to set standards.	6	4.83	0	0	0	17	83
The facilitator helped me understand the process.	6	4.67	0	0	0	33	67
The materials contained the information needed to set standards.	6	4.67	0	0	0	33	67
I understood how to use the materials provided.	6	4.5	0	0	0	50	50
The performance level descriptors were clear.	6	4	0	0	17	67	17
I understood how to make the cut score judgments.	6	4.33	0	0	0	67	33
I understood how to use the feedback provided after each round.	6	4.5	0	0	0	50	50
I understood how to use the impact data.	6	4	0	0	17	67	17
I understood how the cut scores were calculated.	6	4	0	0	17	67	17
The facilitator was able to get answers to my questions.	6	4.5	0	0	0	50	50
Sufficient time was allotted for training on the standard setting tasks.	6	4.5	0	0	0	50	50
Sufficient time was allotted to complete the standard setting tasks.	6	4.5	0	0	0	50	50
The facilitator helped the standard setting process run smoothly.	6	4.5	0	0	0	50	50

Usefulness of	*N*	*Mean*	*%Low*				*%High*
The opening session.	6	3.67	0	17	17	50	17
The small group activities.	6	4.83	0	0	0	17	83
Becoming familiar with the assessment.	6	4.17	0	17	0	33	50
Articulating the differences between the performance levels.	5	5	0	0	0	0	100
Discussions with other participants.	6	5	0	0	0	0	100
Providing additional details to the performance level descriptors.	5	4.8	0	0	0	20	80

Influence of	*N*	*Mean*	*%Low*				*%High*
The performance level descriptors.	6	4.83	0	0	0	17	83
My expectations of students.	6	4.5	0	0	0	50	50
The difficulty of the test materials.	6	4.67	0	0	0	33	67
The student responses.	6	4.83	0	0	0	17	83
My experience in the field.	6	5	0	0	0	0	100
Discussions with other participants.	6	4.83	0	0	0	17	83
Cut scores of other participants.	6	4.83	0	0	0	17	83
Impact data.	6	4	0	0	33	33	33

Recommended Cut Scores	*N*	*% Too Low*	*% About Right*	*% Too High*
	Math			
Advanced/Proficient	6	0	100	0
Proficient/Nearing Proficiency	6	0	100	0
Nearing Proficiency/Novice	6	0	100	0
	Reading			
Advanced/Proficient	6	0	100	0
Proficient/Nearing Proficiency	6	0	100	0
Nearing Proficiency/Novice	6	0	100	0

Please provide any additional comments about the standard setting process or suggestions as to how the training and process could be improved.

On the tasklets overview pages, indicate which questions came from each tasklet, so from a quick glance you can review the tasklet overview for that question.

Most people have already given this assessment so the beginning of the training, although it was smooth, was very redundant of things we already knew.

Grade 10

Training Evaluation

	N	*Mean*	*% SD*	*% D*	*% U*	*% A*	*% SA*
I understand the goals of the standard setting meeting.	7	4.29	0	0	0	71	29
I understand the procedures we are using to set standards.	7	4.43	0	0	0	57	43
I understand how to use the standard setting materials.	7	4.57	0	0	0	43	57
I understand the differences between the performance levels.	7	4.29	0	0	0	71	29
I understand how to make the cut score judgment.	7	4.43	0	0	0	57	43
I know what tasks to expect for the remainder of the meeting.	7	4.43	0	0	0	57	43
I am confident in my understanding of the standard setting task.	7	4.29	0	0	0	71	29

Please indicate any areas in which you would like more information before you continue.

Presentation materials strongly influence student responses. Do we take that into account? I'm thinking, "No".

Please indicate any questions you may have about the remainder of the standard setting meeting.

Final Evaluation

Panelist Demographics	*N = 7*
Gender:	
Male	1
Female	6
Race/Ethnicity:	
White	6
Black	1
Hispanic	0
Asian	0
Pacific Islander	0
American Indian	0
Years of Educational Experience:	
0-5	1
5-10	0
10-15	2
More than 15	3
Professional Experience (check all that apply):	
Students with Disabilities	7
Students with Limited English Proficiency	2
Economically Disadvantaged Students	4
Gifted and Talented Students	2
General Education	2

	N	Mean	% SD	% D	% U	% A	% SA
I understood the goals of the standard setting meeting.	7	4.43	0	0	0	57	43
I understood the procedures we used to set standards.	7	4.57	0	0	0	43	57
The facilitator helped me understand the process.	7	5	0	0	0	0	100
The materials contained the information needed to set standards.	7	4.43	0	0	0	57	43
I understood how to use the materials provided.	7	4.43	0	0	0	57	43
The performance level descriptors were clear.	7	4.14	0	0	0	86	14
I understood how to make the cut score judgments.	7	4.57	0	0	0	43	57
I understood how to use the feedback provided after each round.	7	4.43	0	0	0	57	43
I understood how to use the impact data.	7	4.57	0	0	0	43	57
I understood how the cut scores were calculated.	7	4.57	0	0	0	43	57
The facilitator was able to get answers to my questions.	7	4.57	0	0	0	43	57
Sufficient time was allotted for training on the standard setting tasks.	7	4.57	0	0	0	43	57
Sufficient time was allotted to complete the standard setting tasks.	7	4	0	14	0	57	29
The facilitator helped the standard setting process run smoothly.	7	4.86	0	0	0	14	86

Usefulness of	*N*	*Mean*	*%Low*				*%High*
The opening session.	7	4	0	0	14	71	14
The small group activities.	7	4.43	0	0	0	57	43
Becoming familiar with the assessment.	7	4.14	0	0	14	57	29
Articulating the differences between the performance levels.	7	4.57	0	0	0	43	57
Discussions with other participants.	7	4.71	0	0	0	29	71
Providing additional details to the performance level descriptors.	7	4.43	0	0	0	57	43

Influence of	*N*	*Mean*	*%Low*				*%High*
The performance level descriptors.	7	4.29	0	0	0	71	29
My expectations of students.	7	4	0	14	0	57	29
The difficulty of the test materials.	7	4.29	0	0	0	71	29
The student responses.	7	4.43	0	0	0	57	43
My experience in the field.	7	4.57	0	0	0	43	57
Discussions with other participants.	7	4.43	0	0	0	57	43
Cut scores of other participants.	7	4.43	0	0	0	57	43
Impact data.	7	4.14	0	0	14	57	29

Recommended Cut Scores	*N*	*% Too Low*	*% About Right*	*% Too High*
	Math			
Advanced/Proficient	7	0	100	0
Proficient/Nearing Proficiency	7	0	86	14
Nearing Proficiency/Novice	7	0	86	14
	Reading			
Advanced/Proficient	7	0	86	14
Proficient/Nearing Proficiency	7	0	86	14
Nearing Proficiency/Novice	7	0	86	14

Please provide any additional comments about the standard setting process or suggestions as to how the training and process could be improved.

Sometimes it took too long to complete tasks; there should be a time frame
Our facilitator, Lynn Albee, was so good (smooth) at resolving impasses and creating clarity. She's very professional

Appendix B—Performance Level Descriptors—Raw & Scaled Scores

Alternate Performance Level Descriptors for Grade 3 Reading

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • follows three-step or more directions • chooses correct choice among the 4 options correctly • asks for clarification/help if needed • gives full attention to literacy materials/selection • communicates using expanded vocabulary • correctly answers who, what, and where questions and contributes own thoughts/ideas • is able to generalize information from one setting to another • responds with a complete thought • recognizes and articulates the main idea
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • follows two-step directions • attends fully to the activity • contributes/elaborates on the response • shows independence/confidence • chooses correctly among three options (verbal, pictures, touch, other stimuli) • participates actively • understands what he/she is doing • cooperates with the administrator • addresses responses with Yes or No • communicates and demonstrates words he/she knows and asks for clarification if needed • attends long enough to complete a given task • attempts to answer what and where questions
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • explores literary items (holds book in correct position, recognizes pictures vs. print, uses left to right orientation) • attends with support easily • begins to respond to literacy with varied prompts • responds to others • holds eye contact • begins to communicate with a purpose • communicates the correct choice between two options • follows one-step direction consistently
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • anticipates a reading activity • attends to materials being displayed • responds to own name • attends for a short period of time • begins/attempts to participate with supports • attempts to communicate

Alternate Performance Level Descriptors for Grade 3 Mathematics

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • creates a repeating pattern using objects, shapes, designs, or numbers • carries out a strategy to solve problems involving patterns, relations, or functions • recognizes two-dimensional shapes • carries out a strategy to solve a geometric problem • determines which of two numbers is closer to the quantity in a given set • uses methods and tools to solve a problem, including drawing pictures, modeling with objects, estimating, using paper and pencil, and using a calculator • identifies a reasonable quantity when guessing the amount of a given set
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • extends and explains an alternating pattern of two or more objects, shapes, designs, or numbers • shows a quantity • extends or supplies a missing element in a repeating pattern by attribute or number • reproduces an alternating pattern of two or more objects, shapes, designs, or numbers • recognizes properties of two-dimensional shapes • uses a quantitative label when making a guess • touches and moves shapes toward creating new shapes
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • demonstrates an understanding that numbers, as opposed to letters, are used to express quantity, order, or size/amount • counts with another person • identifies/names shapes as circles, squares, triangles, rectangles, and ovals • matches two-dimensional physical shapes to pictures of the shapes in different orientations • explains/shows spatial reasoning • finds various shapes in the environment • enters numbers correctly on a calculator/writes (communicates) numbers correctly
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • anticipates a math activity • attends to materials being displayed • attends to another person combining and subdividing shapes • attends to another person making patterns and to a person describing patterns • attends to a person demonstrating with concrete materials • attends to objects or pictures of two- and three-dimensional geometric shapes and the relationships among them • attends to another person estimating an amount of a given set

Alternate Performance Level Descriptors for Grade 4 Reading

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • follows three -step or more directions • may read independently • communicates knowledge using expanded vocabulary • communicates a complete thought related to topic or concept • correctly answers who, what, when, and where questions • is able to generalize information from one setting to another • recognizes and articulates the main idea • relates and uses relevant knowledge to make connections
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • arrives at correct answer with limited prompting • follows two-step directions • communicates knowledge of basic vocabulary and familiar words • demonstrates written words have meaning • explores pictures, symbols, and objects • answers yes and no questions • identifies beginning main idea • uses literacy materials appropriately • contributes/elaborates on responses
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • follows one-step directions consistently • understands when response is needed • needs multiple redirection to the test material to respond to a specific item • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • begins to respond to literacy with varied prompts • uses prior knowledge to demonstrate knowledge of basic vocabulary • begins to communicate with a purpose
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • may anticipate a reading activity • responds to own name • attempts to communicate • attends for short periods of time to the teacher, materials, and test items • attends to pictures, symbols, and objects when presented • begins/attempts to participate with support

Alternate Performance Level Descriptors for Grade 4 Mathematics

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • extends a repeating pattern using objects, shapes, designs, or numbers • uses methods and tools to solve a problem involving patterns, relations, or functions • sets up a graph (i.e., labels axes) • carries out a strategy to solve problems involving patterns, relations, or functions • determines which of two numbers is closer to the quantity in a given set • understands and uses comparison words (more, less, some, none) • demonstrates reasoning about probability items • understands words that indicate operations in word problems
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • arrives at correct answer with limited prompting • understands the concept of 1:1 correspondence • sorts objects into sets • understands comparison words (more, less, some, none) • extends or supplies a missing element in a repeating pattern by attribute or number • reads a simple graph • demonstrates a basic understanding of math skills, concepts, and vocabulary
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • understands the concept of “1” • may recognize a simple pattern • demonstrates an understanding that numbers, as opposed to letters, are used to express quantity, order, or size/amount • counts with another person • may recognize quantities • identifies basic shapes (i.e., circles, squares, triangles, and rectangles) and the relationships among them • matches two-dimensional physical shapes to pictures of the shapes in different orientations • may communicate some numbers correctly
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • may anticipate a math activity • attends to materials being displayed • attends to another person making patterns and to a person describing patterns • attends to a person demonstrating with concrete materials • attends to objects or pictures of two- and three-dimensional geometric shapes • attends to another person estimating an amount of a given set

Alternate Performance Level Descriptors for Grade 4 Science

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • ability to independently attend, compare/contrast, sort/categorize, recognize, identify • understands content at higher level • consistent high scores • minimal scaffolding
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • ability to attend • ability to recognize and identify with minimal assistance • ability to compare/contrast and sort/categorize with minimal assistance • occasional scaffolding
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • attending with some assistance • ability to recognize and identify with some assistance • moderate to heavy scaffolding
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • limited to no attending skills • minimal recognition and identification skills • maximum scaffolding required • consistently low scores

Alternate Performance Level Descriptors for Grade 5 Reading

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • relates and uses relevant prior knowledge to make connections • uses pictures, symbols, and objects independently in problem solving • responds to test materials to respond to a specific item • gives correct response among four options • orients text and reads independently and with teacher • communicates the correct choice with multiple options • responds to basic comprehension questions • sounds out unfamiliar words using phonics
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • relates prior knowledge accurately and appropriately • explores pictures, symbols, and objects • needs occasional redirection to the test materials to respond to a specific item • responds to test materials to respond to a specific item • orients text and uses text with limited prompting • communicates the correct choice among three options • responds to basic comprehension questions given three options • sounds out unfamiliar words using phonics with assistance
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • understands when response is needed • displays knowledge of front/back, right side up, page turning and scanning of literacy materials with prompting • communicates the correct choice between two options • explores pictures, symbols, and objects when prompted • needs multiple redirection to the test material to respond to a specific item • relates prior knowledge to present situation • sounds out unfamiliar words using limited phonemic knowledge • responds to basic comprehension questions using yes or no
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • anticipates a reading activity • attends to materials being displayed • explores pictures, symbols, and objects with teacher assistance • responds when given modeling and supports • recognizes phonemic correspondence when modeled • attends and acknowledges literacy activities

Alternate Performance Level Descriptors for Grade 5 Mathematics

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • recognizes 0–100 independently • requires no clarification or prompts • demonstrates mastery of basic math concepts • demonstrates mastery of math vocabulary • solves problems using addition & subtraction • uses measurement tools • responds to test questions
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • recognizes 0–100 • discriminates correctly among three choices • begins to understand words that indicate operations in word problems • demonstrates a basic understanding of sequencing • demonstrates a basic understanding of math skills • demonstrates a basic understanding of math concepts and vocabulary
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • demonstrates a limited understanding of math concepts • demonstrates a limited understanding of math vocabulary • demonstrates a limited ability to generalize • demonstrates a limited ability to master a specific task in a specific environment • uses patterns to copy concrete patterns using manipulatives • recognizes digits 0–20 • demonstrates 1:1 correspondence • demonstrates single-digit addition (i.e., less than 9)
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • anticipates a math activity • attends to materials being displayed • demonstrates an understanding of the concepts of some/more/ less/take away/all gone/no more • selects the appropriate tool to be used in making a measure

Alternate Performance Level Descriptors for Grade 6 Reading

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • orients text and reads independently or with teacher • communicates the correct choice with multiple options • uses diagrams and models to understand text independently • creates diagrams and charts to show understanding of text • relates text to appropriate personal experiences • identifies meaning of unfamiliar words using context clues • responds to basic questions about plot outcome • demonstrates basic understanding of main ideas and some supporting details • recognizes diverse perspectives
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • orients and uses text • communicates the correct choice among three options • uses diagrams and models to understand text with limited prompting • creates diagrams and charts to show understanding of text • relates text to appropriate personal experiences • identifies meaning of unfamiliar words using context clues • responds to basic questions about plot outcome • demonstrates basic understanding of main ideas and some supporting details • recognizes diverse perspectives
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • understands when response is needed • displays knowledge of front/back, right side up, page turning and scanning of literacy materials with prompting • communicates the correct choice between two options • uses diagrams and models to understand text • creates diagrams and charts to show understanding of text • relates text to personal experiences • identifies meaning of unfamiliar words using context clues • responds to basic questions about plot • demonstrates basic understanding of main ideas and some supporting details • recognizes diverse perspectives
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • anticipates a reading activity • attends to materials being displayed • orients text • acknowledges correct choice • attends to teacher-created diagrams and models to understand text • connects text to personal experience only with teacher guidance • acknowledges and attends to literacy activity

Alternate Performance Level Descriptors for Grade 6 Mathematics

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • demonstrates mastery understanding of abstract math concepts and skills • demonstrates mastery of telling time to the one half hour and hour and applies the concepts of time • demonstrates mastery of the ability to perform visual/spatial reasoning • demonstrates mastery of the ability to sequence numbers and/or patterns • demonstrates mastery of the understanding and use of math vocabulary • consistently demonstrates the ability to generalize knowledge and skills to different environments
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • discriminates correctly among three choices • demonstrates a basic understanding of abstract math concepts and skills (addition and subtraction) • tells time to the one half hour and hour and applies concepts of time • demonstrates a basic ability to perform visual/spatial reasoning with minimal prompts • demonstrates a basic understanding of sequencing • student demonstrates a basic understanding of and the ability to use math vocabulary • demonstrates the ability to generalize knowledge and skills to different environments and with some supports
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • responds accurately when choosing between two answers • demonstrates a limited understanding of abstract math concepts and skills • demonstrates a limited ability to tell time or apply the concepts of time • demonstrates a limited ability to perform visual/spatial reasoning • requires concrete manipulatives when creating a pattern • demonstrates a limited understanding of math vocabulary • demonstrates a limited ability to generalize knowledge and skills to different environments
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • anticipates a math activity • attends to materials being displayed • demonstrates the ability to cover a figure with shapes • produces a numeral to 10

Alternate Performance Level Descriptors for Grade 7 Reading

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • makes inferences • sequences beginning, middle, and end and supporting details (specific facts) • differentiates between fact and opinion • understands abstract vocabulary (true/false) • identifies/understands various genre (i.e., cultural lessons, informational, fables/myths, biographies) • understands story lessons/author's purpose • identifies chapter heading (abstract sense) to find/use info • uses reading strategies to gain information (i.e., rereading, use of key words, use of features of text)
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • demonstrates readiness with limited/no prompting • sequences beginning, middle, and end • recalls multiple facts about a reading selection • understands literal vocabulary and the relationships • identifies main idea of the story and some supporting facts/details • identifies purposes of various texts (i.e., map, dictionary, bus schedule, etc.) • identifies title and basic parts of a book • responds with three response options
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • demonstrates readiness by following one-step directions or with teacher modeling/prompting • identifies an object and its function • maintains focus from beginning to end • understands story beginning and ending • understands basic main idea (answer with one picture/short response) • recalls at least one fact about a reading selection • locates name of book and basic print awareness • responds mostly through basic yes/no questions or with two options (or three options with further teacher clarification)
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • anticipates a reading activity • attends to materials being displayed • directs attention to external stimuli when requested (i.e., turns head in direction, sits quietly, etc.) • interacts with stimuli • responds to external stimuli (i.e., nods head, operates switch, points to, etc.) • is assisted through a correct response • attempts to participate in activity • has general awareness of people and activity • responds to own name • responds to words, pictures and symbols

Alternate Performance Level Descriptors for Grade 7 Mathematics

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • engaged in the task • understands 1:1 correspondence • adds/counts money • graphs • sorts and makes decisions based on sorting
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • identifies coins and values • sorts objects by function • makes comparisons (>,<) • makes a statement about the data • adds and subtracts
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • knows 1:1 correspondence, concept of "none" • understands the concept addition (more) • understands the concept subtraction (less) • matches coins • sorts by appearance, various (two or more) characteristics (size, shape, color)
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • anticipates a math activity • attends to materials being displayed • attends to models/prompts • recognizes numbers (symbol or rote recitation) • sorts by one characteristic

Alternate Performance Level Descriptors for Grade 8 Reading

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • connects prior knowledge to make meaning of text • identifies main idea and various supporting details • understands story lessons • locates title and other information from a variety of documents/sources • recognizes vowel letter-sound • uses reading and/or listening strategies when needed to gain information (i.e., rereading, use of key words, use of features of text) • comprehends a simple paragraph
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • arrives at correct answer with limited prompting • has basic word recognition • tracks while reading or being read to • identifies basic words and recognizes some words in different contexts • identifies a word/picture/symbol for content communication • identifies title and basic parts (beginning, middle, and end) of a reading selection • identifies main idea of a story and some supporting facts/details • identifies purposes of various texts (i.e., dictionary, map)
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • requires a high level of rephrasing • shows an understanding of the beginning and end of a story by giving attention to the reader or the text • recognizes that letters have names and is aware of letter sounds • recognizes difference between letters and other symbols (i.e., numerals) • identifies letters by name/sign • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • identifies a word/picture/object of familiar places and people • responds mostly through basic yes/no questions • understands basic main idea (answer with one picture/short response)
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a reading activity • attends to materials being displayed • demonstrates readiness by following one-step directions or with teacher modeling/prompting • responds to name, words, pictures, and symbols • directs attention and responds to external stimuli when requested (i.e., turns head in direction, nods head, operates switch, points to, etc.) • interacts with stimuli (i.e., teacher, words, pictures, and symbols)

Alternate Performance Level Descriptors for Grade 8 Mathematics

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • measures to the inch • measures the distance between two points on a map • completes missing components in basic number sentence • recognizes and understands all operational symbols (+, –, =), measurement symbols (in., cm, etc), monetary symbols ($), and time • uses all comparison words (more, less, some, none, most, least) correctly • understands ordinal numbers beyond 3rd • selects the correct label for a graph (i.e., label axis) • explains conclusions drawn from graph • applies beginning connections between concrete and symbolic representations, operations, measurement, graphing and problem solving strategies
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • arrives at correct answer with limited prompting • reads simple measurements • can subtract single digits • understands first three ordinal numbers (1st, 2nd, 3rd) • uses some comparison words (more, less, some, none, most, least) correctly • understands numbers can represent monetary amounts, measurement, and time • demonstrates basic problem solving skills • fills in data, as directed, to create a representation on a bar graph • recognizes and understands most operational symbols (+, –, =), measurement symbols (in., cm, etc), monetary symbols ($), and time • identifies places on a map • answers questions about a bar graph • makes a statement about data • demonstrates knowledge of basic number sentences
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • identifies and/or recognizes a map and measuring tools • demonstrates solid number concept for 1:1 correspondence (consistently touch counts) • can count single digits • can add single digits • recognizes and understands some operational symbols (+, –, =), measurement symbols (in., cm, etc), and monetary symbols ($) • basic understanding of bar graphs and data • can make general statements about a bar graph
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • anticipates a math activity • attends to materials being displayed • attends to another person reviewing a map with prompting • attends to another person reviewing a graph with prompting • engages with instructor with prompts • recognizes numbers (symbol or rote recitation)

Alternate Performance Level Descriptors for Grade 8 Science

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • independently attends • no scaffolding on most items • best answer majority of the time • shows understanding of content most of the time
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • can attend • when difficult distracters are reworded, student will answer correctly • identifies correct answer out of three choices most of the time
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • can attend • identifies correct answer out of two choices most of the time • guess level performance • limited understanding of content
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • requires assistance to select correct response with maximum scaffolding

Alternate Performance Level Descriptors for Grade 10 Reading

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • identifies main idea and supporting details from various reading selections • identifies appropriate resources for gaining specific information • draws conclusions from a variety of texts (i.e., poem, fiction) • communicates meaning of new and unfamiliar vocabulary • communicates a complete thought related to topic or concept • uses word-recognition skills, context clues, and prior knowledge to understand text • rereads to gain understanding
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • arrives at correct answer with limited prompting • has basic reading and comprehension skills • understands difference between various literacy materials • begins to access prior knowledge to understand text • communicates a basic thought on topic • identifies main ideas and some supporting details/facts • is beginning to identify appropriate resources for gaining specific information • identifies words/pictures/symbols and objects that are new and unfamiliar
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • explores literary items (holds reading material in correct position, recognizes pictures vs. print, uses left to right orientation) • able to match and identify familiar words/pictures/symbols/objects • identifies basic main idea (answers with one picture/short response) • communicates an opinion • identifies familiar literary resources (i.e., newspaper, CDs, Internet, oral histories)
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • attends to materials being displayed • responds to name, words, pictures and symbols • demonstrates readiness by following one-step directions or with teacher modeling/prompting • directs attention and responds to external stimuli when requested (i.e., turns head in direction, nods head, operates switch, points to, etc.) • interacts with stimuli (i.e., teacher, words, pictures, and symbols)

Alternate Performance Level Descriptors for Grade 10 Mathematics

Advanced	The student at the Advanced level accurately and independently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistently and independently arrives at correct answer • generalizes very basic information • completes two to three-step processes of addition and subtraction • completes basic division and multiplication problem • applies beginning connections between concrete and symbolic representations by using a chart/table to draw conclusions • creates graph/tables and explains conclusions drawn from graph • understands and communicates relationship between variables • solves problems using bills and their values • follows navigational directions and recalls shapes and locations
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • arrives at correct answer with limited prompting • completes and/or extends basic patterns of data • sorts items into sets by multiple defining characteristics • demonstrates basic connections between concrete and symbolic representations • identifies basic information from a graph/chart • matches bills and their values • recognizes and identifies two-dimensional shapes • chooses correct procedures to solve simple number problems • adds and subtracts two-digit numbers
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow set of content-specific performance indicators. • arrives at correct answer with moderate prompting • recognizes properties of limited (square/circle) two-dimensional shapes • recognizes distinct categories • recognizes basic patterns of data • sorts items into sets by one defining characteristic • understands quantity • can count single digits • can add/subtract single digits • communicates beginning connections between concrete and symbolic representations
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. • requires high level of prompting/physical assistance to arrive at correct answer • attends to materials being displayed • shows limited understanding of quantity when given two choices • recognizes numbers (symbol or rote recitation)

Alternate Performance Level Descriptors for Grade 10 Science

Advanced	The student at the Advanced level accurately, independently, and consistently demonstrates the ability to carry out comprehensive content-specific performance indicators. • consistent performance across standards • capable of abstract thought/models • understands scientific variables • ability to handle three distracters • ninety-five percent of responses will be “4”
Proficient	The student at the Proficient level, given limited prompting, demonstrates the ability to respond accurately in performing a wide variety of content-specific performance indicators. • less scattered performance across standards • exhibits more abstract thinking • ability to relate cause to effect • recognizes there is a scientific process • majority of responses are “3”+ • ability to handle two or more distracters • expanded exposure to science content
Nearing Proficiency	The student at the Nearing Proficiency level, given moderate prompting, demonstrates the ability to respond accurately in performing a narrow concrete set of content-specific performance indicators. • ability to attend and show compliance • identifies concrete concepts and objects of science • performance on standards may vary • greater understanding/skills related to daily living as related to science • majority of responses will earn a “2”+ • can handle limited distracters • limited exposure to science content
Novice	The student at the Novice level, given physical assistance and/or modeling, is supported to participate in content-specific performance indicators. (none)

Reading and Mathematics Raw and Scaled Scores: Grades 3 through 8, and 10

CRT-ALT Raw Cut Scores 2009				
Grade	content area	cut1	cut2	cut3
03	REA	40	74	95
03	MAT	75	88	98
04	REA	50	77	95
04	MAT	57	77	93
05	REA	48	74	88
05	MAT	72	81	98
06	REA	43	68	93
06	MAT	60	89	98
07	REA	32	59	88
07	MAT	42	69	96
08	REA	47	66	85
08	MAT	51	69	89
10	REA	55	71	92
10	MAT	51	80	93

CRT-ALT Scaled Cut Scores 2009				
grade	content area	cut1	cut2	cut3
03	REA	225	250	265
03	MAT	225	250	269
04	REA	225	250	267
04	MAT	225	250	269
05	REA	225	250	263
05	MAT	225	250	297
06	REA	225	250	275
06	MAT	225	250	258
07	REA	225	250	277
07	MAT	225	250	275
08	REA	225	250	275
08	MAT	225	250	278
10	REA	225	250	283
10	MAT	225	250	261

Science Raw and Scaled Scores: Grades 4, 8, and 10

CRT-ALT Raw Cut Scores 2009				
grade	content area	cut1	cut2	cut3
04	SCI	59	78	96
08	SCI	46	73	96
10	SCI	73	93	108

CRT-ALT Scaled Cut Scores 2009				
grade	content area	cut1	cut2	cut3
04	SCI	225	250	274
08	SCI	225	250	271
10	SCI	225	250	269

Appendix C—SAMPLE TASKLET

SAMPLE TASKLET

Content Standards Addressed: Standard 4: Geometry
4.1 Students will describe, model and classify two- and three-dimensional shapes.

Activity
This activity engages students in demonstrating and understanding of two- and three dimensional shapes by

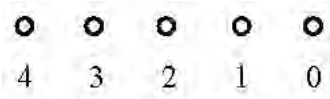
- identifying two congruent shapes from a set of shapes; sorting triangles and squares into groups;
- identifying a circle among four different shapes; and
- using spatial reasoning to match shapes with congruent shapes in different orientations.

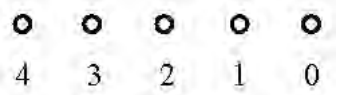
Materials Provided

- Squares: 2 large, 1 medium, 1 small
- Triangles: 1 large, 1 medium, 1 small
- Circles: 1 large, 1 medium, 1 small
- Rectangles: 1 large, 1 medium
- Sorting Template
- Matching Template

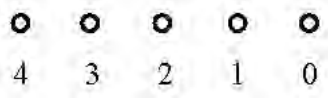
Other Materials Needed

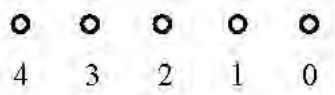
- Materials typically used by the student for reading/writing other that what is provided in this kit
- Materials typically used by the student to communicate (e.g., communication device, objects, switches, eye gaze board, tactile symbols)
- Throughout the activity, make any material substitutions necessary to enable the student to understand test questions (e.g., objects, larger print, different pictures, materials in auditory formats).
- Materials provided may need to be further adapted for students who are hearing or visually impaired. Suggestions for adapting materials are in the CRT-Alternate Administration Manual.

Materials	Activity Steps Teacher will:	Student Work Student will:	Performance Indicators Use Scoring Guide
1. • 1 medium square • 1 medium triangle • 1 medium circle **Communication support strategies:** • Word/picture symbols for “yes” and “no” may be used to indicate readiness to move on. • Throughout the activity, make any material substitutions necessary to enable the student to understand test questions (e.g., objects, larger print, different pictures, materials in auditory formats).	1. Place all the shapes on the work space. ***“Let’s start now. Here are 3 different shapes. This is a square. A square has 4 straight equal sides. This is a triangle. A triangle has 3 straight sides. This is a circle. A circle is a closed shape that is round with no straight sides. Did you see/hear about the 3 shapes I just showed you?”*** Allow the student to touch the shapes.	1. Attend to the teacher naming a square, triangle, and a circle.	1. Attend to objects or pictures of two- and three- dimensional geometric shapes and the relationships among them.  Performance Indicator: 4.1.1.1 Expanded Benchmark: 4.1.1

Materials	Activity Steps Teacher will:	Student Work Student will:	Performance Indicators Use Scoring Guide
2. • 1 large square • 1 large triangle • 1 large circle • 1 large rectangle **Communication support strategies:** • Student may look at/point to task materials to express a choice. • Request may be rephrased to require a yes/no response (e.g., "Is this the circle?") • Student may tell teacher to "stop" at desired response as teacher sequentially points to each of the 4 choices.	2. Place all the shapes in random order on the work space. ***"Show me the circle."*** <u>Scaffold</u>: <u>Level 3</u>: Remove an incorrect response. Repeat task request. <u>Level 2</u>: Remove another incorrect response. Repeat task request. <u>Level 1</u>: *"This is the circle."* Assist the student as needed to identify the circle.	2. Identify a circle.	2. Identify (name) shapes as circles, squares, triangles, rectangles, and ovals.  Performance Indicator: 4.1.1.6 Expanded Benchmark: 4.1.1

Materials	Activity Steps Teacher will:	Student Work Student will:	Performance Indicators Use Scoring Guide
3. • Triangles: 1 large, 1 medium, 1 small • Squares: 1 large, 1 medium, 1 small • Sorting Template **Communication support strategies:** • Student may look at/point to task materials to express a choice. • Request may be rephrased to require a yes/no response (e.g., *"Is this where the square should go?"*) • Student may tell teacher to "stop" at desired location.	3. Place all the shapes in random order on the work space. ***"Here are some squares and triangles. Put all of the squares together and all of the triangles together."*** <u>Scaffold</u>: <u>Level 3</u>: Place the sorting template in front of the student. Review the picture of the square and the triangle on the template. *"Put all of the squares here and all of the triangles here."* <u>Level 2</u>: Place 1 square and 1 triangle on the template. *"I put 1 square and 1 triangle on the paper. Now, you finish putting the squares together and the triangles together."* <u>Level 1</u>: Place the rest of the triangles and the squares on the paper. *"All of the squares are here. All of the triangles are here."* Assist the student as needed to identify the group of triangles.	3. Indicate that all the triangles belong together and all the squares belong together.	3. Sort 2-dimensional physical shapes according to their shape. Performance Indicator: 4.1.1.5 Expanded Benchmark: 4.1.1

Materials	Activity Steps Teacher will:	Student Work Student will:	Performance Indicators Use Scoring Guide
4. • 1 large triangle • 1 small triangle • 2 congruent large squares • 1 small square **Communication support strategies:** • Student may look at/point to task materials to express a choice. • Request may be rephrased to require a yes/no response (e.g., *"Is this shape the same size and shape as this shape?"*) • Student may tell teacher to "stop" at desired location.	4. Place all the shapes on the work space. ***"Show me the 2 shapes that are the same shape and size."*** Note: When removing shapes, only remove the triangles and small square. <u>Scaffold</u>: <u>Level 3</u>: Remove an incorrect response. Repeat task request. <u>Level 2</u>: Remove another incorrect response. Repeat task request. <u>Level 1</u>: *"These 2 shapes are the same shape and size. They both are squares."* Assist the student as needed to identify the congruent squares.	4. Identify congruent squares.	4. Recognize 2-dimensional physical shapes as being the same (congruent) or different.  Performance Indicator: 4.1.1.4 Expanded Benchmark: 4.1.1

Materials	Activity Steps Teacher will:	Student Work Student will:	Performance Indicators Use Scoring Guide
5. • 1 medium square • 1 medium triangle • 1 medium rectangle • Matching Template **Communication support strategies:** • Student may look at/point to task materials to express a choice. • Request may be rephrased to require a yes/no response (e.g., *"Does this shape match this shape?"*) • Student may tell teacher to "stop" at desired location.	5. Place the matching template and 4 shapes on the work space. ***"Match each of these shapes with its picture."*** <u>Scaffold</u>: <u>Level 3</u>: Remove incorrect responses from the template and validate the correct responses. If student did not have a correct response, place a shape with its picture. *"I matched the ____ with its picture. Now, you finish matching the shapes with their pictures."* <u>Level 2</u>: Remove incorrect responses from the template and validate the correct responses. Match 2 shapes with their pictures. *"I matched the ____ and the ____ with their pictures. Now, you finish matching the shapes with their pictures."* <u>Level 1</u>: Remove the incorrect responses. Match the remaining shapes with their pictures. "Each shape is with its picture." Assist the student as needed to match the 4 shapes to their pictures.	5. Match 4 shapes with their pictures in different orientations.	5. Match 2-dimensional physical shapes to pictures of the shapes in different orientations.  Performance Indicator: 4.1.1.7; 4.5.1.5 Expanded Benchmark: 4.1.1, 4.5.1 **End of Sample Tasklet**

Appendix A—INTERRATER RELIABILITY REPORT

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

Examining the Interrater Reliability of Montana's CRT-Alternate

Gail McGregor, Ed.D.
University of Montana-Missoula

Submitted
by
The Montana Office of Public Instruction
to
The United States Department of Education
for
Peer Review
by
Judy Snow
State Assessment Director

As an outcome of the U.S. Department of Education's review of Montana's assessment system, the state was asked to submit *evidence* of the interrater reliability of its alternate assessment, the CRT-Alt. Dr. Stanley Rabinowitz, a consultant made available to Montana's Office of Public Instruction by the U.S. Department of Education because of his role with the Assessment and Accountability Comprehensive Center, provided guidance that led to the design of a study to respond to this requirement. This design was shared with Montana's Technical Advisory Committee at its January, 2007 meeting. With their endorsement, the study was implemented during the spring, 2007 testing window. This report summarizes the results of this effort.

Design of the Study

As suggested by Dr. Rabinowitz, this study was designed to gather multiple sources of data that, collectively, would produce a "preponderance of evidence" supporting the overall integrity as well as the interrater reliability of the CRT-Alt. This broader view is based on the belief that scoring will not be meaningful if the assessment is not administered as required. This approach is responsive to the unique characteristics of Montana, and the small number of students with disabilities who take this form of the test. During the March, 2007 assessment period, a total of 698 students were tested using the CRT-Alt across grades 3, 4, 5, 6, 7, 8, and 10. The number of students tested per grade ranged from a low of 84 students in Grade 5 to a high of 133 students in Grade 6.

The study encompassed plans to gather data relative to five test characteristics. These focus areas, and the data sources used to evaluate them, are summarized in Table 1 below.

Table 1: Test Characteristics and Sources of Evidence for CRT-Alt Interrater Reliability Study

Test Characteristic	Source of Data
1. Evidence-Base for Practices used in Test Design	• Review of professional literature addressing pedagogical practices for students with severe cognitive disabilities. • Examination of reliability indices in published research using presentation and prompting methodology adopted for the CRT-Alt.
2. Accessibility of Training for Test Administrators	• Test administrator training survey. • Test administrator questions included in the Student Response Booklet.
3. Test Administrator Knowledge and Understanding of Testing Procedures	• Test administrator training survey. • Independent observer ratings of fidelity of test administration.
4. Fidelity of Test Administration	• Independent observer ratings of fidelity of test administration.
5. Level of Agreement: Item Scoring	• Comparison of scores of test administrator with those of a trained independent observer present during test administration. • Sample of Evidence Templates submitted with Student Test Booklet, reviewed and scored by independent reviewer.

In the remainder of this report, the activities that have been undertaken in each of these areas, and the results, are summarized.

Use of Evidence-Based Practices in Test Design

The CRT-Alt is a performance based assessment, measuring a student's response to a series of test items that are presented in the format of short instructional tasks. Given the heterogeneity of the students who are eligible to be assessed with this instrument in terms of their motor, sensory, language, and cognitive skills, the test builds in considerable flexibility in regard to the materials used to present test items, and the response modalities used by students to communicate and interact throughout the assessment. For example, real objects may be substituted for the pictures provided in the test materials kit to accommodate students with visual limitations. In sharp contrast to this flexibility, all other aspects of the administration and scoring of this assessment are tightly controlled.

Administration of the CRT-Alt incorporates a response prompting methodology known as the "system of least prompts" (Wolery, Ault & Doyle, 1992). This is a well-established strategy that has been found to be effective as a teaching procedure for students with severe disabilities across a wide range of applications (Doyle, Wolery, Ault & Gast, 1988). The rationale for its use in this testing context is based on the information summarized below.

- Students with severe disabilities often demonstrate skill gains in small increments that would be lost if performance was scored with a dichotomous correct/incorrect response system. For this population of students, learning is typically measured in terms of the amount of support required to produce a correct response. When responses do not occur independently, a structured sequence of prompts allows teachers to consistently present and systematically control the amount of external support provided in a teaching situation. Student learning is measured in terms of increasing levels of independence (i.e., decreased reliance upon external prompts).

  The CRT-Alt uses a "least to most" prompt hierarchy. As described by Wolery et al. (1992), the system of least prompts consists of a hierarchy of at least three levels. The first level is the opportunity for a student to respond independently, without external prompts. If that does not occur, a planned sequence of prompts, arranged from the least intrusive to the most intrusive in terms of amount of assistance, is implemented. The final level of the prompt sequence results in an assisted, correct response. For the CRT-Alt, a four level hierarchy has been developed for each test item.

  With origins in an applied behavior analysis model of teaching that dates back to the late 1960's and 70's, the prevalence and value of this methodology for students with severe disabilities is unquestioned in the research and practice literature (e.g., Alberto & Troutman, 1995; Demchak, 1990; Falvey, 1986). While much has been learned about effective instruction for students who experience significant challenges to learning since that time, the value of systematic instructional procedures continues to be recognized. The sixth edition of one of the most popular textbooks on teaching students with severe disabilities (Snell & Janney, 2006) continues to emphasize the importance of these very procedures in working with students with severe disabilities.

- Since prompt response systems are a common teaching approach for students with severe disabilities, teachers are familiar with this methodology and use it on a regular basis. University coursework focused on the needs of students with severe disabilities emphasizes systematic instructional procedures that are grounded in the science of applied behavior analysis. A national review of preservice programs (Ryndak, Clark, Conroy & Stuart, 2001) verifies the importance of this skill set in teacher preparation programs focused on the needs of students with severe disabilities. Because this is an effective and common teaching methodology, the approach to test administration is relatively easy to understand and implement for those experienced in teaching students with severe cognitive disabilities. Most recent data available from the Office of Public Instruction indicate that for the 2005-06 school year, 98.5% of the state's 750 special educators were reported to be Highly Qualified, suggesting their familiarity with this methodology.

- In the extensive research base about response prompting systems, acceptable levels of interrater reliability have been achieved. The use of this and other response prompting methods has been a strategy used in special education research for over thirty-five years. This body of research utilizes single subject research methods (Tawney & Gast, 1984) due to the low incidence and unique characteristics of the participants in these studies. Direct observational data are collected, requiring the use of independent observers to verify the reliability of the observational data. A standard rule of thumb in this type of research is that an average reliability index of 80% is acceptable. Results typically are reporting in the 85-95% range (e.g., Colyer & Collins, 1996; McDonnell, 1987; West & Billingsley, 2005), as the prompting procedures are clearly spelled out, easy to implement, and readily observable. This evidence provides a strong foundation for the selection of this methodology for this assessment context, especially under conditions of tight controls for the training and administration of the measure, as is the case in Montana.

The administration of the CRT-Alt is based upon systematic procedures that are time-tested and evidence-based with the population of students for whom this test is designed. In this application, *scaffolding* is the term used to describe the least to most prompting process that is consistently and predictably used in the administration of each item. Each test item is carefully scripted, eliminating the need for teachers to determine how to present a question or what should be said. The scaffolding sequence is also scripted, guiding the teacher in a step-by-step manner through the administration of each test item.

This same predictable and consistent structure is applied to the scoring of each item. The scaffolding sequence is directly aligned with the scoring rubric for each test item. Finally, there is a requirement that test administrators submit selected pieces of evidence for each student in all subject areas tested. Submission of concrete evidence of student's performance relative to a specifically designated test item provides a means of checking whether information recorded on evidence templates are consistent with item scores entered on student scoring forms.

Collectively, these design features create a standardized structure intended to provide teachers with sufficient support to implement the CRT-Alt with integrity. Other components of OPI's implementation approach, described in the next section, further support this goal.

Accessibility of Training

For the 2006-07 test administration, the OPI implemented a training plan designed to address the limitations of large group training formats, conducted over the state's compressed video system and the internet, used in previous years. There was a general consensus that this training did not reach the intended audience – the actual test administrators. To address this concern, a training package was prepared and included in the Test Materials Kit provided to every test administrator. An Implementation Checklist (see Appendix A) was included in this Kit, indicating that reviewing the test training CD was the first thing that was to be done in preparing for test administration. System Test Coordinators were also alerted to the expectation that test administrators access these training materials prior to test administration.

In order to measure the success of this approach, two questions were included in the teacher-only section at the end of the test administration booklet. Additional questions were asked in a separate survey document distributed with the test materials, designed to gather information about the level of experience of the test administrators and the source of their training. These questions, and a summary of the responses received, are provided in Tables 2 through 4. In viewing these data, the total possible number of respondents is 632. This number represents the total number of students tested. However, some test administrators tested more than one student, meaning that they may have responded to the questions each time they administered the test.

Table 2: Test Administrator Responses to Yes/No Training Questions (N=632)

Training Question	Response (number/percent of respondents)		
	Yes	No	No Response
Have you given the CRT-Alt before this year, 2007?	317 (50%)	109 (17%)	206 (33%)
Did you view the teacher training CD provided with the test materials before administering the test?	462 (73%)		170[1] (27%)

[1]NOTE: "No" was not a response option. Respondents answered in the affirmative if they DID view the training CD, so it is not possible to distinguish between those who did not view the CD and those who skipped the question.

As seen in Table 2, at least half of the test administrators responding to this question reported having given the CRT-Alt before. Given the fact that this questions was left blank on the test booklets for one third of students, the actual percentage could very well be higher. It is reasonable to conclude that the population of CRT-Alt test administrators in 2007 was mostly experienced with this test. This provides a context in which to view the data about the number of test administrators who viewed the CD before administering the test.

Interpreting the responses given to the question "*was the training CD used?*", must be done with caution. The only choice on the scan form for respondents to fill in for this question was an affirmative option, indicating that they did view the CD. The assumption in the design of the response form was that those who did not view the CD would leave this blank. Unfortunately, the proportion of other items left blank on this survey makes it impossible to distinguish between true "no" responses and those that were simply skipped. With this caveat, affirmative responses to this question were made by test administrators for almost three-fourths of the students tested. The CD was a training format that did make the information accessible to those who needed it.

Information reported in Table 3 places the use of the training CD within the larger context of test administrator experience and other supports that might be provided on the local level. It was possible to mark more than one option for the question "*Describe the training you received to give this test.*" As seen in this table, the largest percentage of respondents reported receiving training through the use of the CD provided by OPI either in the current year (58%) or in a previous year (22%). Twenty percent of the respondents reported attending a training session, while 11% indicated watching the CD and attending training. A single respondent

reported having never accessed training materials prior to test administration.

Table 3: Test Administrator Training Access (N=492)

Source of Training	Response (number/percent of respondents)[1]
Used training CD in 2007	285 (58%)
Attended a training in 2007	100 (20%)
Used CD and attended training in 2007	53 (11%)
Received training or viewed CD in previous year(s)	106 (22%)
Have never accessed training materials	1 (.002%)

[1]Respondents were instructed to check all responses that apply.

The final dimension of the training that was considered was the test administrator's perception of its value. They were asked to rate its value on a four-point rating scale, with a rating of "1" indicating that it was not very valuable, and "4" indicating that it was extremely valuable. Since this question was included in the back of the Student Response Booklet, a total of 632 responses were possible.

As seen in Table 4, forty-five percent of the respondents felt the training was "valuable" or "extremely valuable". The meaning rating among respondents was 2.68. This item was left blank in 25% of the Student Response Booklets. It is not possible to know whether these were left blank because the test administrator did not view the CD this year (see results above), had already responded to this question when completing the test booklet for another student, or simply chose not to respond to this question. Nevertheless, available data suggest that the training format was generally seen as helpful.

Table 4: Test Administrator Ratings of Training CD (N=632)

1 (not very valuable)	2	3	4 (extremely valuable)	No Response	Mean Rating
51 (08%)	133 (21%)	204 (32%)	84 (13%)	160 (25%)	2.68

Test Administrator Knowledge and Understanding of Testing Procedures

The next component of the research plan focused on the impact of the training materials on test administrator knowledge and understanding of the testing procedures. A series of questions was posted on a website, which test administrators were directed to access, after they had finished reviewing the training materials. For those teachers without ready access to the internet, a Word document was included on the training CD, enabling teachers to complete this training post-test, and submit it via e-mail or FAX. In order to encourage responses, teachers

were not required to identify themselves.

A total of 35 responses were received. Of this total, 9 were received via e-mail, 1 was received via FAX, and the remaining 25 surveys were completed online. While this was a disappointing rate of response, it is not possible to pinpoint exactly what percent of respondents are represented by these data. As the testing contractor for Montana's CRT-Alt, Measured Progress adds these questions to the end of the test administration booklet for each student and subject area. As a result, there is some duplication in respondents since many teachers administer the assessment to more than one student. Information provided by Measured Progress indicates that 288 unique teachers were identified as test administrators for the March, 2007 assessment. Unfortunately, the teacher identification field was not completed in a number of surveys. Given this situation, the best approximation of the response rate is 12%.

As illustrated in Table 5, those that did respond to the survey correctly answered questions about the training content. The proportion of those responding correctly to the questions ranged from 89% to 100%. The questions asked, and results for each, are provided in Table 5.

Table 5: CRT-Alt Training Evaluation Questionnaire Summary (N=35)

Question **[correct response]**	**Number (%) Correct**	**Number (%) Incorrect**	**Number (%) Missing**
1. The CRT-Alt should be administered by a certified teacher who is familiar with the student being tested. [TRUE]	**32 (91%)**	**2 (6%)**	**1 (3%)**
2. It is not permissible for another person to assist in the administration of the test. [FALSE]	**33 (94%)**	**2 (6%)**	**0 (0%)**
3. The skills assessed in the CRT-Alt are aligned with Montana's Curriculum Standards, with benchmarks that have been expanded to measure skills that lead to the acquisition of grade level skills. [TRUE]	**35 (100%)**	**0 (0%)**	**0 (0%)**
4. All materials required to administer the CRT-Alt are provided in the Test Materials Kit. [FALSE]	**34 (97%)**	**1 (3%)**	**0 (0%)**
5. Test administrators can modify the script provided for the test questions, using language that the student will understand, if the intent of the statement remains the same. [TRUE]	**32 (91%)**	**3 (9%)**	**0 (0%)**

Question [correct response]	Number (%) Correct	Number (%) Incorrect	Number (%) Missing
6. Scaffolding refers to the careful placement of test materials on the work space. [FALSE]	**32 (91%)**	**3 (9%)**	**0 (0%)**
7. The score a student receives for each test item is unrelated to the amount of assistance required for the student to produce a correct response. [FALSE]	**33 (94%)**	**2 (6%)**	**0 (0%)**
8. The Halting Rule describes when it is permissible to discontinue the test due to student resistance. [TRUE]	**32 (91%)**	**3 (9%)**	**0 (0%)**
9. Introductory items in each task/tasklet are scored on a simplified rubric of 4 and 0. [TRUE]	**33 (94%)**	**2 (6%)**	**0 (0%)**
10. A magnifying glass indicates that evidence must be collected to document the response made by the student. [TRUE]	**34 (97%)**	**1 (3%)**	**0 (0%)**
11. Scores from the student Test Booklet must be transferred to a scanning form that is part of the Student Kit. [TRUE]	**31 (89%)**	**3 (9%)**	**1 (3%)**
12. A score of “4" indicates that the test administrator provided complete assistance to the student to make the response. [FALSE]	**34 (97%)**	**1 (3%)**	**0 (0%)**
13. Students are not allowed to use specialized communication devices during testing. [FALSE]	**34 (97%)**	**1 (3%)**	**0 (0%)**

Fidelity of Implementation

While the initial areas of investigation focused on the training and preparation of test administrators, the remainder of the study examined implementation and scoring practices. An Implementation Checklist (see Appendix A) was developed to serve as a self-check for test administrators to ensure that they performed all test administration steps accurately and completely. A question was included in the test administrator survey to determine the extent to which this tool was actually used. As shown in Table 6, test administrators responsible for implementing the assessment for 56% of the students tested reported that they did use the Checklist. While only 11% said they did not, this question was left blank in the test booklets of 33% of the students.

Table 6: Test Administrator Responses to Implementation Checklist Question (N=632)

Training Question	Response (number/percent of respondents)		
	Yes	No	No Response
Did you check your test administration procedures against the Implementation Checklist that was provided with the 2007 training CD sent with the materials kit/ replacement materials?	357 (56%)	69 (11%)	206 (33%)

The second method of assessing fidelity of test implementation was through the direct observation of test administrators. During a December, 2006 phone consultation with Dr. Stanley Rabinowitz, the issue of sampling size and composition for an interrater reliability study was discussed. Given the few number of students in the testing pool, the size of the state, and the limited resources available to train and deploy qualified observers, his recommendation was that we begin with a sample of no less than 5 students per grade, with observations focused on both math and reading. If initial findings with this limited sample size showed mixed results in terms of scoring reliability and implementation fidelity, he indicated that additional observations would be required until more definitive findings were obtained. Further, the study should be repeated over multiple years to provide more cumulative evidence supporting the technical adequacy of the assessment.

When statewide information was available to indicate where students registered for the CRT-Alt were located, a sampling plan was developed that balanced statewide distribution with the practical reality of where students registered to take the CRT-Alt were clustered. The final plan, contained in Appendix B, included observation of 5 students each in Grades 3, 4, 5, 6, 7, 8 and 10. Half of the students were observed being tested in Reading, while the other half were observed during the Math Assessment. Students in the sample attended schools in the Bozeman, Helena, Billings, Great Falls, and Missoula and the small towns in the surrounding areas. Beyond the steps taken to stratify the sample to get equal representation of students at each grade level, across subject areas, and within each region of the state, the other steps taken to finalize student selection were driven by logistics. A list was compiled to indicate the location of students within each grade level Final student selection was driven by matching test administration scheduling with the availability of independent observers to travel to a school at these scheduled times.

During January and February of 2007, independent observers were recruited and trained to implement the CRT-Alt. They were also introduced to the specific observation procedures that had been developed for this study. Four experienced educators were found to observe in the Helena, Bozeman, Great Falls and Billings area school districts. In the region around Missoula, five graduate students in school psychology were recruited to serve as observers, receiving the same training as the other observers. All observers conducted a “test run” to ensure the procedures were understood before moving into the actual observations for the purposes of this study.

During each school visitation, observation focused the fidelity issues listed below. The forms used to structure and these observations are contained in Appendix C.

- Teacher interview – teacher report of test preparation activities
- Observation of test implementation practices – occurred for an entire tasklet (Grades 3, 5, 6, 7) or 5 consecutive items in a Task (Grades 4, 8, 10)

Results of the test fidelity observations are summarized in Table 7. Information in this table is based upon observation protocols coded for 40 student/teacher pairs, a slightly larger sample than the lower limit recommended by Dr. Rabinowitz. Results indicated a consistently high level of fidelity in each key procedure that is part of the testing procedures. Test administrators observed presented the materials as described in the test booklet, and accurately followed by scripted scaffolding procedures. Introductory items, implemented in a slightly different way than other test items, were implemented correctly 95% of the time. Similarly, as described in the test booklet, students were given an opportunity to respond independently before the test administrator moved on to the use of the sequential scaffolding procedures. When these were required, they were used with fidelity 97% of the time. The only implementation practice falling below the 95% fidelity level involved the documentation of evidence. Most observers wrote explanatory notes that when these items came up, the teacher often elected to actually fill out the evidence recording form after the test administration was halted in order to maintain attention to the student and maintain the pace of the assessment.

Table 7: Fidelity of Implementation Results

Test Administration Practice	% of Observations Practice Observed
Test Preparation	
Teacher reported that they had participated in training about test administration	95%
All materials for test administration not included in test kit have been located	95%
Test materials are organized and easily accessible for test administration	95%
Test is administered in a location in which student can work without interruption	90%
Implementation Practices	
Introductory items were implemented without scaffolding, scored as either a “4" or “0"	95%
Teacher presented the materials as described in the Test Booklet.	95%
Student was given an opportunity to respond independently before any scaffolding was provided	95%
Teacher implemented the scaffolding as described in the Test Booklet.	97%
Teacher scored student response based on the level of scaffolding necessary	97%
Teacher documented evidence for those items that required it.	85%

Level of Agreement

Direct observation of test administration was conducted to gather data to assess the level of agreement between the test administrator and an independent observer. This involved the independent scoring of a minimum of 5 consecutive test items (Grades 4, 8, 10) or an entire tasklet for students assessed in grades 3, 5, 6 and 7. No interaction occurred between observer and test administrator relative to the scoring of these items. The test administrator submitted the student scores to Measured Progress, following established procedures for returning materials. The independent observers submitted their observation materials to OPI. These materials were sent to Measured Progress for analysis.

Results of the comparison in scoring between test administrators and independent observers are summarized in Table 7. An overall agreement index of 88% is based on data gathered in nineteen observations of students taking the Reading assessment, and twenty-one observations of students taking the Math assessment. The agreement level for Reading assessment items was 83%, while the level of agreement for math tasks was 91%. A breakdown of this information by grade and subject is provided in Table 8.

Table 8. Interrater Reliability Indices By Subject and Grade

Grade	Reading Results		Math Results		Combined Results	
	# of Items	% Agreement	# of Items	% Agreement	# of Items	% Agreement
3	29	69%	10	100%	39	77%
4	21	100%	38	90%	59	93%
5	16	69%	35	97%	51	88%
6	24	92%	20	100%	44	95%
7	4	100%	40	88%	44	89%
8	20	100%	20	90%	40	95%
10	27	70%	28	82%	55	76%
Total	141	83%	191	91%	332	88%

Analysis of Evidence Templates

In one or more tasklets at each grade level, there is a test item that is flagged as requiring further documentation of the student response in the form of an evidence template and Evidence Template Recording Sheet. A sample of these documents is provided in Appendix E. The Evidence Template Recording Form requires the test administrator to document the student's response to each attempt to elicit a correct response to an item, following the prescribed scaffolding process. If test administration procedures are followed correctly, there should be a direct correspondence between the information recorded on the Evidence Template Recording Form and the score given to the student on the item.

Evidence Templates from the sample of students who were independently observed for the fidelity and level of agreement analysis were used as another source of data about the accuracy of scoring by test administrators. Templates for test items that were implemented when independent observers were present were identified by Measured Progress, duplicated, and provided to an independent person to score. The reviewer had access only to the Templates, and was asked to provide, for each, the score that the template data indicate should have been given to the student for that item. These data were sent to Measured Progress where they were compared with the score given to this item by the test administrator.

Data for this analysis encompasses an examination of 64 items in Reading and 55 items in Math, for a total of 119 items. There is variability in the number of items reviewed per grade, since they are embedded at different points in the testing process and observations captured varying numbers of these "evidence" items. Results of this analysis are provided in Table 9. As seen in this table, the level of agreement based on an aggregation of all responses across content areas is 92%, indicating a consistent correspondence between the documented sequence of response and the final score given to a student for an individual item.

Table 9. Analysis of Evidence Templates

Grade Level	Reading		Math		Combined Subjects	
	# Items	% Exact Agreement	# Items	% Exact Agreement	# Items	% Exact Agreement
3	14	100	4	75	18	94.44
4	15	100	20	90	35	94.29
5	7	71.43	2	100	9	77.78
6	5	100	3	100	8	100
7	9	100	4	75	13	92.31
8	7	85.71	9	100	16	93.75
10	7	71.43	13	92.31	20	95
Total/ Mean	64	92.19%	55	90.91%	119	91.60%

Feedback from Technical Advisory Committee

Feedback about this study was solicited from Montana's Technical Advisory Committee (TAC) at two points in time. In January of 2007, the plan was presented to the TAC for their suggestions and input. They concurred that the approach of gathering as much information as possible across the different steps of the test training and implementation process was appropriate given the limitations of the size of the student population and available resources. This approach created the opportunity to evaluate multiple sources of evidence collected at these various steps in the process.

The initial results of the study were shared with the TAC in July, 2007. The feedback received at that time was that the process implemented was sound, representing more than a study of the CRT-Alt's inter-rater reliability. The picture that emerges from putting together all of the information gathered during this study is that the process and procedures used for Montana's CRT-Alt appear sound. Comments suggested that the level of scripting provided for the item implementation and scaffolding was very good, likely contributing to the positive results in relation to both implementation fidelity and scoring reliability of the CRT-Alt.

Summary and Conclusions

This study examined the entire process involved in the implementation of the CRT-Alt by test administrators in Montana. From the point at which materials are received and reviewed by the test administrators through the actual implementation and scoring of the test, data were gathered to evaluate current procedures and associated outcomes. Concluding remarks, including recommendations for future evaluation, are provided relative to each area examined in this study.

- The test design incorporates evidence-based implementation approaches that are appropriate for the group of students who are eligible for an alternate assessment under NCLB guidelines. The format achieves a good and necessary balance between the flexibility needed to address the individual needs of students and the structured, scripted method used to guide the test administrator through the item presentation, scaffolding, and scoring processes.

- The current format of the training, available on a CD that can used by a test administrator at his/her convenience, appears to be a viable method of getting the basic information about test administration out to the people who need it. While the static nature of this form of training is not ideal, test administrator ratings indicate that it is seen an efficient way of imparting necessary information. Since the data indicate that only a small proportion of test administrators receive training in any other form, additional opportunities for training that is more interactive merits consideration as a supplement to the Training CD approach, demonstrated to be effective in reaching test administrators.

- There are some mechanical issues about the way in which the training and teacher survey data are collected that need to be examined for future administrations. Given the number of test administrators that give the test to multiple students, it would be beneficial to identify a way to collect survey data so that these test administrators see and/or respond to the questions only once. This would help to reduce the loss of information when a sizeable proportion of questions are left blank.

- Self-check tools such as the Implementation Checklist appear to be beneficial. They do not have much of an associated "cost" in terms of time or materials, and provide a comprehensive list of the entire process in a single place. Continuation of this practice is recommended.

- The results of the direct observation of a sample of test administrators were very positive. They suggest that the supports built into the current test administration protocols are sufficient to yield consistent implementation practices and scoring. As resources are available, repeating this approach in other parts of the state or with larger samples may be warranted. The next issue to consider is the generalization and maintenance of this level of fidelity across time, as Science assessments are introduced in the next testing cycle. Given the utility of the observation methodology used this year, it is worth considering the use of this methodology to conduct "spot checks" to evaluate maintenance of implementation fidelity and scoring reliability in future years.

- The evaluation of Evidence Templates provides another opportunity for period "spot checks" in a manner that is not too costly in terms of additional time and resources. Conducting this type of analysis on a random sample of students across time is suggested, given the fact that the data are readily available.

In conclusion, the preponderance of evidence gathered in this study confirms the integrity of the CRT-Alt procedures currently in use in Montana. An appropriate "next step" is to determine how to fine tune the collection of the range of data considered in this study to address the identified data collection limitations, and to develop an implementation plan that allows for periodic maintenance probes to verify that these results continue over time.

References

Alberto, P., & Troutman, A. C. (1995). *Applied behavior analysis for teachers* (4th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Colyer, S., & Collins, B. (1996). Using natural cues within prompt levels to teach the next dollar strategy to students with disabilities. *The Journal of Special Education, 30*(3), 305-318.

Demchak, M. (1990). Response prompting and fading methods: A review. *American Journal on Mental Retardation, 94,* 603-615.

Doyle, P. M., Wolery, M., Ault, M. J., & Gast, D. L. (1988). System of least prompts: A literature review of procedural parameters. *Journal of the Association for the Severely Handicapped, 13*(1), 28-40.

Falvey, M. (1986). *Community-based curriculum. Instructional strategies for students with severe handicaps.* Baltimore: Paul H. Brookes Publishing.

McDonnell, J. (1987). The effects of time delay and increasing prompt hierarchy strategies on the acquisition of purchasing skills by students with severe handicaps. *Journal of the Association for the Severely Handicapped, 12*(3), 227-236.

Ryndak, D. L., Clark, D., Conroy, M., & Stuart, C. H. (2001). Preparing teachers to meet the needs of students with severe disabilities: Program configuration and expertise. *Journal of the Association for Persons with Severe Handicaps, 26*(2), 96-105.

Snell, M. E., & Brown, F. (2006). *Instruction of students with severe disabilities* (5th ed.). Upper Saddle River, NJ: Merrill.

Tawney, J. W., & Gast, D. L. (1984). *Single subject research in special education.* Columbus, OH: Charles E. Merrill Publishing Company.

West, E.A., & Billingsley, F. (2005). Improving the system of least prompts: A comparison of procedural variations. *Education and Training in Developmental Disabilities, 40*(2), 131-144.

Wolery, M., Ault, M. J., & Doyle, P. M. (1992). *Teaching students with moderate to severe disabilities.* New York: Longman.

Appendix A
Implementation Checklist for CRT-Alt

Implementation Checklist for CRT-Alt
Spring, 2007

Please review this checklist before you start to administer the CRT-Alt as a final reminder of all components of the test preparation and implementation protocol.

Preparation Activities

I have viewed the Training CD or attended training about the administration of this test.

I have reviewed the student test booklet and testing materials.

If needed for this student, I have modified the testing materials.

If needed for this student, communication supports have been prepared.

Materials not provided for the test have been located, are organized, and available for this test administration.

If needed, I have found a second person to assist with the administration of this test.

I have scheduled test administration for periods of time that match the student's attention span and endurance, breaking it up into multiple sessions as needed.

Test administration will occur in a location in which the student can work without interruptions.

Implementation Practices

Introductory items were implemented without scaffolding, scored as either a "4" or a "0".

For each item, the student was given with an opportunity to respond independently before any scaffolding was provided.

Scores for each item were given based on the level of scaffolding that was necessary in order for the student to make a correct response.

Student responses that required complete teacher assistance were given a score of "1".

If a student actively resisted responding to a test item, this item was given a score of "0".

If a student received a score of "0" for 3 consecutive test items, the halting rule for the designated test grade level was used.

I have completed the Teacher Recording Sheet and Evidence Template for each item requiring evidence (i.e., those marked with a magnifying glass).

I have completed all tasks/tasklets for this student in both Reading and Math OR I have followed the appropriate halting rule in response to active student resistance to participation.

Submission of Student Information

The student's name has been written on the Student Response Booklet, the CRT-Alternate Test Booklet, and all Evidence Templates and Teacher Recording Sheets.

I have placed the student bar code label in the space provided on page 1 of the Student Response Booklet. If no label is available, I entered in the 9 digit student ID instead, entering a zero followed by the 9 digit number in the 10 spaces provided on this form.

I have entered the appropriate information on page 2, Section 1 of the Student Response Booklet, including the last bubble, indicating the student participated in the CRT-Alternate.

I have filled in all required information on four pages of the Student Response Booklet.

I have transferred student scores from the Test Booklets to the appropriate sections of the bubble forms in the Student Response Booklet.

I have responded to the questions about test administration in the area marked Test Administration Activity Information.

I have completed a Material Replacement Form to replenish materials that cannot be used again in the Test Materials Kit used for this administration, returning it with my student's test materials.

I have returned the Test Materials Kit to the System Test Coordinator for secure storage.

I have placed all required materials (CRT-Alt Test Booklet, Evidence Templates, Teacher Recording Sheets, Student Response Booklet, Class Identification Sheet, Material Replacement Order Form and Teacher Questionnaire (grade 3 only), in the white plastic envelope labeled For Return of CRT-Alternate Student Materials.

Materials Kit used for this administration, returning it with my student's test materials.

I have returned the Test Materials Kit to the System Test Coordinator for secure storage.

Appendix B
CRT-Alt Interrater Reliability Study
Sampling Plan

CRT-Alt Interrater Reliability Study
Sampling Plan

During a December, 2006 phone consultation with Dr. Stanley Rabinowitz, the issue of sampling size and composition for the interrater reliability study was discussed. His recommendation was that we begin with a sample of no less than 5 students per grade, with observations focused on both math and reading testing. If initial findings with this sample size showed mixed results in terms of scoring reliability and implementation fidelity, additional observations would be required until more definitive findings were obtained.

Based on this recommendation, an initial sample achieving the minimum distribution would look like this. This does not allow for scheduling difficulties, absences, etc. and does not take into account any changes needed based on the actual distribution of students in each location.

	Grade Level						
Location	**3**	**4**	**5**	**6**	**7**	**8**	**10**
Helena	1	1	1	1	1	1	1
Bozeman	1	1	1	1	1	1	1
Great Falls	1	1	1	1	1	1	1
Billings	1	1	1	1	1	1	1
Missoula	1	1	1	1	1	1	1
Total	5	5	5	5	5	5	5

Adjustments based on actual distribution of students, based on 1/2/07 Excel Spreadsheet

	Grade Level						
Location	**3**	**4**	**5**	**6**	**7**	**8**	**10**
Helena	2	1	1	1	0	1	1
Bozeman	0	1	1	1	1	1	1
Great Falls	1	1	1	1	1	1	1
Billings	1	1	1	1	2	1	1
Missoula	1	1	1	1	1	1	1
Total	5	5	5	5	5	5	5

Appendix C
Independent Observer Protocol

Student ID Number [to be filled in later]: _____ Page __ of __

CRT-Alt Observer Checklist

Please ask teachers the questions below before test administration begins. Questions in ***italics*** should be answered based on your observation.

Test Preparation Activities	**Item Rating**	
1. Have you viewed the Training CD or attended training about the administration of this test? ……	Yes	No
2. Have you reviewed the student test booklet and testing materials? …..	Yes	No
3. Have you modified the testing materials for this student? ..... If yes, please describe what you have to customize the materials for this student.	Yes	No
4. Does this student need any type of communication support in order to be able to respond to test item? If yes, please describe what supports you have available for the student to use.	Yes	No
5. Have you located all materials for test administration that are not provided?	Yes	No
Do materials appear to be organized and easily accessible for test administration?	Yes	No

Teacher Name: _____ School: _____

Observer: _____ Date: _____

Student Name: _____ Grade: _____

6. Is a second person present to assist with the administration of this test?	Yes	No
If no, please indicate whether it appears that a second person would have been helpful in administering the test.	Yes	No
7. ***Does the scheduled test administration period seem to be matched with the student's attention span and endurance?***	Yes	No
8. ***Is the test being administered in a location in which the student can work without interruption?***	Yes	No

Test Administration Activities

Circle information to describe the test activity that you observed.

Grade	Subject	Tasklet # (Grade 3, 5, 6, 7 only)
3 4 5 6 7 8 10	**Reading Math**	**1 2 3 4 5**

Implementation Practices	**Item Rating**	
1. Introductory items were implemented without scaffolding, scored as either a "4" or a "0".	Yes	No
Observe a sequence of 4 test items, following along in the Test Booklet to see the instructions provided.	**1** y n	**2** y n
2. Did the teacher present the materials as described in the Test Booklet?	**3** y n	**4** y n
3. For each item, was the student given an opportunity to respond independently before any scaffolding was provided?	**1** y n	**2** y n
	3 y n	**4** y n
4. Did the teacher implement the scaffolding as described in the Test Booklet?	**1** y n	**2** y n
	3 y n	**4** y n
5. Did the teacher score the student's response based on the level of scaffolding that was necessary in order for the student to make a correct response?	**1** y n	**2** y n
	3 y n	**4** y n

6. Did the student actively resist responding to a test item during this period of observation? ……..……..	Yes	No
If yes, was this item given a score of "0"? …….……..	Yes	No
Did active resistance continue for 3 consecutive test items? ……...	Yes	No
If so, was the halting rule correctly implemented? ……..	Yes	No
7. Did test administration include an item requiring evidence? ……..	Yes	No
If so, was the Teacher Recording Sheet and Evidence Template completed for this item? …….…	Yes	No

Appendix D
Independent Observer Score Recording Form

Student ID Number [to be filled in later]: _____ Page __ of __

Circle responses to describe activity you are observing:

Grade	**Subject**	**Tasklet #** (Grade 3, 5, 6, 7 only)
3 4 5 6 7 8 10	Reading Math	1 2 3 4 5

Test Item: [change test item numbers if necessary for grades 4, 8, 10]	**Observer Score**
Item 1:	4 3 2 1 0
Item 2:	4 3 2 1 0
Item 3:	4 3 2 1 0
Item 4:	4 3 2 1 0
Item 5:	4 3 2 1 0

Circle responses to describe activity you are observing:

Grade	**Subject**	**Tasklet #** (Grade 3, 5, 6, 7 only)
3 4 5 6 7 8 10	Reading Math	1 2 3 4 5

Test Item: [change test item numbers if necessary for grades 4, 8, 10]	**Observer Score**
Item 1:	4 3 2 1 0
Item 2:	4 3 2 1 0
Item 3:	4 3 2 1 0
Item 4:	4 3 2 1 0
Item 5:	4 3 2 1 0

Teacher Name: _____ School: _____

Observer: _____ Date: _____

Student Name: _____ Grade: _____

Appendix E
Evidence Template Example

Number Sentence Evidence Template

Item 5

7 + □ = 10

Place student barcode label here.

EVIDENCE TEMPLATE TEACHER RECORDING SHEET

Math Tasklet 3		Item 5
Describe how the student communicated their response.	• Used words to respond …… • Used communication device/display ... .. • Pointed to/manipulated task materials … • Used auditory scanning …… • Used gestures/sign language …….. • Other form of communication ……	❒ …….. ❒ …….. ❒ …….. ❒ …….. ❒ …….. ❒ _____ _____ _____
Describe student's initial response to the task before scaffolding.	• Correct response…… • No response …… • Incorrect response……..	❒ …….. ❒ …….. ❒ ……..
If applicable, describe the student's response after level 3 scaffolding.	• Correct response…… • No response…… • Incorrect response……	❒ …….. ❒ …….. ❒ ……..
If applicable, describe the student's response after level 2 scaffolding.	• Correct response…… • No response …… • Incorrect response……	❒ …….. ❒ …….. ❒ ……..
If applicable, describe the student's response after level 1 scaffolding.	• Correct response …… • No response…… • Incorrect response……	❒ …….. ❒ …….. ❒ ……..
If applicable, check the box and describe the student's behavior if the student was not responsive to the task.		❒ _____ _____ _____ _____

Place student barcode label here.

Appendix E—Technical Advisory Committee

Table E-1. 2008–09 MT CRT-Alternate: 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
Derek	Briggs, Ph.D.	Assistant Professor	School of Education	University of Colorado
Susan	Brookhart, Ph.D.	President		Brookhart Enterprises, LLC
Ellen	Forte, Ph.D.	President		edCount, LLC
Michael	Kozlow, Ph.D.	Program Director	Assessment Program	
Scott	Marion, Ph.D.	Vice-President		Center for Assessment
Stanley	Rabinowitz, Ph.D.	Program Director	Assessment & Standards Development Services	WestEd

Appendix F—Released Performance Indicators

Reading - Grade 3		
Item	**Performance Indicator**	**Standard**
1	Attend to a person demonstrating with concrete materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
2	Demonstrate an understanding that numbers, as opposed to letters, are used to express quantity, order, or size/amount.	Standard 2: Students apply a range of skills and strategies to read.
3	Count with another person.	Standard 2: Students apply a range of skills and strategies to read.
4	Show a quantity.	Standard 2: Students apply a range of skills and strategies to read.
5	Enter numbers correctly on a calculator/ write numbers correctly.	Standard 2: Students apply a range of skills and strategies to read.
6	Attend to another person combining and subdividing shapes.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
7	Touch and move shapes toward creating new shapes.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
8	Recognize properties of 2-dimensional shapes.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
9	Find various shapes in the environment.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
10	Produce 2-dimensional shapes. Carry out a strategy to solve a geometric problem.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
11	Attend to objects or pictures of two- and three-dimensional geometric shapes and the relationships among them.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
12	Identify (name) shapes as circles, squares, triangles, rectangles, and ovals.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
13	Sort 2-dimensional physical shapes according to their shape.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.

Reading - Grade 3		
Item	**Performance Indicator**	**Standard**
14	Recognize 2-dimensional physical shapes as being the same (congruent) or different.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
15	Match 2-dimensional physical shapes to pictures of the shapes in different orientations. Explain/show spatial reasoning.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
16	Attend to another person estimating an amount in a given set.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
17	Use a quantitative label when making a guess.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
18	Identify a reasonable quantity when guessing the amount in a given set.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
19	Use methods and tools to solve a problem, including drawing pictures, modeling with objects, estimating, using paper and pencil, and using a calculator.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
20	Determine which of two numbers is closer to the quantity in a given set.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
21	Attend to another person making patterns and to a person describing patterns.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
22	Extend or supply a missing element in a repeating pattern by attribute or number.	Standard 2: Students apply a range of skills and strategies to read.
23	Extend and explain an alternating pattern of two or more objects, shapes, designs, or numbers.	Standard 2: Students apply a range of skills and strategies to read.
24	Reproduce an alternating pattern of two or more objects, shapes, designs, or numbers.	Standard 2: Students apply a range of skills and strategies to read.
25	Create a repeating pattern using objects, shapes, designs, or numbers. Carry out a strategy to solve problems involving patterns, relations, or functions.	Standard 2: Students apply a range of skills and strategies to read.

Mathematics Grade 3		
Item	**Performance Indicator**	**Standard**
1	Attend to a person demonstrating with concrete materials.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
2	Demonstrate an understanding that numbers, as opposed to letters, are used to express quantity, order, or size/amount.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
3	Count with another person.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
4	Show a quantity.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
5	Enter numbers correctly on a calculator/ write numbers correctly.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
6	Attend to another person combining and subdividing shapes.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
7	Touch and move shapes toward creating new shapes.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
8	Recognize properties of 2-dimensional shapes.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
9	Find various shapes in the environment.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
10	Produce 2-dimensional shapes. Carry out a strategy to solve a geometric problem.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 4: Students demonstrate understanding of shape and ability to use geometry.
11	Attend to objects/pictures of two- and three-dimensional geometric shapes and their relationships;	Standard 4: Students demonstrate understanding of shape and ability to use geometry.

Mathematics Grade 3		
Item	**Performance Indicator**	**Standard**
12	Identify (name) shapes as circles, squares, triangles, rectangles, and ovals.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
13	Sort 2-dimensional physical shapes according to their shape.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
14	Recognize 2--dimensional physical shapes as being the same (congruent) or different.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
15	Match 2-dimensional physical shapes to pictures of the shapes in different orientations. Explain/show spatial reasoning.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 4: Students demonstrate understanding of shape and ability to use geometry.
16	Attend to another person estimating an amount in a given set.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
17	Use a quantitative label when making a guess.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
18	Identify a reasonable quantity when guessing the amount in a given set.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.

Mathematics Grade 3		
Item	**Performance Indicator**	**Standard**
19	Use methods and tools to solve a problem, including drawing pictures, modeling with objects, estimating, using paper and pencil, and using a calculator.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
20	Determine which of two numbers is closer to the quantity in a given set.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
21	Attend to another person making patterns and to a person describing patterns.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
22	Extend or supply a missing element in a repeating pattern by attribute or number.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
23	Extend and explain an alternating pattern of two or more objects, shapes, designs, or numbers.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
24	Reproduce an alternating pattern of two or more objects, shapes, designs, or numbers.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
25	Create a repeating pattern using objects, shapes, designs, or numbers. Carry out a strategy to solve problems involving patterns, relations, or functions.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.

Reading - Grade 4		
Item	**Performance Indicator**	**Standard**
1	Attends to people and objects in the environment.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
2	Locates a picture/symbol/object when named or signed.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
3	Selects literacy materials/books by character or topic.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
4	Uses word recognition skills and context clues to comprehend text.	Standard 2: Students apply a range of skills and strategies to read.
5	Identifies words/pictures/symbols/objects that are new and unfamiliar.	Standard 2: Students apply a range of skills and strategies to read.
6	Anticipates the beginning of literacy activity (looks toward reader, tolerates headphones, locates literacy materials).	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
7	Located a picture/symbol/object when named or signed.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
8	Provides details about perspective.	Standard 4: Student select, read, and respond to print and nonprint material for a variety of purposes.
9	Identifies events or steps from a functional text.	Standard 4: Student select, read, and respond to print and nonprint material for a variety of purposes.
10	Uses a timeline to provide information about an event.	Standard 4: Student select, read, and respond to print and nonprint material for a variety of purposes.
11	Attends to literacy materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
12	Identifies components related to the beginning of a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
13	Analyzes supporting details in order to draw conclusions from a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
14	Identifies the main character in a story.	Standard 2: Students apply a range of skills and strategies to read.
15	Answers "what" questions about objects in story.	Standard 2: Students apply a range of skills and strategies to read.

Reading - Grade 4		
Item	**Performance Indicator**	**Standard**
16	Attends to literacy materials from beginning to end.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
17	Identifies a preferred resource to gain information.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
18	Identifies supporting details from a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
19	Identifies words/pictures/symbols/objects that are new and unfamiliar.	Standard 2: Students apply a range of skills and strategies to read.
20	Demonstrates understanding of a new word based on context of a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
21	Attends to literacy materials from beginning to end.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
22	Answers “who” questions about characters in stories.	Standard 2: Students apply a range of skills and strategies to read.
23	Responds to yes/no questions about information in print and nonprint materials.	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.
24	Identifies supporting details from a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
25	Identifies components related to the end of a story.	Standard 2: Students apply a range of skills and strategies to read.

Mathematics - Grade 4		
Item	**Performance Indicator**	**Standard**
1	Attends to another person reviewing counters; anticipates the beginning f the math activity; and attends to materials being displayed.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
2	Demonstrates the concept of one.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
3	Applies a number (word) to a quantity of objects in a collection.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
4	Determines which of two numbers is closer to the quantity in a given set.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
5	Computes addition and subtraction problems with single digits.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
6	Attends to another person counting; anticipated the beginning of the math activity; and attends to materials being displayed.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
7	Counts using a sequential order of numbers (e.g., 1, 2, 3, 4; rote counting).	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
8	Demonstrates one-to-one correspondence among up to 12 objects and counting numbers with no recounting (rational counting).	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
9	Demonstrates an understanding that the final number said when counting objects is the quantity of the set.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
10	Computes addition and subtraction problems with single digits.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.

Mathematics - Grade 4		
Item	**Performance Indicator**	**Standard**
11	Attends to another person reviewing table; anticipates the beginning of the math activity; and attends to materials being displayed.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
12	Given a class of objects, sorts into categories and subcategories.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
13	Sets up graph (table), (i.e., labels axes); provides title.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
14	Uses symbols to represent data; creates a simple graph, frequency plat, or frequency table using real objects and/or symbols.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
15	Explains/shows how decisions were made, using a table or graph.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
16	Attends to another person reviewing a graph; anticipates the beginning of the match activity; and attends to materials being displayed.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
17	Determines which category has the most/least.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
18	Compares categories of data using comparison words.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
19	Communicates the relationships between categories of collected data.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
20	Predicts the outcome of a chance event using a chance device.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
21	Attends to another person reviewing two difference sets of counters; anticipates the beginning of the math activity; and attends to materials being displayed.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.

Mathematics - Grade 4		
Item	**Performance Indicator**	**Standard**
22	Groups/sorts objects into two sets.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
23	Reproduces (matches) a repeated event.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
24	Creates a growing pattern or attribute or number.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.

Science - Grade 4		
Item	**Performance Indicator**	**Standard**
1	Attend to common substances or objects.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
2	Recognize a mixture.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
3	Recognize a mixture.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
4	Identify the different components of a mixture.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
5	Identify how a given mixture can be separated.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
6	Attends to pictures being shown.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
7	Recognize animals.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate thinking skills associated with this knowledge.

Science - Grade 4		
Item	**Performance Indicator**	**Standard**
8	Recognize plants.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
9	Recognize arms, legs, heads, bodies, antennae, eyes, nose, mouths and tails of animals.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
10	Recognize which is living when given a choice between something that is living and something that is nonliving. Identify which components in a group are living and which are nonliving.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
11	Sort plants and animals according to their similarities and differences.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
12	Attend to the weather.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
13	Recognize that rain is liquid water.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.

Science - Grade 4		
Item	**Performance Indicator**	**Standard**
14	Recognize that rain is liquid water.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
15	Identify parts of the water cycle. Recognize that lakes and rivers have water in them.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
16	Recognize that winter is usually the colder time of year.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
17	Attend to the seasons.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
18	Recognize that fall is the time that the weather begins to become colder.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
19	Recognize that summer is usually the hottest time of the year.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
20	Recognize that winter is usually the colder time of year.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.

Science - Grade 4		
Item	**Performance Indicator**	**Standard**
21	Identify a question that would increase knowledge about the world.	Standard 6: Students understand historical developments in science and technology.
22	Attend to tools being shown.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
23	Compare the common physical properties.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical and chemical systems, and demonstrate the thinking skills associated with this knowledge.
24	Identify tools needed to solve a problem.	Standard 5: Students understand how scientific knowledge and technological developments impact today's societies and cultures.
25	Attend to common tools to measure length.	Standard 1: Students design, conduct, evaluate, and communicate processes and results of scientific investigations, and demonstrate the thinking skills associated with this procedural knowledge.
26	Recognize technology as tools and techniques to solve problems.	Standard 5: Students understand how scientific knowledge and technological developments impact today's societies and cultures.

Reading - Grade 5		
Item	**Performance Indicator**	**Standard**
1	Attend to literacy materials from beginning to end.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
2	Use a resource to solve a problem or gain needed information.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
3	Use a resource to solve a problem or gain needed information.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
4	Accurately order steps from a functional text.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
5	Demonstrate understanding of the difference between an information resource and literature.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
6	Attend to person and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
7	Make an appropriate prediction.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
8	Compare and contrast the impact of setting.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
9	Identify environmental print in context.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
10	Follow directions that contain a preposition.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
11	Attend to person and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
12	Recall the name of a common object when given the function of the object.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
13	Select important details from reading materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
14	Identify a resource to gain information.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
15	Identify the main message of an expository reading selection.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.

Reading - Grade 5		
Item	**Performance Indicator**	**Standard**
16	Attend to a literacy activity in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
17	Identify components related to the beginning of a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
18	Answer “where” questions about the story.	Standard 2: Students apply a range of skills and strategies to read.
19	Sequence events in simple stories.	Standard 2: Students apply a range of skills and strategies to read.
20	Draw conclusions based on facts in the story.	Standard 2: Students apply a range of skills and strategies to read.
21	Attend to person and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
22	Match pictures to printed words.	Standard 2: Students apply a range of skills and strategies to read.
23	Recognize consonant sounds.	Standard 2: Students apply a range of skills and strategies to read.
24	Use simple letter-sound association to decode unfamiliar words.	Standard 2: Students apply a range of skills and strategies to read.
25	Identify syllables.	Standard 2: Students apply a range of skills and strategies to read.

Mathematics - Grade 5		
Item	**Performance Indicator**	**Standard**
1	Attend to teacher placing numbers in order from least/smallest to greatest/largest.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
2	Position numbers on a number line.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
3	Identify first and last.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
4	Indicate ordinal position.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
5	Arrange a set of objects, up to ten, from least to most. Carry out a strategy to solve a number problem.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
6	Attend to another person combining objects to add.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
7	Demonstrate an understanding of the concepts of some/more/ less/take away/all gone/ no more.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
8	Connect plus and minus symbols to operations.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
9	Demonstrate an understanding that adding 0 to any number equals the same number. Carry out a strategy to solve a number problem.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
10	Model a written addition problem using sets of objects, combining the sets, and counting the objects, either counting all or counting on.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
11	Attend to another person showing the relationship between two variables using objects, pictures, symbols, or numbers.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.

Mathematics - Grade 5		
Item	**Performance Indicator**	**Standard**
12	Recognize a cause-effect relationship between two elements.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
13	Choose correct strategies or procedures to solve an algebraic problem in algebra.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
14	Demonstrate/ communicate what the relationship is between two elements.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
15	Use methods and tools to solve a measurement problem, including drawing pictures, modeling with objects, estimating, using paper and pencil, and using a calculator.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
16	Attend to another person reading temperature.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
17	Select the appropriate tool to be used in making a measure.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
18	Read temperatures from a thermometer to the accuracy of the labeled numbers.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.

Mathematics - Grade 5		
Item	**Performance Indicator**	**Standard**
19	Carry out a strategy to solve a measurement problem.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
20	Attend to real world problems that require measurement.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
21	Attend to another person measuring capacity.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
22	Select the appropriate tool to be used in making a measure.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
23	Use methods and tools to solve a measurement problem, including drawing pictures, modeling with objects, estimating, using paper and pencil, and using a calculator.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
24	Use nonstandard tools and units to determine the capacity of a container.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
25	Use standard tools and standard units of capacity to measure the capacity of a container.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.

Reading - Grade 6		
Item	**Performance Indicator**	**Standard**
1	Attend to people and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
2	Display knowledge of front and back, right-side up, page turning, and scanning when exploring literacy materials.	Standard 2: Students apply a range of skills and strategies to read.
3	Use listening/observing strategies to comprehend a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
4	Based on the context of a reading selection, identify appropriate definition of multiple-meaning words.	Standard 2: Students apply a range of skills and strategies to read.
5	Use word recognition skills and context clues to comprehend text.	Standard 2: Students apply a range of skills and strategies to read.
6	Attend to person and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
7	Identify the main idea in a selection.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
8	Identify details related to the main idea.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
9	Select important details/facts from reading materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
10	Creates an illustration/photo essay/ object box/ specific to the text.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
11	Attend to person and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
12	Identify the main message of an expository reading selection.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
13	Retell key events in sequence.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
14	Identify common object/symbol when given the function of the object or symbol.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.

Reading - Grade 6		
Item	**Performance Indicator**	**Standard**
15	Select important details/facts from reading materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
16	Attend to person and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
17	Answer “who” questions about characters in stories.	Standard 2: Students apply a range of skills and strategies to read.
18	Answer “what” questions about objects in stories.	Standard 2: Students apply a range of skills and strategies to read.
19	Answer “why” questions about issues in a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
20	Identify cultural elements in a reading selection.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
21	Attends to literacy materials from beginning to end.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
22	Identify details of characters that are the same.	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.
23	Compare/contrast information in reading materials.	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.
24	On an organizer, make a graphic representation of similarities and differences from a topic in the text.	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.
25	Make connections between reading materials and personal experiences.	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.

Mathematics Grade 6		
Item	**Performance Indicator**	**Standard**
1	Attend as another person demonstrates an understanding that written numerals represent number (quantities).	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
2	Match a numeral to a quantity of a set of objects.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
3	Produce a numeral to 10.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
4	Use methods and tools to solve a number problem, including modeling with objects.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
5	Carry out a strategy to solve a number problem.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
6	Attend to another person removing objects or comparing sets to subtract.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
7	Employ strategies to recall simple subtraction facts for single-digit differences from 10 (e.g., counting back; comparison/addition— add to the smaller number to get the larger one).	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
8	Demonstrate understanding that subtracting 0 from any number equals the number.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
9	Use a calculator for whole-number computation.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
10	Use methods and tools to solve a number problem, including drawing pictures, modeling with objects, estimating, using paper and pencil, and using a calculator.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.

Mathematics Grade 6		
Item	**Performance Indicator**	**Standard**
11	Attend to another person demonstrating congruence.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
12	Recall shapes and their relative positions after they have been viewed for only a brief period of time.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
13	Demonstrate transformations of shapes, e.g., sliding.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
14	Cover a figure with shapes.	Standard 4: Students demonstrate understanding of shape and ability to use geometry.
15	Use methods and tools to solve a geometric problem, including modeling with objects.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 4: Students demonstrate understanding of shape and ability to use geometry.
16	Attend to another person telling time.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
17	Tell time to the hour using an analog clock.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
18	Use methods and tools to solve a measurement problem.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
19	Read time using a digital clock.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
20	Read time using a digital clock (e.g., “It is two twenty-five.”).	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.

Mathematics Grade 6		
Item	**Performance Indicator**	**Standard**
21	Attend to another person modeling mathematical relationships (e.g., modeling different numbers).	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
22	Model sets that contain nothing or one or more items (some, none).	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
23	Demonstrate that objects defined by a shared attribute form a set to which a number can be applied. (For example, make a set of red triangles. How many are there?)	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
24	Model sets of the same/different amounts and compare them.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
25	Use methods and tools to solve a problem involving patterns, relations, or functions, including modeling with objects.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.

Reading - Grade 7		
Item	**Performance Indicator**	**Standard**
1	Attend to people and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
2	Based on the context of a reading selection, identify appropriate definition of multiple-meaning words.	Standard 2: Students apply a range of skills and strategies to read.
3	Identify antonyms.	Standard 2: Students apply a range of skills and strategies to read.
4	Explain the meaning of vocabulary words in the context of a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
5	Identify cultural elements in a reading selection.	Standard 4: Students select, read, and respond to print and nonprint materials for a variety of purposes.
6	Attend to people and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
7	Identify the main message of an expository reading selection.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
8	Retell key events in sequence.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
9	Identify common object/symbol when given the function of the object or symbol.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
10	Select important details/facts from reading materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
11	Attend to literacy materials from beginning to end.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
12	Identify details related to the main idea.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
13	Identify the main idea of a reading selection.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
14	Identify details related to the main idea.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.

Reading - Grade 7		
Item	**Performance Indicator**	**Standard**
15	Identify common object/symbol when given the function of the object or symbol.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
16	Attend to people and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
17	Locate title.	Standard 2: Students apply a range of skills and strategies to read.
18	Use chapter headings to locate information.	Standard 2: Students apply a range of skills and strategies to read.
19	Use text features to move through text in appropriate sequence.	Standard 2: Students apply a range of skills and strategies to read.
20	Answer questions about the main idea of the text.	Standard 2: Students apply a range of skills and strategies to read.
21	Attend to people and literacy materials in a purposeful manner.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
22	Attend to people and literacy materials in a purposeful manner.	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.
23	Defend an author's point of view.	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.
24	Identify facts in text.	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.
25	Identify non-truths within a text.	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.

Mathematics Grade 7		
Item	**Performance Indicator**	**Standard**
1	Attend as another person demonstrates an understanding of the concept of some and none.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
2	Associate the number 0 with empty sets in different settings.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
3	Use a quantitative label when making a guess (e.g., a few, many, and seventeen).	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
4	Determine which of two numbers is closer to the quantity in a given set.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
5	Identify a reasonable quantity when guessing the amount in a given set.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
6	p coins by attributes (metal color, size, weight, texture).	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
7	Match coins to like coins and bills to like bills.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
8	Match coins and their values.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
9	Count out an exact amount of money.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.

Mathematics Grade 7		
Item	**Performance Indicator**	**Standard**
10	Round numbers to the nearest 10 (e.g., 27 rounds to 30) or nearest 100.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
11	Attend to another person setting up a number sentence with a box as a placeholder.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
12	Recognize that a box is used as a placeholder in a number sentence.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
13	Find a simple missing addend represented by a box in a number sentence.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
14	Choose correct strategies or procedures to solve an algebraic problem in algebra.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
15	Use methods and tools to solve a problem, including drawing pictures, modeling with objects, estimating, using paper and pencil, and using a calculator.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.

Mathematics Grade 7		
Item	**Performance Indicator**	**Standard**
16	Attend to another person showing relationships between two variables using objects.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
17	Recognize a cause-effect relationship between two elements.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
18	Choose correct strategies or procedures to solve an algebraic problem.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
19	Use methods and tools to solve a problem, including modeling with objects.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
20	Demonstrate/ communicate what the relationship is between two elements.	Standard 3: Students use algebraic concepts, process, and language to model and solve a variety of real-world and mathematical problems.
21	Attend to another person collecting data.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
22	Given a class of objects, sort into categories.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
23	Display data using concrete objects.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.

Mathematics Grade 7		
Item	**Performance Indicator**	**Standard**
24	Determine which category has the most/ least.	Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.
25	Make decisions based on data, a table, or a graph.	Standard 1: Students engage in the mathematical process of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology. Standard 6: The students demonstrate understanding of an ability to use data analysis, probability, and statistics.

Reading - Grade 8		
Item	**Performance Indicator**	**Standard**
1	Anticipates the beginning of a literacy activity (looks toward reader, tolerates headphones, locates literacy materials).	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
2	Identifies resource materials to gain information about words.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
3	Identify fiction.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
4	Selects important details/facts from reading materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
5	Identifies words/pictures/symbols/objects used for content communication.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
6	Anticipates the beginning of a literacy activity (looks toward reader, tolerates headphones, locates literacy materials).	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
7	Locates title, chapter, glossary, etc.	Standard 2: Students apply a range of skills and strategies to read.
8	Recognizes vowel letter-sound association.	Standard 2: Students apply a range of skills and strategies to read.
9	Recognizes word order in simple sentences.	Standard 2: Students apply a range of skills and strategies to read.
10	Recognizes familiar printed words.	Standard 2: Students apply a range of skills and strategies to read.
11	Attends to literacy materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
12	Identifies words/pictures/symbols/objects to name familiar people.	Standard 2: Students apply a range of skills and strategies to read.
13	Identifies words/pictures/symbols/objects to name familiar places.	Standard 2: Students apply a range of skills and strategies to read.
14	Creates an illustration/photo essay/object box specific to the text.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.

Reading - Grade 8		
Item	**Performance Indicator**	**Standard**
15	Selects literacy materials/books by character or topic.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
16	Responds to own name, spoken/signed, print/picture.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
17	Identifies items on a map.	Standard 4: Student select, read, and respond to print and nonprint material for a variety of purposes.
18	Identifies an appropriate information resource to gain specific information.	Standard 4: Student select, read, and respond to print and nonprint material for a variety of purposes.
19	Explains the meaning of new vocabulary words in the context of a story/reading selection/activity.	Standard 2: Students apply a range of skills and strategies to read.
20	Demonstrates understanding of the difference between an information resource and literature.	Standard 4: Student select, read, and respond to print and nonprint material for a variety of purposes.
21	Anticipates routines or patterns connected to a literacy activity.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
22	Uses listening/observing strategies to comprehend a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
23	Explains the meaning of new vocabulary words in the context of a story/reading selection/activity.	Standard 2: Students apply a range of skills and strategies to read.
24	Uses word recognition skills ad context clues to comprehend text.	Standard 2: Students apply a range of skills and strategies to read.
25	Identifies facts in text.	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.

Mathematics - Grade 8		
Item	**Performance Indicator**	Standard
1	Attends to teacher and materials in environment.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
2	Positions numbers on a number line.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
3	Demonstrate an understanding of the concepts of some/more/less/take away/ all gone/no more/less.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
4	Chooses correct strategies or procedures to solve a number problem.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
5	Produces fractional parts of whole unit and vice versa.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
6	Attends to teacher and materials in environment.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
7	Demonstrates/communicates what the relationship is between elements.	Standard 3: Students use algebraic concepts, processes, and language to model and solve a variety of real-world and mathematical problems.
8	Shows a relationship between two variables, using ordered pairs or a table; then makes a table.	Standard 3: Students use algebraic concepts, processes, and language to model and solve a variety of real-world and mathematical problems.
9	Given a numerical relationship between two variables, finds the value of one given the other.	Standard 3: Students use algebraic concepts, processes, and language to model and solve a variety of real-world and mathematical problems.
10	Supplies the missing number represented by a blank number sentence, in which the operation might be +, -, or ×.	Standard 3: Students use algebraic concepts, processes, and language to model and solve a variety of real-world and mathematical problems.

Mathematics - Grade 8		
Item	**Performance Indicator**	Standard
11	Attends to teacher and materials in environment.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
12	Identifies tools associated with measurement.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
13	Uses rulers to measure objects that area whole number of inches or centimeters.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
14	Uses words to compare distances or lengths.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
15	Chooses correct strategies or procedures t solve a measurement problem, measured correctly.	Standard 5: Students demonstrate understanding of measurable attributes and an ability to use measurement processes.
16	Attends to teacher and materials in environment.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
17	Determines which questions to ask to gain information.	Standard 6: The students demonstrate understanding of and an ability to use data analysis, probability, and statistics.
18	Creates a simple graph, frequency plot, or frequency table using real objects and/or symbols.	Standard 6: The students demonstrate understanding of and an ability to use data analysis, probability, and statistics.
19	Sets up a graph (i.e., labels axes, provides title).	Standard 6: The students demonstrate understanding of and an ability to use data analysis, probability, and statistics.
20	Uses simple tables, charts, or graphs to represent meaningful data.	Standard 6: The students demonstrate understanding of and an ability to use data analysis, probability, and statistics.
21	Attends to teacher and materials in environment.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.

Mathematics - Grade 8		
Item	**Performance Indicator**	Standard
22	Describes features of the data.	Standard 6: The students demonstrate understanding of and an ability to use data analysis, probability, and statistics.
23	Determines which category had the most/least.	Standard 6: The students demonstrate understanding of and an ability to use data analysis, probability, and statistics.
24	Makes decisions based on data, a table or graph.	Standard 6: The students demonstrate understanding of and an ability to use data analysis, probability, and statistics.
25	Explains/shows how decisions were made using a table or graph.	Standard 6: The students demonstrate understanding of and an ability to use data analysis, probability, and statistics.

Science - Grade 8		
Item	**Performance Indicator**	**Standard**
1	Attend to an inclined plane, wheel and axle, lever, and a pulley.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
2	Identify a lever.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
3	Identify that a pulley can raise an object easier.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
4	Identify a force as a push or pull.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
5	Identify and predict the results of an investigation.	Standard 1: Students design, conduct, evaluate, and communicate processes and results of scientific investigations, and demonstrate the thinking skills associated with this procedural knowledge.
6	Identify a variable.	Standard 1: Students design, conduct, evaluate, and communicate processes and results of scientific investigations, and demonstrate the thinking skills associated with this procedural knowledge.
7	Attend to common substances or objects.	Standard 2: Students demonstrate knowledge of properties, forms, changes, interactions of physical; and chemical systems, and demonstrate thinking skills associated with knowledge.

Science - Grade 8		
Item	**Performance Indicator**	**Standard**
8	Identify something that needs energy from food.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
9	Identify an animal as something that breathes.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
10	Identify a plant as something that breathes.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
11	Recognize that plants make their own food.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
12	Attend to what the pictures are showing.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
13	Identify whether a person or a representation of a person is a baby, child, or adult.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.

Science - Grade 8		
Item	**Performance Indicator**	**Standard**
14	Sequence baby, child, young adult, and adult as the life cycle of a human.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
15	Sequence seed, seedling, young plant, mature plant as the life cycle of a flowering plant.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
16	Sequence an egg, caterpillar, chrysalis, and butterfly as the life cycle of a butterfly.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
17	Attend to Earth's changing features.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
18	Identify an island.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
19	Identify a slow change. Identify that the surface of Earth is made of many pieces that move.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
20	Identify a hill or mountain. Identify a slow change. Recognize that mountains can form where pieces collide.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.

Science - Grade 8		
Item	**Performance Indicator**	**Standard**
21	Identify a slow change.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
22	Attend to teacher, soil, rock, air, and water.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
23	Distinguish rocks from other objects or materials.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
24	Describe rocks using one to two physical properties. (e.g. color, size, and shape of particles, texture, weight/density).	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
25	Distinguish water from other objects or materials.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
26	Identify a rock or mineral being used.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.

Reading - Grade 10		
Item	**Performance Indicator**	**Standard**
1	Attends to people and objects in the environment.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
2	Locates a picture/symbol/object when named or signed.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
3	Identifies words/pictures/symbols/objects used for content communication.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
4	Follows directions that contain verbs (points to, looks at, turns page, hits switch).	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
5	Identifies a variety of resources.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
6	Anticipates the beginning of a literacy activity (looks toward reader, tolerates headphones, locates literacy materials).	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
7	Communicates ideas generated from reading.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
8	Explains the meaning of new vocabulary words in the context of a story/reading selection/activity.	Standard 2: Students apply a range of skills and strategies to read.
9	Communicates feelings generated from reading.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
10	Communicates preferred mode for reading/comprehending literacy materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
11	Attends to literacy materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
12	Identifies details of characters that are the same.	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.
13	Explains the meaning of new vocabulary words in the context of a story/reading selection/activity.	Standard 2: Students apply a range of skills and strategies to read.
14	Communicates ideas generated from reading.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.

Reading - Grade 10		
Item	**Performance Indicator**	**Standard**
15	Selects important details/facts from reading materials.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
16	Previews/explores literacy material (looks at, touches, holds, listens),	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
17	Uses text features to comprehend content-area texts; and uses word recognition skills and context clues to comprehend text.	Standard 2: Students apply a range of skills and strategies to read.
18	Demonstrates understanding of a new word based on context of a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
19	Identifies synonyms.	Standard 2: Students apply a range of skills and strategies to read.
20	Uses one course to organize information.	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.
21	Responds to own name, spoken/signed, print/picture.	Standard 1: Students construct meaning as they comprehend, interpret, and respond to what they read.
22	Uses listening/observing strategies to comprehend a reading selection.	Standard 2: Students apply a range of skills and strategies to read.
23	Identifies an appropriate information resource to gain specific information.	Standard 4: Students select, read, and respond to print and nonprint material for a variety of purposes.
24	Identifies items on a graph or table.	Standard 4: Students select, read, and respond to print and nonprint material for a variety of purposes.
25	Demonstrates understanding of the difference between an information resource and literature.	Standard 4: Students select, read, and respond to print and nonprint material for a variety of purposes.

Mathematics - Grade 10		
Item	**Performance Indicator**	**Standard**
1	Attends to another person demonstrating concrete materials.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
2	Demonstrates the concept of one.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
3	Demonstrates that a collection of objects has a quantity.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
4	Demonstrates an understanding of addition as combining collections of things.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
5	Determines whether the numbers of identical objects in two structured groups are the same or different; which group has more; and chooses the correct strategies or procedures to solve a number problem.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
6	Attends to another person reviewing a weekly budget chart; anticipates the beginning of a math activity; and attends to materials being displayed.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
7	Matches bills and their values.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
8	Uses different bill combinations to show equivalent amounts.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
9	Determines how much more money is needed when funds are insufficient.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
10	Determines change when funds are more than cost.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
11	Attends to another person showing relationships between two variables, using objects, pictures, symbols, or numbers.	Standard 3: Students use algebraic concepts, processes, and language to model and solve a variety of real-world and mathematical problems.

Mathematics - Grade 10		
Item	**Performance Indicator**	**Standard**
12	Shows a relationship between two variables, using ordered pairs or a table; then, makes a table.	Standard 3: Students use algebraic concepts, processes, and language to model and solve a variety of real-world and mathematical problems.
13	Given a numerical relationship between two variables, finds the value of one given the other.	Standard 3: Students use algebraic concepts, processes, and language to model and solve a variety of real-world and mathematical problems.
14	Uses or extends a T-table to find the value of a variable.	Standard 3: Students use algebraic concepts, processes, and language to model and solve a variety of real-world and mathematical problems.
15	Demonstrates an understanding of division, using concrete materials.	Standard 2: Students demonstrate understanding of and an ability to use numbers and operations.
16	Attends to another person reviewing a series of functional signs representing different shapes; anticipates the beginning of a math activity; and attends to materials being displayed.	Standard 1: Students engage in the mathematical processes of problem solving and reasoning, estimation, communication, connections and applications, and using appropriate technology.
17	Recognizes properties of two-dimensional shapes.	Standard 4: Students demonstrate understanding of shape and an ability to use geometry.
18	Identifies circles, squares, triangles, ovals, and rectangles regardless of their orientation or general shape.	Standard 4: Students demonstrate understanding of shape and an ability to use geometry.
19	Follows navigational directions.	Standard 4: Students demonstrate understanding of shape and an ability to use geometry.
20	Recalls shapes and their relative positions after they have been viewed for only a brief period of time.	Standard 4: Students demonstrate understanding of shape and an ability to use geometry.
21	Attends to another person making patterns and to a person describing patterns.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
22	Groups/sorts objects into sets.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
23	Demonstrates that objects defined by a shared attribute form a set to which a number can be applied.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.

Mathematics - Grade 10		
Item	**Performance Indicator**	**Standard**
24	Models mathematical problems.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.
25	Uses models, tables, and graphs to make decisions.	Standard 7: Students demonstrate understanding of and an ability to use patterns, relations and functions.

Science - Grade 10		
Item	**Performance Indicator**	**Standard**
1	Attend to temperature changes (heat) being produced by rubbing.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
2	Identify that temperature changes (heat) can be produced by a heat source (e.g. burner, fire).	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
3	Identify that temperature changes (heat) can move from one object to another.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
4	Identify the changes in matter from solid to liquid to gas as temperature increases or from gas to liquid to solid as temperature decreases.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
5	Identify the changes in matter from solid to liquid to gas as temperature increases or from gas to liquid to solid as temperature decreases.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
6	Recognize that the model represents an element.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
7	Attend to something moving.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
8	Recognize that motion is caused by outside forces.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical, chemical systems, and demonstrate thinking skills associated with this knowledge.

Science - Grade 10		
Item	**Performance Indicator**	**Standard**
9	Recognize that motion is caused by outside forces. (e.g. a push causes something to move)	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
10	Demonstrate that some objects are attracted or repelled by magnets, and some objects are not affected by magnets.	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
11	Recognize that motion is caused by outside forces. (e.g. a push causes something to move).	Standard 2: Students demonstrate knowledge of properties, forms, changes and interactions of physical; and chemical systems, and demonstrate the thinking skills associated with this knowledge.
12	Attend to cells.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
13	Recognize bacteria/germs.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
14	Identify a microscope.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
15	Identify one or two places where bacteria/germs might be found.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.

Science - Grade 10		
Item	**Performance Indicator**	**Standard**
16	Identify that bacteria/germs cause some diseases.	Standard 3: Students demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment, and demonstrate the thinking skills associated with this knowledge.
17	Recognize that medical treatment received is a benefit of scientific or technological innovation.	Standard 5: Students understand how scientific knowledge and technological developments impact today's societies and cultures.
18	Attend to weather measurement instruments.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
19	Identify the thermometer in preparation for reading the temperature from it.	Standard 1: Students design, conduct, evaluate, and communicate processes and results of scientific investigations, and demonstrate Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
20	Read a thermometer.	Standard 1: Students design, conduct, evaluate, and communicate processes and results of scientific investigations, and demonstrate the thinking skills associated with this procedural knowledge. Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
21	Identify the tools and resources needed for the investigation.	Standard 1: Students design, conduct, evaluate, and communicate processes and results of scientific investigations, and demonstrate the thinking skills associated with this procedural knowledge.

Science - Grade 10		
Item	**Performance Indicator**	**Standard**
22	Get information about the weather from a weather report.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
23	Attend to the Sun, Moon, and stars.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
24	Identify the Sun.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
25	Recognize a simple telescope.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
26	Identify that light and heat come from the Sun.	Standard 1: Students design, conduct, evaluate, and communicate processes and results of scientific investigations, and demonstrate the thinking skills associated with this procedural knowledge. Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.
27	Given an investigation, identify the things that change in the investigation.	Standard 1: Students design, conduct, evaluate, and communicate processes and results of scientific investigations, and demonstrate the thinking skills associated with this procedural knowledge.
28	Identify that light and heat come from the sun.	Standard 4: Students demonstrate knowledge of the composition, processes and interactions of Earth's systems and other objects in space, and demonstrate the thinking skills associated with this knowledge.

Appendix G—Item-Level Classical Statistics Results

Table G-1. 2008–09 MONTANA ALT: Item Level Classical Stats by Grade, Content Area and Form—Mathematics

Grade	*Item Position*	*N*	*Mean Rubric Score*	*SD*	*Difficulty*	*Discrimination*
	1	84	3.86	0.75	0.72	0.97
	2	84	3.62	0.98	0.85	0.91
	3	84	3.58	1.12	0.81	0.90
	4	82	3.40	1.14	0.73	0.85
	5	82	3.55	0.98	0.79	0.89
	6	84	3.81	0.86	0.76	0.95
	7	84	3.26	1.18	0.62	0.82
	8	83	3.24	1.24	0.75	0.81
	9	82	3.56	0.96	0.76	0.89
	10	82	3.00	1.24	0.67	0.75
	11	84	3.81	0.86	0.73	0.95
	12	84	3.73	0.78	0.77	0.93
3	13	84	3.44	1.03	0.83	0.86
	14	82	2.85	1.29	0.63	0.71
	15	82	3.77	0.71	0.59	0.94
	16	84	3.76	0.95	0.62	0.94
	17	84	3.29	1.19	0.67	0.82
	18	83	2.60	1.07	0.44	0.65
	19	82	2.73	1.12	0.50	0.68
	20	81	2.81	1.34	0.62	0.70
	21	84	3.86	0.75	0.72	0.97
	22	84	3.55	1.08	0.81	0.89
	23	84	3.30	1.20	0.79	0.83
	24	81	3.46	1.05	0.80	0.87
	25	81	2.99	1.26	0.61	0.75
	1	100	3.80	0.88	0.61	0.95
	2	99	3.61	0.99	0.77	0.90
	3	100	3.38	1.24	0.87	0.85
	4	97	3.05	1.24	0.69	0.76
	5	97	3.35	1.11	0.75	0.84
	6	99	3.88	0.69	0.58	0.97
	7	99	3.57	1.03	0.76	0.89
	8	99	2.91	1.39	0.75	0.73
	9	97	3.14	1.30	0.82	0.79
	10	96	2.55	1.37	0.57	0.64
	11	98	3.88	0.69	0.58	0.97
	12	97	3.16	1.20	0.81	0.79
4	13	99	2.45	1.21	0.64	0.61
	14	96	2.68	1.17	0.67	0.67
	15	96	2.97	1.18	0.58	0.74
	16	98	3.88	0.69	0.42	0.97
	17	99	3.23	1.15	0.75	0.81
	18	99	2.82	1.26	0.66	0.71
	19	97	3.09	1.09	0.64	0.77
	20	96	3.16	1.20	0.60	0.79
	21	99	3.92	0.57	0.32	0.98
	22	99	3.67	0.89	0.71	0.92
	23	99	3.04	1.26	0.72	0.76
	24	98	2.69	1.18	0.55	0.67
	25	97	2.97	1.23	0.62	0.74

continued

Grade	Item Position	N	Mean Rubric Score	SD	Difficulty	Discrimination
	1	95	3.62	1.18	0.61	0.91
	2	95	2.98	1.49	0.90	0.75
	3	95	3.00	1.48	0.83	0.75
	4	89	3.24	1.21	0.85	0.81
	5	87	2.90	1.34	0.79	0.73
	6	95	3.66	1.12	0.63	0.92
	7	95	3.26	1.29	0.86	0.82
	8	95	2.84	1.45	0.84	0.71
	9	90	3.13	1.26	0.76	0.78
	10	89	2.89	1.42	0.79	0.72
	11	95	3.62	1.18	0.70	0.91
	12	95	3.15	1.40	0.89	0.79
5	13	95	3.14	1.40	0.89	0.79
	14	90	2.83	1.28	0.75	0.71
	15	88	2.61	1.41	0.75	0.65
	16	95	3.66	1.12	0.72	0.92
	17	95	3.28	1.27	0.75	0.82
	18	95	3.04	1.41	0.90	0.76
	19	90	2.89	1.38	0.81	0.72
	20	89	3.11	1.21	0.63	0.78
	21	95	3.75	0.98	0.64	0.94
	22	95	3.00	1.38	0.77	0.75
	23	95	2.94	1.37	0.77	0.74
	24	91	2.88	1.43	0.80	0.72
	25	88	3.10	1.32	0.77	0.78
	1	106	3.85	0.77	0.56	0.96
	2	106	3.74	0.80	0.77	0.94
	3	106	3.72	0.88	0.75	0.93
	4	104	3.34	1.12	0.74	0.84
	5	104	3.30	1.23	0.77	0.83
	6	106	3.77	0.93	0.57	0.94
	7	106	3.39	1.08	0.83	0.85
	8	106	3.24	1.24	0.75	0.81
	9	104	3.28	1.15	0.73	0.82
	10	104	2.77	1.35	0.66	0.69
	11	106	3.85	0.77	0.56	0.96
	12	106	3.29	1.10	0.66	0.82
6	13	106	3.54	1.11	0.77	0.89
	14	104	3.53	1.01	0.70	0.88
	15	104	3.41	1.12	0.78	0.85
	16	106	3.89	0.67	0.64	0.97
	17	106	3.52	1.11	0.79	0.88
	18	106	2.58	1.46	0.60	0.65
	19	104	3.52	1.14	0.75	0.88
	20	102	3.42	1.18	0.72	0.86
	21	106	3.85	0.77	0.56	0.96
	22	106	3.54	1.03	0.81	0.89
	23	106	3.45	1.11	0.82	0.86
	24	104	3.51	0.96	0.77	0.88
	25	104	3.13	1.17	0.75	0.78

continued

Grade	Item Position	N	Mean Rubric Score	SD	Difficulty	Discrimination
	1	69	4.00	0.00		1.00
	2	69	3.38	1.21	0.70	0.85
	3	69	2.93	1.31	0.82	0.73
	4	69	3.03	1.27	0.76	0.76
	5	67	3.00	1.26	0.63	0.75
	6	69	3.94	0.48	0.37	0.99
	7	69	3.42	1.06	0.79	0.86
	8	69	3.20	1.28	0.84	0.80
	9	68	2.28	1.40	0.77	0.57
	10	67	1.75	1.31	0.56	0.44
	11	69	3.88	0.68	0.11	0.97
	12	69	2.75	1.42	0.80	0.69
7	13	69	2.64	1.40	0.81	0.66
	14	69	2.68	1.47	0.79	0.67
	15	66	2.70	1.31	0.64	0.68
	16	69	3.94	0.48	0.37	0.99
	17	69	3.16	1.30	0.85	0.79
	18	69	3.09	1.33	0.84	0.77
	19	68	3.16	1.27	0.88	0.79
	20	65	3.14	1.22	0.82	0.79
	21	69	3.94	0.48	0.37	0.99
	22	69	3.19	1.25	0.78	0.80
	23	69	3.33	1.18	0.82	0.83
	24	68	3.35	1.12	0.81	0.84
	25	67	3.03	1.30	0.84	0.76
	1	98	3.92	0.57	0.04	0.98
	2	98	3.53	0.95	0.63	0.88
	3	98	2.61	1.15	0.56	0.65
	4	98	2.74	1.17	0.60	0.69
	5	97	2.61	1.25	0.61	0.65
	6	98	3.96	0.40	0.31	0.99
	7	98	3.56	0.97	0.71	0.89
	8	98	2.79	1.30	0.65	0.70
	9	97	2.96	1.23	0.64	0.74
	10	95	3.08	1.28	0.71	0.77
	11	98	3.96	0.40	0.31	0.99
	12	97	3.27	1.14	0.73	0.82
8	13	97	2.62	1.46	0.70	0.66
	14	96	2.76	1.30	0.56	0.69
	15	94	2.39	1.27	0.58	0.60
	16	98	3.92	0.57	0.28	0.98
	17	97	2.85	1.29	0.69	0.71
	18	98	3.09	1.24	0.74	0.77
	19	97	2.63	1.27	0.68	0.66
	20	96	2.74	1.36	0.73	0.69
	21	98	3.88	0.69	0.22	0.97
	22	98	2.83	1.30	0.71	0.71
	23	98	3.22	1.16	0.74	0.81
	24	97	3.04	1.31	0.70	0.76
	25	94	2.83	1.36	0.65	0.71

continued

Grade	Item Position	N	Mean Rubric Score	SD	Difficulty	Discrimination
	1	126	3.90	0.61	0.38	0.98
	2	126	3.66	0.96	0.81	0.92
	3	126	3.51	1.09	0.82	0.88
	4	125	3.11	1.30	0.79	0.78
	5	123	3.30	1.17	0.70	0.83
	6	126	3.87	0.70	0.61	0.97
	7	125	3.42	1.11	0.88	0.86
	8	126	3.36	1.10	0.80	0.84
	9	123	2.94	1.33	0.67	0.74
	10	122	3.01	1.20	0.68	0.75
	11	126	3.78	0.92	0.65	0.95
	12	126	3.13	1.20	0.80	0.78
10	13	125	3.14	1.32	0.78	0.79
	14	121	3.15	1.15	0.55	0.79
	15	120	2.79	1.35	0.55	0.70
	16	125	3.97	0.36	0.35	0.99
	17	126	3.60	1.03	0.73	0.90
	18	126	3.55	1.02	0.77	0.89
	19	125	3.31	1.18	0.69	0.83
	20	123	3.20	1.16	0.70	0.80
	21	126	3.94	0.50	0.47	0.99
	22	125	3.54	1.10	0.85	0.89
	23	125	3.49	1.11	0.87	0.87
	24	124	3.52	1.09	0.82	0.88
	25	119	3.45	1.05	0.72	0.86

Table G-2. 2008–09 MONTANA ALT: Item Level Classical Stats by Grade, Content Area and Form—Reading

Grade	*Item Position*	*N*	*Mean Rubric Score*	*SD*	*Difficulty*	*Discrimination*
	1	82	3.80	0.87	0.58	0.95
	2	83	3.10	1.25	0.65	0.78
	3	82	2.98	1.28	0.67	0.75
	4	80	3.15	1.08	0.40	0.79
	5	80	2.91	1.17	0.56	0.73
	6	83	3.86	0.75	0.68	0.97
	7	83	3.24	1.26	0.66	0.81
	8	81	3.23	1.29	0.76	0.81
	9	80	3.29	1.20	0.60	0.82
	10	78	2.90	1.20	0.37	0.73
	11	82	3.80	0.87	0.68	0.95
	12	81	3.19	1.26	0.60	0.80
3	13	82	3.71	0.85	0.78	0.93
	14	80	3.36	1.11	0.72	0.84
	15	80	3.64	0.96	0.75	0.91
	16	82	3.85	0.76	0.69	0.96
	17	83	2.60	1.33	0.73	0.65
	18	83	2.71	1.33	0.61	0.68
	19	79	2.94	1.23	0.49	0.74
	20	79	3.29	1.26	0.65	0.82
	21	83	3.95	0.44	0.42	0.99
	22	82	3.20	1.21	0.69	0.80
	23	81	3.60	0.98	0.77	0.90
	24	81	3.41	1.09	0.76	0.85
	25	81	3.40	1.07	0.55	0.85
	1	94	3.96	0.41	0.45	0.99
	2	94	3.65	0.91	0.81	0.91
	3	94	3.18	1.15	0.68	0.80
	4	93	3.59	0.94	0.75	0.90
	5	92	3.45	0.92	0.64	0.86
	6	93	3.96	0.41	-0.02	0.99
	7	94	3.46	1.07	0.81	0.87
	8	93	3.42	1.02	0.71	0.86
	9	94	3.22	1.14	0.62	0.81
	10	91	3.03	1.09	0.65	0.76
	11	94	4.00	0.00		1.00
	12	93	2.67	1.30	0.57	0.67
4	13	93	3.24	1.18	0.71	0.81
	14	93	3.26	1.17	0.76	0.82
	15	91	3.62	0.85	0.76	0.91
	16	94	3.87	0.71	0.19	0.97
	17	94	3.53	0.95	0.74	0.88
	18	94	2.85	1.24	0.65	0.71
	19	94	3.04	1.13	0.71	0.76
	20	90	3.23	1.10	0.62	0.81
	21	94	3.96	0.41	0.10	0.99
	22	94	3.35	1.10	0.77	0.84
	23	93	3.65	0.93	0.65	0.91
	24	94	3.34	1.12	0.68	0.84
	25	93	3.43	1.03	0.78	0.86

continued

Grade	Item Position	N	Mean Rubric Score	SD	Difficulty	Discrimination
	1	93	3.61	1.19	0.60	0.90
	2	93	2.60	1.32	0.68	0.65
	3	93	2.96	1.37	0.76	0.74
	4	88	2.80	1.21	0.58	0.70
	5	85	2.95	1.25	0.49	0.74
	6	92	3.65	1.13	0.75	0.91
	7	93	3.12	1.36	0.81	0.78
	8	93	3.11	1.36	0.81	0.78
	9	88	3.28	1.25	0.82	0.82
	10	86	3.17	1.19	0.74	0.79
	11	93	3.61	1.19	0.78	0.90
	12	93	2.68	1.32	0.80	0.67
5	13	93	2.52	1.37	0.66	0.63
	14	87	3.29	1.20	0.65	0.82
	15	85	2.33	1.26	0.43	0.58
	16	93	3.70	1.06	0.75	0.93
	17	93	2.82	1.42	0.80	0.71
	18	93	3.22	1.29	0.84	0.81
	19	88	3.17	1.25	0.69	0.79
	20	87	3.29	1.23	0.65	0.82
	21	93	3.74	0.99	0.73	0.94
	22	93	3.10	1.38	0.79	0.78
	23	93	2.94	1.37	0.76	0.74
	24	88	3.38	1.16	0.78	0.85
	25	87	2.91	1.21	0.60	0.73
	1	103	3.88	0.68	0.45	0.97
	2	102	3.57	0.97	0.77	0.89
	3	103	3.55	0.98	0.78	0.89
	4	103	2.66	1.22	0.58	0.67
	5	102	3.29	1.10	0.57	0.82
	6	102	3.84	0.78	0.60	0.96
	7	102	3.54	0.97	0.74	0.89
	8	102	3.27	1.06	0.79	0.82
	9	100	3.02	1.13	0.54	0.76
	10	100	3.17	1.22	0.63	0.79
	11	103	3.84	0.78	0.62	0.96
	12	103	3.42	0.98	0.70	0.86
6	13	103	3.24	1.17	0.68	0.81
	14	101	3.64	0.86	0.78	0.91
	15	101	3.58	0.93	0.77	0.90
	16	103	3.88	0.68	0.47	0.97
	17	103	3.37	1.10	0.74	0.84
	18	103	3.50	1.03	0.72	0.88
	19	102	3.31	1.15	0.73	0.83
	20	101	3.55	0.94	0.74	0.89
	21	103	3.81	0.86	0.60	0.95
	22	103	3.33	1.05	0.75	0.83
	23	103	3.29	1.17	0.70	0.82
	24	101	3.03	1.25	0.69	0.76
	25	101	3.37	1.09	0.54	0.84

continued

Grade	Item Position	N	Mean Rubric Score	SD	Difficulty	Discrimination
	1	69	4.00	0.00		1.00
	2	69	2.94	1.24	0.62	0.74
	3	69	2.39	1.29	0.65	0.60
	4	69	2.81	1.30	0.69	0.70
	5	67	2.64	1.43	0.65	0.66
	6	69	4.00	0.00		1.00
	7	69	3.46	1.05	0.77	0.87
	8	69	3.10	1.23	0.82	0.78
	9	69	3.48	1.09	0.83	0.87
	10	67	3.36	1.04	0.76	0.84
	11	69	3.88	0.68	0.48	0.97
	12	69	3.41	1.17	0.84	0.85
7	13	69	3.36	1.26	0.76	0.84
	14	68	3.28	1.16	0.69	0.82
	15	67	3.60	1.02	0.76	0.90
	16	69	3.94	0.48	0.43	0.99
	17	69	3.14	1.25	0.80	0.79
	18	69	2.90	1.37	0.68	0.73
	19	68	3.32	1.16	0.78	0.83
	20	66	3.35	1.22	0.65	0.84
	21	69	3.94	0.48	0.43	0.99
	22	69	2.97	1.32	0.75	0.74
	23	69	3.07	1.26	0.79	0.77
	24	68	2.81	1.38	0.60	0.70
	25	66	3.59	0.89	0.68	0.90
	1	96	3.96	0.41	0.09	0.99
	2	96	3.21	1.10	0.62	0.80
	3	96	2.68	1.16	0.51	0.67
	4	96	3.09	1.31	0.79	0.77
	5	94	3.20	1.14	0.65	0.80
	6	96	4.00	0.00		1.00
	7	96	3.31	1.19	0.70	0.83
	8	96	2.81	1.25	0.59	0.70
	9	95	2.94	1.24	0.71	0.74
	10	95	3.20	1.20	0.68	0.80
	11	96	3.92	0.57	0.17	0.98
	12	96	3.09	1.21	0.68	0.77
8	13	96	3.21	1.09	0.77	0.80
	14	96	3.04	1.22	0.69	0.76
	15	94	3.20	1.17	0.73	0.80
	16	96	3.96	0.41	-0.03	0.99
	17	96	3.50	0.88	0.61	0.88
	18	96	3.20	1.16	0.70	0.80
	19	96	3.40	1.13	0.79	0.85
	20	95	2.67	1.20	0.58	0.67
	21	96	4.00	0.00		1.00
	22	96	3.30	1.15	0.70	0.83
	23	96	3.17	1.24	0.78	0.79
	24	96	3.25	1.25	0.85	0.81
	25	95	3.11	1.25	0.74	0.78

continued

Grade	Item Position	N	Mean Rubric Score	SD	Difficulty	Discrimination
	1	128	3.91	0.61	0.36	0.98
	2	128	3.23	1.18	0.83	0.81
	3	128	3.48	1.12	0.85	0.87
	4	126	3.40	1.05	0.70	0.85
	5	123	3.50	0.94	0.69	0.88
	6	128	3.84	0.78	0.70	0.96
	7	128	3.07	1.22	0.75	0.77
	8	128	3.46	1.10	0.81	0.87
	9	124	2.84	1.23	0.61	0.71
	10	122	3.52	0.96	0.66	0.88
	11	128	3.84	0.78	0.70	0.96
	12	128	3.35	1.25	0.83	0.84
10	13	128	2.91	1.29	0.74	0.73
	14	123	3.16	1.18	0.68	0.79
	15	121	3.01	1.17	0.57	0.75
	16	128	3.81	0.85	0.67	0.95
	17	127	2.98	1.32	0.72	0.75
	18	127	2.94	1.27	0.70	0.74
	19	123	3.41	1.11	0.73	0.85
	20	122	3.11	1.18	0.64	0.78
	21	127	3.87	0.70	0.55	0.97
	22	127	3.40	1.17	0.75	0.85
	23	127	3.21	1.26	0.82	0.80
	24	124	3.54	0.87	0.72	0.89
	25	122	3.19	1.16	0.71	0.80

Table G-3. 2008–09 MONTANA ALT: Item Level Classical Stats by Grade, Content Area and Form—Science

Grade	*Item Position*	*N*	*Mean Rubric Score*	*SD*	*Difficulty*	*Discrimination*
4	1	104	3.92	0.55	0.54	0.98
	2	104	3.31	1.17	0.73	0.83
	3	103	3.12	1.24	0.75	0.78
	4	102	3.35	1.20	0.78	0.84
	5	100	3.14	1.17	0.55	0.79
	6	103	3.92	0.55	0.57	0.98
	7	102	3.59	1.00	0.70	0.90
	8	102	3.41	1.21	0.83	0.85
	9	101	3.62	0.87	0.67	0.91
	10	100	3.35	1.12	0.80	0.84
	11	100	3.18	1.25	0.68	0.80
	12	102	3.92	0.56	0.57	0.98
	13	102	3.69	0.90	0.85	0.92
	14	102	3.57	1.02	0.84	0.89
	15	101	3.66	0.89	0.80	0.92
	16	99	3.58	0.90	0.73	0.90
	17	102	3.92	0.56	0.57	0.98
	18	102	3.36	1.09	0.83	0.84
	19	101	3.34	1.08	0.80	0.84
	20	100	3.54	1.04	0.76	0.89
	21	99	2.79	1.28	0.53	0.70
	22	103	3.96	0.39	0.42	0.99
	23	103	3.19	1.17	0.71	0.80
	24	103	3.17	1.25	0.74	0.79
	25	102	3.09	1.19	0.67	0.77
	26	99	3.42	1.13	0.79	0.86
8	1	99	3.96	0.40	0.45	0.99
	2	100	2.95	1.24	0.55	0.74
	3	100	3.16	1.24	0.79	0.79
	4	99	3.37	1.18	0.69	0.84
	5	97	3.18	1.21	0.44	0.80
	6	95	2.85	1.30	0.54	0.71
	7	100	3.96	0.40	0.45	0.99
	8	100	3.28	1.08	0.59	0.82
	9	99	3.53	0.97	0.73	0.88
	10	99	3.45	1.10	0.75	0.86
	11	97	3.39	1.02	0.71	0.85
	12	100	3.96	0.40	0.45	0.99
	13	100	3.74	0.86	0.70	0.94
	14	100	2.92	1.28	0.72	0.73
	15	99	2.99	1.35	0.75	0.75
	16	98	2.51	1.24	0.64	0.63
	17	100	3.96	0.40	0.45	0.99
	18	100	3.49	1.11	0.67	0.87
	19	100	3.08	1.20	0.66	0.77
	20	99	2.95	1.30	0.76	0.74
	21	96	2.71	1.26	0.64	0.68
	22	100	3.92	0.56	0.43	0.98
	23	100	3.76	0.83	0.71	0.94

continued

Grade	Item Position	N	Mean Rubric Score	SD	Difficulty	Discrimination
8	24	100	3.58	1.03	0.65	0.90
	25	98	3.71	0.80	0.74	0.93
	26	97	3.53	0.97	0.70	0.88
10	1	128	3.88	0.70	0.65	0.97
	2	128	3.13	1.15	0.60	0.78
	3	128	3.60	1.01	0.79	0.90
	4	125	3.54	0.88	0.69	0.89
	5	125	3.20	1.26	0.70	0.80
	6	125	2.96	1.24	0.62	0.74
	7	128	3.91	0.61	0.36	0.98
	8	128	3.34	1.17	0.84	0.84
	9	128	3.30	1.21	0.68	0.83
	10	126	3.51	1.00	0.79	0.88
	11	123	3.53	1.02	0.71	0.88
	12	128	3.88	0.70	0.44	0.97
	13	128	3.45	1.11	0.78	0.86
	14	128	3.54	1.11	0.87	0.89
	15	126	3.48	1.06	0.63	0.87
	16	123	3.79	0.73	0.74	0.95
	17	123	3.67	0.89	0.73	0.92
	18	128	3.91	0.61	0.42	0.98
	19	127	3.52	1.10	0.80	0.88
	20	128	3.42	1.19	0.82	0.86
	21	126	3.60	0.98	0.80	0.90
	22	124	3.69	0.83	0.73	0.92
	23	127	3.94	0.50	0.53	0.99
	24	128	3.72	0.90	0.76	0.93
	25	128	3.52	1.06	0.87	0.88
	26	126	3.44	1.13	0.66	0.86
	27	125	3.22	1.21	0.76	0.81
	28	123	3.57	0.99	0.79	0.89

Appendix H—Decision Accuracy and Consistency Results

Table H-1. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 3 Mathematics

Accuracy and Consistency of Classification Indices				
Overall Indices	Accuracy	Consistency	Kappa (κ)	
	0.804	0.747	0.651	
Indices Conditional on Level		Accuracy	Consistency	
	Novice	0.898	0.873	
	Nearing Proficiency	0.659	0.561	
	Proficient	0.658	0.595	
	Advanced	0.930	0.842	
Indices for Dichotomous Decisions Around Cutpoints		Accuracy		Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.957	0.025	0.018	0.941
	NP : P	0.941	0.038	0.022	0.919
	P : A	0.905	0.072	0.023	0.882

Table H-2. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 4 Mathematics

Accuracy and Consistency of Classification Indices				
Overall Indices	Accuracy	Consistency	Kappa (κ)	
	0.841	0.783	0.699	
Indices Conditional on Level		Accuracy	Consistency	
	Novice	0.888	0.851	
	Nearing Proficiency	0.757	0.678	
	Proficient	0.745	0.676	
	Advanced	0.939	0.879	
Indices for Dichotomous Decisions Around Cutpoints		Accuracy		Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.966	0.018	0.016	0.952
	NP : P	0.945	0.032	0.023	0.924
	P : A	0.930	0.048	0.022	0.906

Table H-3. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 5 Mathematics

Accuracy and Consistency of Classification Indices				
Overall Indices	Accuracy	Consistency	Kappa (κ)	
	0.859	0.818	0.741	
Indices Conditional on Level		Accuracy	Consistency	
	Novice	0.939	0.925	
	Nearing Proficiency	0.536	0.423	
	Proficient	0.769	0.731	
	Advanced	0.939	0.871	
Indices for Dichotomous Decisions Around Cutpoints		Accuracy		Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.966	0.020	0.014	0.953
	NP : P	0.961	0.023	0.016	0.947
	P : A	0.929	0.052	0.019	0.911

Table H-4. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 6 Mathematics

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.844	0.796	0.704		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.900	0.868		
	Nearing Proficiency	0.828	0.784		
	Proficient	0.650	0.578		
	Advanced	0.942	0.881		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.978	0.012	0.010	0.969
	NP : P	0.951	0.030	0.018	0.933
	P : A	0.915	0.061	0.024	0.892

Table H-5. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 7 Mathematics

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.856	0.805	0.715		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.862	0.808		
	Nearing Proficiency	0.806	0.743		
	Proficient	0.834	0.808		
	Advanced	0.928	0.838		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.978	0.011	0.011	0.968
	NP : P	0.953	0.027	0.021	0.934
	P : A	0.926	0.055	0.019	0.902

Table H-6. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 8 Mathematics

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.828	0.764	0.663		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.842	0.777		
	Nearing Proficiency	0.728	0.639		
	Proficient	0.783	0.725		
	Advanced	0.926	0.857		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.967	0.016	0.016	0.954
	NP : P	0.939	0.034	0.027	0.916
	P : A	0.922	0.053	0.026	0.894

Table H-7. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 10 Mathematics

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.863	0.816	0.706		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.872	0.825		
	Nearing Proficiency	0.799	0.739		
	Proficient	0.669	0.586		
	Advanced	0.958	0.917		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.978	0.011	0.010	0.970
	NP : P	0.953	0.028	0.019	0.935
	P : A	0.931	0.047	0.022	0.909

Table H-8. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 3 Reading

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.843	0.790	0.692		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.835	0.763		
	Nearing Proficiency	0.820	0.768		
	Proficient	0.777	0.736		
	Advanced	0.933	0.855		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.982	0.009	0.009	0.975
	NP : P	0.946	0.031	0.023	0.926
	P : A	0.915	0.062	0.023	0.889

Table H-9. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 4 Reading

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.864	0.816	0.724		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.867	0.812		
	Nearing Proficiency	0.813	0.755		
	Proficient	0.786	0.741		
	Advanced	0.946	0.891		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.982	0.009	0.009	0.975
	NP : P	0.956	0.025	0.019	0.939
	P : A	0.926	0.051	0.023	0.902

Table H-10. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 5 Reading

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.869	0.821	0.736		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.903	0.872		
	Nearing Proficiency	0.796	0.728		
	Proficient	0.689	0.594		
	Advanced	0.956	0.920		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.971	0.016	0.014	0.959
	NP : P	0.955	0.027	0.019	0.937
	P : A	0.944	0.036	0.020	0.923

Table H-11. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 6 Reading

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.886	0.844	0.742		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.853	0.789		
	Nearing Proficiency	0.789	0.716		
	Proficient	0.817	0.780		
	Advanced	0.955	0.914		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.987	0.006	0.007	0.982
	NP : P	0.968	0.018	0.015	0.955
	P : A	0.931	0.046	0.023	0.908

Table H-12. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 7 Reading

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.897	0.858	0.757		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.825	0.743		
	Nearing Proficiency	0.811	0.745		
	Proficient	0.832	0.791		
	Advanced	0.957	0.923		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.990	0.005	0.005	0.986
	NP : P	0.969	0.016	0.014	0.957
	P : A	0.938	0.040	0.023	0.915

Table H-13. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 8 Reading

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.863	0.811	0.676		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.814	0.729		
	Nearing Proficiency	0.733	0.643		
	Proficient	0.747	0.673		
	Advanced	0.950	0.912		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.981	0.009	0.010	0.973
	NP : P	0.956	0.024	0.020	0.938
	P : A	0.926	0.046	0.028	0.898

Table H-14. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 10 Reading

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.884	0.842	0.749		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.900	0.867		
	Nearing Proficiency	0.694	0.591		
	Proficient	0.795	0.742		
	Advanced	0.961	0.927		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.976	0.013	0.011	0.967
	NP : P	0.966	0.019	0.015	0.952
	P : A	0.942	0.038	0.020	0.922

Table H-15. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 4 Science

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.888	0.847	0.744		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.899	0.867		
	Nearing Proficiency	0.727	0.635		
	Proficient	0.758	0.692		
	Advanced	0.966	0.935		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.978	0.012	0.010	0.969
	NP : P	0.966	0.020	0.015	0.952
	P : A	0.944	0.036	0.019	0.924

Table H-16. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 8 Science

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.864	0.814	0.709		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.834	0.756		
	Nearing Proficiency	0.795	0.727		
	Proficient	0.805	0.761		
	Advanced	0.941	0.886		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.986	0.007	0.008	0.980
	NP : P	0.957	0.023	0.020	0.941
	P : A	0.921	0.053	0.026	0.894

Table H-17. 2008–09 Montana CRT-ALT: Accuracy and Consistency—Grade 10 Science

Accuracy and Consistency of Classification Indices					
Overall Indices	Accuracy	Consistency	Kappa (κ)		
	0.872	0.828	0.720		
Indices Conditional on Level		Accuracy	Consistency		
	Novice	0.904	0.876		
	Nearing Proficiency	0.685	0.586		
	Proficient	0.714	0.648		
	Advanced	0.963	0.922		
Indices for Dichotomous Decisions Around Cutpoints		Accuracy			Consistency
		Accuracy	*False Positives*	*False Negatives*	
	N : NP	0.975	0.014	0.011	0.966
	NP : P	0.963	0.022	0.015	0.949
	P : A	0.934	0.047	0.020	0.912

Appendix I—Analysis and Reporting Decision Rules

Analysis and Reporting Decision Rules
Montana Comprehensive Assessment System (MontCAS) CRT and CRT-Alternate
Spring 08–09 Administration

This document details rules for analysis and reporting. The final student level data set used for analysis and reporting is described in the "Data Processing Specifications." This document is considered a draft until the Montana Office of Public Instruction (OPI) signs off. If there are rules that need to be added or modified after said sign-off, OPI sign-off will be obtained for each rule. Details of these additions and modifications will be in the Addendum section.

I. General Information

A. Tests administered

Grade	*Content Area*	*Items Included in Raw Score*		*IABS Reporting Categories (Standards) (Not Applicable for CRT-Alternate)*
		CRT	*CRT-Alt*	
03	Reading & Mathematics	Common	All	Cat2
04	Reading & Mathematics	Common	All	Cat2
	Science	Common	All	Cat3
05	Reading & Mathematics	Common	All	Cat2
06	Reading & Mathematics	Common	All	Cat2
07	Reading & Mathematics	Common	All	Cat2
08	Reading & Mathematics	Common	All	Cat2
	Science	Common	All	Cat3
10	Reading & Mathematics	Common	All	Cat2
	Science	Common	All	Cat3

B. Reports produced
1. Student Labels
2. Student Report
3. Roster & Item Level Report (online system)
   - by grade, subject, and class/group
4. Summary Report
   Consists of sections:
   I. Distribution of Scores
   II. Subtest Results
   III. Results for Subgroups of Students
   - by grade, subject, and school
   - by grade, subject, and system
   - by grade and subject (state level)

C. Files produced (comma delimited format)
1. One state file for each grade
   a. Consists of student level results
   b. Alternately assessed students are in separate files by grade.
2. Naming convention
   a. CRT All subjects- Studentdatafile[2 digit grade].csv
   b. CRT-Alternate All subjects- altStudentdatafile[2 digit grade].csv

D. School Type

Schtype	*Source*	*Description*	*Included in Aggregations*		
			School	System	State
"Pras"	Data file provided by state	Private accredited school. They are their own system	Yes. Same information for school & system but both sets of reports produced	Yes. Same information for school & system but both sets of reports produced	No
"Prnas"	Data file provided by state	Private non-accredited school. They are their own system	Yes. Same information for school & system but both sets of reports produced	Yes. Same information for school & system but both sets of reports produced	No
"SNE"	Scanned data/updated by OPI	Student not enrolled	No	No	No
"Oth"		Non-private school	Yes	Yes	Yes

E. Other information

1. CRT tests are constructed with a combination of common and embedded field test items.
2. The CRT-Alternate consists of a set of performance tasks. The tasks are grouped into five (5) sets of tasklets for each subject. The number of activities in each tasklet varies.
3. Braille students:
   a. There are two common items and three field test items that are not included in the Braille form of the test at grade 3. (See Appendix A.1 for a list of the items not included in the Braille form).
   b. If a student is identified as taking the Braille test, these items are not included in the student's raw score. The student is scaled on a separate form based on the items that are available to him or her. See the Calculations section for more information.
4. Students using JAWS[1]:
   a. There is one reading common item and two math common items which were removed from a grade 7 student booklet in order for JAWS to be used. (See Appendix A.2 for a list of these items.)
   b. If a student is identified as using the JAWS, these items are not included in the student's raw score. The student is scaled on a separate form based on the items that are available to him or her. See the Calculations section for more information.

II. Student Participation/Exclusions

A. Test attempt rules

1. A valid response to a multiple-choice item is A, B, C, or D. An asterisk (multiple marks) is not considered a valid response.
2. Incomplete (CRT): The student has fewer than two (2) but at least one (1) valid response to common multiple choice items.

[1] JAWS (job access with speech) is a Windows© compatible screen reader application.

3. Incomplete (CRT-Alternate): The student responded to fewer than three (3) items.
4. The student is classified as Did Not Participate (DNP) in CRT if the student does not have any valid responses for that subject in either CRT or CRT-Alternate.

B. Not tested reasons

N/A

C. Student participation status

1. The following students are excluded from all aggregations.
    a. foreign exchange students (FXS)
    b. homeschooled students (schtype = "SNE")
    c. part-time students (PSNE)
    d. DNP (for that subject)
    e. First year LEP
    f. student tested with non-standard accommodations (NSA for that subject)

2. If any of the nonstandard accommodations are bubbled, the student is considered tested with non-standard accommodations (NSA) in that subject.
3. If the student has not been in that school for the entire academic year the student is excluded from school level aggregations (NSAY).
4. If the student has not been in that system for the entire academic year the student is excluded from system and school level aggregations (NDAY).
5. If the student took the alternate assessment the student is not counted as participating in the general assessment. Alternate assessment students receive their results on an Alternate Assessment Student Report. They are reported according to participation rules stated in this document.
6. (CRT-Alternate): If the teacher halted the administration of the assessment after the student scored zero (0) for three (3) consecutive items within tasklets, the student is classified as Halted in that subject. If the student was halted within a tasklet, the rest of the items within the tasklet are blanked out and do not count toward the student's score. If the other tasklets are complete, those items will be counted toward the student's score.
7. If the student took the Braille form of the grade 3 test the raw scores are not included in raw score school, system, or state averages. They are not included in group averages on the interactive roster.
8. If the student used JAWS the raw scores are not included in raw score school, system, or state averages. They are not included in group averages on the interactive roster.

D. Student participation summary

Participation Status	*Part. Flag*	*Raw Score*	*Scaled Score*	*Perf. Level*	*Included on Roster*	*Included in Aggregations*		
						Sch	Sys	Sta
FXS	E	✓	✓	✓				
SNE	E	✓	✓	✓				
PSNE	E	✓	✓	✓				
NSA (by subject)	A	✓	✓	✓	✓			
First Year LEP	A	✓	see report specific rules	see report specific rules	✓	only in count of First year LEP		
NSAY Only	B	✓	✓	✓	✓		✓	✓
NDAY	C	✓	✓	✓	✓			✓
ALT*	A	✓	✓	✓	✓	see footnote below		
Incomplete	A	✓	✓	✓	✓			
DNP (non-participants)	F	✓	✓	✓	✓			
Halted (CRT-Alt only by subject)	D	✓	✓	✓	✓	✓	✓	✓
Tested	Z	✓	✓	✓	✓	✓	✓	✓

* Alternate assessment students are included only in the count of alternate assessment students in general assessment reports. They are included in summary data only for alternate assessment reports (according to participation rules).

If a student has conflicting participation statuses the following hierarchy is applied to determine how the student is reported:

F (student attempted no common items and is not alt)
E (FXS, SNE, or PSNE)
A (NSA, first year LEP, ALT or INC)
C (NDAY)
B (NSAY)
Z (completed CRT and none of the above conditions apply)

III. Calculations

A. Raw scores
   1. (CRT) Raw scores are calculated using the scores on common multiple-choice and open-response items.
   2. (CRT-Alternate) Raw score is the sum of the individual item scores.

B. Scaling
   1. Scaling is accomplished by defining the unique set of test forms for each grade/subject combination. This is accomplished as follows:
      a. Translate each form and position into the unique item number assigned to the form/position.
      b. Order the items by
         I. Type—multiple-choice, short-answer, constructed-response
         II. Form—common, then by ascending form number.
         III. Position
      c. If an item number is on a form, set the value for that item number to “1”, otherwise set to “.”. Set the exception field to “0” to indicate this is an original test form.
      d. If an item number contains an “X” (item is not included in scaling), set the item number to “.”. Set the exception field to “1” to indicate this is not an original test form.

e. Compress all of the item numbers together into one field in the order defined in step II to create the test for the student.
f. Select the distinct set of tests from the student data and order by the exception field and the descending test field.
g. Check to see if the test has already been assigned a scale form by looking in the tblScaleForm table. If the test exists, assign the existing scale form. Otherwise assign the next available scale form number. All scale form numbering starts at 01 and increases by increments of 1 up to 99.

2. Psychometrics provides a lookup table for each scale form. These lookup tables are used to assign scaled scores, performance levels, and standard errors.
3. The scaled score cuts for all three subjects and all grades have been fixed and are the same as last year for the CRT.

C. The class code is created using the following steps:
   1. The following students are not included when creating the class codes.
      - SNE
      - ALT (CRT-only)
      - FXS
      - PSNE
   2. The dataset (by grade) is sorted by schcode and class/group name
   3. The records are then numbered consecutively starting at 1. This number is then padded with zeros (in front) to create a three-digit number.

D. Performance level coding:

Numeric Performance Level	*Performance Level Name*	*Abbreviation*
1 (lowest)	Novice	N
2	Nearing Proficient	NP
3	Proficient	P
4 (highest)	Advanced	A

E. Rounding table

Calculation	*Rounded*
Static reports: percents and averages	Rounded to the nearest whole number
Item averages : multiple-choice items	The average is multiplied by 100 and rounded to the nearest whole number.
Item averages: open-response items	Open-response item averages are rounded to the nearest tenth.

F. Minimum N size
   1. The number of included students (N) in a subject is the number of students in the school/system/state minus FXS minus PRAS minus PRNAS minus PSNE minus SNE minus First year LEP minus Incomplete minus NSA minus DNP.
   2. Minimum N size is 10.
   3. School/system reports are produced regardless of N-size.

G. The common items are used in reporting the average number of points for each standard.

H. Assignment of rperflevel
   1. If the student is marked as taking the CRT-Alt, the rperflevel = “A” otherwise
   2. If the student is classified as did not participate (DNP), the rperfleve l = “D” otherwise
   3. If the student is Incomplete in a subject and not marked First year LEP, the rperflevel = “I” otherwise
   4. If the student does not complete the reading test and is marked First year LEP, the rperflevel = “L” for all subjects; otherwise
   5. If the student does not meet any of the above conditions, the rperflevel = perflevel.

IV. Report Specific Rules

A. Student Label
   1. If a student is First year LEP and incomplete in reading, the reading performance level is ‘LEP’. The reading scaled score is blank.
   2. If a student is First year LEP, the mathematics and science performance levels are the same as the earned performance level and the scaled scores are the student’s earned score.
   3. If the student is not First year LEP, the performance level name corresponding to the student’s earned score is displayed.
   4. If the student is First year LEP but is not Incomplete in reading, the the student receives his or her earned scaled score and performance level.
   5. If the student is DNP, the student receives a student label. The student receives scaled score = 200 and performance level = Novice.

B. Student Report
   1. If a student is First year LEP and incomplete in reading the reading performance level is ‘LEP’ and the scaled score is blank.
   2. If the student is First year LEP but is not Incomplete in reading, the student receives his or her earned scaled score and performance level.
   3. If a student is First year LEP, the mathematics and science performance levels are the same as the earned performance level and the scaled score is the student’s earned score.
   4. If the student is not First year LEP, the performance level name corresponding to the student’s earned score is displayed.
   5. If the student is Incomplete the student receives the scores with a footnote (†) “Student did not complete the assessment.”
   6. If the student is NSA the student receives his or her scores with the footnote (§) “Student took non-standard accommodation.”
   7. If there is no last name or first name for the student, the name displayed is “Name Not Provided.”
   8. Alt students who are halted receive their scores and performance level with a footnote (§)
      “Teacher halted the administration of one or more of the five tasklets after the student scored a 0 for three consecutive items within a tasklet on two different test administrations. Any completed tasklets have been scored and are reflected in the student’s scaled score.”
   9. If the student is DNP, the student receives a Student Report. The student receives scaled score = 200 and performance level = Novice. The standards will not be reported. The student receives a footnote (**) “Student did not participate.”

10. Total Points Possible, Student % of points earned, and Average state % are suppressed for students who took Braille test (Braille = "1") or who used JAWS (JAWS = "1"). This suppression is applied only to the standards which contain the items not on the student's form.
11. In section 1 a diamond represents the student's earned scaled score for that subject. The bar is drawn using the high and low scaled score values provided by psychometrics.
12. In section 2, a check mark (✓) is placed in the gray area corresponding to the student's earned performance level for that subject.
13. A % sign is printed in each cell for the state percentages in section 2.
14. If science is not assessed at the grade of the student, the science sections are left blank.
15. The following standards are not reported for either CRT or CRT-Alt:
    a. reading standard 3
    b. mathematics standard 1
    c. science standards 5 and 6

C. Roster & Item Level Report—Alternate Assessment only
1. If a student is First year LEP and the student is not Incomplete in reading:
    a. The mathematics (and science) performance level is the abbreviation of the earned performance level, and the scaled score is the student's earned score.
    b. The reading performance level is the abbreviation of the earned performance level and the scaled score is the student's earned score.
    c. The student is excluded from reading, mathematics and science aggregations.
2. If the student is First year LEP and Incomplete in reading
    a. The student's reading, mathematics (and science) performance levels are 'LEP'.
    b. The student's mathematics (and science) scaled score is the student's earned scaled score, and the reading scaled score is blank.
    c. The student's responses for all subjects are displayed.
    d. The student is excluded from mathematics, reading (and science) aggregations.
3. If the student is not First year LEP, the performance level abbreviation corresponding to the student's earned score is displayed.
4. If the student is Incomplete, the student receives the scores with a footnote (†) "Student did not complete the assessment."
5. If the student is NSA the student will receive his or her scores with the footnote (§) "Student took non-standard accommodation."
6. If there is no last name or first name for the student, the name displayed is "Name Not Provided".
7. If class/group information is missing, the roster is done at the school level.
8. Alternate Assessment students are reported only on their class/group/school's alternate *Roster & Item Level Report.*
9. If the student is a Non-Participant, the student is listed on the *Roster & Items level Report.* All responses and scores will be blank. The scaled score = 200 and performance level = N. The student will receive the footnote "Student did not participate in assessment."

D. Interactive Roster—CRT only

1. Students who test with Non-Standard Accommodations (NSA) are included in school, system, and state level aggregations.
2. Students who are NSAY are included in school, system, and state level aggregations.
3. Students who are NDAY are included in school, system, and state level aggregations.
4. Students who are DNP in a subject are reported with scaled score = 200 and performance level = "DNP".
5. Students who are Incomplete in a subject are reported with their earned scaled score and performance level = "INC" on the interactive roster.
6. Students who are First year LEP and who complete the reading test are reported with their earned scaled score and performance level and are included in school, system, and state level aggregations for all subjects unless otherwise excluded based on completeness in mathematics or science.
7. Students who are First year LEP and who do not complete the reading test are reported with their earned scaled score and performance level = "LEP" for all subjects. These students are excluded from school, system and state level aggregations.
8. Students who participated in Alternate assessment are listed on the rosters. Their scaled score is blank and the performance level = "ALT". These students are not included in aggregations.
9. The items are reported using the released item number.
10. Students who took the Braille form or who used JAWS are not included in any aggregations. These students have a scaleform other than 01.
11. Students to be included in roster aggregations are marked with a "1" in the "Included" field in tblscoreditem. Otherwise, included = "0."
12. Students with participation status E are not available on the interactive site.
13. State level item averages do not include students with school type PRAS, PRNAS or SNE.
14. District level item averages do not include students who are marked nday = "1".

E. School Summary

1. Section III (Results for Subgroups of Students)
    a. Performance level results for subgroups with N less than 10 are suppressed. N is always reported. Footnote * "Less than 10 students were assessed."
    b. Count of students who are considered NSA for that subject excluding those students who are incomplete, nsay (at school level), nday (at school and system level), or FXS or SNE or PSNE or First year LEP or alt (general assessment report).
    c. Count of students who are alt excludes those students who are nsay (at school level), nday (at school or system level), or incomplete or FXS or SNE or PSNE or NSA or First year LEP.
    d. Count of First year LEP students excludes those students who are nsay (at school level), nday (at school or system level), or incomplete or FXS or SNE or PSNE or NSA or alt (general assessment).
    e. Students with scaleform other than 01 are not included in Subtest Results for standards containing items not available to them on their test.

V. Data File Rules (comma delimited format)

1. The following students are not included in the state file
    a. Alternate Assessment students (in CRT)
    b. Homeschooled students (SNE)
    c. Part-Time students (PSNE)
2. If the student receives a performance level 'LEP' on the student report in Reading, the student receives LEP for the reading performance level in the state files.
3. Alt students who are halted are marked '1' in the halted field for that subject.
4. Students who take the Braille form of the test are flagged Braille = "1" in the state and system level files.
5. Students who use JAWS are flagged JAWS = "1" in the state and system level files.
6. In the system level files only the released scored items are included.

VI. Shipping Product Code Summary

1. School (ReportFor = "1")

Grade	*Report Name*	*Report Type*	*Content Area*	*Content Code*	*Quantity*
03	Student Labels(CRT)	03	Reading & Mathematics	00	1 for each student
04	Student Labels (CRT)	03	Reading, Mathematics, and Science	00	1 for each student
05	Student Labels (CRT)	03	Reading & Mathematics	00	1 for each student
06	Student Labels (CRT)	03	Reading & Mathematics	00	1 for each student
07	Student Labels (CRT)	03	Reading & Mathematics	00	1 for each student
08	Student Labels (CRT)	03	Reading Mathematics and Science	00	1 for each student
10	Student Labels (CRT)	03	Reading Mathematics and Science	00	1 for each student
03	Student Report (CRT)	02	Reading & Mathematics	00	1 for each student
04	Student Report (CRT)	02	Reading Mathematics and Science	00	1 for each student
05	Student Report (CRT)	02	Reading & Mathematics	00	1 for each student
06	Student Report (CRT)	02	Reading & Mathematics	00	1 for each student
07	Student Report (CRT)	02	Reading & Mathematics	00	1 for each student
08	Student Report (CRT)	02	Reading Mathematics and Science	00	1 for each student
10	Student Report (CRT)	02	Reading Mathematics and Science	00	1 for each student
03	Student Labels (CRT-Alt)	03	Reading & Mathematics	00	1 for each student
04	Student Labels (CRT-Alt)	03	Reading, Mathematics and Science	00	1 for each student
05	Student Labels (CRT-Alt)	03	Reading & Mathematics	00	1 for each student
06	Student Labels (CRT-Alt)	03	Reading & Mathematics	00	1 for each student
07	Student Labels (CRT-Alt)	03	Reading & Mathematics	00	1 for each student
08	Student Labels (CRT-Alt)	03	Reading Mathematics and Science	00	1 for each student
10	Student Labels (CRT-Alt)	03	Reading Mathematics and Science	00	1 for each student
03	Student Report (CRT-Alt)	02	Reading & Mathematics	00	1 for each student
04	Student Report (CRT-Alt)	02	Reading Mathematics and Science	00	1 for each student
05	Student Report (CRT-Alt)	02	Reading & Mathematics	00	1 for each student
06	Student Report (CRT-Alt)	02	Reading & Mathematics	00	1 for each student
07	Student Report (CRT-Alt)	02	Reading & Mathematics	00	1 for each student
08	Student Report (CRT-Alt)	02	Reading Mathematics and Science	00	1 for each student
10	Student Report (CRT-Alt)	02	Reading Mathematics and Science	00	1 for each student
00	Interp. Guide	04		00	1 per school

Appendix A

1. Items not available on the Braille form

Grade	*Subject*	*Form*	*Position*	*Reporting Category*
03	Rea	01	12	5
03	Rea	01	50	2
03	Rea	00	71	2
03	Mat	00	67	4
03	Mat	01	73	5

2. Common items removed from Grade 7 student using JAWS

Grade	*Subject*	*Form*	*Position*	*Reporting Category*
07	Rea	00	77	2
07	Mat	00	23	7
07	Mat	00	24	4

All field test items were removed.