Missouri

End-of-Course Assessments

Missouri Department of Elementary and Secondary Education

Technical Report
Phase I Assessments
2009–2010

English II
Algebra I
Biology

Developed and published by The Riverside Publishing Company, a division of Houghton Mifflin Harcourt, 3800 Golf Road, Suite 200, Rolling Meadows, IL 60008-4015. Copyright © 2010 by the Missouri Department of Elementary and Secondary Education. All rights reserved. Only Missouri state educators and citizens may copy and/or download and print the document, located online at http://dese.mo.gov/divimprove/assess/tech/index.html.

Table of Contents

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

LIST OF TABLES

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

List of Figures

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

List of Abbreviations

Below is a list of abbreviations that appear frequently in this technical report.

ALD ....Achievement-Level Descriptor
AMO ....Additional Materials Order
ARC ....Assessment Resource Center
AYP...................Adequate Yearly Progress
CCSSO.............Council of Chief State School Officers
CLE...................Course-Level Expectation
CMS ....Content Management System
CSEM................Conditional Standard Error of Measurement
CTT...................Classical Test Theory
DESE................Department of Elementary and Secondary Education
DIF ....Differential Item Functioning
DOK..................Depth of Knowledge
EFT ....Embedded Field Test
EOC...................End-of-Course
FRL ....Free and Reduced Lunch
FT......................Field Test
GLE...................Grade-Level Expectation
GRF...................General Research File
IAP ....Individualized Accommodation Program
IDEA.................Individuals with Disabilities Education Act
IEP.....................Individualized Education Program
IPASS................Image-Based Performance Assessment Scoring System
IRT ....Item Response Theory
ITS.....................Internet Testing Systems
LEP ....Limited English Proficient
LOSS................Lowest Obtainable Scale Score
MAP..................Missouri Assessment Program
MH ....Mantel-Haenszel procedure
MOSIS ....Missouri School Information System
MS ....Mean Square
NCLB................No Child Left Behind
PCA...................Principal Components Analysis
PE......................Performance Event
RIF ....Regional Instructional Facilitator
RPC...................Riverside Publishing Company
RSS ....Riverside Scoring Service®
SE......................Standard Error
SEM...................Standard Error of Measurement
SR......................Selected Response
TAC...................Technical Advisory Committee
TCC...................Test Characteristic Curve
TDS...................Test Development Specialist
WP.....................Writing Prompt

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

EXECUTIVE SUMMARY

This document provides a technical summary of the 2009–2010 administrations of the Missouri End-of-Course (MO EOC) Assessments in English II, Algebra I, and Biology. The criterion-referenced MO EOC Assessments are designed to assess students' knowledge of Missouri's Course-Level Expectations (CLEs) in these three content areas. The 2009–2010 school year marked the second operational administration of the assessments. For a technical summary of the first operational administration of the assessments, see the *2009 Missouri End-of-Course Assessments Technical Report*.

E.1 Background

In 1993, the Missouri legislature passed the Outstanding Schools Act (Senate Bill 380), requiring the Missouri State Board of Education to adopt challenging academic performance standards that define the skills and competencies necessary for students to successfully advance through the public school system, prepare for post-secondary education and the workplace, and participate as citizens in a democratic society. The Missouri State Board of Education formally adopted the academic standards known as the Show-Me Standards in January 1996.

In addition to mandating the development of rigorous academic standards, the Outstanding Schools Act of 1993 required the development and implementation of a comprehensive, primarily performance-based assessment program to measure student proficiency in the knowledge, skills, and competencies identified in the standards. Upon adoption of the standards in 1996, Missouri began developing the Missouri Assessment Program (MAP).

In January 2007, the Missouri State Board of Education approved a plan to replace the MAP for high school students with MO EOC Assessments beginning with English II, Algebra I, and Biology in the 2008–2009 school year.

E.2 Administration

The EOC Assessments are administered in three different assessment windows each year. Test windows are available for summer, fall, and spring, but reports are provided only after the spring testing window. Because the No Child Left Behind Act (NCLB) goal for every school in the state is Proficient, as defined by the Missouri State Board of Education, EOC testing is conducted as close as possible to the end of each course to allow school staff and students the greatest opportunity to achieve the goal of proficiency.

The scope of this technical report includes the Summer 2009, Fall 2009, and Spring 2010 assessments. Data analyses for the total assessed population, which include students who have not yet reached the secondary level, are based on a combination of assessment results as well as Missouri Department of Elementary and Secondary Education (DESE) demographic criteria as required under NCLB.

Paper score reports for the MO EOC Assessments are produced and distributed following each Spring administration. The score reports for the 2009–2010 assessment year contained information from the Summer 2009, Fall 2009, and Spring 2010 assessments.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

In future years, reports will continue to include information from the previous year's Summer, Fall, and Spring administrations.

E.3 Student Performance

A MO EOC Assessment score describes the relationship of student performance to a defined level of achievement. Achievement-level descriptors (ALDs) associated with each level provide details about the content expectations that students at that level meet or exceed. Missouri uses four achievement levels for the EOC Assessments: Below Basic, Basic, Proficient, and Advanced.

Table E.1 displays the percentage of students at each achievement level for the Summer 2009, Fall 2009, and Spring 2010 MO EOC Assessments. The NCLB Act requires states to assess all students at least once in high school in mathematics, English/communication arts, and science. Students who take the MO EOC Assessment but are not yet in high school are not included in Missouri's high school accountability data. Rather, their scores are "banked" until they actually reach high school, at which time they are rolled into the high school accountability data for that year. However, the data for *all* tested students are used each year for purposes of item analysis, scaling, and equating. For this reason, the numbers and/or percentages of tested students reported in Table E.1 and elsewhere in this technical report do not match the numbers of students reported by DESE for accountability purposes.

Table E.1 Percentage of Students at Each Achievement Level

Test Period	Achievement Level	English II	Algebra I	Biology
Summer 2009	Below Basic	10.5	21.1	20.2
	Basic	45.0	49.0	55.0
	Proficient	40.5	24.9	21.2
	Advanced	4.0	5.0	3.7
Fall 2009	Below Basic	1.6	8.4	8.8
	Basic	22.9	38.7	33.3
	Proficient	62.2	37.9	40.9
	Advanced	13.4	15.0	17.1
Spring 2010	Below Basic	3.0	6.2	6.2
	Basic	23.0	34.0	34.9
	Proficient	51.1	41.9	46.7
	Advanced	22.9	17.9	12.2

E.4 Evidence Supporting the Validity of Inferences from the MO EOC Assessment Scores

The MO EOC Assessments are part of an integrated program of testing, accountability, and curricular and instructional support. This technical report provides extensive detail about the development and operation of EOC Assessments. While a section of this report is devoted specifically to the documentation of validity evidence for the MO EOC Assessment scores, all information contained in the report ultimately contributes to the argument for the validity of the scores for their intended purposes.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

A summary of the information contained in this report follows.

Chapter 1: Introduction

Chapter 1 provides background information about MAP in general as well as some context for the MO EOC Assessments. Additionally, the chapter provides information about the organizational support provided by each contractor and subcontractor for the MO EOC Assessment program. The chapter ends with a statement of purpose for this technical report.

Chapter 2: Test Development

Chapter 2 contains thorough descriptions of each step in the development process for the MO EOC Assessments, including test design, test specifications and target point distributions, test blueprints, item writing, content and bias review procedures, test form assembly, and statistical item review. The evidence provided in this chapter is important to the content-related validity of the MO EOC Assessment scores. Additionally, the chapter covers principles of universal design and outlines the quality control processes employed throughout the test development process.

Chapter 3: Achievement-Level Setting

Chapter 3 details each step in the planning and execution of the 2009 standard-setting event that resulted in the cut scores for each of the MO EOC achievement levels. While this chapter was included in the *2009 Missouri End-of-Course Assessments Technical Report*, it is repeated here for the convenience of the reader because the results are relevant to the current test administrations. Chapter 3 covers selection of participants, development of ALDs, an overview of the methodology and considerations for the data that were available at the time of the standard-setting event, detailed information about each step in the process, and standard-setting results. Additionally, the chapter contains many appendices with examples of the materials that participants used during the standard-setting event.

Chapter 4: Item Analysis

Chapter 4 contains summary information, including item difficulty and discrimination indices, at the item level for each content area. The chapter also contains information on omit rates for the Summer 2009, Fall 2009, and Spring 2010 operational items, as well as differential item functioning (DIF) analyses performed on the Spring 2009 field-test item data.

Chapter 5: Test Administration

Chapter 5 contains information about the paper-and-pencil and online administration of the MO EOC Assessments, beginning with a description of students for whom the assessments are appropriate. Following this, the details of the administration are summarized. This summary includes a description of how the materials are distributed and how Test Examiners are trained, as well as information about the organization of the assessments, preparation of students to take the assessments, and directions for administration. Next, the chapter includes information about the accommodations

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

allowed on the MO EOC Assessments. Finally, the chapter briefly describes how materials are submitted for processing and scoring.

Chapter 6: Scanning, Scoring, and Quality Control Procedures

Chapter 6 covers the processes involved with scanning, scoring, and controlling the quality of the resulting score information for both the selected response and Performance Event/Writing Prompt (PE/WP) items on the MO EOC Assessments. The first part of Chapter 6 addresses the selected response items. The chapter contains detailed information on how the Riverside Publishing Scoring Service® (RSS) prepared for processing the MO EOC selected response items, including a check of scanning procedures prior to receipt of materials. Next, it details how the materials were handled from the time they were received and processed at the RSS on through to report generation.

The second part of Chapter 6 relates to the scanning, scoring, and quality control procedures undertaken by the Assessment Resource Center (ARC) for the PE/WP items. Information includes a description of range-finding activities and scoring materials development, project staffing and training, qualification of scorers, scoring procedures, and monitoring for quality assurance. Also included are the results of the inter-rater reliability study.

Chapter 7: Scaling and Equating

Chapter 7 begins with an introduction to the item response theory (IRT) model used for the scaling and equating of the MO EOC Assessments. The actual scaling and equating procedures are described in detail, including the calibration of the 2008 standalone field-test items, steps undertaken to establish a base scale for the MO EOC Assessments, examination of the stability of the linking items, steps taken to recenter the 2008 item bank, and steps taken to bring Spring 2009 field-test items onto the base scale. This chapter also includes information about the Summer 2009, Fall 2009, and Spring 2010 operational forms, a description of the IRT model assumptions and evidence of data-to-model fit, and a description of a post-equating check procedure.

Chapter 8: Reporting

Chapter 8 contains information about the reports Riverside Publishing produced for the MO EOC Assessments, including the Individual Student Report and Student Score Label. A brief summary of state-produced reports is also included.

Chapter 9: Summary Statistics

Chapter 9 provides descriptive statistics for raw scores and scale scores for the MO EOC Assessments. Raw score statistics are summarized by test administration, content area, and cluster. Scale score statistics are summarized for each content area and are also broken down by gender, ethnicity, migrant status, free and reduced lunch (FRL), limited English proficiency, Title I, individualized education program, and accommodations.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Chapter 10: Reliability

Chapter 10 begins by defining reliability and providing an overview of reliability estimation techniques. Raw-score internal consistency reliability coefficients are presented for all students and for each demographic group. Conditional standard errors of measurement (*CSEM*s) are presented at each scale-score cut point. Finally, the section provides information about the inter-rater reliability for the scoring of the PE/WP operational items in Summer 2009, Fall 2009, and Spring 2010, as well as for the Spring 2009 field-test PE/WP items.

Chapter 11: Validity

Chapter 11 provides evidence supporting the validity of the MO EOC Assessments for their intended purposes. After a brief introduction to the validity evidence for the MO EOC Assessments, the chapter documents more specific evidence related to test content, the internal structure of the assessments, and other types of validity evidence proposed by the *Standards for Educational and Psychological Testing* (AERA, APA, and NCME 1999). The chapter summarizes and reiterates validity evidence presented in earlier chapters in addition to providing new information not presented elsewhere. It provides a thorough argument supporting the validity of the MO EOC Assessments for measuring Missouri students' mastery of the CLEs, for identifying students' strengths and weaknesses, for serving as a basis for evaluating accountability plans, and for program evaluation.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 1: INTRODUCTION

1.1 History of Missouri's End-of-Course Assessments

In 1993, the Missouri legislature passed the Outstanding Schools Act (Senate Bill 380), requiring the Missouri State Board of Education to adopt challenging academic performance standards defining the skills and competencies necessary for students to successfully advance through the public school system, prepare for post-secondary education and the workplace, and participate as citizens in a democratic society. The Missouri State Board of Education formally adopted the academic standards known as the Show-Me Standards in January 1996.

These 73 standards are organized around four broad goals that address application, communication, problem solving, and responsible decision making. Thirty-three process standards emphasize the importance of engaging students of all ages in hands-on, active learning and integrating practical, challenging learning across all content areas. An additional 40 content standards define the academic skills and knowledge that provide the foundation for student learning in six content areas: Communication Arts, Mathematics, Science, Social Studies, Fine Arts, and Health/Physical Education. Content standards serve as the vehicle through which students demonstrate proficiency in the broader process standards. The Show-Me Standards are available for review on the Missouri Department of Elementary and Secondary Education's (DESE) website at http://dese.mo.gov/standards/index.html.

In 2001, DESE developed Grade-Level Expectations (GLEs) to assist districts in articulating the Show-Me Standards across grade levels and content areas. GLEs have been developed for Mathematics, Communication Arts, Science, Social Studies, Physical Education, Health, Music, Visual Arts, and Theater. GLEs are available for review on the DESE website at http://dese.mo.gov/divimprove/curriculum/GLE/index.html.

In addition to mandating the development of rigorous academic standards, the Outstanding Schools Act of 1993 also required the development and implementation of a comprehensive assessment program to measure student proficiency in the knowledge, skills, and competencies identified within the standards. Upon adoption of the standards in 1996, Missouri began developing the Missouri Assessment Program (MAP) in collaboration with the statewide assessment contractor, CTB/McGraw-Hill.

The Missouri State Board of Education adopted the purposes listed below to serve as guiding principles for development of the MAP:

- Improving students' acquisition of important knowledge, skills, and competencies
- Monitoring the performance of Missouri's educational system
- Empowering students and their families to improve their educational prospects
- Supporting the teaching and learning process

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

The first MAP assessments administered to students statewide were grade-span Mathematics assessments in grades 4, 8, and 10 in spring 1998. A voluntary grade-span Communication Arts Assessment for students in grades 3, 7, and 11 was also administered in spring 1998 and became mandatory in spring 1999. Voluntary Science and Social Studies grade-span assessments (grades 3, 7, and 10, and grades 4, 8, and 11, respectively) were added to the program in subsequent years. A voluntary Health/Physical Education assessment was available in 2000, and a Fine Arts assessment was field tested in 2001. Budget constraints prevented Science, Social Studies, and Health/Physical Education assessments from being added to the required components of the assessment program. Likewise, lack of funding prevented the completion of Fine Arts assessment development.

Through the spring 2005 administration, the MAP statewide assessment program included grade-span tests in the following grade levels/subject areas:

- Mathematics at grades 4, 8, and 10
- Communication Arts at grades 3, 7, and 11
- Science at grades 3, 7, and 10
- Social Studies at grades 4, 8, and 11 (Districts had the opportunity to administer grade-span science and social studies assessments voluntarily at the designated grade levels.)

All MAP assessments included three types of items: selected response (SR), constructed response (CR), and performance events (PE). For all content areas, MAP assessments included selected response items from the TerraNova Survey Edition. Constructed response items and performance events were custom-developed with significant input from Missouri educators.

During the initial MAP development/implementation period, DESE developed two to four equivalent forms for each content area/grade level assessment, using the first form for a voluntary testing cycle and administering the next form(s) in subsequent years. Early in the development phase, DESE tried out new items using separate field tests that usually occurred in the fall of the school year. As the program continued, each test form contained embedded field test items. Small-scale pilots continued as well.

As each content area/grade level assessment was administered, DESE used the Bookmark approach to set achievement levels, defining student performance through spring 2005 as Advanced, Proficient, Nearing Proficiency, Progressing, or Step 1.

After nearly a decade of MAP administration, new federal and state legislation prompted change in the program. To comply with requirements of No Child Left Behind (NCLB) legislation, Missouri's assessment program needed to incorporate Mathematics and Communication Arts assessments at all elementary and middle school grade levels (grades 3 through 8) and at one high school grade level. As a result, new grade-level assessments were developed for both content areas. These assessments were administered for the first time in spring 2006.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Additional NCLB requirements necessitated the addition of a mandatory science assessment once in the elementary grade range, once in the middle school grade range, and once in the high school grade range, beginning in spring 2008. The voluntary Science assessment in grades 3, 7, and 10 became a requirement and was moved to grades 5, 8, and 11. The voluntary Social Studies MAP Assessment was eliminated following the spring 2007 administration. Missouri's assessment system changed further in 2008–2009, when high school content area MAP assessments were replaced by End-of-Course (EOC) Assessments.

1.2 Brief Description of Missouri's Current Assessment System

The current MAP system includes the following assessment components for elementary and middle school:

- Grades 3–8 Communication Arts
- Grades 3–8 Mathematics
- Grades 5 and 8 Science

The EOC Assessments administered in 2009–2010 included:

- English II
- Algebra I
- Biology

In addition, the statewide assessment program currently includes the Missouri Assessment Program–Alternate (MAP-A) for students with severe cognitive disabilities, the LAS Links for English Language Learners (ELL), and a Personal Finance assessment for high school students who do not enroll in a personal finance course or who are receiving personal finance credit for embedded coursework.

1.3 Summary of the MO EOC Assessments

In response to feedback from Missouri districts regarding large-scale assessments for high school, the Missouri End-of-Course (MO EOC) Assessments were developed and first administered in 2008 for English II, Algebra I, and Biology. The MO EOC Assessments were created to address the needs of Missouri districts, schools, teachers, and students, while also meeting state and federal requirements. The Missouri State Board of Education identified the following purposes for the Missouri EOC Assessments:

- Measuring and reflecting students' mastery toward post-secondary readiness
- Identifying students' strengths and weaknesses
- Communicating expectations for all students
- Serving as the basis for state and national accountability plans
- Evaluating programs

Course-Level Expectations (CLEs) outline the ideas, concepts, and skills that form the foundation for an assessed EOC subject area, regardless of student grade level. Because a course such as Algebra I could be delivered at any grade level, CLEs replace the GLEs. This replacement is necessary because each EOC Assessment is more specific and

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

tailored to each EOC subject area. Districts can offer courses with different titles that cover the same CLEs.

Each MO EOC Assessment includes two types of test items: selected response items and Performance Events (PEs), which include Writing Prompts (WPs). An SR item presents students with a question followed by four response options. The PE items are constructed response items that require students to perform more complicated work. A PE often allows more than one approach to arrive at a correct response. The advantage of this type of item is that it provides insight into a student's ability to apply knowledge and understanding in real-life situations. The WP, a special type of PE that appears in the English II EOC Assessment, is an open-ended item that requires students to demonstrate their writing proficiency.

The MO EOC Assessments are offered in both paper-and-pencil and online administration modes.

1.4 Testing, Reporting, and Accountability

Evidence of students' progress in meeting the Show-Me Content Standards/CLEs is obtained from the MO EOC Assessments. These assessments provide the data that DESE uses to inform students, parents, the public, and the state legislature about students' performance, to help make informed decisions about educational issues, and to drive student services throughout the state.

The MO EOC Assessment reports provide useful information for determining the performance of students in a particular school and classroom. These reports help identify students who are below Proficient in a particular test area so that the school may determine a course of action that will meet the students' specific needs. Additionally, districts may use locally designed assessments, aligned to the Show-Me Content Standards/CLEs, to provide more detailed information for each student in specific test areas.

Testing for the MO EOC Assessments is conducted during three state-designated windows each year. Test windows are available for Summer, Fall, and Spring. Per contract requirements, however, paper reports for all administrations are provided only after the Spring testing window each year. (To aid in course grading, teachers may use an online interface to access student raw scores for the selected response items and to score student Performance Events shortly after the district's testing materials have been returned for processing.) Because the NCLB goal for every school in the state is Proficient, as defined by the Missouri State Board of Education, MO EOC testing is conducted as close as possible to the end of each course to allow school staff and students the greatest opportunity to achieve that goal.

Data for this technical report were collected during the Summer 2009, Fall 2009, and Spring 2010 operational administrations. Data analyses for the total assessed population, which includes students who have not yet reached the secondary level, are based on a combination of assessment results as well as DESE-provided demographic criteria required under NCLB.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

1.5 MO EOC Assessments Organizational Support

DESE coordinates the development and implementation of the MO EOC Assessments. In addition to planning, scheduling, and directing all EOC activities, the staff is extensively involved in numerous test reviews, security, and quality assurance procedures. Riverside Publishing is the primary contractor working in partnership with Questar, the Assessment Resource Center (ARC), Internet Testing Systems (ITS), Bookette, and others. The main activities for each of these groups are outlined in Table 1.1.

Table 1.1: Main Activities for Groups Involved in MO EOC Organizational Support

Group	Responsibilities
Riverside Publishing	• Provides program management, including primary contact with DESE; coordinates all meetings; handles all administrative costs/activities; generates all program management reports and status reports • Works with DESE to develop items with Missouri educators • Creates *Test Coordinator's Manual*, *Test Examiner's Manuals*, and other ancillary materials • Facilitates all review meetings with Missouri teachers and DESE • Conducts all psychometric analyses, reporting, linking/equating studies, and associated tasks, including participating in achievement-level setting • Provides all needed prepress work for program materials through camera-ready art • Produces all materials, including online, paper-and-pencil, Braille, and Large Print versions of test • Accounts for secure test books received after testing • Provides a direct customer service line, including technical support and general support to the program and customer interactions • Stores materials after testing • Participates in and presents at TAC meetings • Scores all selected response items • Produces and distributes all score reports and the *Guide for Interpreting Results* • Completes the technical report for DESE • Completes additional research studies
Questar	• Provides online enrollment and pre-ID system for use by Missouri districts • Packages and distributes all materials • Barcodes test books with security IDs • Leads facilitation and planning of achievement-level setting and provides members for the achievement-level-setting team • Contributes to the technical report • Participates in meetings with DESE, contributes to status reports, etc.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 1.1: Main Activities for Groups Involved in MO EOC Organizational Support (continued)

Group	Responsibilities
Assessment Resource Center (ARC)	• Receives and scans test books containing student responses to Performance Events and Writing Prompts (English II, Algebra I, and Biology only) • Scores the operational and field -test Performance Events and Writing Prompts • Develops training materials for Performance Event and Writing Prompt scoring • Provides scoring rubrics, anchor papers, annotated instructions, and practice papers to Bookette for software training development • Provides facilities for item writing if contracted by DESE • Contributes to the technical report • Stores materials after testing • Participates in meetings with DESE, contributes to status reports, etc.
Internet Testing Services (ITS)	• Sets up a Missouri DESE-branded website for access to the online testing system • Provides the online test delivery of one complete form for each administration for the following content areas: English II, Algebra I, and Biology beginning in 2008, and Integrated Math II, Integrated Math III, Geometry, Algebra II, English I, American History, and Government beginning in 2009 • Provides system documentation for test administrators and the DESE website • Provides technical support from 8 A.M. to 6 P.M., Monday through Friday, for the Riverside Publishing help desk • Produces and hosts practice tests for the English II, Algebra I, and Biology content areas • Provides online tools for graphing and table creation/editing and provides an equation editor • Offers ruler and reference sheets in tests • Provides three administrations per contract year in fall, spring, and summer for all content areas • Supplies a data feed of results from ITS to Riverside Publishing • Transfers student images from the Phase I Session II testing events for the teacher interface and for ARC to score
Bookette	• Provides a web-based interactive, software-based tutorial to help teachers learn how to score PE and WP items • Provides customer support as needed
Districts	• Distribute materials to the school buildings, track all secure materials, and promptly return all materials, including answer documents, for scoring • Assist in the timely resolution of scoring alerts • Act as liaison between Riverside Publishing and buildings

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 1.1: Main Activities for Groups Involved in MO EOC Organizational Support (continued)

Group	Responsibilities
School Buildings	• Administer tests, track all secure materials, and promptly return materials to districts for scoring
RR Donnelly	• Prints all nonscannable testing materials
Techniforms	• Prints all scannable test books and answer documents
Region IV, 3X	• Prints Braille and Large Print versions, respectively

1.6 Purpose of the Technical Report

The purpose of this technical report is to provide information about the technical characteristics of the 2009 embedded field-test administration and 2009–2010 operational administration of the MO EOC Assessments. Because this report is technical in nature and the intended audience is psychometric and educational research experts, it is best understood with a working knowledge of measurement concepts such as reliability and validity and statistical concepts such as correlation and central tendency. For some chapters, the reader is presumed to have basic familiarity with advanced topics in measurement and statistics such as item response theory (IRT).

This technical report provides extensive detail about the development and operation of the MO EOC Assessments. The empirical reliability of the assessments and validity of intended uses of the scores are reported explicitly in this document. While Chapter 10: Reliability is relatively straightforward, the steps in creating and operating the program are all aspects of validity, which is discussed in Chapter 11. The validity of score use and interpretation for any assessment stems from the statement of the test's purpose and the intended use of the scores; the steps taken in designing the test; and the processes of developing the content of the test, consulting with stakeholders, communicating about the test to users, scoring and reporting, and data analysis. The careful documentation of each of these steps is a necessary piece of a comprehensive, defensible validity argument for the intended uses of the assessment scores. In short, while there is a specific chapter devoted to validity, other parts of this document provide evidence necessary to assess the validity of the MO EOC Assessment scores for their intended purposes.

In reading this technical report, it is critical to remember that the testing program does not exist in a vacuum; it is not just a test. It is one part of a complex network intended to help schools focus their energies on improving student learning. The MO EOC Assessment is an integrated program of testing and accountability as well as curricular and instructional support. It can be evaluated properly only within its full context.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 2: TEST DEVELOPMENT

2.1 Introduction

The English II, Algebra I, and Biology End-of-Course (EOC) Assessments were first administered operationally during the 2008–2009 school year. This chapter provides an overview of the development of the Missouri End-of-Course (MO EOC) Assessments, including the test specifications, item development, item review, and test forms development. Forms development is described for the following administrations: Summer 2009, Fall 2009, Spring 2010, and Summer 2010. According to the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999) (hereafter referred to as the Standards), "Important validity evidence can be obtained from an analysis of the relationship between a test's content and the construct it is intended to measure" (p. 11). Accordingly, the thorough descriptions of the test development procedures included in this chapter provide evidence to support the construct validity of the MO EOC Assessments.

2.2 Design of the MO EOC Assessments

Figure 2.1 details the design of the Spring 2008 standalone field test, Fall 2008 operational administration, and Spring 2009 operational administration with embedded field test for English II. Additionally, Figure 2.1 displays the design of the Spring 2009 standalone Writing Prompt (WP) field test.

Figure 2.2 details the design of the Spring 2008 standalone field test, Fall 2008 operational administration, and Spring 2009 operational administration with embedded field test for Algebra I and Biology.

Figure 2.3 details the design of the linking forms for the 2009–2010 administration year. The Fall 2009 operational administration was linked to both the Spring 2009 and the Spring 2010 operational administrations. Additionally, the Spring 2010 administration was linked to the Summer 2010 administration. Besides being linked to the Spring 2010 administration, the Summer administration was also linked to the Spring 2009 administration.

2.2.1 Spring 2008 Standalone Field Test

The Spring 2008 standalone field test provided item data to inform the 2008–2009 operational forms selection process. There were two sessions in the Spring 2008 field test. For each assessment, Session I included 10 unique forms of selected response (SR) items, with each form containing 19 items. Session II included 10 unique forms of Performance Event/Writing Prompt (PE/WP) items. For English II, the PE forms in Session II each consisted of one 4-point Writing Prompt. For Algebra I, each Session II form consisted of one 4-point PE. For Biology, each Session II form consisted of 10–12 constructed response (CR) items, for a total of 20 points on each form. Forms within each session were spiraled at the student level across the state.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

2.2.2 Fall 2008 Operational Administration

The Fall 2008 administration consisted of three operational assessments. English II consisted of one 35-item SR form and one WP. Algebra I and Biology consisted of one 35-item SR form and one PE form each. The Algebra I PE consisted of one 4-point item. The Biology PE consisted of 10 CR items, ranging from 1 to 4 points each, for a total of 20 points.

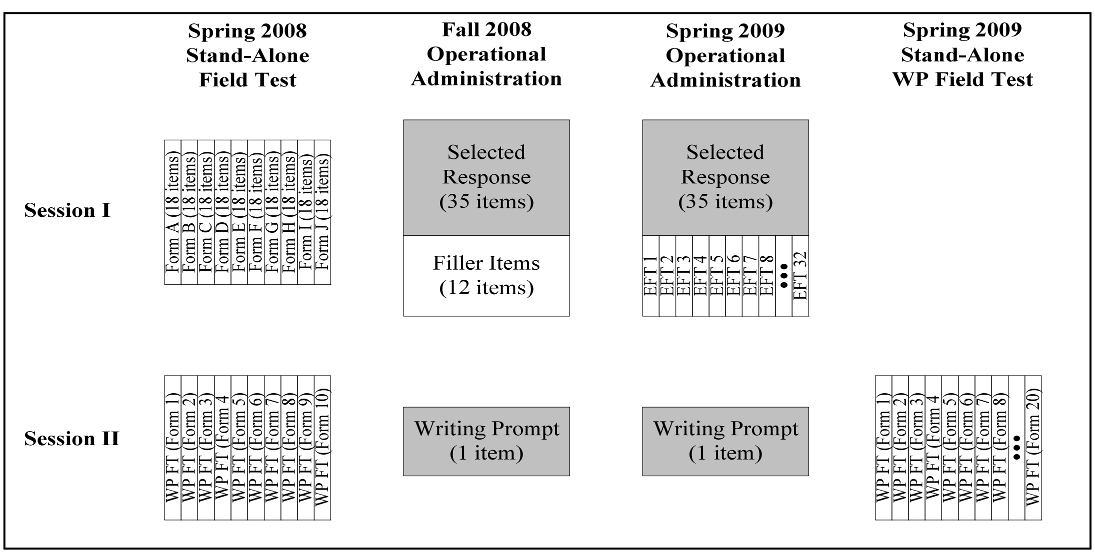
In addition to the 35 scored items, each Fall 2008 Algebra I and Biology Session I test book contained a set of 12 linking items (designated as M1 in Figures 2.2 and 2.3). These linking items were used for the post-equating check of the pre-equating results following the Spring 2009 operational administration (indicated by an arrow in Figure 2.2). For the English II assessment, the 12 additional SR items in Session I were filler (nonscored) items.

2.2.3 Spring 2009 Operational Administration

The Spring 2009 assessments consisted of three operational assessments. For all three content areas, Session I consisted of 35 operational SR items. For English II, there were 32 unique sets of 12 embedded field test items (labeled as EFT 1 through EFT 32 in Figure 2.1). For Algebra I and Biology, there were 24 unique sets of 12 embedded field test items (labeled as EFT 1 through EFT 24 in Figure 2.2). Additionally, the sets of items used to link the Spring 2009 form to the Fall 2008 (M1) and Summer 2009 (M2) assessments for the post-equating check occupied two of the embedded field test slots on the Algebra I and Biology assessments.

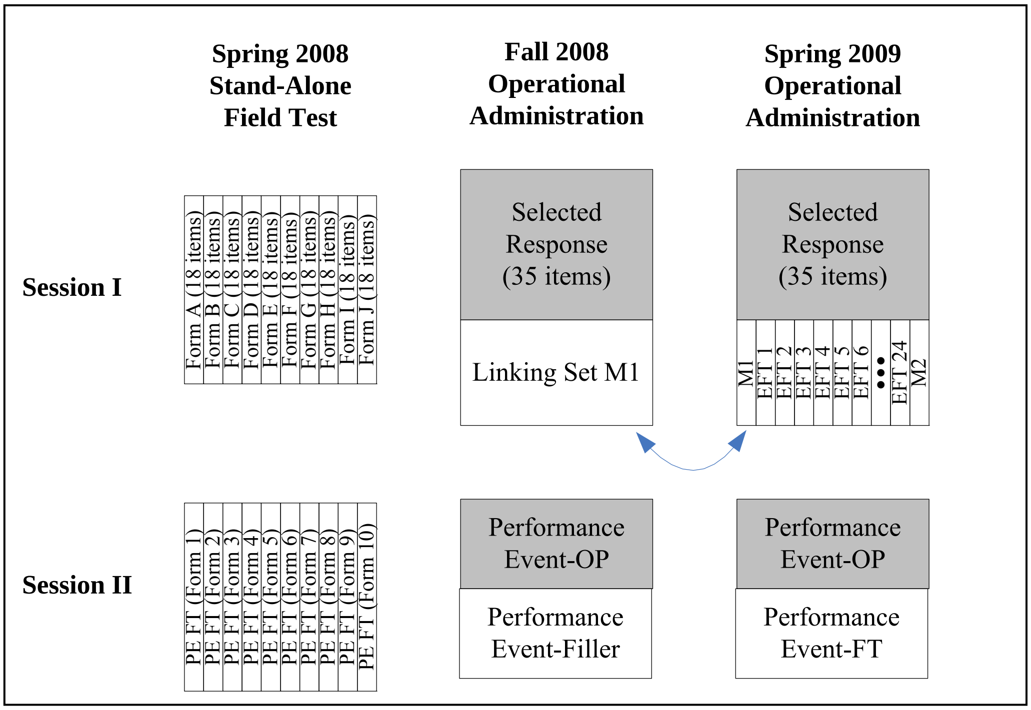
Session II of the Spring 2009 English II assessment contained one 4-point WP item. Session II of the Algebra I assessment contained one 4-point PE. Finally, Session II of the Biology assessment contained 11 CR items, ranging from 1 to 4 points each, for a total of 20 points. Session II of the Algebra I and Biology assessments also contained an embedded field test PE.

Figure 2.1: Field-Test and Operational Assessment Design, English II



Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 2.2: Field-Test and Operational Assessment Design, Algebra I and Biology



2.2.4 Spring 2009 Standalone WP Field Test

There was a separate standalone WP field test in which 20 WPs were administered statewide, each on its own form.

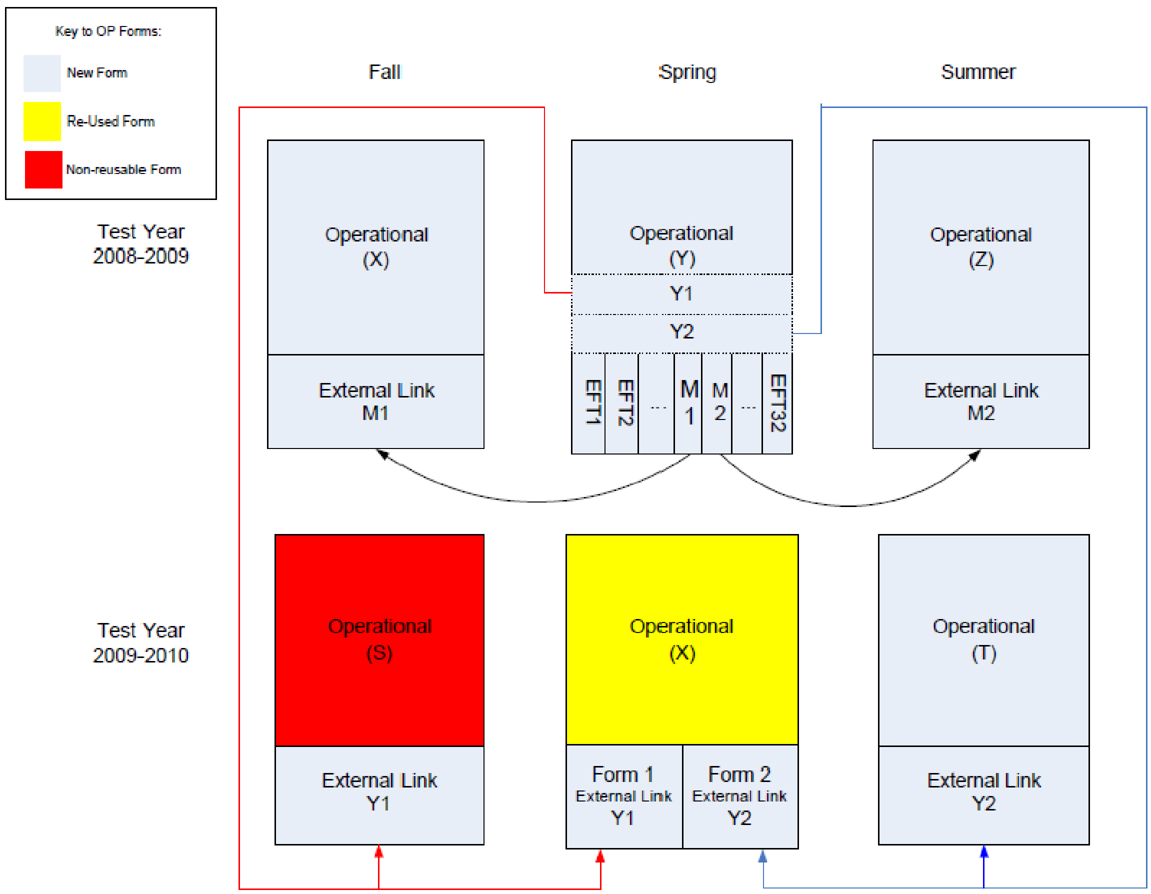
In Figures 2.1 and 2.2, the shaded portions designate the operational (scored) items.

2.2.5 Released Forms

In addition to the operational forms that were constructed for 2008–2009, the Missouri Department of Elementary and Secondary Education (DESE) and Riverside Publishing also worked together to construct “released” forms for each operational assessment. These forms were posted on the DESE website in August 2008. They were constructed to mirror the test content of the actual operational forms (minus the EFT items) to allow Missouri teachers, parents, and students the opportunity to review the new format and representative content of the EOC Assessments. Although these forms were constructed to parallel the operational forms, the items in these released forms were never used on an operational EOC Assessment.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 2.3: Linking Design for Post Equating the Missouri EOC Assessments



2.2.6 Fall 2009 Operational Administration

The Fall 2009 assessments consisted of three operational assessments. For all three assessments, Session I consisted of 35 operational SR items. Each form also contained 12 linking items used to equate the Fall 2009 forms to the Spring 2009 forms as shown in Figure 2.3.

Session II of the Fall 2009 English II assessment contained one 4-point WP item. Session II of the Algebra I assessment contained one 4-point PE. Finally, Session II of the Biology assessment contained 12 CR items, ranging from 1 to 4 points each, for a total of 20 points.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

2.2.7 Spring 2010 Operational Administration

The Fall 2008 operational form was reused for the Spring 2010 operational SR assessment. There were two forms each of the English II, Algebra I, and Biology assessments, which contained 12 linking items. As shown in Figure 2.3, there were two unique sets of linking items per course. Set Y1 linked the Spring 2010 Form A to the Fall 2009 and the Spring 2008 forms. Set Y2 linked the Spring 2010 Form B to the Spring 2008 and the Summer 2009 forms.

Session II of the Spring 2010 English II assessment contained one 4-point WP item. Session II of the Algebra I assessment contained one 4-point PE item. Finally, Session II of the Biology assessment contained 12 CR items, ranging from 1 to 4 points each, for a total of 20 points. Session II of the Algebra I and Biology assessments also contained an embedded field test PE.

2.2.8 Summer 2010 Operational Administration

The Summer 2010 assessments consisted of three operational assessments. For all three assessments, Session I consisted of 35 operational SR items. Each form also contained 12 linking items used to equate the Summer 2010 forms to the Spring 2009 and 2010 forms as shown in Figure 2.3.

Session II of the Fall 2009 English II assessment contained one 4-point WP item. Session II of the Algebra I assessment contained one 4-point PE item. Finally, Session II of the Biology assessment contained 12 CR items, ranging from 1 to 4 points each, for a total of 20 points. Session II of the Algebra I and Biology assessments also contained an embedded field test PE.

2.3 Test Blueprints

The test blueprint specifies the relative percentages of items in each high-level content strand. This document helps ensure that each strand is represented by the minimum number of points (8) for student score reports.

Riverside Publishing content experts worked with DESE to develop blueprints for each course before item writing began in fall 2007. Blueprint development was guided by the Missouri Show-Me Standards.

Tables 2.1 through 2.3 outline the test construction blueprints for English II, Algebra I, and Biology.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.1: Test Construction Blueprint for English II

Big Idea	Target # of Points	Point Range*	Target % Total	Minimum Emphasis	Maximum Emphasis
1. Develop and apply skills and strategies to the reading process	12	10–14	31%	**26%**	**36%**
2. Develop and apply skills and strategies to comprehend, analyze, and evaluate fiction, poetry, and drama	9	8–11	23%	**23%**	**28%**
3. Develop and apply skills and strategies to comprehend, analyze, and evaluate nonfiction	9	8–11	23%	**23%**	**28%**
Writing 1. Apply a writing process in composing text 2 Compose well-developed text 3 Write effectively in various forms and types of writing	9	8–9	23%	**23%**	**23%**
Total	39		100%		

Note: Total score points for each content strand may vary depending on which passages are selected for a particular administration. The percentage of total score points from each content strand (emphasis) will fall within the blueprint range described above.

*****The minimum number of points in each strand will be 8.

This blueprint was built under the following assumptions:

1. The operational test will be composed of two sessions. Session I will have thirty-five (35) 1-point selected response items, and Session II will have one (1) 4-point WP.
2. The reading passages will generally be balanced between nonfiction and fiction. A slight imbalance may occur if an odd number of passages appears on the operational test.
3. Content strand 1 has a larger percentage of total points because it can be assessed using both fiction and nonfiction passages.
4. The writing form/type will vary depending on the WP selected for a particular administration. Writing prompts will be aligned to a primary CLE; however, multiple writing CLEs may be assessed to reflect the holistic rubric.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.2: Test Construction Blueprint for Algebra I

Content Strand	Target # of Points	10% Tolerance	Point Range	Target % Total	Minimum Emphasis	Maximum Emphasis
Number and Operations	8	0.8	7–9	21%	**19%**	**23%**
Algebraic Relationships	23	2.3	21–25	58%	**53%**	**63%**
Geometric and Spatial Relationships*	0	0	0	0%	**0%**	**0%**
Measurement*	0	0	0	0%	**0%**	**0%**
Data and Probability	8	0.8	7–9	21%	**19%**	**23%**
Total	39			100%		

Note: Total score points for the operational tests may vary depending on which PE is selected for a particular administration. Regardless of the total score points on a particular operational test, the percentage of total score points from each content strand (emphasis) will fall within the blueprint range described above.

*These strands are not included on the EOC Assessment but are assessed locally.

This blueprint was built under the following assumptions:

1. The operational test will be composed of two sessions. Session I will have thirty-five (35) 1-point selected response items, and Session II will have one (1) 4-point PE item.
2. Each PE will be aligned to one CLE from the Algebraic Relationships strand.

Table 2.3: Test Construction Blueprint for Biology

Content Strand	Target # of Points	10% Tolerance	Point Range	Target % Total	Minimum Emphasis	Maximum Emphasis
Strand 1: Properties and Principles of Matter and Energy*	0	0	0	0%	**0%**	**0%**
Strand 2: Properties and Principles of Force and Motion*	0	0	0	0%	**0%**	**0%**
Strand 3: Characteristic and Interactions of Living Organisms	22	2.2	20–24	40%	**36%**	**44%**
Strand 4: Changes in Ecosystems and Interactions of Organisms with Their Environments	13	1.3	12–14	24%	**22%**	**27%**
Strand 5: Processes and Interactions of the Earth's Systems (Geosphere, Atmosphere, and Hydrosphere)*	0	0	0	0%	**0%**	**0%**
Strand 6: Composition and Structure of the Universe and the Motion of the Objects Within It*	0	0	0	0%	**0%**	**0%**
Strand 7: Scientific Inquiry	20	(2.0)	20	36%	**36%**	**36%**

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.3: Test Construction Blueprint for Biology (continued)

Content Strand	Target # of Points	10% Tolerance	Point Range	Target % Total	Minimum Emphasis	Maximum Emphasis
Strand 8: Impact of Science, Technology, and Human Activity	0	0	0	0%	**0%**	**0%**
Total:	55			100%		

Note: Total score points for the operational tests may vary depending on which PE prompts are selected for a particular administration. Regardless of the total score points on a particular operational test, the percentage of total score points from each content strand (emphasis) will fall within the blueprint range described above.

*These strands are not included on the EOC Assessment but are assessed locally.

This blueprint was built under the following assumptions:

1. The operational test will be composed of two sessions. Session I will have thirty-five (35) 1-point selected response items, and Session II will have one (1) 20-point performance task that is made up of a main context and several prompts.
2. Prompts within PEs will be aligned to CLEs from strand 7 only.

The actual 2008–2010 English II, Algebra I, and Biology point distributions for each form fell within the blueprint targets.

2.4 Test Specifications

Standard 1.6[1] specifically addresses the appropriateness of test content and its relationship to a solid validity argument. Additionally, Standard 3.3[2] defines “test specifications” and provides examples of the type of information that should be included in a specification document. The test specifications describe the content and format of the test and delineate the ideal number of items and points assessed for each Course-Level Expectation (CLE). This section details the development and use of the test specification documents for the MO EOC Assessments.

[1] **Standard 1.6:** When the validation rests in part of the appropriateness of test content, the procedures followed in specifying and generating test content should be described and justified in reference to the construct the test is intended to measure or the domain it is intended to represent. If the definition of the content sampled incorporates criteria such as importance, frequency, or criticality, these criteria should also be clearly explained and justified (p. 18).

[2] **Standard 3.3:** The test specifications should be documented, along with their rationale and the process by which they were developed. The test specifications should define the content of the test, the proposed number of items, the item formats, the desired psychometric properties of the items, and the item and section arrangement. They should also specify the amount of time for testing, directions to the test takers, procedures to be used for test administration and scoring, and other relevant information (p. 43).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

In 2007, Riverside Publishing content experts developed draft test specifications for each course. These draft test specifications were subsequently reviewed and approved by DESE. The specifications were finalized in fall 2007, before the development of items for field-test forms.

The test specification document served as the foundation for all test item development. The material in the test specifications was designed for use by Riverside Publishing content experts and DESE to construct tests containing the following items:

- Aligned to Missouri CLEs
- Aligned to Norman Webb's depth of knowledge (DOK) cognitive levels
- SR and PE/WPs
- Standalone and passage-based

Detailed descriptions of the test content measured in English II, Algebra I, and Biology are presented in the following sections.

2.4.1 English II

The English II MO EOC Assessment measures students' achievement in reading and writing. Session I of the test contains commissioned passages that comprise both fiction and nonfiction and cover a wide range of genres, including poems, short stories, newspaper articles, historical fiction, functional texts, and webpages. The questions associated with each passage are in SR format. There are 35 SR items on the English II Assessment. The English II EOC Assessment also contains standalone SR items that assess grammar and language usage. Session II of the English II EOC Assessment comprises a Writing Prompt, which could cover one or more of the following genres: narrative, expository, persuasive, or informative. The Writing Prompt is scored based on a holistic 4-point rubric.

Table 2.4 contains targets for the CLE point distribution on the English II operational forms. Some of the CLE point targets may not be met because the use of a passage or scenario is not conducive to items written to the CLE. Some Big Ideas are not represented in this chart because they are not assessed at this course level.

Tables 2.5–2.10 contain actual point distributions for the Fall 2008 through the Summer 2010 English II operational forms.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.4: Target Point Distributions for the English II Operational Forms

READING STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Develop and apply skills and strategies to the reading process	C. Phonics	*Apply decoding strategies to “problem-solve” unknown words when reading when needed		Assessed locally
	D. Fluency	*Read grade-level instructional text a. with fluency: accuracy, comprehension and appropriate expression b. adjusting reading rate to difficulty and type of text		Assessed locally
	E. Vocabulary	Develop vocabulary through text, using a. roots and affixes b. context clues c. glossary, dictionary, and thesaurus	2	3–4
	F. Prereading	*Apply prereading strategies to aid comprehension a. access prior knowledge b. preview c. predict with text support or rationale d. set a purpose and rate for reading		Assessed locally
	G. During Reading	*During reading, utilize strategies to a. determine meaning of unknown words b. self-monitor comprehension c. question the text d. infer e. visualize f. paraphrase g. summarize	2	Assessed locally
	H. Post Reading	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: *a. question to clarify *b. reflect c. draw conclusions d. paraphrase e. summarize	3	5–6
	I. Making Connections	Compare, contrast, analyze and evaluate connections: a. text to text (information and relationships in various fiction and nonfiction works) *b. text to self (text ideas and own experiences) *c. text to world (text ideas and the world by analyzing and evaluating the relationship between literature and its historical period and culture)	3	1–2

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.4: Target Point Distributions for the English II Operational Forms (continued)

READING STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
2. Develop and apply skills and strategies to comprehend, analyze, and evaluate fiction, poetry and drama from a variety of cultures and times	A. Text Features	Analyze and evaluate the text features in grade-level text	3	2–3
	B. Literary Techniques	Identify and explain literary techniques, emphasizing a. understatement b. parallelism c. allusion d. analogy e. analyze and evaluate literary techniques previously introduced	3	2–3
	C. Literary Elements	Use details from text(s) to a. demonstrate comprehension skills previously introduced b. analyze character, plot, setting, point of view c. analyze the development of a theme across genres d. identify and analyze tone	3	4–5
3. Develop and apply skills and strategies to comprehend, analyze, and evaluate nonfiction (such as biographies, newspapers, technical manuals) from a variety of cultures and times	A. Text Features	Explain, analyze, and evaluate the author's use of text features to clarify meaning	3	1–2
	B. Literary Techniques	Identify, explain, and analyze literary techniques in nonfiction, emphasizing a. understatement b. parallelism c. allusion d. analogy e. figurative language and sound devices previously introduced	3	2–3
	C. Text Structures	Use details from informational and persuasive text(s) to a. analyze and evaluate the organizational patterns b. identify and analyze faulty reasoning and unfounded inferences c. evaluate proposed solutions d. evaluate for accuracy and adequacy of evidence e. evaluate effect of tone on the overall meaning of work f. analyze and evaluate point of view g. analyze and evaluate author's viewpoint/perspective h. demonstrate comprehension skills previously introduced	3	3–4

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.4: Target Point Distributions for the English II Operational Forms (continued)

READING STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
	D. Understanding Directions	*Read and apply multistep directions to perform complex procedures and/or tasks	2	Assessed locally

WRITING STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Apply a writing process in composing text	A. Writing Process	*Apply a writing process to write effectively in various forms and types of writing (W3A)		Assessed locally
2. Compose well-developed text	A. Audience and Purpose	*Compose text: a. showing awareness of audience b. choosing a form and point of view appropriate to purpose and audience		Assessed locally
	B. Ideas and Content	*Compose text with a. strong controlling idea b. relevant specific details c. complex ideas d. freshness of thought		Assessed locally
	C. Organization and Sentence Structure	*Compose text with a. effective beginning, middle, and end b. a logical order c. effective paragraphing d. cohesive devices e. varied sentence structure f. clarity of expression g. active voice		Assessed locally
	D. Word Choice	*Compose text using a. precise and vivid language b. writing techniques such as imagery, humor, voice, and figurative language		Assessed locally
	E. Conventions	In written text, apply a. conventions of capitalization b. conventions of punctuation c. standard usage	1	5

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.4: Target Point Distributions for the English II Operational Forms (continued)

WRITING STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
3. Write effectively in various forms and types of writing	A. Forms/ Types/Modes of Writing	Compose a variety of texts a. using narrative, descriptive, expository, and/or persuasive features b. in various formats, including workplace communication c. including summary d. including literary analysis e. including reflective writing	3	4

* These CLEs are locally assessed.

Table 2.5: Actual Point Distributions for the Fall 2008 English II Operational Form

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**
Reading Process	12				11		11	
Reading (fiction)	9				8		8	
Reading (nonfiction)	9				11		11	
Writing	5	1		4	5	1	5	4
Total Items/Points	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Table 2.6: Actual Point Distributions for the Spring 2009 English II Operational Form

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**
Reading Process	12				11		11	
Reading (fiction)	9				10		10	
Reading (nonfiction)	9				9		9	
Writing	5	1		4	5	1	5	4
Total Items/Points	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Table 2.7: Actual Point Distributions for the Summer 2009 English II Operational Form

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**
Reading Process	12				10		10	
Reading (fiction)	9				11		11	
Reading (nonfiction)	9				9		9	
Writing	5	1		4	5	1	5	4
Total Items/Points	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.8: Actual Point Distributions for the Fall 2009 English II Operational Form

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**
Reading Process	12				12		12	
Reading (fiction)	9				9		9	
Reading (nonfiction)	9				9		9	
Writing	5	1		4	5	1	5	4
Total Items/Points	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Table 2.9: Actual Point Distributions for the Spring 2010 English II Operational Forms A and B

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**
Reading Process	12				12		12	
Reading (fiction)	9				9		9	
Reading (nonfiction)	9				9		9	
Writing	5	1		4	5	1	5	4
Total Items/Points	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Table 2.10: Actual Point Distributions for the Summer 2010 English II Operational Form

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**	**SR**	**WP**
Reading Process	12				10		10	
Reading (fiction)	9				11		11	
Reading (nonfiction)	9				9		9	
Writing	5	1		4	5	1	5	4
Total Items/Points	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

2.4.2 Algebra I

The Algebra I EOC Assessment measures students' ability to solve problems by applying mathematical concepts. There are three strands assessed on the Algebra I Assessment:

- Numbers and Operations
- Algebraic Relationships
- Data and Probability

The 35 selected response questions in Session I are aligned to the strands listed above. Session II contains a PE aligned to the Algebraic Relationships strand. The PE is a mathematical scenario in which the student is required to respond to several CR items. The student may be asked to construct a graph and/or provide equations. On some items, the student is required to show his or her work for full credit. The PE is worth a total of 4 points and is scored on an item-specific rubric.

Table 2.11 contains targets for the CLE point distribution on the Algebra I operational forms. Some Big Ideas are not represented in this table because they are not assessed at this course level. Tables 2.12–2.17 contain actual point distributions for the Fall 2008 through the Summer 2010 Algebra I operational forms.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.11: Target Point Distributions for the Algebra I Operational Forms

NUMBERS AND OPERATIONS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Understand numbers, ways of representing numbers, relationships among numbers, and number systems	A. Read, write, and compare numbers	Compare and order rational and irrational numbers, including finding their approximate locations on a number line	1	3–5
	B. Represent and use real numbers	Use real numbers and various models, drawing, etc. to solve problems	3	3–5
	C. Compose and decompose numbers	*Use a variety of representations to demonstrate an understanding of very large and very small numbers	2	Assessed locally
2. Understand meanings of operations and how they relate to one another	B. Describe effects of operations	*Describe the effects of operations, such as multiplication, division, and computing powers and roots on the magnitude of quantities	2	Assessed locally
	D. Apply operations on real and complex numbers	*Apply operations to real numbers, using mental computation or paper-and-pencil calculations for simple cases and technology for more complicated cases	2	Assessed locally
3. Compute fluently and make reasonable estimates	D. Estimate and justify solutions	*Judge the reasonableness of numerical computations and their results	3	Assessed locally
	E. Use proportional reasoning	*Solve problems involving proportions	2	Assessed locally

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.11: Target Point Distributions for the Algebra I Operational Forms (continued)

ALGEBRAIC RELATIONSHIPS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Understand patterns, relations, and functions	B. Create and analyze patterns	Generalize patterns using explicitly or recursively defined functions	2	1–2
	C. Classify objects and representations	Compare and contrast various forms of representations of patterns	3	1–2
	D. Identify and compare functions	Understand and compare the properties of linear and nonlinear functions	2	2–3
	E. Describe the effects of parameter changes	Describe the effects of parameter changes on linear, exponential growth/decay, and quadratic functions including intercepts	2	2–3
2. Represent and analyze mathematical situations and structures using algebraic symbols	A. Represent mathematical situations	Use symbolic algebra to represent and solve problems that involve linear and quadratic relationships, including equations and inequalities	3	2–3
	B. Describe and use mathematical manipulation	Describe and use algebraic manipulations, including factoring and rules of integer exponents, and apply properties of exponents, including order of operations, to simplify expressions	2	2–3
	C. Use equivalent forms	Use and solve equivalent forms of equations (linear, absolute value, and quadratic)	2	1–2
	D. Use systems	Use and solve systems of linear equations or inequalities with 2 variables	2	1–2
3. Use mathematical models to represent and understand quantitative relationships	A. Use mathematical models	Identify quantitative relationships and determine the type(s) of functions that might model the situation to solve the problem	2	3–4
4. Analyze change in various contexts	A. Analyze change	Analyze linear and quadratic functions by investigating rates of change, intercepts, and zeros	3	3–4

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.11: Target Point Distributions for the Algebra I Operational Forms (continued)

GEOMETRIC AND SPATIAL RELATIONSHIPS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Analyze characteristics and properties of two- and three-dimensional geometric shapes and develop mathematical arguments about geometric relationships	B. Apply geometric relationships	*Apply geometric properties such as similarity and angle relationship to solve multistep problems in two dimensions	2	Assessed locally
4. Use visualization, spatial reasoning, and geometric modeling to solve problems	B. Draw and use visual models	*Draw or use visual models to represent and solve problems	3	Assessed locally
MEASUREMENT STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
2. Apply appropriate techniques, tools, and formulas to determine measurements	D. Analyze precision	*Describe the effects of operations, such as multiplication, division, and computing powers and roots, on magnitudes of quantities and the effects of computation on precision, which include the judging of reasonable numerical computations and their results	2	Assessed locally
	E. Use relationships within a measurement system	*Use unit analysis to solve problems	2	Assessed locally
DATA AND PROBABILITY STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them	A. Formulate questions	Formulate questions and collect data about a characteristic, which include sample spaces and distributions	3	1–2
	C. Represent and interpret data	Select and use appropriate graphical representation of data and, given one-variable quantitative data, display the distribution and describe its shape	3	2–3

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.11: Target Point Distributions for the Algebra I Operational Forms (continued)

DATA AND PROBABILITY STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
2. Select and use appropriate statistical methods to analyze data	A. Describe and analyze data	Apply statistical measures of center to solve problems	2	2–3
	C. Represent data algebraically	Given a scatterplot, determine an equation for a line of best fit	2	1–2
3. Develop and evaluate inferences and predictions that are based on data	A. Develop and evaluate inferences	Make conjectures about possible relationships between 2 characteristics of a sample on the basis of scatterplots of the data	3	2–3

* These CLEs are locally assessed.

Table 2.12: Actual Point Distributions for the Fall 2008 Algebra I Operational Form

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**PE**	**SR**	**PE**	**SR**	**PE**	**SR**	**PE**
Number and Operation	8		8		8		8	
Algebraic Relationships	19	1	19	4	19	1	19	4
Data and Probability	8		8		8		8	
Total Items/Points	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Table 2.13: Actual Point Distributions for the Spring 2009 Algebra I Operational Form

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**PE**	**SR**	**PE**	**SR**	**PE**	**SR**	**PE**
Number and Operation	8		8		8		8	
Algebraic Relationships	19	1	19	4	19	1	19	4
Data and Probability	8		8		8		8	
Total Items	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Table 2.14: Actual Point Distributions for the Summer 2009 Algebra I Operational Form

	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
Reporting Categories	**SR**	**PE**	**SR**	**PE**	**SR**	**PE**	**SR**	**PE**
Number and Operation	8		8		8		8	
Algebraic Relationships	19	1	19	4	19	1	19	4
Data and Probability	8		8		8		8	
Total Items	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.15: Actual Point Distributions for the Fall 2009 Algebra I Operational Form

Reporting Categories	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
	SR	PE	SR	PE	SR	PE	SR	PE
Number and Operation	9		9		9		9	
Algebraic Relationships	17	1	17	4	17	1	17	4
Data and Probability	9		9		9		9	
Total Items	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Table 2.16: Actual Point Distributions for the Spring 2010 Algebra I Operational Forms A and B

Reporting Categories	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
	SR	PE	SR	PE	SR	PE	SR	PE
Number and Operation	9		9		9		9	
Algebraic Relationships	17	1	17	4	17	1	17	4
Data and Probability	9		9		9		9	
Total Items	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

Table 2.17: Actual Point Distributions for the Summer 2010 Algebra I Operational Form

Reporting Categories	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
	SR	PE	SR	PE	SR	PE	SR	PE
Number and Operation	9		9		9		9	
Algebraic Relationships	17	1	17	4	17	1	17	4
Data and Probability	9		9		9		9	
Total Items	**35**	**1**	**35**	**4**	**35**	**1**	**35**	**4**

2.4.3 Biology

The Biology EOC Assessment measures students' achievement in the following content and process strands:

- Characteristics and Interactions of Living Organisms
- Changes in Ecosystems and Interactions of Organisms with Their Environments
- Scientific Inquiry

The 35 SR questions in Session I are aligned to the first two strands listed above. Session II contains a PE aligned to the Scientific Inquiry strand, in which the student is required to respond to several open-ended items. The student may be asked to construct a data table, measure, and/or graph scientific results. Individual items within the PE may be worth 1, 2, 3, or 4 points and are scored on item-specific rubrics. The total point value of each operational PE is 20 points.

Table 2.18 is used as a target for the CLE point distribution for the Biology operational forms. Some Big Ideas are not represented in this table because they are not assessed at this course level. Tables 2.19–2.24 contain actual point distributions for the Fall 2008 through the Summer 2010 Biology operational forms.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms

CHARACTERISTICS AND INTERACTIONS OF LIVING ORGANISMS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. There is a fundamental unity underlying the diversity of all living organisms	B. Organisms progress through life cycles unique to different types of organisms	a. Recognize cells both increase in number and differentiate, becoming specialized in structure and function, during and after embryonic development	1	1–2
	C. Cells are the fundamental units of structure and function of all living things	b. Describe the structure of cell parts (e.g., cell wall, cell membrane, cytoplasm, nucleus, chloroplast, mitochondrion, ribosome, vacuole) found in different types of cells (e.g., bacterial, plant, skin, nerve, blood, muscle) and the functions they perform (e.g., structural support, transport of materials, storage of genetic information, photosynthesis and respiration, synthesis of new molecules, waste disposal) that are necessary to the survival of the cell and organism	2	1–2
2. Living organisms carry out life processes in order to survive	A. The cell contains a set of structures called organelles that interact to carry out life processes through physical and chemical means	c. Explain physical and chemical interactions that occur between organelles (e.g., nucleus, cell membrane, chloroplast, mitochondrion, ribosome) as they carry out life processes	2	1–2
	B. Photosynthesis and cellular respiration are complementary processes necessary to the survival of most organisms on Earth	a. Explain the interrelationship between the processes of photosynthesis and cellular respiration (e.g., recycling of oxygen and carbon dioxide), comparing and contrasting photosynthesis and cellular respiration reactions (Do NOT assess intermediate reactions.)	2	1–2
	B. Photosynthesis and cellular respiration are complementary processes necessary to the survival of most organisms on Earth	b. Determine what factors affect the processes of photosynthesis and cellular respiration (i.e., light intensity, availability of reactants, temperature)	2	1–2

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

CHARACTERISTICS AND INTERACTIONS OF LIVING ORGANISMS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
2. Living organisms carry out life processes in order to survive	D. Cells carry out chemical transformations that use energy for the synthesis or breakdown of organic compounds	*a. Summarize how energy transfer occurs during photosynthesis and cellular respiration as energy is stored in and released from the bonds of chemical compounds (i.e., ATP)	2	Assessed locally
	D. Cells carry out chemical transformations that use energy for the synthesis or breakdown of organic compounds	*b. Relate the structure of organic compounds (e.g., proteins, nucleic acids, lipids, carbohydrates) to their role in living systems		Assessed locally
	D. Cells carry out chemical transformations that use energy for the synthesis or breakdown of organic compounds	*d. Explain how protein enzymes affect chemical reactions (e.g., the breakdown of food molecules, growth and repair, regulation)	1	Assessed locally
	F. Cellular activities and responses can maintain stability internally while external conditions are changing (homeostasis)	a. Explain the significance of the selectively permeable membrane to the transport of molecules	2	1–2
	F. Cellular activities and responses can maintain stability internally while external conditions are changing (homeostasis)	b. Predict the movement of molecules across a selectively permeable membrane (i.e., diffusion, osmosis, active transport) needed for a cell to maintain homeostasis given concentration gradients and different sizes of molecules	2	1–2
	F. Cellular activities and responses can maintain stability internally while external conditions are changing (homeostasis)	c. Explain how water is important to cells (e.g., is a buffer for body temperature, provides a soluble environment for chemical reactions, serves as a reactant in chemical reactions, provides hydration that maintains cell turgidity, maintains protein shape)	2	1–2

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

CHARACTERISTICS AND INTERACTIONS OF LIVING ORGANISMS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
3. There is a genetic basis for the transfer of biological characteristics from one generation to the next through reproductive processes	B. All living organisms have genetic material (DNA) that carries hereditary information	a. Describe the chemical and structural properties of DNA (e.g., DNA is a large polymer formed from linked subunits of four kinds of nitrogen bases; genetic information is encoded in genes based on the sequence of subunits; each DNA molecule in a cell forms a single chromosome) (Assess the concepts; do NOT memorize the nitrogen base pairs.)	1	1–2
	B. All living organisms have genetic material (DNA) that carries hereditary information	b. Recognize the DNA codes for proteins, which are expressed as the heritable characteristics of an organism.	1	1–2
	B. All living organisms have genetic material (DNA) that carries hereditary information	a. Identify possible external causes (e.g., heat, radiation, certain chemicals) and effects of DNA mutations (e.g., altered proteins which may affect chemical reactions and structural development)	2	1–2
	C. Chromosomes are components of cells that occur in pairs and carry hereditary information from one cell to daughter cells and from parent to offspring during reproduction	a. Recognize the chromosomes of daughter cells, formed through the processes of asexual reproduction and mitosis, the formation of somatic (body) cells in multicellular organisms, are identical to the chromosomes of the parent cell	1	1–2
	C. Chromosomes are components of cells that occur in pairs and carry hereditary information from one cell to daughter cells and from parent to offspring during reproduction	b. Recognize that during meiosis, the formation of sex cells, chromosomes are reduced to half the number present in the parent cell	1	1–2

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

CHARACTERISTICS AND INTERACTIONS OF LIVING ORGANISMS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
3. There is a genetic basis for the transfer of biological characteristics from one generation to the next through reproductive processes	C. Chromosomes are components of cells that occur in pairs and carry hereditary information from one cell to daughter cells and from parent to offspring during reproduction	c. Explain how fertilization restores the diploid number of chromosomes	2	1–2
	D. There is heritable variation within every species of organism	a. Describe the advantages and disadvantages of asexual and sexual reproduction with regard to variation within a population	2	1–2
	D. There is heritable variation within every species of organism	*c. Recognize that new heritable characteristics can only result from new combinations of existing genes or from mutations of genes in an organism's sex cells	1	Assessed locally
	E. The pattern of inheritance for many traits can be predicted by using the principles of Mendelian genetics	b. Predict the probability of the occurrence of specific traits, including sex-linked traits, in an offspring by using a monohybrid cross	2	1–2

CHANGES IN ECOSYSTEMS AND INTERACTIONS OF ORGANISMS WITH THEIR ENVIRONMENTS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Organisms are interdependent with one another and with their environment	A. All populations living together within a community interact with one another and with their environment in order to survive and maintain a balanced ecosystem	a. Explain the nature of interactions between organisms in predator/prey relationships and different symbiotic relationships (i.e., mutualism, commensalism, parasitism)	1	1–3

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

CHANGES IN ECOSYSTEMS AND INTERACTIONS OF ORGANISMS WITH THEIR ENVIRONMENTS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Organisms are interdependent with one another and with their environment	A. All populations living together within a community interact with one another and with their environment in order to survive and maintain a balanced ecosystem	b. Explain how cooperative (e.g., symbiotic) and competitive (e.g., predator/prey) relationships help maintain balance within an ecosystem	2	1–2
	B. Living organisms have the capacity to produce populations of infinite size, but environments and resources are finite	a. Identify and explain the limiting factors (biotic and abiotic) that may affect the carrying capacity of a population within an ecosystem	2	1–3
	D. The diversity of species within an ecosystem is affected by changes in the environment, which can be caused by other organisms or outside processes	a. Predict the impact (beneficial or harmful) a natural environmental event (e.g., forest fire, flood, volcanic eruption, avalanche) or human caused change (e.g., acid rain, global warming, pollution, deforestation, introduction of an exotic species) may have on the diversity of different species in an ecosystem	2	1–2
	D. The diversity of species within an ecosystem is affected by changes in the environment, which can be caused by other organisms or outside processes	*b. Predict the impact (beneficial or harmful) a natural or human caused environmental event (e.g., forest fire, flood, volcanic eruption, avalanche, acid rain, global warming, pollution, deforestation, introduction of an exotic species) may have on the biodiversity of a community	2	Assessed locally
2. Matter and energy flow through the ecosystem	A. As energy flows through the ecosystem, all organisms capture a portion of that energy and transform it to a form they can use	c. Predict how the use and flow of energy will be altered due to changes in a food web	2	1–2

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

CHANGES IN ECOSYSTEMS AND INTERACTIONS OF ORGANISMS WITH THEIR ENVIRONMENTS STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
3. Genetic variation sorted by the natural selection process explains evidence of biological evolution	A. Evidence for the nature and rates of evolution can be found in anatomical and molecular characteristics of organisms and in the fossil record	*b. Evaluate the evidence that supports the theory of biological evolution (e.g., fossil records, similarities between DNA and protein structures, similarities between developmental stages of organisms, homologous and vestigial structures)	2	Assessed locally
	B. Reproduction is essential to the continuation of every species	*a. Define a species in terms of the ability to mate and produce fertile offspring	1	Assessed locally
	B. Reproduction is essential to the continuation of every species	b. Explain the importance of reproduction to the survival of a species (i.e., the failure of a species to reproduce will lead to extinction of that species)	1	1–2
	C. Natural selection is the process of sorting individuals based on their ability to survive and reproduce within their ecosystem	a. Identify examples of adaptations that may have resulted from variations favored by natural selection (e.g., long-necked giraffes, long-eared jack rabbits) and describe how that variation may have provided populations an advantage for survival	2	1–2
	C. Natural selection is the process of sorting individuals based on their ability to survive and reproduce within their ecosystem	c. Explain how environmental factors (e.g., habitat loss, climate change, pollution, introduction of non-native species) can be agents of natural selection	2	1–2

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

SCIENTIFIC INQUIRY STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Science understanding is developed through the use of science process skills, scientific knowledge, scientific investigation, reasoning, and critical thinking	A. Scientific inquiry includes the ability of students to formulate a testable question and explanation, and to select appropriate investigative methods in order to obtain evidence relevant to the explanation	a. Formulate testable questions and hypotheses	3	2–3
	A. Scientific inquiry includes the ability of students to formulate a testable question and explanation, and to select appropriate investigative methods in order to obtain evidence relevant to the explanation	b. Analyzing an experiment, identify the components (i.e., independent variable, dependent variables, control of constants, multiple trials) and explain their importance to the design of a valid experiment	3	2–4
	A. Scientific inquiry includes the ability of students to formulate a testable question and explanation, and to select appropriate investigative methods in order to obtain evidence relevant to the explanation	c. Design and conduct a valid experiment	4	2–6
	A. Scientific inquiry includes the ability of students to formulate a testable question and explanation, and to select appropriate investigative methods in order to obtain evidence relevant to the explanation	d. Recognize it is not always possible, for practical or ethical reasons, to control some conditions (e.g., when sampling or testing humans, when observing animal behaviors in nature)	2	0–1

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

SCIENTIFIC INQUIRY STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Science understanding is developed through the use of science process skills, scientific knowledge, scientific investigation, reasoning, and critical thinking	A. Scientific inquiry includes the ability of students to formulate a testable question and explanation, and to select appropriate investigative methods in order to obtain evidence relevant to the explanation	g. Evaluate the design of an experiment and make suggestions for reasonable improvements	3	1–2
	B. Scientific inquiry relies upon gathering evidence from qualitative and quantitative observations	b. Measure length to the nearest millimeter, mass to the nearest gram, volume to the nearest milliliter, force (weight) to the nearest Newton, temperature to the nearest degree Celsius, time to the nearest second	1	0–1
	B. Scientific inquiry relies upon gathering evidence from qualitative and quantitative observations	c. Determine the appropriate tools and techniques to collect, analyze, and interpret data	2	1–2
	B. Scientific inquiry relies upon gathering evidence from qualitative and quantitative observations	d. Judge whether measurements and computation of quantities are reasonable	2	1–2
	B. Scientific inquiry relies upon gathering evidence from qualitative and quantitative observations	e. Calculate the range, average/mean, percent, and ratios for sets of data	1	0–6

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

SCIENTIFIC INQUIRY STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Science understanding is developed through the use of science process skills, scientific knowledge, scientific investigation, reasoning, and critical thinking	C. Scientific inquiry includes evaluation of explanations (laws/principles, theories/models) in light of evidence (data) and scientific principles (understandings)	a. Use quantitative and qualitative data as support for reasonable explanations (conclusions)	3	1–2
	C. Scientific inquiry includes evaluation of explanations (laws/principles, theories/models) in light of evidence (data) and scientific principles (understandings)	b. Analyze experimental data to determine patterns, relationships, perspectives, and credibility of explanations (e.g., predict/extrapolate data, explain the relationship between the independent and dependent variable)	3	1–2
	C. Scientific inquiry includes evaluation of explanations (laws/principles, theories/models) in light of evidence (data) and scientific principles (understandings)	c. Identify the possible effects of errors in observations, measurements, and calculations, on the validity and reliability of data and resultant explanations (conclusions)	3	2–3
	C. Scientific inquiry includes evaluation of explanations (laws/principles, theories/models) in light of evidence (data) and scientific principles (understandings)	d. Analyze whether evidence (data) and scientific principles support proposed explanations (laws/principles, theories/models)	3	0–3

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

SCIENTIFIC INQUIRY STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
1. Science understanding is developed through the use of science process skills, scientific knowledge, scientific investigation, reasoning, and critical thinking	D. The nature of science relies upon communication of results and justification of explanations	A. Communicate the procedures and results of investigations and explanations through: • oral presentations • drawings and maps • data tables (allowing for the recording and analysis of data relevant to the experiment such as independent and dependent variables, multiple trials, beginning and ending times or temperatures, derived quantities) • graphs (bar, single, and multiple line) • equations and writings	3	4–8
	D. The nature of science relies upon communication of results and justification of explanations	c. Explain the importance of the public presentation of scientific work and supporting evidence to the scientific community (e.g., work and evidence must be critiqued, reviewed, and validated by peers; needed for subsequent investigations by peers; results can influence the decisions regarding future scientific work)	2	0–1

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.18: Target Point Distributions for the Biology I Operational Forms (continued)

IMPACT OF SCIENCE, TECHNOLOGY AND HUMAN ACTIVITY STRAND				
Big Idea	**Concept**	**CLE**	**DOK Limit**	**Range of Points per CLE on the Operational Test**
2. Historical and cultural perspectives of scientific explanations help to improve understanding of the nature of science and how science knowledge and technology evolve over time	B. Scientific theories are developed based on the body of knowledge that exists at any particular time and must be rigorously questioned and tested for validity	*a. Identify and describe how explanations (laws/principles, theories/models) of scientific phenomena have changed over time as a result of new evidence (e.g., cell theory, theories of spontaneous generation and biogenesis, theories of extinction, evolution theory, structure of the cell membrane, genetic theory of inheritance)	2	Assessed locally
	B. Social, political, economic, ethical and environmental factors strongly influence, and are influenced by, the direction of progress of science and technology	*e. Analyze and evaluate the drawbacks (e.g., design constraints, unintended consequences, risks), benefits, and factors (i.e., social, political, economic, ethical, and environmental) affecting progress toward meeting major scientific and technological challenges (e.g., limitations placed on stem-cell research or genetic engineering, introduction of alien species, deforestation, bioterrorism, nuclear energy, genetic counseling, use of alternative energies for carbon fuels, use of pesticides)	3	Assessed locally

* These CLEs are locally assessed.

Table 2.19: Actual Point Distributions for the Fall 2008 Biology Operational Form

	Blueprint Target				**Actual**			
	# Items		**# Points**		**# Items**		**# Points**	
Reporting Categories	**SR**	**PE**	**SR**	**PE**	**SR**	**PE**	**SR**	**PE**
Characteristics and Interactions of Living Organisms Total	22		22		22		22	
Changes in Ecosystems and Interactions of Organisms with Their Environments Total	13		13		13		13	13
Scientific Inquiry		10–16		20		10		20
Total Items/Points	**35**	**10–16**	**35**	**20**	**35**	**10**	**35**	**20**

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.20: Actual Point Distributions for the Spring 2009 Biology Operational Form

Reporting Categories	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
	SR	PE	SR	PE	SR	PE	SR	PE
Characteristics and Interactions of Living Organisms Total	22		22		22		22	
Changes in Ecosystems and Interactions of Organisms with Their Environments Total	13		13		13		13	
Scientific Inquiry		10–16		20		12		20
Total Items/Points	**35**	**10–16**	**35**	**20**	**35**	**12**	**35**	**20**

Table 2.21: Actual Point Distributions for the Summer 2009 Biology Operational Form

Reporting Categories	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
	SR	PE	SR	PE	SR	PE	SR	PE
Characteristics and Interactions of Living Organisms Total	22		22		22		22	
Changes in Ecosystems and Interactions of Organisms with Their Environments Total	13		13		13		13	
Scientific Inquiry		10–16		20		10		20
Total Items/Points	**35**	**10–16**	**35**	**20**	**35**	**10**	**35**	**20**

Table 2.22: Actual Point Distributions for the Fall 2009 Biology Operational Form

Reporting Categories	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
	SR	PE	SR	PE	SR	PE	SR	PE
Characteristics and Interactions of Living Organisms Total	20–24		20–24		22		22	
Changes in Ecosystems and Interactions of Organisms with Their Environments Total	12–14		12–14		13		13	
Scientific Inquiry		11–15		20		11		20
Total Items/Points	**35**	**11–15**	**35**	**20**	**35**	**11**	**35**	**20**

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 2.23: Actual Point Distributions for the Spring 2010 Biology Operational Forms A and B

Reporting Categories	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
	SR	PE	SR	PE	SR	PE	SR	PE
Characteristics and Interactions of Living Organisms Total	20–24		20–24		22		22	
Changes in Ecosystems and Interactions of Organisms with Their Environments Total	12–14		12–14		13		13	
Scientific Inquiry		10–16		20		12		20
Total Items/Points	**35**	**10–16**	**35**	**20**	**35**	**12**	**35**	**20**

Table 2.24: Actual Point Distributions for the Summer 2010 Biology Operational Form

Reporting Categories	Blueprint Target				Actual			
	# Items		# Points		# Items		# Points	
	SR	PE	SR	PE	SR	PE	SR	PE
Characteristics and Interactions of Living Organisms Total	20–24		20–24		22		22	
Changes in Ecosystems and Interactions of Organisms with Their Environments Total	12–14		12–14		13		13	
Scientific Inquiry		10–16		20		13		20
Total Items/Points	**35**	**10–16**	**35**	**20**	**35**	**13**	**35**	**20**

2.5 Development of Test Items

Content-related evidence of validity supporting test interpretation is presented in terms of how the 2008–2010 MO EOC Assessments were assembled for English II, Algebra I, and Biology. Detailed information regarding both item-development procedures and content coverage is included in this section.

The forms for the Fall 2008 through Summer 2010 administrations were constructed using items that were field tested in spring 2008. During the process of building the forms for the 2008–2010 operational test administrations, statistical characteristics (i.e., *p*-values and point-biserial correlations) were monitored to ensure that the statistical properties of the forms were similar within each content area and across operational test forms for fall, spring, and summer.

Riverside Publishing Test Development Specialists (TDSs) created a detailed item and passage development plan based on the blueprints for each content area. The plans included the number of items necessary for each assessable CLE, as well as an outline of the review process for developed items and passages. This process included internal Riverside Publishing reviews, a DESE review on a percentage of the items, and a content and bias review by Missouri educators.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

2.5.1 Item Writing

Missouri educators, DESE staff members, Regional Instructional Facilitators, and Riverside Publishing TDSs created all the test items, including the PEs. English II passages and WPs were developed by item writers trained by Riverside Publishing, Riverside Publishing TDSs, and DESE staff. These passages were developed and refined prior to the item-writing workshops. Requirements to be an item writer included experience in classroom teaching and expert content knowledge.

In September 2007 and June 2008, Riverside Publishing conducted item-writing workshops to develop SR items for English II, Algebra I, and Biology as well as PEs for Algebra I and Biology. These workshops were conducted at the Assessment Resource Center (ARC) in Columbia, Missouri. Participants in the workshops included Missouri educators, DESE staff, Regional Instructional Facilitators, and Riverside Publishing TDSs. The workshops were held over a five-day period and were conducted with 15–20 teacher participants per content area. Teacher participants were selected by DESE to represent school districts throughout Missouri. The content developed at the workshops was based on the Missouri Show-Me Standards CLEs.

The English II participants wrote SR items associated with the passages that had been developed prior to the item-writing workshops. The Algebra I and Biology participants wrote SR items and PEs along with rubrics. Biology PEs consist of a science investigation scenario and several associated constructed response items. The Biology PE items were written based on an existing science PE development template that specifies the types of tasks and numbers of items that compose a PE.

During the item-writing workshops, Riverside Publishing TDSs conducted training sessions with the item writers and provided instructions on avoiding bias and stereotyping of groups and individuals on the basis of gender, race, ethnicity, religion, age, language, socioeconomic group, and disability. Riverside Publishing TDSs also trained item writers to write items that adhere to the principles of universal design, making the items accessible to the widest range of students. For example, items and passages were written using clear and concise language, and all art, graphs, and tables were labeled and were not overly crowded with extraneous information. Instruction was also provided on developing items at particular cognitive levels based on Norm Webb’s Depth of Knowledge (DOK) levels.

Riverside Publishing TDSs trained item writers to enter content into the company’s electronic Content Management System (CMS). During training, each item writer wrote several items and received feedback on them. Participants also received feedback through CMS, as Riverside Publishing TDSs responded to teachers’ items as they were submitted. As items were produced, they were continuously reviewed, revised, edited, and evaluated by Riverside Publishing TDSs and DESE staff. Item writers who generated high-quality work on or ahead of schedule were given additional assignments.

As items were written, they were tracked according to the item development plan. Riverside Publishing kept careful records to maintain a workflow that generated items in assessment strands and CLEs as required by the test blueprint. All items and passages went through several rounds of internal reviews, including content and editorial reviews.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Riverside Publishing TDSs reviewed each item with respect to alignment, clarity, and correspondence with item specifications.

2.5.2 Universal Design

Riverside Publishing TDSs are experienced in employing the principles of universal design in item development so that all students have equal access to the assessments. Riverside Publishing included these principles when training Missouri teachers to write the test items.

According to the *NCEO Synthesis Report 44* (Thompson, Johnstone, and Thurlow, 2002), there are seven elements of universally designed assessments:

1. Inclusive assessment population
2. Precisely defined constructs
3. Accessible, nonbiased items
4. Amenable to accommodations
5. Simple, clear, and intuitive instructions and procedures
6. Maximum readability and comprehensibility
7. Maximum legibility

All items for the MO EOC Assessments were developed with these elements in mind. Riverside Publishing ensured the development of MO EOC items in accordance with these principles in the following manner:

- Items were developed to include a wide array of contexts and cultures. These item types may make students feel more included, increase motivation, and avoid bias.
- The test and item specifications served as a model for precisely defining the constructs that the tests would measure. These specifications indicated to the item writer, content reviewer, and TDS exactly what was to be measured. The item could assess a particular part of a standard or a combination of elements within a standard. The reviews served as a method for eliminating items that included assessment of knowledge outside the standard. For example, a mathematics item should have nonmathematical vocabulary below grade level; otherwise, the item might also be assessing reading ability, introducing construct-irrelevant variance.
- The review of items, which included Missouri teachers from diverse ethnic and geographic backgrounds, served to ensure that all items were accessible to as many students as possible.
- Riverside Publishing staff members trained Missouri teachers to create clear and simple instructions so that students would have a clear understanding of the task needed to answer an item. Teacher review committees had an opportunity to review the instructions to ensure that they were appropriate for the grade levels and subject areas. To ensure the appropriateness of the level of the vocabulary, *Children's Writer's Word Book* and *EDL Core Vocabulary* were employed by test developers and item review committees.
- Finally, items with text, art, tables, maps, and diagrams were constructed with maximum legibility.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

2.5.3 Content and Bias Review Process

Standard 3.6[3] specifically addresses the importance of item review by both an examination of the item statistics and the use of expert panels of judges. This section details the steps that were taken to ensure that the items chosen for the operational forms of the MO EOC Assessments were of high technical quality and were free from bias. Content and bias reviews were conducted in November 2007 and July 2008 in Columbia, Missouri. The content review committees included DESE staff, Missouri educators from around the state, Regional Instructional Facilitators, and Riverside Publishing staff.

The content and bias review committees reviewed selected response items, PEs, and WPs using the following criteria:

- Overall quality and syntactical clarity
- Content coverage and content appropriateness
- Alignment to the specified CLE
- Appropriate contexts
- One clearly correct answer and plausible distractors for selected response items
- Freedom from bias or any racial, socioeconomic, gender, or other sensitivity issues

Before reviewing the items, a group training session was held with all committee members. Riverside Publishing presented a PowerPoint that described the MO EOC program, the test development process, and the content and bias review procedures. After the large-group session, the committee members went to their respective break-out rooms to discuss the week's activities in more detail. The committee members were provided with copies of the CLEs and Item Specifications for the courses for the items they were to review. Each Riverside Publishing content facilitator reviewed these documents with the committee and answered any questions. The committee members were given the following checklists that could be referenced throughout the review process:

For SR items:

- ☐ Does the item assess the assigned CLE?
- ☐ Is the item clear, concise, and complete?
- ☐ Does the item contain accurate and sufficient content information?
- ☐ Is the item grade-level appropriate, and are the vocabulary and syntax appropriate for the students at the intended grade? (Reference the *EDL Core Vocabularies*.)
- ☐ Is the item fair to all students and free of bias and sensitivity issues?
- ☐ Does the item have correct punctuation, and is it grammatically correct?

---

[3] **Standard 3.6:** The type of items, response formats, scoring procedures, and test administration procedures should be selected based on the purposes of the test, the domain to be measured, and the intended test takers. To the extent possible, test content should be chosen to ensure that intended inferences from test scores are equally valid for members of different groups of test takers. The test review process should include empirical analyses and, when appropriate, the use of expert judges to review items and response formats. The qualifications, relevant experiences, and demographic characteristics of expert judges should also be documented (p. 44).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- ☐ Is the item free from spelling and typographical errors?
- ☐ Is clueing avoided within an item stem and options, as well as among items?
- ☐ Does the item stand alone? (The answer to one item should not be dependent on the content of another item.)
- ☐ Are the equations, tables, charts, graphs, and other art clear, accurate, and necessary?
- ☐ Does the item have only one correct answer?
- ☐ Does the item have unique, plausible distractors containing common errors students would make?
- ☐ Are all the options parallel in form and arranged in logical order?
- ☐ Do all distractors contain clear rationale statements? (Math and Science only)
- ☐ Is the item free from absolutes ("none of the above," "all of the above") as options and free from the use of negatives ("not," "none," "except") in the stem?
- ☐ Does the item avoid repeating words from the stem in the options?
- ☐ Does the item pose a single problem (although the solution may require more than one step)?

For PE/WP items:

- ☐ Does the item assess the assigned CLE?
- ☐ Does the item clearly specify how the student should respond?
- ☐ Does the item allow for a variety of acceptable responses for the student to get full credit?
- ☐ Is the item grade-level appropriate, and are the vocabulary and syntax appropriate for the students at the intended grade? (Reference the *EDL Core Vocabularies*.)
- ☐ Is the item rich enough to elicit an appropriate range of responses covering all possible score points?
- ☐ Is the item fair to all students and free of bias and sensitivity issues?
- ☐ Does the rubric clearly define an acceptable answer or answers at each score point level?

Twenty Missouri educators participated in the review process for each content area. The committee members read and reviewed each item. Discussions were held about whether the items met the criteria listed above. The committees then rejected or revised any items they deemed unsatisfactory. If there was disagreement about how to proceed with an item, the Riverside Publishing facilitator polled the group and followed the direction of the majority. Approximately 95% of the items were accepted (as–is or with edits) by the content and bias committees. Table 2.25 shows the number of items that were reviewed in 2007. The accepted items were placed in a pool of items from which the 2008 standalone field-test forms were built. Table 2.26 shows the number of items that were

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

reviewed in 2008 and placed on embedded field test forms in the Spring and Summer 2009 operational administrations.

To further preserve validity, all item review sessions were held in secure meeting rooms, and all materials were confidential. Committee members were required to sign confidentiality agreements so that the integrity of the test content was not compromised. Although educators were encouraged to share information with their colleagues about the process of the item review, they were made fully aware of the expectation that any information about specific items and passages was to remain secure and confidential.

Table 2.25: 2007 Content/Bias Item Review Acceptance Rates

	Total Number of Items Presented for Review	Number of Items Accepted (As Is or With Edits	Acceptance Rate (Items Accepted As Is or With Edits)
English II	404	398	99%
Algebra I	239	233	97%
Biology	402	365	91%

Table 2.26: 2008 Content/Bias Item Review Acceptance Rates

	Total Number of Items Presented for Review	Number of Items Accepted (As Is or With Edits	Acceptance Rate (Items Accepted As Is or With Edits)
English II	298	298	100%
Algebra I	288	288	100%
Biology	164	161	98%

2.6 Test Form Assembly

2.6.1 Field-Test Selection and Administration

The items accepted at the content/bias review were used to build the standalone field-test forms that were administered in spring 2008. Field-test items were selected so that each form met the established operational blueprint requirements for content coverage as closely as possible. For any standalone field-test form that deviated slightly from the blueprint, another field-test form made up for that difference so that the entire pool of field-tested items met the blueprint requirements.

The MO EOC Spring 2008 field test consisted of 10 SR forms per course, 10 English II WPs, 10 Algebra I PE forms, and 10 Biology PE forms. All field-test forms were reviewed and approved by DESE. They were administered to a sample of Missouri students in April 2008.

2.6.2 Statistical Item Review

After the 2008 field-test item scoring was completed, Riverside Publishing TDSs and psychometricians reviewed the statistical characteristics of the items. Riverside Publishing used classical item statistics, including *n*-counts, *p*-values, percentage choosing each response option, point-biserial correlations, and differential item

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

functioning (DIF) analysis for the SR items. Additionally, the Rasch model was used for distractor analysis for the SR items and for DIF analysis for the PE/WP items.

During the data review, which was held June 6–8, 2008, Riverside Publishing Research and Test Development staff and DESE staff reviewed students' performance on the Spring 2008 field-test items. Items were carefully reviewed with respect to their statistical characteristics. Item reviewers from DESE and Riverside Publishing were provided with the following information:

- Form
- Position
- Item as it appeared in the printed books
- Item alignment to the Missouri Show-Me State Standards
- The *p*-value of the correct answer and percentage of students who selected each distractor (for SR items only)
- Mean and *SD* of item score (for PE/WP items only)
- Point-biserial correlation of correct response and point-biserial for each distractor (for SR items only)
- Total number of students who attempted to answer each question
- DIF using the Mantel-Haenszel (MH) (1959) procedure and Educational Testing Service (ETS) classification (for SR items only)

Riverside Publishing and DESE staff reviewed items that were flagged because of statistics that fell outside the parameters determined by the Riverside Publishing Research staff. Table 2.27 contains the guidelines that were used for data review.

Table 2.27: Criteria for Flagged Items

Item Flagging Criteria	Indicates
If *p*-value of keyed response < 0.35	Difficult item
If *p*-value of keyed response > 0.95	Easy item
If *p*-value of keyed response < *p*-value of distractor	Possible miskey
If *p*-value of distractor > 0.35	Possible second correct option
If point-biserial of keyed response < 0.20	Poorly discriminating item
If point-biserial of a distractor is > 0.00	Possible second correct option
If ETS classification is B or C (from DIF analysis)	Possible bias in item

Each flagged item was reviewed, and then Riverside Publishing and DESE made a decision about whether the item should be accepted or rejected. The review included items flagged with moderate to severe DIF (an ETS classification of B or C). A flagged item was accepted if the review team determined that the item was strong and tested students on content that they were expected to know. Accepted items were then made available in the pool of items that could be used to create the operational forms. Items that the review team felt were biased or inappropriate for the Missouri EOC Assessments were rejected. Rejected items were removed from the item pool, making them invalid for the MO EOC Assessments. Of the 690 total items reviewed, 91% were accepted.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

2.6.3 Operational Test Selection and Administration

In June and July 2008, Riverside Publishing TDSs selected operational items for test forms for use in 2008 and 2009. Items for the 2009–2010 administration cycle were selected in June and July of 2009. Using item response theory (IRT) item difficulty values, four equivalent operational forms were selected for each content area. These four forms are the operational component of the Fall, Spring, and Summer EFT forms, as well as the released form. Three forms were selected for the 2009–2010 administrations. The Fall forms were administered in November 2008 and 2009, the Spring forms in April 2009 and 2010, and the Summer forms in June 2009 and 2010.[4]

The operational form construction process was based on content requirements and statistical criteria. The steps associated with assembling the test forms included the following:

1. *Determine form design*. Each form consists of operational items and embedded field-test items.
2. *Select items that meet content specifications*. Each form was constructed based on the test specifications for that content area. The test specifications delineate the item distribution across assessment strands. They also outline the test length, type of items, and number of points to be assessed at each CLE.
3. *Evaluate statistical specifications and select items to meet these specifications*. Spreadsheets (form matrices) are used to ensure that the test forms meet statistical specifications. These matrices contain the following statistics: average *p*-values, point-biserial correlations, and DIF statistics. Riverside Publishing psychometricians conducted a review of the test forms to ensure equivalence of test difficulty across forms.[5]
4. *Review and approve test forms*. Once the content and statistical specifications were met for each content area, the forms were reviewed and approved by MO DESE. The forms were then released for production and additional content and editorial reviews.

2.7 Braille and Large Print Versions

Beyond employing the principles of universal design, all operational assessments were offered in Braille and Large Print versions for visually impaired students taking the EOC Assessments in Missouri. The Fall 2008 operational paper-and-pencil version was converted into Braille and large print to accommodate these students.

Once the Braille and Large Print forms were created, two separate reviews were held with educators from Missouri who had specialized training in working with visually impaired students.

[4] The Summer 2009 administration is part of the 2009–2010 assessment year, and its results are not included in this technical report.
[5] Rasch values were not available for all items when the 2008–2009 operational forms were built.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

The Large Print form review was held in Jefferson City, Missouri, at the DESE offices on September 29, 2008. The Braille review was held in St. Louis, Missouri, at the Missouri School for the Blind on October 10, 2008.

The teachers consulted the *Large Print and Braille Style Guide*, which was also used during form composition, and relied on their own expertise to determine whether changes to directions, passages, or items were needed, or whether items should be omitted. Riverside Publishing Braille vendor (Region IV) also reviewed the forms and made recommendations based on how items, passages, and directions would be transcribed to Braille.

Riverside Publishing and DESE reviewed the recommendations from all of these sources. It was determined that no items had to be omitted to accommodate Large Print students. For the Braille version of the form, one item from English II was removed because the content of the item prohibited transcription to Braille. Students taking the Braille form were given credit for this item. The embedded field test items were eliminated from both versions of these forms due to the irregular testing conditions and the small sample sizes for these groups. These versions of the Large Print and Braille forms were used through the Summer 2010 test administration.

2.8 Online Forms Construction

All items were written so that they could be presented in the online delivery system without any alterations.

2.9 Quality Control for Test Construction

Checklists and quality control procedures accompany each stage of form development. Following is a list of some quality control procedures used during the assembly of the MO EOC Assessment forms:

- Construct forms based on all content requirements noted in the test blueprint and test specifications.
- Verify correct number of items per standard or reporting category based on test blueprint.
- Review items to ensure a wide sampling of the knowledge and skills being measured.
- Ensure that all items have been through the appropriate review procedures and are approved for use by DESE.
- Check for a variety of item topics, equal distribution of males and females, ethnicities, etc.
- Verify appropriate portions of items with and without artwork.
- Check for clueing across all items on each form.
- Verify equal or nearly equal distribution of answer choices for SR items.
- Ensure that the test meets the required statistical specifications (i.e., that as many items as possible have *p*-values between .35 and .90 and as many items as possible have point-biserial correlations above .20).
- Consider any statistical flags or problems.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Check statistics to ensure that the collection of items on a given form yields an overall difficulty that falls within the specified range.
- Verify that items have not been released to the public.
- Verify correct answer key for each item.
- Perform content review of form (senior staff).
- Perform statistical review of form (psychometrician/statistician).
- Send form to DESE for review and approval.

2.10 Summary

The MO EOC Assessments in English II, Algebra I, and Biology provide an indication of student progress toward achieving the knowledge and skills identified in the Missouri Show-Me Standards. Just as the Show-Me Standards guided the item development and selection process, the consideration of content played an equally important role in form development. Form development required a balance of both content coverage and item difficulty. As items were selected for inclusion on particular forms, every effort was made to balance the content coverage to ensure the items aligned to the Missouri Show-Me Standards/CLEs being assessed while simultaneously considering the overall difficulty of the forms.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 3: ACHIEVEMENT-LEVEL SETTING

3.1 Introduction

To provide full documentation for the Missouri End-of-Course (MO EOC) Assessments, important information from the 2009 standard-setting workshop is provided as the cut scores established at the 2009 standard-setting sessions were applied to 2010 test scores. Reports of past standard-setting activities provide continuity and will help readers more fully understand the MO EOC program and its impact on student learning.

One purpose of assessment is to establish clear guidelines for educational decision making. By assigning meaning to test scores, standard setting allows policymakers, administrators, teachers, parents, and students to make statements about the level of proficiency of individual students and groups of students. The purpose of this chapter is to provide documentation of the achievement-level-setting (or standard-setting) event conducted for the MO EOC Assessments on November 3, 4, and 5, 2008. These activities were undertaken for three MO EOC Assessments: English II, Algebra I, and Biology. These three assessments were administered operationally for the first time during the 2008–2009 school year.

3.2 Goal of the Standard Setting

The main goal of the standard-setting event was to establish three cut scores for each test in the MO EOC Assessments:

1. The cut score that differentiates Below Basic performance from Basic performance
2. The cut score that differentiates Basic performance from Proficient performance
3. The cut score that differentiates Proficient performance from Advanced performance

The determination of three cut scores yields four performance categories for each assessment.

3.3 Staff and Participants

Staff from Questar Assessment, Inc., a subcontractor to Riverside Publishing, planned and facilitated the standard-setting workshops. Questar's most-experienced facilitators—Michael Beck, Sheila Potter, and Martha Caswell—served as facilitators for the workshops. Each of these individuals has facilitated standard-setting sessions for multiple clients for both elementary level and high school level assessments.

In addition to the staff from Questar, two psychometricians from Riverside Publishing attended the workshops. Their function was to enter panelist data, produce tables and reports, and oversee data quality control. A Riverside Publishing program manager was present for the entire workshop to assist Missouri Department of Elementary and Secondary Education (DESE) staff and the panelists with logistical issues. Content area specialists from Riverside Publishing's Content Development group were present in the three panel rooms to serve as

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

resources for content-related questions. Finally, curriculum staff from DESE attended the standard-setting workshops to serve as content resources to the appropriate panels.

A total of 46 panelists participated in the standard-setting workshop: 14 in English II, 15 in Algebra I, and 17 in Biology. One to three members of each panel had participated in an earlier achievement-level-setting workshop for other Missouri assessments. The significant majority of panelists had not been members of any of the assortment of committees for MO EOC development activities. More than half of each panel was made up of active classroom teachers in the relevant content area; several other panel members were other professional educators, such as administrators and curriculum coordinators. One or two members of each panel were business professionals with expertise in the relevant field. Lists of the standard-setting participants appear in Appendix A.

3.4 Development of Achievement-Level Descriptors

The MO EOC Assessments utilize the same achievement-level labels used for previous high school Missouri Assessment Program (MAP) assessments: Advanced, Proficient, Basic, and Below Basic. For each of these levels, the achievement-level descriptor (ALD) describes the specific knowledge and skills that a student at that level is able to demonstrate. As suggested by Missouri's Technical Advisory Committee (TAC), DESE conducted sessions devoted to developing draft ALDs prior to the standard-setting workshop.

Riverside Publishing staff used the ALDs from the Missouri Assessment Program (MAP) and the End-of-Course Course-Level Expectations (CLEs) to create the initial draft ALDs for English II. This document was then reviewed and approved by DESE. This content area was used as a template for the Algebra I and Biology ALDs. This template was provided to committees comprised of people representing higher education in Missouri. These committees reviewed the CLEs and added skills associated with each level. These ALDs were reviewed and approved by DESE.

At the standard-setting workshop, participants devoted a significant portion of time to fine-tuning the draft ALDs for each assessment. The facilitators provided the panelists with draft copies of the appropriate ALDs, copies of the MO EOC Assessment blueprint, and the appropriate CLEs. Using these materials as references and drawing on the expertise of the panelists, the Questar facilitators led each panel in an extended discussion and exercise to refine and elaborate each of the ALDs. Once this activity was complete, the panels relied on the resulting ALDs as a reference during the actual standard-setting activities. In addition, the panelists were allowed to make appropriate, though generally minor, revisions and refinements to the ALDs during and after the standard-setting activities.

3.5 Overview of Standard-Setting Activities

3.5.1 Methodology and Data Considerations

The specific methodology used for the standard-setting activities was a modified Angoff procedure, as recommended by the state's TAC. The Angoff procedure and its modifications are well-recognized and heavily researched methods for establishing student performance standards for tests such as the EOC. Prior Missouri standard-setting

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

workshops utilized an item-mapping procedure commonly known as Bookmark standard setting; however, that method requires placing the items in a difficulty-ordered item book, which necessitates that the item difficulty parameters be known. In the case of the MO EOC Assessments, because the Spring 2009 operational assessment had not yet been administered at the time of the standard-setting workshop, parameter estimates for the operational test forms were not available. The modified Angoff method does not require placing the items in difficulty order; it was, therefore, a suitable choice of methods for this event.

Consistent with the methods used for prior MAP standard-setting events, the modified Angoff method allows three distinct rounds of panelist judgments. Between the first and second rounds, Riverside Publishing provided the panelists with item-difficulty data for their consideration. Because operational data were not available in November, the item data were derived from the 2008 field test. Panelists were appropriately cautioned about the limitations of such data. (For more information about the similarity of the data from the Spring 2008 standalone field test and the Spring 2009 operational assessment, the reader is referred to Section 7.3.3 of this technical report.)

Before the last round of judgments, Riverside Publishing staff provided the panelists with statewide impact data for the assessment. These data were intended to serve as an anchor for the panelists' recommendations. Again, because actual performance data were not available, the data were based on projected statewide score distributions generated from the field test. It is likely that a standalone field test would produce lower-than-expected results due to decreased student effort; therefore, Riverside Publishing psychometricians would consider the field test data "lower-bound" estimates of actual student performance in an operational event. As with the item-level data estimates, the facilitators cautioned the panelists about relying too much on these impact data.

Despite the limitations of the field test data for the standard-setting activities, Riverside Publishing psychometricians believed that providing panelists with even tentative data was desirable, both to mirror procedures used for establishing standards for previous Missouri assessments and to provide panelists with an "external reality check" on their evolving recommendations. Missouri TAC discussions confirmed the use of these statewide impact data.

In addition to the caveats about item level and impact data, panel facilitators clearly communicated to the panelists that the results of their standard-setting activities would be purely advisory to DESE. DESE would consider the recommendations and select the final cut scores for each assessment.

3.5.2 Description of the Test Forms and Considerations

DESE chose to use the MO EOC Spring 2009 operational forms for the standard-setting event. These forms were selected from the several available operational forms because they would be the most widely used in the 2008–2009 test administration year.

The MO EOC Assessments are comprised of selected response (multiple choice) items and a Performance Event/Writing Prompt (PE/WP). Each English II and Algebra I form includes a single PE/WP worth 4 points. On the Biology assessment, the Performance

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Event consists of 11 open-ended items, each worth between 0 and 4 points (for a total of 20 points).

3.6 Specific Standard-Setting Activities

The Standard-Setting Session Agenda provided a general guide regarding the time devoted to each activity. This agenda is included as Appendix B. Questar facilitators held closely to the times contained in the agenda. They used identical processes, including presentation slides and scripts, across all sessions to minimize any intersession differences related to facilitator or session variance.

The following sections provide details about the processes that Questar and Riverside Publishing followed during the course of the standard-setting workshop.

3.6.1 General Process Overview

The first 90 minutes of the three-day session served as an introduction and overview to the general standard-setting processes. First, Michael Muenks, Coordinator of Curriculum and Assessment for DESE, oriented the panelists to the MO EOC program and briefly outlined the session purpose and intended outcomes.

Next, Michael Beck of Questar led a brief general overview, "What Is Standard Setting?" Its purpose was to ensure a common understanding of the fundamental elements of the process. Mr. Beck included a brief overview of the general process of establishing student performance standards, ground rules for panelist activities, and some key elements for the panelists to focus on when attempting to set standards. Mr. Beck also advised the panelists that their work was advisory to DESE. This introduction was a high-level overview of the standard-setting process; individual facilitators provided more detail about each step in the process after the panels broke into content-specific groups. The PowerPoint slides presented during the opening session are included as Appendix C of this report.

Finally, Dr. Sheila Potter of Questar provided a general overview of ALDs and their importance to the standard-setting process. Since the panels would be reviewing, editing, and expanding on draft versions of the ALDs provided by the state, it was important for panelists to understand the critical role of ALDs in the standard-setting process. Following this activity, panelists divided into the three content-specific panel break-out rooms, where all remaining work for the sessions took place.

3.6.2 Panelists Take the Operational Assessments

After reconvening in the content-area panels, panelists first introduced themselves and signed DESE-provided confidentiality forms. Facilitators introduced themselves and reiterated the high-level standard-setting processes that Mr. Beck had discussed during the opening session. Facilitators then allowed the panelists time to take and score the Spring 2009 form of the operational assessment. For this activity, panelists had access to the test administration procedures, the actual test content, and all relevant scoring materials. Field test items that were included in these forms were removed from the test books seen by the panelists. Because these were "live" materials, facilitators stressed the confidentiality of all of the items.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

The primary purpose of this activity was to familiarize panelists with the actual, complete assessment content prior to beginning the standard-setting judgments. Following this review of the tests, each panel spent a short time reacting to the assessment content: difficulty, sources of challenge, scoring issues, and general and specific reactions. This exercise provided the panelists, especially those not familiar with the MO EOC Assessments, with a context concerning the definition of *Proficient* as conveyed by the assessments.

3.6.3 Panelists Discuss and Fine-Tune the ALDs

All three panels began this activity with a review of the draft ALDs for the particular content area. Separate panels of Missouri educators had developed these draft ALDs during DESE-led sessions several weeks earlier. The ALD review activity was highly interactive, with panelists suggesting changes and other refinements—both substantive and editorial—to the draft ALDs. The ultimate task was to operationalize specific behaviors indicating performance at the Advanced, Proficient, Basic, and Below Basic levels in the content area. Panel suggestions were discussed until consensus was reached and were then recorded on the draft ALDs, a copy of which was given to each panelist or placed on chart paper displayed around the room. Panelists were later able to refer to these pages, along with the original drafts, during the actual judgment activities. The thoroughness of the ALD refinement activities and the extent to which the panelists, individually and as a group, internalized the ALDs significantly impacted the soundness of the subsequent standard-setting activities. For this reason, approximately two hours was devoted to this activity in the session agenda.

At the conclusion of the standard-setting sessions, DESE collected the panelist recommendations for ALD revisions for consideration in the wording of the final ALDs. Appendix D contains a copy of the draft ALDs that were distributed to the panelists at the outset of the standard-setting workshop. Appendix E contains a copy of the final ALDs.

3.6.4 Orientation to the Modified Angoff Procedures

After the ALD activity was complete, facilitators oriented the panels to the specific tasks involved with the modified Angoff standard-setting process. The modified Angoff process requires panelists to read and make judgments about each successive item in the test book using the following procedures. When reading an item, panelists were to consider the item's importance in the context of the underlying CLE, the task(s) required of the student, and the item's difficulty. They were to decide what percentage of minimally Proficient students should be able to answer the item correctly. Panelists were then to decide what percentage of minimally Advanced students would answer the item correctly. Finally, they were to decide what percentage of minimally Basic students would answer the item correctly. (While the MO EOC Assessments contain four levels of student performance, cuts are made at only three locations on the score distribution.) The panelists were instructed to consider their judgments in this order—Proficient, Advanced, and Basic—as it anchors the item judgments on the most important cut, Proficient. In addition, once panelists made their judgment for the Proficient students, they had a clearer, more defined range of values to consider for the other two cuts.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

For the constructed response item(s), panelists were to consider the average item score of minimally Proficient, minimally Advanced, and minimally Basic students. In other words, judgments for the constructed response items were made as whole-point values (i.e., 1, 2, 3, etc.) rather than as percentages of students answering correctly.

The facilitators included the following important points in their presentations:

- Panelists should focus on the *threshold* of performance in each category.
- Panelists should review and recall what each performance descriptor means.
- Panelists should focus on MO EOC students statewide, not just in the school or district in which they work.

Finally, the facilitators explained that the panelists' judgments should be made independently and anonymously and that security of the testing materials should be maintained at all times.

The steps outlined in Sections 3.6.1 through 3.6.4 composed the activities of the first day of the workshop. The second day began with an overview of the previous day's activities and outcomes, after which panelists took a five-item multiple choice qualifying test concerning the standard-setting procedures they were about to use. A copy of this instrument is provided as Appendix F. This instrument was used to ensure that all panelists understood the importance of the ALDs and selected elements of the modified Angoff procedure before beginning the process of making item judgments.

Next, the facilitators led their panels in a practice exercise using the modified Angoff rating procedures. The practice test contained five multiple choice items and one performance item (with score points comparable to the performance items on the operational test). The practice items were released Grade 4 NAEP Mathematics items selected to distance the practice exercise from the content area and grade levels of the EOC Assessments. During this exercise, panelists had the opportunity to practice the mechanical aspects of the modified Angoff judgment process and procedures for recording their recommendations before beginning work on the real test. Additionally, the practice test allowed the facilitators to check the panelists' understanding of the mechanics of the technique and corresponding recording of judgments. Facilitators allowed the panelists about 20 minutes to complete the practice activity. The practice judgments were reviewed on a group basis by discussing the range of judgments made about each item.

Following completion of the practice exercise, facilitators asked the panelists to complete and sign a form indicating that they understood the information they had received and discussed and that they felt prepared to make their Round 1 judgments. All panelists so indicated.

3.6.5 Round 1 Judgments

At this point, panelists were ready to make their Round 1 judgments. This work was completed anonymously (via judge numbers known only to Riverside Publishing staff) and independently. Panelists indicated their judgments on specially designed scannable rating sheets developed for each content area. These rating sheets contained three fields for each test item: one for Basic, one for Proficient, and one for Advanced. For the

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

selected response items, each field contained a set of bubbles corresponding to the percentage of students expected to choose the correct answer. As panelists made their judgments for each item, facilitators instructed them to “bubble in” one value for each achievement level. In other words, for Item 1, the panelist entered a number corresponding to the percentage of students expected to choose a correct answer at the minimally Basic level, a number for the minimally Proficient level, and a number for the minimally Advanced level. Panelists were constrained to choosing multiples of 5 (i.e., 5%, 10%, 15%, etc.), as they appeared on the Rating Forms. Panelists then followed this same procedure for all the remaining selected response items. For the PE/WP items, each field contained bubbles corresponding to the various point values possible for the item. Facilitators instructed the panelists to make a judgment about how many points a borderline student at each achievement level would score on that item (in other words, how many points a Basic student would score, how many points a Proficient student would score, and how many points an Advanced student would score).

Most panelists completed their first round of judgments within 60 minutes; however, there was no time limit for this activity, and some panelists required 90 minutes to complete their judgments. This is not unusual for the first round of judgments in a modified Angoff workshop; often some panelists are still struggling to understand the task at this point, thus requiring more time to make their judgments. After panelists completed their judgments, they turned in their rating sheets and were excused for a lunch break.

3.6.6 Feedback and Discussion of Round 1 Judgments

While the panelists were at lunch, the Riverside Publishing psychometricians prepared reports of the Round 1 judgment results. The afternoon session began with an overview of these reports. The first report was a table displaying all three raw score cuts as determined individually by each panelist’s judgments. This table also contained the entire panel’s average, median, highest, and lowest raw score cuts, as well as the standard deviation of all the panelists’ judgments for each of the three raw score cuts. The second report contained a frequency display of all three cut scores (Basic, Proficient, and Advanced) recommended by each panelist. This bar graph displayed all the panelists’ judgments on a single graph so that areas of dispersion or overlap in the raw cut scores would be apparent. These reports are anonymous; ID numbers, rather than names, are used to identify individual panelists.

Facilitators spent time reviewing these reports with the participants to ensure that everyone understood how to interpret the information contained in them. Using the Round 1 results, facilitators then led an extended discussion of the Round 1 judgments. This discussion focused primarily on the panelists’ judgments of individual items. Facilitators actively engaged all the panelists in the discussion to gauge whether they had indicated the item percentage values that they intended, that the reasoning processes they followed in making their judgments were consistent with good practice, and that the panelists clearly understood the mechanics of making item judgments. Throughout these discussions, facilitators focused on the key elements of the standard-setting process: establishing the *threshold* of each cut, projecting the cuts for a *statewide* population of these students, and focusing on the particular *course* and *performance level* of the target populations.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Much like a jury deliberation, this discussion also allowed the panelists to hear their peers' comments and rationales for their judgments. This phase took around two hours, depending on the session; facilitators permitted discussion to continue until they perceived that all panelists were prepared to make their second round of judgments.

Next, facilitators distributed statewide item difficulty data derived from the 2008 field test. For the selected response items, the derived item difficulties were item *p*-values. For the PE/WP, the item difficulties were average item scores. Facilitators advised the panelists that caution should be taken in interpreting the item difficulty data, since the data were collected during a standalone field test (and student motivation may not have been the same as it would be on an operational assessment). Facilitators also explained that these data were relevant, but not critical, to the process of setting standards.

Before making Round 2 judgments, panelists again signed a short form indicating that they understood the procedures and were prepared to make Round 2 recommendations.

3.6.7 Round 2 Judgments

During Round 2, panelists again worked independently to make judgments about the percentage of students at the threshold of each achievement level who would answer each item correctly. Facilitators explained to the panelists that they were free to maintain their Round 1 judgments or to revise them as they deemed appropriate. Before beginning this round of judgments, panelists were once more reminded of the key elements of the process and were asked to focus specifically on the ALDs for their assessment. Again, there was no time limit, although this round required significantly less time than did Round 1 because the panelists more clearly understood the judgment process. In addition, they were increasingly familiar with the specific items for which they were making the judgments. Further, many panelists had begun to formulate some or all of their Round 2 item judgments during the discussion of the Round 1 results.

After panelists completed their Round 2 judgments and recorded their recommendations on their rating sheets, they submitted the forms and were excused for the evening. After all rating sheets were collected, Riverside Publishing psychometricians prepared the reports of the Round 2 judgments.

3.6.8 Feedback and Discussion of Round 2 Judgments

When the panels convened on the third day of the standard-setting workshop, facilitators presented the results of the Round 2 judgments. The reports showing the Round 2 results were used to guide another discussion of specific items. The presentation and discussion at this stage were similar to, although more focused than, those following Round 1.

Following this discussion, facilitators provided panelists with estimated statewide impact data—that is, the percentages of students statewide whose performance would likely be labeled Below Basic, Basic, Proficient, or Advanced if the panels' Round 2 judgments adopted. The panels' median Round 2 judgments were used to determine cut scores for this report. Again, facilitators advised the panelists that the impact data were relevant to, but not essential for, setting performance standards. (This cautionary information was especially important in the case of MO EOC, as the data were not grounded in an operational administration of the assessments.)

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

When facilitators were comfortable that all panelists were prepared to make their final recommendations, they proceeded to Round 3 of judgments.

3.6.9 Round 3 of Judgments, Meeting Evaluation, and Final Inspection of ALDs

For Round 3, the panelists' judgments consisted of one recommended cut score for each achievement level; panelists were not required to make item-level judgments. Panelists were given unlimited time to complete their Round 3 (final) recommendations, although most completed their judgments within 20 minutes. All panelists clearly understood that only the Round 3 judgments counted as their recommendations and that the three rounds were not combined in any way to form the proposed cuts.

After completing their final round of judgments, individual panelists were excused for lunch. Following the lunch break, facilitators provided the panels with the results of their Round 3 judgments.

Immediately following the final presentation of cut scores, panelists completed a written evaluation of the process. This evaluation covered the panelists' opinions of the adequacy of the training provided and their comfort with and confidence in their judgments on a round-by-round basis. The form also contained spaces for the panelists to write other comments concerning the workshop. A copy of this evaluation is included as Appendix G of this report.

After facilitators collected the panelist evaluations, they allowed the panels 30 to 40 minutes for a final review of the ALDs. During this time, panelists were allowed to discuss and, if necessary, fine-tune or revise the ALDs. Finally, panelists were thanked for their participation and dismissed.

3.7 Session Results by Panel and Round

Appendices H, I, and J contain the feedback reports by round for English II, Algebra I, and Biology, respectively. Selected data from these graphs and tables are summarized below for ease of cross-round and cross-content-area comparison.

The standard-setting literature typically considers the *median* recommendation to be the best indicator of a panel's judgment, as the median would not be impacted by the judgments of a few outlying panelists. In the case of this standard-setting event, as a review of Appendices H, I, and J indicates, all median and mean cut scores are within a single rounded raw-score point for all of the content areas. Therefore, the choice of a measure of central tendency for these particular panels would not markedly impact the resulting cut scores.

Table 3.1 contains the median recommended cut scores for all rounds and content areas. As data in Table 3.1 indicate, the panels did not markedly change their typical recommended cut scores across the three rounds of judgments. This is not to say that individual panelists made the same recommendations across rounds. In fact, across the nine sets of judgments between rounds (three content areas with three cut scores each), the mean change in median raw cut scores was −0.5 between Rounds 1 and 2, −1.0 between Rounds 2 and 3, and −1.5 between Rounds 1 and 3. (The median raw-score change between any pair of rounds was 0.) Though the mean changes were minimal from round to round, individual panelists changed their round-to-round recommendations by as

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

much as 17 raw-score points. Across all panels, the mean absolute value of raw cut score changes made was 1.5 between Rounds 1 and 2, 1.9 between Rounds 2 and 3, and 2.6 between Rounds 1 and 3.

Table 3.1: Median Recommended Cut Scores by Content Area and Round

	Content Area								
	English II			Algebra I			Biology		
Cut*	BB–B	B–P	P–A	BB–B	B–P	P–A	BB–B	B–P	P–A
Round 1	16.5	24.5	32	13	23	32	20	35	48
Round 2	16.5	25.5	33	13	23	31	19	34	46
Round 3/Final	15.5	24	33	13	22	31	18	32	45
No. Points Possible	39			39			55		

* BB = Below Basic; B = Basic; P = Proficient; A = Advanced

As is typically the case with standard-setting activities conducted over multiple rounds, the standard deviations of panelists' recommendations got smaller across rounds, indicating both an increasing level of panelist understanding of the process and increasing interpanel agreement based on group discussions between rounds of judgments. This is illustrated graphically through an examination of the frequency bar charts in Appendices H, I, and J, as well as statistically in the tabled results. The colored bars representing Basic, Proficient, and Advanced cut scores clearly become taller and more compact over each round of judgments. While panelists came closer to their peers in judging the most appropriate cut scores, even in Round 3—not unexpectedly—there was still a fair amount of spread in the recommended scores. That variability is especially notable in the Biology assessment; however, this assessment is significantly longer than the others, which may partially account for the larger Round 3 variability.

Standard errors of the median judgments were computed for all cut scores across all panels. In no case did the Round 3 standard error reach a whole raw-score unit. Most were lower than half of a raw-score point. This indicates that the final median judgments are highly stable. These standard errors are, of course, at least partially a function of the relatively small standard deviations and large panel sizes.

Table 3.2 summarizes the projected statewide percentages of students whose EOC scores will fall in each of the four performance categories. These data are based on the 2008 field test results and may be viewed as "lower-bound" estimates of the likely statewide results that will be obtained at the end of the 2008–2009 school year.

Table 3.2: Projected Statewide Percentages of Students Scoring in the Various Performance Categories on the EOC Assessments, 2008

	Performance Category*			
Assessment	Below Basic	Basic	Proficient	Advanced
English II	15%	31%	39%	16%
Algebra I	18%	38%	33%	11%
Biology	12%	39%	39%	10%

*Percentages may not sum to 100% due to rounding.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

After the standard-setting event, the Missouri State Board of Education adopted the panels' median cut scores without revision.

3.8 Results of Participant Evaluations

Appendix K contains the data collected from panelists on the evaluation form. For the questions pertaining to the organization and adequacy of information provided in the opening session, the panelists generally provided ratings of 4 or 5 (on a scale of 1 to 5, with 5 being the highest). For the evaluation questions pertaining to the discussions of the achievement-level descriptors and the panelists' understanding of each of the ALDs after the discussions, in all cases at least 70% of the panelists provided ratings of 4 or 5. The questions pertaining to the panelists understanding of the judgment process and feedback on the results of each round received similar scores. Overall, these data indicate that the panelists generally understood what was expected of them, were comfortable with the process, and were comfortable with the resulting cut scores.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 4: ITEM ANALYSIS

4.1 Introduction

Item analyses were conducted for Missouri End-of-Course (MO EOC) Assessments in English II, Algebra I, and Biology. In this chapter, the summary information, which includes mean item score and discrimination indices, is presented at the item level for each content area. The item summary statistics presented in this section (*p*-values, point-biserial correlations, and omit rates) are based on the operational administrations that included responses from 6,195 students for Fall 2009, 2,666 students for Summer 2009, and 183,194 students for Spring 2010 across the three content areas. The differential item functioning (DIF) analyses are based on the Spring 2008 standalone field and the Spring 2009 embedded field test data.

For selected response (SR) items, the mean score is simply the proportion of students who gave correct responses to the item (usually referred to as item difficulty, or *p*-value), and the discrimination index is the point-biserial correlation between the item score and the total score based on the remaining items. The total score included both SR and Performance Event/Writing Prompt (PE/WP) items.

For PE/WPs, the mean score is the mean of student scores on a scale of 0 to 4 for English II and Algebra I, and 0 to 20 for Biology. The discrimination index is the correlation between the item score and the total score based on the remaining items.

4.2 Analysis of Forms for Each End-of-Course Assessment

Tables 4.1 through 4.9 summarize item difficulty, discrimination, and omit rates for the SR and PE/WP items that composed each assessment for the Fall 2009, Summer 2009, and Spring 2010 operational administrations. For SR items, the *p*-value is the proportion of students who answered the item correctly. For PE/WP items, the mean value is the average student score on that item. The item discrimination, or corrected point-biserial correlation, is the correlation between students' item scores and their total scores on the remaining test items. Both item difficulty and item discrimination are expressed in the raw score metric.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.1: Item Statistics for English II, Summer 2009 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.85	0.19	0.00
2	0.82	0.24	0.00
3	0.65	0.40	0.00
4	0.91	0.35	0.00
5	0.32	0.07	0.00
6	0.31	0.24	0.00
7	0.81	0.21	0.00
8	0.60	0.25	0.00
9	0.79	0.37	0.00
10	0.58	0.31	0.00
11	0.37	0.18	0.00
12	0.41	0.17	0.00
25	0.87	0.37	0.00
26	0.64	0.51	0.00
27	0.40	0.13	0.00
28	0.46	0.34	0.00
29	0.51	0.29	0.00
30	0.70	0.47	0.00
31	0.26	0.07	0.00
32	0.45	0.09	0.00
33	0.69	0.45	0.00
34	0.44	0.26	0.00
35	0.70	0.39	0.00
36	0.57	0.43	0.00
37	0.63	0.47	0.00
38	0.55	0.28	0.00
39	0.49	0.40	0.00
40	0.55	0.40	0.00
41	0.73	0.49	0.00
42	0.58	0.53	0.00
43	0.39	0.23	0.00
44	0.26	0.04	0.00
45	0.52	0.34	0.00
46	0.25	0.12	0.00
47	0.60	0.27	0.00
PE	2.57	0.35	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.2: Item Statistics for Algebra I, Summer 2009 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.77	0.28	0.01
2	0.53	0.30	0.01
3	0.71	0.35	0.00
4	0.56	0.42	0.00
5	0.56	0.30	0.00
10	0.62	0.32	0.00
11	0.61	0.28	0.00
12	0.70	0.41	0.00
13	0.75	0.34	0.00
14	0.58	0.38	0.01
15	0.70	0.33	0.00
16	0.71	0.33	0.00
17	0.48	0.30	0.01
18	0.64	0.41	0.00
19	0.61	0.41	0.00
20	0.45	0.40	0.00
21	0.42	0.30	0.00
26	0.49	0.35	0.00
27	0.49	0.34	0.00
28	0.50	0.29	0.00
29	0.48	0.34	0.00
30	0.42	0.40	0.00
31	0.38	0.10	0.00
32	0.30	0.27	0.00
33	0.55	0.38	0.00
34	0.35	0.17	0.00
35	0.30	0.16	0.00
36	0.31	0.28	0.00
37	0.22	0.09	0.00
38	0.39	0.23	0.00
43	0.25	0.10	0.00
44	0.31	0.17	0.00
45	0.21	0.22	0.00
46	0.22	0.19	0.00
47	0.39	0.22	0.00
PE	1.13	0.61	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.3: Item Statistics for Biology, Summer 2009 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.80	0.27	0.00
2	0.41	0.13	0.00
3	0.78	0.41	0.00
4	0.34	0.05	0.00
5	0.40	0.41	0.00
10	0.89	0.23	0.00
11	0.59	0.36	0.00
12	0.40	0.32	0.00
13	0.50	0.36	0.00
14	0.45	0.37	0.00
15	0.59	0.36	0.00
16	0.57	0.32	0.01
17	0.50	0.29	0.00
18	0.36	0.11	0.00
19	0.56	0.27	0.00
20	0.75	0.25	0.00
21	0.47	0.41	0.00
26	0.41	0.23	0.00
27	0.56	0.37	0.00
28	0.30	0.33	0.00
29	0.65	0.38	0.00
30	0.49	0.28	0.01
31	0.33	0.15	0.00
32	0.41	0.27	0.00
33	0.52	0.31	0.00
34	0.60	0.41	0.00
35	0.69	0.18	0.00
36	0.43	0.17	0.00
37	0.48	0.34	0.00
38	0.47	0.25	0.00
43	0.46	0.27	0.01
44	0.29	0.21	0.00
45	0.43	0.24	0.00
46	0.46	0.33	0.00
47	0.31	0.23	0.00
PE	9.12	0.67	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.4: Item Statistics for English II, Fall 2009 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.72	0.26	0.00
2	0.67	0.25	0.00
3	0.64	0.30	0.00
4	0.85	0.34	0.00
5	0.81	0.24	0.00
6	0.81	0.34	0.00
7	0.36	0.35	0.00
8	0.51	0.27	0.00
9	0.58	0.36	0.00
10	0.39	0.17	0.00
11	0.58	0.28	0.00
12	0.82	0.26	0.00
24	0.53	0.18	0.00
25	0.76	0.41	0.00
26	0.58	0.18	0.00
27	0.52	0.34	0.00
28	0.68	0.30	0.00
29	0.55	0.39	0.00
30	0.60	0.37	0.00
31	0.31	0.22	0.00
32	0.35	0.22	0.00
33	0.47	0.21	0.00
34	0.43	0.33	0.00
35	0.46	0.28	0.00
36	0.55	0.27	0.00
37	0.47	0.33	0.00
38	0.49	0.08	0.00
39	0.63	0.48	0.00
40	0.75	0.38	0.00
41	0.51	0.32	0.00
43	0.32	0.18	0.00
44	0.49	0.31	0.00
45	0.72	0.30	0.00
46	0.64	0.39	0.00
47	0.70	0.25	0.00
PE	2.76	0.50	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.5: Item Statistics for Algebra I, Fall 2009 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.67	0.44	0.00
2	0.89	0.33	0.00
3	0.68	0.50	0.00
4	0.72	0.44	0.00
5	0.37	0.34	0.00
10	0.62	0.45	0.00
11	0.47	0.27	0.00
12	0.73	0.36	0.00
13	0.29	0.41	0.00
14	0.36	0.23	0.00
15	0.27	0.35	0.00
16	0.41	0.22	0.00
17	0.39	0.25	0.00
18	0.33	0.48	0.00
19	0.42	0.26	0.00
20	0.41	0.14	0.00
21	0.44	0.48	0.00
26	0.40	0.13	0.00
27	0.24	0.10	0.00
28	0.67	0.42	0.00
29	0.50	0.42	0.00
30	0.67	0.50	0.00
31	0.47	0.40	0.00
32	0.52	0.49	0.00
33	0.62	0.31	0.00
34	0.49	0.29	0.00
35	0.43	0.32	0.00
36	0.49	0.25	0.00
37	0.40	0.17	0.01
38	0.69	0.42	0.00
43	0.73	0.38	0.00
44	0.46	0.46	0.00
45	0.68	0.48	0.00
46	0.49	0.36	0.00
47	0.75	0.32	0.00
PE	1.34	0.59	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.6: Item Statistics for Biology, Fall 2009 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.81	0.47	0.00
2	0.47	0.49	0.00
3	0.60	0.56	0.00
4	0.68	0.50	0.00
5	0.85	0.27	0.00
10	0.58	0.45	0.00
11	0.47	0.38	0.00
12	0.43	0.38	0.00
13	0.57	0.44	0.00
14	0.68	0.26	0.00
15	0.61	0.39	0.00
16	0.54	0.46	0.00
17	0.37	0.32	0.00
18	0.83	0.46	0.00
19	0.92	0.35	0.00
20	0.61	0.49	0.00
21	0.51	0.20	0.00
26	0.70	0.50	0.00
27	0.33	0.43	0.00
28	0.51	0.44	0.00
29	0.49	0.37	0.00
30	0.54	0.49	0.00
31	0.70	0.39	0.00
32	0.73	0.43	0.00
33	0.33	0.26	0.00
34	0.74	0.38	0.00
35	0.73	0.39	0.00
36	0.85	0.43	0.00
37	0.52	0.30	0.00
38	0.55	0.45	0.00
43	0.43	0.30	0.00
44	0.72	0.43	0.00
45	0.58	0.36	0.00
46	0.70	0.39	0.00
47	0.58	0.35	0.00
PE	11.07	0.75	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.7: Item Statistics for English II, Spring 2010 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.93	0.29	0.00
2	0.89	0.36	0.00
3	0.49	0.13	0.00
4	0.72	0.18	0.00
5	0.76	0.33	0.00
6	0.94	0.40	0.00
7	0.95	0.29	0.00
8	0.62	0.27	0.00
9	0.54	0.27	0.00
10	0.53	0.22	0.00
11	0.55	0.24	0.00
12	0.37	0.21	0.00
25	0.79	0.47	0.00
26	0.69	0.24	0.00
27	0.53	0.30	0.00
28	0.88	0.49	0.00
29	0.80	0.41	0.00
30	0.78	0.40	0.00
31	0.68	0.34	0.00
32	0.73	0.37	0.00
33	0.81	0.55	0.00
34	0.77	0.49	0.00
35	0.70	0.33	0.00
36	0.76	0.44	0.00
37	0.46	0.37	0.00
38	0.75	0.42	0.00
39	0.74	0.39	0.00
40	0.72	0.39	0.00
41	0.68	0.23	0.00
42	0.84	0.43	0.00
43	0.79	0.42	0.00
44	0.70	0.31	0.00
45	0.36	0.16	0.00
46	0.84	0.36	0.00
47	0.57	0.32	0.00
PE	2.96	0.45	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.8: Item Statistics for Algebra I, Spring 2010 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.74	0.34	0.00
2	0.83	0.40	0.00
3	0.65	0.22	0.00
4	0.88	0.28	0.00
5	0.79	0.33	0.00
10	0.91	0.37	0.00
11	0.81	0.44	0.00
12	0.84	0.38	0.00
13	0.51	0.47	0.00
14	0.77	0.37	0.00
15	0.78	0.35	0.00
16	0.48	0.27	0.00
17	0.59	0.36	0.00
18	0.56	0.35	0.00
19	0.68	0.41	0.00
20	0.59	0.38	0.00
21	0.64	0.37	0.00
26	0.59	0.34	0.00
27	0.50	0.36	0.00
28	0.68	0.48	0.00
29	0.66	0.35	0.00
30	0.63	0.43	0.00
31	0.57	0.21	0.00
32	0.76	0.36	0.00
33	0.52	0.27	0.00
34	0.69	0.57	0.00
35	0.62	0.46	0.00
36	0.66	0.40	0.00
37	0.28	0.16	0.00
38	0.28	0.27	0.00
43	0.28	0.30	0.00
44	0.24	0.01	0.00
45	0.41	0.18	0.00
46	0.43	0.31	0.00
47	0.43	0.39	0.00
PE	1.93	0.60	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.9: Item Statistics for Biology, Spring 2010 Operational Administration

Item	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate %
1	0.93	0.33	0.00
2	0.93	0.22	0.00
3	0.84	0.26	0.00
4	0.79	0.41	0.00
5	0.84	0.32	0.00
10	0.82	0.34	0.00
11	0.76	0.49	0.00
12	0.80	0.43	0.00
13	0.72	0.21	0.00
14	0.60	0.36	0.00
15	0.64	0.25	0.00
16	0.83	0.35	0.00
17	0.62	0.33	0.00
18	0.63	0.47	0.00
19	0.65	0.37	0.00
20	0.81	0.38	0.00
21	0.56	0.36	0.00
26	0.65	0.34	0.00
27	0.70	0.29	0.00
28	0.69	0.50	0.00
29	0.62	0.40	0.00
30	0.76	0.48	0.00
31	0.59	0.36	0.00
32	0.44	0.16	0.00
33	0.50	0.30	0.00
34	0.56	0.30	0.00
35	0.43	0.03	0.00
36	0.53	0.35	0.00
37	0.53	0.34	0.00
38	0.43	0.16	0.00
43	0.53	0.36	0.00
44	0.51	0.36	0.00
45	0.37	0.30	0.00
46	0.39	0.21	0.00
47	0.38	0.21	0.00
PE	9.17	0.71	N/A

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

4.3 Speededness

The consequence of time limits on examinees' scores is called speededness. A test is speeded if examinees taking it score lower than they would have had the test not been timed. Most speededness statistics are based on the number of items that were not attempted by students. For the purpose of this analysis, if a student did not attempt the last item on any of the separately timed subsections of the test, it was assumed that the student might not have reached the item because he or she ran out of time.

The MO EOC Assessments were not designed to be speeded tests. Rather, they were intended to be "power tests"; that is, all students were expected to have ample time to finish all items and prompts.

The last column in Tables 4.1 through 4.9 shows the percentage of students who omitted each SR item for each MO EOC Assessment. It is clear from the tables that the omit rates are negligible or zero for the majority of items.

4.4 Item Bias Statistics

DIF occurs when an item has difficulty measures that vary across contexts for similarly able subgroups of examinees. Using the Spring 2008 standalone field test data, DIF was examined with the Mantel-Haenszel (MH) (1959) procedure for the SR items and a Rasch DIF analysis using WINSTEPS (v3.64, Linacre, 2006b) for the PE/WP items.

The Mantel-Haenszel method is a nonparametric approach to DIF. In the MH procedure, total raw scores are held constant while the odds ratio is estimated. In practice, the odds ratio is generally converted to the delta metric, and the Educational Testing Service (ETS) categorization is applied to flag the significance of DIF effects (Dorans and Holland, 1993).

With the groups matched on raw score, the comparable examinees can be placed in *j* 2×2 tables of group by item response, where *j* equals the number of levels of the matching variable. For these analyses, *j* equals each observed score category of the *k*-item tests, with $j = 0, 1, 2, \ldots, k$, then one 2×2 table for a given item with score category *j* can be represented as

	Correct	Incorrect	Total
Reference	y_j	x_j	m_j
Focal	y'_j	x'_j	m'_j
Total	n_j	n'_j	N_j

The Delta MH test statistic and variance have the following form:

$$DeltaMH = 2.35 \ln \frac{\left[\sum_{j=0}^{K} \frac{\left(y_j x'_j - y'_j x_j \right)}{N_j} \right]}{\sum_{j=0}^{K} \frac{y'_j x_j}{N_j}} ,$$

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Where y_j, x_j, y'_j, and x'_j are the frequency counts of cells of the 2 × 2 tables, and N_j is the total n for the cells.

The critical values of the ETS categorizations are 1.00 and 1.50 on the delta scale for categories A, B, and C. Specifically, if the absolute value of delta is smaller than 1.00, the item is categorized as A. If the absolute value of delta is larger than or equal to 1.50, the item is classified as C. Otherwise, items are categorized as B. In both the A and C categories, statistical significance is set at the 5% level for a single item.

DIF detection with WINSTEPS is a Rasch-model-based approach. According to Linacre (2006a), detecting DIF using WINSTEPS requires the following steps:

- A joint run with all persons and all items is used to produce anchor values (i.e., ability and rating [or partial credit] scale structure).
- A subgroup run (reference group) with person abilities (or partial credit) scale structure anchored is used to produce group *R* item difficulties (D_R).
- Another subgroup run (focal group) with person abilities (or partial credit) scale structure anchored is used to produce group *F* item difficulties (D_F).
- DIF contrast ($D_F - D_R$) is obtained using the different DIF measures of the two subgroup runs.
- A *t*-test statistic provides significance values as a unit normal deviate.

The Rasch and Mantel-Haenszel procedures for DIF are equivalent under certain conditions (Linacre and Wright, 1989; Schulz et al., 1996). Similar to the ETS classifications, the DIF output yielded by WINSTEPS is classified as negligible (A), slight to moderate (B), or moderate to severe (C). If a *t*-value is smaller than 2.58 or the DIF contrast is smaller than 0.45 logits, the item is flagged as A. If a *t*-value is larger than 2.58 and the DIF contrast is larger than 0.65 logits, the item is flagged as C. Otherwise, items are flagged as B. This categorization seems to be slightly more conservative than the ETS categorizations using the MH odds ratio (Liu and Mix, 2006).

Results of the DIF analyses for the items contained in the Summer 2009, Fall 2009, and Spring 2010 operational administrations are summarized in Table 4.10. Tables 4.11 and 4.12 contain DIF statistics for the entire pool of MO EOC Assessment items.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.10: Frequency Distribution of DIF Categories by Item Type for the Summer 2009, Fall 2009, and Spring 2010 Operational Assessments

Test	Group***	Selected Response Items*						PE/WP Items*					
		A**	A–**	B**	B–**	C**	C–**	A**	A–**	B**	B–**	C**	C–**
Summer 2009													
English II	M/F	34	0	1	0	0	0	0	0	1	0	0	0
	W/B	35	0	0	0	0	0	1	0	0	0	0	0
	W/H	32	0	2	1	0	0	1	0	0	0	0	0
Algebra I	M/F	35	0	0	0	0	0	1	0	0	0	0	0
	W/B	33	0	0	2	0	0	1	0	0	0	0	0
	W/H	34	0	1	0	0	0	1	0	0	0	0	0
Biology	M/F	32	0	1	2	0	0	10	0	0	0	0	0
	W/B	33	0	2	0	0	0	9	0	0	1	0	0
	W/H	31	0	1	3	0	0	10	0	0	0	0	0
Fall 2009													
English II	M/F	35	0	0	0	0	0	1	0	0	0	0	0
	W/B	34	0	1	0	0	0	1	0	0	0	0	0
	W/H	34	0	0	1	0	0	1	0	0	0	0	0
Algebra I	M/F	34	0	0	1	0	0	1	0	0	0	0	0
	W/B	35	0	0	0	0	0	1	0	0	0	0	0
	W/H	35	0	0	0	0	0	1	0	0	0	0	0
Biology	M/F	35	0	0	0	0	0	11	0	0	0	0	0
	W/B	33	0	1	1	0	0	10	0	1	0	0	0
	W/H	32	0	2	1	0	0	11	0	0	0	0	0
Spring 2010													
English II	M/F	35	0	0	0	0	0	1	0	0	0	0	0
	W/B	33	0	1	1	0	0	1	0	0	0	0	0
	W/H	33	0	1	1	0	0	1	0	0	0	0	0
Algebra I	M/F	35	0	0	0	0	0	1	0	0	0	0	0
	W/B	30	0	0	5	0	0	1	0	0	0	0	0
	W/H	34	0	1	0	0	0	1	0	0	0	0	0
Biology	M/F	35	0	0	0	0	0	12	0	0	0	0	0
	W/B	33	0	1	1	0	0	10	0	0	2	0	0
	W/H	34	0	1	0	0	0	12	0	0	0	0	0

Note: Classifications with a negative sign ("–") favor the reference group, while classifications with no sign favor the focal group.

* The Mantel-Haenszel procedure is applied for the SR items and WINSTEPS for the PE/WP items.

** DIF categories: A, negligible; B, slight to moderate; and C, moderate to severe.

*** DIF contrast groups: M/F, male versus female; W/B, white versus black; and W/H, white versus Hispanic.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 4.11: Frequency Distribution of DIF Categories by Item Type for the Entire Pool of MO EOC Assessment Items (Spring 2008 Field Test)

Test	Group***	Selected Response Items*						PE/WP Items*					
		A**	A–**	B**	B–**	C**	C–**	A**	A–**	B**	B–**	C**	C–**
English II	M/F	281	0	8	7	1	3	7	0	3	0	0	0
	W/B	285	0	8	6	0	1	10	0	0	0	0	0
	W/H	285	0	5	10	0	0	10	0	0	0	0	0
Algebra I	M/F	178	0	0	2	0	0	10	0	0	0	0	0
	W/B	161	0	7	10	1	1	8	0	0	2	0	0
	W/H	167	0	9	4	0	0	10	0	0	0	0	0
Biology	M/F	173	0	3	4	0	0	105	0	0	0	0	0
	W/B	169	0	3	6	0	2	97	0	2	6	0	0
	W/H	169	0	5	6	0	0	104	0	0	1	0	0

Note: Classifications with a negative sign ("–") favor the reference group, while classifications with no sign favor the focal group.
* The Mantel-Haenszel procedure is applied for the SR items and WINSTEPS for the PE/WP items.
** DIF categories: A, negligible; B, slight to moderate; and C, moderate to severe.
*** DIF contrast groups: M/F, male versus female; W/B, white versus black; and W/H, white versus Hispanic.

Table 4.12: Frequency Distribution of DIF Categories by Item Type for the Entire Pool of MO EOC Assessment Items (Spring 2009 Embedded Field Tests)

Test	Group***	Selected Response Items*						PE/WP Items*					
		A**	A–**	B**	B–**	C**	C–**	A**	A–**	B**	B–**	C**	C–**
English II	M/F	365	0	5	12	0	2	1	0	7	0	2	0
	W/B	344	0	11	24	0	5	10	0	0	0	0	0
	W/H	365	0	7	9	1	2	10	0	0	0	0	0
Algebra I	M/F	277	0	2	9	0	0	24	0	0	0	0	0
	W/B	257	0	14	15	0	2	24	0	0	0	0	0
	W/H	271	0	9	7	1	0	23	0	0	1	0	0
Biology	M/F	279	0	4	4	0	1	172	0	2	3	0	2
	W/B	263	0	14	10	0	1	150	0	10	12	2	5
	W/H	274	0	7	6	0	1	172	0	4	2	1	0

Note: Classifications with a negative sign ("–") favor the reference group, while classifications with no sign favor the focal group.
* The Mantel-Haenszel procedure is applied for the SR items and WINSTEPS for the PE/WP items.
** DIF categories: A, negligible; B, slight to moderate; and C, moderate to severe.
*** DIF contrast groups: M/F, male versus female; W/B, white versus black; and W/H, white versus Hispanic.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

4.5 Summary

The item analyses provided in this chapter show that the MO EOC Assessments have sound psychometrics properties. For example, *p*-values show that MO EOC Assessment items measure achievement across a broad range of difficulty. Also, item discrimination values show that most items are appropriately correlated with the total test score and thus contribute to distinguishing between lower-performing and higher-performing students. In addition, very few students omitted items during testing. The low percentage of students omitting selected response items provides evidence that the test is a power test of the students' skills and not a speeded test. Finally, item bias statistics based on data from the 2008 standalone field test and the 2009 embedded field test administrations show the items to be generally free from statistical bias.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 5: TEST ADMINISTRATION

5.1 Introduction

This chapter contains information about the Missouri Department of Elementary and Secondary Education (DESE) and Riverside Publishing processes that ensure the standardized administration of the Missouri End-of-Course (MO EOC) Assessments. The *Standards* (AERA, APA, and NCME, 1999) state that, "For tests designed to assess the examinee's knowledge, skills, or abilities, standardization helps to ensure that all examinees have the same opportunity to demonstrate their competencies" (p. 61). In other words, careful attention to the details of information dissemination, Test Examiner training, accommodations and modifications, and test security help ensure that students taking the EOC Assessments in different locations have equal opportunities for success.

The EOC *Test Coordinator's Manual* and *Test Examiner's Manual* contain detailed information about the testing guidelines, materials handling, and standardized administration instructions for the EOC Assessments. While those manuals are not included here, much of the information contained in this chapter can be found in them.

For the MO EOC Assessments, districts can choose either a paper-and-pencil or online delivery format. The *Test Coordinator's Manual* and the *Online Test Examiner's Manual* contain information specific to the registration for and administration of the online version of the MO EOC Assessments. Relevant information related to the online delivery, where it differs from the paper-and-pencil format, is included in this chapter.

5.2 Students for Whom the EOC Assessments Are Appropriate

The responsibility and authority for testing students in the Missouri EOC Assessments at the appropriate time in the course of instruction belongs to the local district. The EOC Assessments are based on Course-Level Expectations (CLEs) rather than on Grade-Level Expectations (GLEs). Therefore, when the content of the CLEs is covered in the local school district's curriculum, the test may be administered regardless of student grade level or course name.

5.2.1 Students with Individualized Education Program

A student with disabilities, as classified under the Individuals with Disabilities Education Act (IDEA), has an Individualized Education Program (IEP) that, in part, governs whether a particular assessment is appropriate for the student. In the case of the EOC Assessments, decisions about whether a student with a disability will participate in the assessments are made by the student's IEP team and are documented in the IEP. All students must take the three Phase I EOC Assessments (English II, Algebra I, and Biology), plus the Government EOC Assessment from Phase II. If, however, a student's disability qualifies him or her to take the MAP-Alternate Assessment (MAP-A), that student will not be required to participate in the EOC Assessment.

5.2.2 Students with Individual Accommodation Programs

Students with Individual Accommodation Programs (IAPs) are considered disabled under Section 504 of the 1973 Rehabilitation Act. These students are not served under IDEA

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

and are not documented with a particular designation for the EOC Assessment. However, professionals who are knowledgeable about a student's disability and educational needs should make accommodation decisions for the student as they would for a student with an IEP.

5.2.3 English Language Learner Students

Students who have been in the United States for 12 consecutive months or less at the time of test administration may be exempted by the local school district from taking EOC Communication Arts assessments.

5.3 Students for Whom a School or District Is Accountable

For accountability purposes, Missouri must include the results for any student who is eligible to take the EOC Assessments and has been enrolled at least one full academic year in a school (for school accountability) or district (for district accountability) without transferring out of the building or district for a significant period of time and re-enrolling. A full academic year is defined as the last Wednesday in September through the EOC Assessment administration. A significant period of time is considered "one more than half of the eligible days between the last Wednesday in September and the test administration." DESE obtains enrollment information from the Missouri School Information System (MOSIS) data that are reported by school districts. This rule applies to the building and district summary levels independently. For example, a student who is coded as "In building less than a year," but was in the district a full academic year, is excluded from the building totals but is included in the district totals.

5.4 Dissemination of Testing Materials and Information

Riverside Publishing works with Questar Assessment, a subcontractor for the EOC Assessment program, to gather all enrollment counts and distribute all paper-and-pencil testing materials. Riverside Publishing distributes all password information for the online system. Before the start of the test window, districts enter their enrollment counts and scheduled testing window into ServicePoint, an online enrollment and materials ordering system. From those enrollment counts, Questar generates each district's order. All paper-and-pencil materials are shipped one week before the district's designated testing window. Districts that administer the assessments online receive an e-mail message with password information one week prior to test administration. The District Test Coordinator (DTC) is responsible for inventorying all paper-and-pencil materials, as well as for distributing the online test information to the test administrators. If additional materials are needed, the Test Coordinator is responsible for placing an Additional Materials Order (AMO) through ServicePoint.

5.5 District and Test Examiner Training

DESE is responsible for training the Test Coordinators on EOC test administration. The Regional Instructional Facilitators (RIFs) are first trained by the Assistant Director of Assessment on all information covered in the *Test Coordinator's Manual* and the *Test Examiner's Manual*. The RIFs then conduct training sessions for the districts within their region. The RIFs also provide assistance with test administration and serve as a liaison

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

between DESE and the districts. Both DESE and Riverside Publishing are available to answer any questions the districts may have about the EOC Assessment administration.

Riverside Publishing provides training to districts that administer the EOC Assessments online. The hour-long training session is conducted via WebEx and gives an overview of both the administrative and student sides of the online system.

5.6 Test Security

The EOC Assessment test books and online assessment are secure. Test Coordinators are instructed to keep the materials in a locked room or cabinet at all times when not in use. No testing materials may be photocopied, duplicated, scanned, or made accessible to personnel who are not responsible for testing. Additionally, written or oral discussion of specific EOC Assessment items breaches the security and integrity of the test. In accordance with the *Standards*, the *Test Coordinator's Manual* and *Test Examiner's Manual* contain explicit instructions about test security for Test Coordinators and Test Examiners.[6] When the tests are delivered online, Test Examiners do not have access to the student screens for the online assessment, only to the test administrator features. In addition, a secure browser must be installed on each student computer prior to administration of the online assessments. Test items, as well as student responses, are encrypted during transmission to and from student computers.

5.7 Test Administration

5.7.1 Test Organization

Students take the EOC Assessments in two sessions. Session I contains only selected response (SR) items. Each item consists of a stem followed by four response options. Session II contains the Performance Event/Writing Prompt (PE/WP). PEs allow insight into the student's ability to apply knowledge and understanding to real-life situations. The WP, a special type of Performance Event that appears in the English II Assessment, is an open-ended item that requires students to demonstrate their on-demand writing proficiency. The amount of time per session varies with the content area; however, the tests are not timed.

Session I and II items are contained in separate test books for the paper-and-pencil version. Session I test books contain only SR items. Answers are marked on a separate answer sheet. Session II test books contain the WP (for English II) or the PE items (for Algebra I and Biology). The Session II test books are scannable; students write their responses directly in the test books.

For the online assessment, the MO EOC Assessments also comprise two sessions. The sessions are designed to be administered in approximately two testing periods with times varying by content area. However, as in the case of the paper-and-pencil administration, the tests are not timed. Students are required to complete the practice tests on the DESE website prior to testing. These practice tests include instructions on how to use the tools in the system and practice questions for the students.

---

[6] **Standard 5.7:** Test users have the responsibility of protecting the security of test materials at all times (p. 64).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

For each SR item in Session I, the student clicks an answer choice. For Session II, the student completes the PE/WP with the help of online tools contained in the system.

5.7.2 Test and Ancillary Materials

District Test Coordinators or School Test Coordinators are responsible for distributing all EOC Assessment materials to Test Examiners. The materials provided by Riverside Publishing and/or DESE include the following:

- *Test Examiner's Manual* (online and paper-and-pencil)
- *Test Coordinator's Manual*
- Building Identification Sheets
- Group Identification Sheets
- Student barcode labels
- Session I test books
- Session I answer sheets
- Session II test books
- Math reference sheets (if applicable)
- Return kit materials

Students need the following additional materials for the paper-and-pencil assessment; these materials are not provided by Riverside Publishing or DESE:

- No. 2 pencils
- Scratch paper
- Metric ruler (for Biology)
- Dictionary, thesaurus, grammar book (for English II)

For the online assessment, each student needs a computer with a monitor, a mouse, and a keyboard. Adequate space should be left between workstations. Students can use scratch, grid, or draft paper and a writing utensil while taking the online assessment. The Test Examiner needs the following:

- A computer for logging on to the proctor interface
- A writing board and utensil

Additionally, students taking either the paper-and-pencil or online version may use a four-function calculator for the Algebra I assessment. (This is not required.)

5.7.3 Preparing the Classroom and the Students

The *Test Examiner's Manual* contains specific instructions for teachers and other test administrators regarding how the classroom should be prepared for testing. These include the following:

- Planning for the distribution and collection of materials
- Planning the seating arrangement to prevent students from seeing other students' responses
- Eliminating distractions such as bells and telephones
- Using a "Do Not Disturb" sign on the door

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Removing from students' view any classroom maps, charts, or other materials that relate to the test content
- Making arrangements for students who may not finish testing in the allotted time

Before students begin the assessment using the online system, a representative of the district or school must do the following:

- Read the entire *Online Test Examiner's Manual*
- Run a system check on each workstation used for testing
- Ensure that the MO EOC browser is downloaded to each workstation for test delivery
- Read the frequently asked questions from the link on the Test Examiner's login page
- Input identification information for students who were not included in the MOSIS precode file
- Contact Riverside Publishing if any changes need to be made to the student roster
- Create a test session immediately before testing

Additionally, while students await proctor approval, the Test Examiner must set and verify class information and set students' testing status codes and/or accommodations information in the online system.

The *Test Examiner's Manual* and *Online Test Examiner's Manual* explain some ways teachers may prepare their students for testing, including the following:

- Helping students approach the testing with a relaxed, positive attitude
- Encouraging and motivating students to do their best work
- Explaining test strategies, such as skipping harder items and coming back to them later
- Reassuring students that they will be given ample time to do their best work

Students are not allowed to use electronic devices, such as cellular phones, digital cameras, gaming devices, or scanners during the testing session. However, students may use four-function calculators during the Algebra I test session.

5.7.4 Directions for Administration

In accordance with Standard 5.1,[7] specific standardized directions for administration are printed in the *Test Examiner's Manual*. Directions that are to be read aloud to the students are printed in **bold** type and have a callout arrow in the margin for clarity. Information for the teacher that should not be read aloud is in *italic* type. Figure 5.1 provides an example of the type styles used in the *Test Examiner's Manual* to differentiate between spoken and unspoken instructions. Figure 5.2 provides an example of a script from the English II EOC Assessment. Figure 5.3 provides an example of a script from the online English II EOC Assessment.

---

[7] **Standard 5.1:** Test administrators should follow carefully the standardized procedures for administration and scoring specified by the test developer, unless the situation or a test taker's disability dictates that an exception should be made (p. 63).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 5.1: Examples of Type Styles Used to Differentiate between Spoken and Unspoken Instructions in the *Test Examiner's Manual*

SAY **The directions you are to read aloud to the students are preceded by the word "SAY" in a box.**

Information that is only for you and not to be read aloud is printed in italic type.

Figure 5.2: Example Script from the *Test Examiner's Manual* for the Algebra I EOC Assessment

TEACHER DIRECTIONS:

Before administering the test, be sure that students understand what each picture means and make sure each student has the appropriate materials.

means that a student may want to use the reference sheet.

Before administering the test, take a moment to have the students look through the test book. Point out different "STOPs" (words or signs). Tell the students that whenever they see one of the "STOPs," they should not go on.

Distribute the test book, reference sheet, and scratch paper. Ensure that all students use a non-mechanical No. 2 pencil. If you have decided the students should use calculators for this part of the test, make sure all students have a working calculator. Remember, the use of a calculator is not necessary.

Instruct students to bubble in the appropriate test window (Fall, Spring, Summer) on the Student Information Sheet located on the back of the Session I answer sheet. For students testing in Spring 2009 only, please fill in the FORM number that corresponds to the student's Session I Test Book.

SAY **For the questions in this session, you will select an answer from a list of given choices. Use scratch paper or graph paper to work the problems. Do not include ANY of your work in the test book for Session I. Remember to fill in the circle on the answer document that goes with the answer you chose. Your score on these questions will depend on how well you follow directions and show your understanding of Algebra I. The reference sheet and calculator can be used in Session I.**

SAY **Open your Session I test book to page 2.**

Check to see that all students are on the correct page in their test books.

SAY **When you come to the word "STOP," you have finished with Session I. You may go back over Session I of the test and check your answers. When you have finished checking your answers, close your test book and sit quietly until everyone has finished. Do you have any questions?**

When you are sure that all students understand the directions, continue.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 5.3: Example Script from the *Online Test Examiner's Manual* for the Online English II EOC Assessment

SAY

For the questions in this session, you will select an answer from a list of given choices. Remember to check that the circle that goes with the answer you chose is filled in after you click it. Your score on these questions will depend on how well you follow directions and show your understanding of what you read. You may choose to look over the questions before reading the passage. You may NOT use a dictionary, thesaurus, or grammar handbook during this session of the test.
See the Help button for instructions on how to use the system tools.

There are several important things to remember:

1. Read each question carefully and think about the answer. Then choose the one answer that you think is best.
2. If you do not know the answer to a question, mark it for review, skip it, and go on. You may return to it later.
3. When you finish the test, you may check over your work.

When you have finished checking your answers, click the End button,. Exit the browser, and sit quietly until everyone has finished. Are there any questions?

5.8 Accommodations and Modifications

A student's IEP team has the responsibility and authority to determine individual accommodations to support and ensure his or her participation in the EOC Assessments. Allowable accommodations are intended to assist the student by reducing the effects of his or her disability without reducing performance expectations. Allowable accommodations for the EOC Assessments include the following:

- A student may receive a Braille or Large Print edition.
- A teacher may present the test content to a student in a nonstandard way, such as by reading it aloud in English or in the student's native language, paraphrasing it, or using sign language. For the English II Assessment, this will result in the lowest obtainable scale score (LOSS).
- A student may be allowed additional time to complete one or more sessions of the assessment.
- A student may use an assistive communicative device.
- A student may be tested individually or in a small group.
- A student may be allowed to use a computer, another word-processing device, or a teacher scribe to record his or her responses.
- A student may use other assistive materials such a calculator (on the English II or Biology Assessment) or a bilingual dictionary.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Modifications are alterations in the test that change construct-related requirements. The resulting information may not be equal to the information that might be obtained without modifications. While modifications invalidate the use of student scores for No Child Left Behind (NCLB) accountability determinations, the following modifications for the EOC Assessments can be provided:

- Oral reading of the English I and English II Assessment, including paraphrasing questions
- Oral reading in a student's native language
- Use of a bilingual dictionary for any part of the English I or English II, Session I, Assessment

As noted above, the modifications listed will result in the lowest obtainable scale score (LOSS) on the EOC Assessments. For more information on accommodations and modifications and their effects on the interpretation of the EOC Assessment scores, see the Appendix to the *Test Examiner's Manual*.

In accordance with Standard 5.2,[8] Test Examiners indicate an accommodation, when allowed by a student's IEP and used for the EOC Assessment, by filling in the bubble corresponding to the accommodation on page 1 of the Session I answer sheet.

Table 5.1 contains information about the percentage of students who received each type of allowable accommodation for each EOC Assessment. The most prevalent type of accommodation across all three EOC Assessments was testing in a small group (provided to between 1.81% and 3.88% of students across assessments).

Table 5.1: Frequency and Percentage of Students Receiving Each Type of Allowable Accommodation on the EOC Assessments

Accommodation	English II		Algebra I		Biology	
	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Braille	4	0.01	6	0.01	5	0.01
Large Print	18	0.03	20	0.03	21	0.03
Oral Reading	17	0.03	1311	1.95	1758	2.70
Oral Reading—Blind/Partial Sight	2	0.00	52	0.08	47	0.07
Signing of Assessment	0	0.00	18	0.03	4	0.01
Paraphrasing	1	0.00	6	0.01	2	0.00
Other Administrations	2	0.00	3	0.00	4	0.01
Oral Reading in Native Language	1	0.00	41	0.06	37	0.06
Extended Time	1368	2.09	1264	1.88	1284	1.97
Administered Using More Than Allotted Periods	321	0.49	325	0.48	295	0.45
Other Timing	28	0.04	27	0.04	33	0.05
Use of Scribe	160	0.24	102	0.15	149	0.23

[8] **Standard 5.2:** Modifications or disruptions of standardized test administration procedures or scoring should be documented (p. 63).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 5.1: Frequency and Percentage of Students Receiving Each Type of Allowable Accommodation on the EOC Assessments (continued)

	English II		Algebra I		Biology	
Accommodation	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Use of Calculator, Math Tables, etc.	73	0.11	812	1.21	545	0.84
Using Bilingual Dictionary	0	0.00	36	0.05	42	0.06
Other Response	12	0.02	14	0.02	11	0.02
Testing Individually	207	0.32	206	0.31	223	0.34
Testing in Small Group	2966	4.54	2841	4.23	3068	4.72
Other Setting	79	0.12	90	0.13	91	0.14

5.9 Materials Handling and Return

The *Test Coordinator's Manual* and *Test Examiner's Manual* contain detailed instructions for how schools and districts should collect and package the paper-and-pencil testing materials at the end of the test administration. For Test Examiners, these activities include, but are not limited to, the following:

- Collecting test books and answer sheets from the students
- Counting the test books and answer sheets and comparing the number to the totals from pretesting
- Returning all used and unused test books and answer sheets to the School Test Coordinator
- Collecting all scratch paper used during testing
- Properly handling all contaminated test books (i.e., books having contact with bodily fluids such as blood or with any potentially hazardous material)
- Verifying that the barcode labels are affixed properly to the answer sheets and Session II test books
- Verifying that the information contained on the Student Information Sheet (SIS) is accurate and compete

For School Test Coordinators, these activities include, but are not limited to, the following:

- Collecting testing materials from the Test Examiners
- Counting all test books and verifying against the pretesting total
- Completing Group Identification Sheet for each class
- Verifying that the Building Identification Sheets are correct or completing new Building Identification Sheets if incorrect
- Returning all answer sheets and test books (scorable and nonscorable) to the District Test Coordinator
- Destroying all unused answer sheets and other nonsecure testing materials

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

After receiving the answer sheets and scorable and nonscorable test books from the School Test Coordinators, District Test Coordinators complete the following steps:

- Verify 100% return of test books
- Complete the Test Book Accountability Form and fax it to Riverside Publishing
- Verify that each group of scorable materials is accompanied by a Group Identification Sheet
- Verify that Group Identification Sheets are used consistently for Session I and Session II scorables
- Return all EOC Assessment materials to the Assessment Resource Center (ARC) following the packaging and shipping instructions outlined in the *Test Coordinator's Manual*

For the online system, the student needs to click the End button once he or she has finished testing to submit the test for scoring. No additional information is needed from the Test Examiner after the student has completed the test. All demographic information is edited or added by the test administrator before the student starts the assessment.

5.10 Summary

The distribution, administration, and collection of the EOC Assessments is carefully communicated and executed in the detailed *Test Examiner's Manual* and *Test Coordinator's Manual*. All standards related to test security, administration, and accommodations are adhered to throughout the process. The most important steps and procedures have been covered in this chapter. Readers interested in further detail should consult the *Test Examiner's Manual* and *Test Coordinator's Manual* for the EOC Assessments.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 6: SCANNING, SCORING, AND QUALITY CONTROL PROCEDURES

6.1 Introduction

This chapter describes the processes used to scan, score, and provide quality control for the Missouri End-of-Course (MO EOC) Assessments. The EOC Assessment forms containing the selected response (SR) items were processed and scored by Riverside Publishing. The Performance Event (PE) and Writing Prompt (WP) items were processed and scored by the Assessment Resource Center (ARC) at the University of Missouri. This chapter is divided into two main parts. Sections 6.2 through 6.6 pertain to Riverside Publishing's scanning, scoring, and quality control processes for the selected response (SR) items. Sections 6.7 through 6.15 outline the processes ARC used to develop scoring materials for the PE/WP items, receive and scan student responses, hire and train scorers, score the PE/WP items, and maintain control of the quality of the scoring processes.

6.2 Quality Control Overview

Riverside Publishing adheres to the guidelines listed in the SCASS/TILSA *Quality Control Checklist for Processing, Scoring, and Reporting* provided by the Council of Chief State School Officers (2003). Quality assurance in processing, scoring, and reporting is the highest consideration in all stages of score report delivery. Additionally, Standard 5.8[9] of the *Standards* (AERA, APA, and NCME 1999) specifically addresses the issue of quality control in the scoring process. To comply with this standard, Riverside Publishing employed a set of checks at each stage in the process of scoring and reporting the SR items to ensure a zero error rate for the MO EOC Assessments. Riverside Publishing documented the various quality control procedures through a variety of reports and checklists during both the preproduction and post-production phases. Documentation took the form of issues logs and quality audit reports.

6.3 Preparation and Materials Check-In

6.3.1 Preparation for Processing

Before any MO EOC Assessment answer documents were processed for operational testing, Riverside Publishing programming staff conducted a complete check of scanning programs using the program specifications and a transfer file. A test set of documents was gridded to include all response ranges, ID ranges, blanks, double grids, all correct responses, all incorrect responses, and other scenarios, depending on the specified scoring rules. These mock data were then processed through the scanning program, the editing programs, and the scoring system. The resulting file was thoroughly hand checked to ensure that the machine was scanning correctly, that the pre-edit program was picking up the proper errors, that the post-edit program was accepting corrections properly, and that the scoring system was applying the answer keys correctly. If any errors were found in the programs, the programmer was notified to make the corrections, and quality control checks were run again.

[9] **Standard 5.8**: Test scoring services should document the procedures that were followed to assure accuracy of scoring. The frequency of scoring errors should be monitored and reported to users of the service on reasonable request. Any systematic source of scoring errors should be corrected (p. 64).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Once the programs were found to be functioning correctly, a batch of live data was processed. This pilot run involved test results from one Missouri district. The resulting data file was put through the same quality control procedures described above, and documents were hand checked against the transfer file created. Riverside Publishing quality control staff checked and verified the live data. All quality control checks were completed successfully before the rest of the live documents were released for processing. This procedure ensured that the scanning programs were accurate and reliable.

When the MO EOC Assessment documents were first checked in at the Riverside Scoring Service® (RSS), they were issued a barcode number and a color-coded sheet that included vital information about the school. The documents were put into barcoded containers that were scanned at each stage of processing to constantly track the location of a client's documents in the Scoring Center.

Next, RSS staff checked that document counts matched the Return Packing Form and that all submitted materials were complete and included fully completed header sheets. If not, the documents were tagged to alert the Scoring Project Manager and Riverside Publishing Customer Service that resolution was necessary.

6.3.2 Materials Check-In

When a shipment of MO EOC Assessment documents was delivered to the Riverside Publishing Scoring Center, the arrival date, time, carrier type, and number of boxes delivered was immediately recorded in the RSS database, thus starting the clock for processing and delivering score reports. As an additional quality step, one of Riverside Publishing's trained receiving clerks hand counted the boxes and entered the number into the RSS tracking system. Any discrepancies were entered into the alert system for resolution.

Box contents (answer sheets) were verified against the Order for Scoring Services forms, and any discrepancies were entered into the RSS alert resolution system. Each order was issued a unique barcoded number that enabled the order to be tracked as it was processed through the RSS.

6.4 Materials Scanning

All documents were scanned using Scan Optics 9000M scanners, which use four mounted cameras (two on top and two on the bottom) to capture both the grayscale and the bitonal images of each page. As each document was scanned, a Print After Scan (PAS) number was printed on the edge of the document. The first six digits in the PAS were identical to the numbers identifying the container in which the documents moved through the Scoring Center. The last digits represented the order of the document in the stack. The PAS number was used by RSS staff to identify the location of an answer document in the processing system. The scanner read the skunk codes at the top of the page to determine which document code should be used for editing and scoring. Image scanners captured the entire test page as if it were a photocopy.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

6.4.1 Handling of Unscannable Documents

The scanner is programmed to detect anchor points and zones to capture the image. Occasionally, a page cannot be scanned and is automatically sent by the scanner to the rejection bin. When this occurs, the scanner stops. The scanning operator follows procedures to either scan the document correctly or insert an Unscannable Document Header along with the document or page that is unscannable. Some reasons that a document might be unscannable include manipulation during the test administration or pages missing or removed from the answer document before it was submitted for scoring. Photocopied documents are also unscannable.

6.4.2 Resolution of **n***-Count Discrepancies*

Throughout the scanning of the MO EOC Assessment documents, the scanning station was monitored to ensure that images were gathered for all answer documents submitted with each school's or district's materials. A Scan Integrity Report compared the scanned *n*-count with the expected *n*-count on each Group/Class Header Sheet. Any discrepancies were logged into the system and resolved through a physical check of the documents before the container passed to the next station. If a resolution could not be reached, the order was entered into the alert system.

6.4.3 Application of Editing Rules

Riverside Publishing has numerous quality control procedures in place to ensure the accuracy of the scanning of the MO EOC Assessment answer documents. The scoring process applied editing rules to each document as it completed the scanning stage. The editing rules identified conflicts caused either by the student or by the scanner. Examples of these conflicts are double marks, excessive omits, or light marks. Based on these rules, documents were placed in the editing queue for an editor to resolve the conflicts. To ensure that the scanners and the editing rules were working properly, a small percentage of documents from each batch were randomly selected to go to editing, even without any mistakes or errors. If an issue could not be resolved in the editing process, an alert was sent, and a Riverside Publishing alerts specialist contacted the MO EOC Assessments program manager, who worked with the particular school or district to resolve the issue as soon as possible.

Documents that could not be read by the scanner (for instance, because the images were too light, pages were bent, etc.) were manually entered. In these instances, the first editor manually key-entered the student responses. A different editor then manually keyed the student responses a second time. The second editor was not able to see the work of the first editor. Upon completion of the two separate key entries, the system notified the second editor if there were differences in the two entries. If discrepancies were identified, the document was reviewed to determine the correct response.

6.5 Quality Control in Report Production

Riverside Publishing uses OCE V7400, OCE PS372, OCE PS88, and IBM 1245 printers. A trained Riverside Publishing print operator inspects all reports for print quality according to defined tolerances and then reprints any documents that fail this inspection.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

A product assembler collated and sorted the reports for each order into folders. As the reports were put into folders, the assembler conducted a final quality control check, specifically looking at print quality, data integrity, and stray or extra sheets. Using the customer packing list, the assembler tracked the foldering process. The compiled folders were organized according to customer specifications, and the collated order was returned to Quality Control for a final check.

6.6 Quality Assurance Product Review

Riverside Publishing worked with the Missouri Department of Elementary and Secondary Education (DESE) to determine which districts were to be used in the review of the first live order. The Process and Quality Engineering department in the RSS reviewed each score report deliverable. The techniques and procedures followed in the quality assurance plan are defined below.

6.6.1 Techniques

Score reports and data were reviewed for accuracy and completeness in the following ways:

- To verify the accuracy of the data, RSS staff hand scored a sampling of student responses to ensure that the scoring system was functioning according to specification.
- To validate the completeness of the data, RSS staff verified that all records were accounted for in the district General Research File (GRF) and score reports (based on the reporting requirements).
- RSS staff reviewed the score reports to ensure that they met the reporting requirements defined for the MO EOC Assessment program.

6.6.2 Procedures

The quality review was documented in the form of a Quality Audit Report, which outlined the data elements of each score report deliverable that was audited. Nonconformance issues were documented in an issues log and were communicated to the project team.

6.7 Scoring Requiring Human Judgment

Standard 5.9[10] relates specifically to item scoring that requires human judgment. The *Standards* suggest specific procedures that should be followed to ensure that hand scoring of open-ended items is consistent and fair. The following sections outline the processes that were established and followed for hand scoring of the PE/WPs in the MO EOC Assessments.

[10] **Standard 5.9:** When test scoring involves human judgment, scoring rubrics should specify criteria for scoring. Adherence to established scoring criteria should be monitored and checked regularly. Monitoring procedures should be documented (pp. 64–65).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

6.8 Performance Events and Writing Prompts

Riverside Publishing contracted with ARC at the University of Missouri to score student responses to PE/WP items for the MO EOC Assessments. The EOC Assessments for Algebra I, Biology, and English II were field tested in Missouri schools in Spring 2008 and Spring 2009; operational testing for the 2009–2010 school year began in Fall 2009. ARC's specific responsibilities were range-finding, development of training materials for performance scoring, scoring of student responses, and reporting of data to Riverside Publishing.

The EOC Assessments were administered in two sessions: one for SR items and one for PE/WP items. ARC scored only the PE/WP item responses for field tests and operational tests.

The MO EOC Assessments for English II contain a WP, while the Algebra I and Biology Assessments contain PE items. The PE/WP items require students to respond with extended written answers to questions on given topics or to a series of questions regarding specific events.

ARC, in collaboration with DESE, developed *End-of-Course Field-Test Range-Finding Activities* and the *Scoring Guides* and *Training Materials* for each content area. DESE reviewed and approved these documents prior to their use by ARC during scoring.

6.9 Processing Documents and Image Quality Control

After receiving and checking in the testing materials, ARC scanned and processed images of all student responses. ARC used Scantron Insight 150 scanners to scan the Missouri EOC student response test books. Each day before scanning, the scanners were cleaned and calibrated according to the manufacturer's specifications. In addition to the automatic quality control routines built into the scanners, ARC added a more stringent procedure to ensure proper calibration. A special verification program was used to scan a sequence of precision reference documents. The optical read levels from the bubbles on the reference documents were compared to baseline reference tables, and any anomalies were flagged as errors. This procedure eliminated variations in read levels that would pass the hardware quality control procedures but could still affect mark thresholds and demographic identification and/or scoring.

6.9.1 Document Tracking

ARC uses an internal tracking system to document the flow of materials through receiving, scanning, scoring, reporting, and shipping. A tracking ID sheet is used at each step to log the order's progress through the workflow. When daily shipments are received by carrier, all boxes are immediately sorted by district and entered into the ARC tracking system by district and box numbers. The system automatically populates the district information, and the data are compared with the box label information. The number of boxes is entered into the tracking system. Any discrepancies are noted and set aside for resolution before moving to the next stage. ARC begins timing its processing from this point forward to meet requirements for a five-day turnaround of images back to classroom teachers.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Documents are checked a second time during the “cleaning” process. Staff open the boxes to verify the counts and contents of each district’s material. Each test book is counted and entered into the ARC tracking system by school building and subject area. If a count is beyond the Riverside Scoring Counts Tolerance Rule, the boxes are flagged for resolution. An order sheet is created that indicates all district information, date of receipt, due date for imaging, number of documents per building, and subject area. The counts submitted by the school and ARC are recorded.

At the start of scanning, the scanner operator is prompted to scan the barcode on the internal tracking ID sheet. The scanner then logs the order and scan counts for comparison with hand counts from the receiving department. As each Missouri EOC student book is scanned, an entry is made in the tracking database. The demographic and test book identification data are written to ARC’s scoring database when the test book is completely scanned. Each document has a security barcode that is logged and checked for uniqueness at the moment it is scanned. If a duplicate security barcode is encountered, an error is reported and the test book is set aside for resolution.

All receipt of material, scanning of documents, and discrepancies are reviewed by the supervisor. The supervisor verifies the accuracy and completeness of all data and resolves any outstanding discrepancies by independent judgment, examination of material, and/or discussions with RSS.

6.9.2 Imaging

Images of 256-level grayscale quality are saved for every student response test book scored. The demographic and test book identification data are saved to ARC’s scoring database and also with the image in TIFF header fields. This process allows the identifying information to be saved with the image to check it against the database for consistency.

6.9.3 Editing Rules (User Exits)

During scanning, a number of checks are made on each document. The presence and uniqueness of each security barcode is checked. A check is made for a pre-identification barcode. If it is missing, the scanning program is checked to ensure that the demographic bubbles are marked. If they are not marked, the scanner operator is prompted to enter the information if it is available. All available information is entered, and scanning is delayed for missing demographic data. In these cases, images of all pages of the test book are made so that documents with incomplete demographic data can be identified later.

6.10 Range-Finding and Development of Scoring Materials

Following development of the field test and operational test forms, scoring rubrics were written by Riverside Publishing for each individual Algebra I and Biology PE test form. The *Writing Scoring Guide for 11th Grade*, revised and adopted by DESE in 1999, was the designated rubric for scoring English II WP items.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

6.10.1 Spring 2009 Field Tests

Algebra I: In spring 2009, 12 new Algebra I items were field tested. The number of student responses to be scored for each item ranged from approximately 3,500 to 10,500. In summer and fall 2009, the DESE Mathematics Consultant, assisted by two to three Regional Instructional Facilitators (RIFs), conducted range-finding on the 12 items. ARC prepared for and facilitated the range-finding activities.

Prior to range-finding, ARC staff randomly selected 350 student responses from each field test (FT) item and conducted "rough cuts" to determine if the responses exhibited the characteristics of a particular score point and whether the content was appropriate for use in range-finding. During range-finding, the DESE Mathematics Consultant and the RIFs verified the scores and selected student responses that exemplified performance at each score-point level. Student responses that were deemed inappropriate for training were removed from the range-finding process. This process of narrowing papers continued until an appropriate number of samples were chosen to meet requirements for anchors, training papers, qualifying papers, and check-set papers.

Anchor papers are chosen as clear examples of performance at each score point and are used as benchmarks for scoring all papers. One anchor was chosen at each score point for Algebra I.

Training papers illustrate the range of responses within each score point. For Algebra I, two training sets are provided with 10 papers each. One of the two training sets is chosen particularly for use on the Bookette website to train teachers to score their students' papers. In addition, check-set papers are chosen and prescored for inclusion in the scoring process as a quality assurance and training measure.

During range-finding, the anchors and training papers were annotated. Annotations provide clarification of why a student response was given a specific score with descriptive information about the correct answer. The Algebra I annotations are based on each of the elements contained within the scoring rubric.

ARC content staff observed the range-finding activities for training purposes and assisted in the logistics and recording of annotations. Following range-finding, content staff prepared the materials for the training of scorers. Sets of training material were checked for completeness and accuracy and organized into packets for training. All materials were approved by the DESE Consultant prior to use in training.

Biology: In spring 2009, 12 new Biology PEs were field tested; each PE was divided into 2 forms each (24 forms total) with approximately 6,500 responses each for Forms 1 and 2. About 2,300 responses were received for each of the remaining 22 forms. Each form contained eight items with point values ranging from 0–1 to 0–4. ARC randomly selected 400 student responses for each FT item for use in range-finding. In summer 2009, the DESE Science Consultant, assisted by 1–3 RIFs, conducted range-finding on 6 of the 12 PEs. The outlined process for range-finding and training materials construction was repeated in fall and winter 2009–2010 for the remaining 6 PEs.

Each sample response was read to determine if the response exhibited the characteristics of a particular score point and whether the content was appropriate for use in training. Responses for this PE were compared with one another and with responses already

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

scored in other PEs. The comparisons ensured that all scoring would continue to be calibrated equally. The DESE Consultant and the RIFs recorded their scores on each sample and included annotation text when appropriate. Anchors, trainers, qualifiers, and check sets were selected from among these samples.

Anchor samples are chosen as clear examples of performance at each score point and are used as benchmarks for scoring all papers. One anchor was chosen at each score point for all Biology items.

Training samples illustrate the range of responses within each score point. For each Biology item, a training set is provided with 5–12 annotated papers, depending on the score point value and comparative complexity of the item in question. The training set is chosen particularly for use on the Bookette website to train teachers to score their students' papers. In addition, check-set papers are chosen and prescored for inclusion in the scoring process as a quality assurance and training measure.

During range-finding, the anchors and training papers were annotated. Annotations provide clarification for why a student response was given a specific score with descriptive information about the correct answer. The Biology annotations are based on the elements contained within each item's scoring rubric.

ARC content staff facilitated range-finding activities, communicated with the Science Consultant and RIFs during the range-finding activities for training purposes, and assisted in the logistics of sample selection and recording of annotations. Following range-finding, content staff prepared the materials for training scorers. The training materials are checked for completeness and accuracy and organized into packets for training. All materials are approved by the DESE Consultant prior to use in training.

English II: In spring 2009, 20 new English II WPs were field tested, with approximately 2,000 responses received for each. ARC randomly selected 400 student responses for each of the first four prompts selected for range-finding. In summer 2009, the DESE Communications Arts Consultant, assisted by 1–3 RIFs, conducted range-finding on the four prompts.

Each sample paper was read to determine if the paper exhibited the characteristics of a particular score point and whether the text was appropriate for use in training. Responses for this writing task were compared with one another and with responses already scored in earlier assessments. The comparisons ensured that all scoring would continue to be calibrated equally. The DESE Consultant and the RIFs read the responses and selected a first round of papers that exemplified performance at each score-point level. The first round of papers were read again and compared with papers at other performance levels, with additional papers being cut from the process. This process of winnowing papers continued until an appropriate number of samples was chosen to meet the stated number needed for anchors, training papers, qualifying papers, and check-set papers.

The scarcity of 4-level and 1-level papers among the original samples required ARC content staff to randomly select more student papers for consideration. In some cases, ARC prescreened student papers for 4-level and 1-level papers to be presented to the range-finding group.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Anchor papers are chosen as clear examples of performance at each score point and are used as benchmarks for scoring all papers. Three anchors are chosen at each score point for English II, illustrating high, medium, and low performance at each possible score point.

Training papers illustrate the range of responses within each score point. For English II, two training sets are provided with 10 papers each. The first training set is chosen particularly for use on the Bookette website to train teachers to score their students' papers. The second set of papers is chosen to provide scorers with an opportunity for more discussion of borderline papers and split scores. In addition, check-set papers are chosen and prescored for inclusion in the scoring process as a quality assurance and training measure.

During range-finding the anchors and training papers were annotated. Annotations provide clarification of why a student response was given a specific score, with descriptive information about the correct answer. The English II annotations are based on each of the 12 elements contained within the scoring rubric.

ARC content staff observed the range-finding activities for training purposes and assisted in the logistics and recording of annotations. Following range-finding, content staff prepared the materials for training scorers. Training materials are checked for completeness and accuracy and organized into packets for training. All materials were approved by the DESE Consultant prior to use in training.

Range-finding was completed on an additional six field-test prompts in fall 2009. The same procedures described above were used to complete the training and scoring material.

6.10.2 Materials Development for the Summer 2009 Operational Test

The PEs and WP used for operational testing in summer 2009 were field tested in spring 2008. The range-finding activities and materials development are described in the *2009 Missouri EOC Technical Report*. DESE directed ARC to revise the scoring materials that were used to score the field-test PE/WPs as needed so that they were appropriate for use in scoring the Summer 2009 operational PE/WPs.

Algebra I: The DESE Mathematics Consultant revised the rubric language used in Spring 2008 field-test scoring slightly for clarity and revised annotations for the anchors and training papers. The number of anchors changed from three per score point to one per score point at DESE's direction. All changes to the Algebra I training materials were reviewed and approved by the DESE Mathematics Consultant. Materials for the Algebra I scoring included the item's rubric, comment codes, red flag language, course-level expectations (CLEs), anchors, two sets of training samples, and two sets of qualifying papers. No additional sample papers were pulled from the spring 2008 pool of student responses.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Biology: Biology materials for the Summer 2009 operational test were divided into individual scoring training guides for each item (10 items in all). Materials for each item included the item's rubric, comment codes, red flag language, CLEs, anchors, training samples, and two qualifying sets. DESE requested that the approved field-test anchors for each item be reduced to only one anchor per score point (for example, field-test anchors for a 2-point item were trimmed to three anchors only: 0 point, 1 point, and 2 points). DESE also requested that the field-test training sets for each item be supplemented with additional samples. Qualifying sets for all items were standardized to include 10 items per set. All materials produced or revised for use in operational scoring of Biology were reviewed and approved by DESE.

English II: Changes in the format and content of English II annotations were requested by the DESE Communication Arts Consultant, who began working at DESE subsequent to the completion of the Spring 2008 field-test range-finding and scoring. Additional sample papers were drawn from the pool of field-test papers to supplement or replace samples chosen during the original range-finding activities in spring 2008. Changes were developed by the Communications Arts Consultant with input from RIFs and were submitted to ARC for finalization and inclusion in the scoring materials. Anchor and training paper annotations addressed each element of the rubric and cited specific examples from the student paper, when appropriate, to illustrate a specific rubric point.

ARC content staff observed the range-finding activities for training purposes and assisted in the logistics and recording of annotations. Following range-finding, content staff prepared the materials for training scorers. Training materials were checked for completeness and accuracy and organized into packets for training. All materials were approved by the DESE Consultant prior to use in training.

Table 6.1 lists the training materials used for the Summer 2009 operational tests.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.1: Training Materials for Summer 2009 Operational Testing

	Algebra I	Biology	English II
Scoring Guide/Notes	Yes	Yes	Yes
Course-Level Expectation(s)	Yes	Yes	Yes
Rubrics	Yes	Yes	Yes
Anchors	1 set of 5	1 pt—1 set of 2 2 pt—1 set of 3 3 pt—1 set of 4 4 pt—1 set of 5	1 set of 12
Training Papers	2 sets of 10	1 pt—1 set of 5 2 pt—1 set of 6 3 pt—1 set of 8 4 pt—1 set of 10 (on average—some counts varied based on item complexity)	2 sets of 10
Qualifying Papers	2 sets of 10	2 sets of 10 per item	2 sets of 10
Definition of Condition Codes*	Yes	Yes	Yes
Definition of Alert or "Red Flag" Papers**	Yes	Yes	Yes

*Condition codes were assigned to nonscorable papers, indicating the primary reason why a paper could not be scored. The condition codes used for scoring the EOC Assessments are essentially the same for all subjects across all testing windows and are provided in Table 6.1.

**Alert or red flag papers are those that contain writing indicating that the student may be involved in, or may be the victim of, inappropriate, unethical, or criminal behavior.

6.10.3 Materials Development for the Fall 2009 Operational Test Forms

The PE/WPs used for the Fall 2009 operational tests were among those for which range-finding was completed in spring 2008. The range-finding and materials development are described in the *2009 Missouri EOC Technical Report*. Additionally, DESE directed ARC to revise the scoring materials that were used to score the field-test PE/WPs as needed so that they were appropriate for use in scoring the Fall 2009 operational PE/WPs.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Algebra I: The DESE Mathematics Consultant revised the rubric language used in Spring 2008 field-test scoring slightly for clarity and revised annotations for the anchors and training papers. The number of anchors changed from three per score point to one per score point at DESE's direction. All changes to the Algebra I training materials were reviewed and approved by the DESE Mathematics Consultant. No additional sample papers were pulled from the Spring 2008 pool of student responses.

Biology: Biology materials for the Fall 2009 operational test were divided into individual scoring training guides for each item (11 items in all). Materials for each item included the item's rubric, comment codes, red flag language, CLEs, anchors, training samples, and two qualifying sets. DESE requested that the approved field-test anchors for each item be reduced to only one anchor per score point (for example, field-test anchors for a 2-point item were trimmed to three anchors only: 0 point, 1 point, and 2 points). DESE also requested that the field-test training sets for each item be supplemented with additional samples. Qualifying sets for all items were standardized to include 10 items per set. All materials produced or revised for use in operational scoring of Biology were reviewed and approved by DESE.

English II: Changes in the format and content of English II annotations were requested by the DESE Communication Arts Consultant, who began working at DESE subsequent to the completion of the Spring 2008 field-test range-finding and scoring. Additional sample papers were drawn from the pool of field-test papers to supplement or replace samples chosen during the original range-finding activities in spring 2008. Changes were developed by the Communication Arts Consultant with input from RIFs and submitted to ARC for finalization and inclusion in the scoring materials. Anchor and training paper annotations addressed each element of the rubric and cited specific examples from the student paper when appropriate to illustrate a specific rubric point.

ARC content staff observed the range-finding activities for training purposes and assisted in the logistics and recording of annotations. Following range-finding, content staff prepared the materials for training of scorers. Training materials were checked for completeness and accuracy and organized into packets for training. All materials were approved by the DESE Consultant prior to use in training.

Table 6.2 lists the training materials used for the Fall 2009 operational tests.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.2: Training Materials for Fall 2009 Operational Testing

	Algebra I	Biology	English II
Scoring Guide/Notes	Yes	Yes	Yes
Course-Level Expectation(s)	Yes	Yes	Yes
Rubrics	Yes	Yes	Yes
Supplemental Training Notes for Supervisor and Team Leaders	No	Yes	Yes
Anchors	1 set of 5	1 pt—1 set of 2 2 pt—1 set of 3 3 pt—1 set of 4 4 pt—1 set of 5	1 set of 12
Training Papers	2 sets of 10	1 pt—1 set of 5 2 pt—1 set of 6 3 pt—1 set of 8 4 pt—1 set of 10 (on average—some counts varied based on item complexity)	2 sets of 10
Qualifying Papers	2 sets of 10	2 sets of 10 per item	2 sets of 10
Definition of Condition Codes*	Yes	Yes	Yes
Definition of Alert or "Red Flag" Papers**	Yes	Yes	Yes

*Condition codes were assigned to nonscorable papers, indicating the primary reason why a paper could not be scored. The condition codes used for scoring the EOC Assessments are essentially the same for all subjects across all testing windows and are provided in Table 6.1.

**Alert or red flag papers are those that contain writing indicating that the student may be involved in, or may be the victim of, inappropriate, unethical, or criminal behavior.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

6.10.4 Materials Development for the Spring 2010 Operational Test Forms

The PE/WPs used for the Spring 2010 operational test were among those field tested in spring 2009.

Algebra I: The DESE Mathematics Consultant revised the rubric language used in Spring 2009 field-test scoring slightly for clarity and revised annotations for the anchors and training papers. All changes to the Algebra I training materials were reviewed and approved by the DESE Mathematics Consultant. No additional sample papers were pulled from the Spring 2009 pool of student responses.

Biology: Biology materials for the Spring 2010 operational test were divided into individual scoring training guides for each item (12 items in all). Materials for each item included the item's rubric, comment codes, red flag language, CLEs, anchors, training samples, and two qualifying sets. Because the Spring 2009 FT materials for this PE were already organized item-by-item, minimal changes were required to develop them into Spring 2010 operational materials. All materials produced or revised for use in operational scoring of Biology were reviewed and approved by DESE.

English II: The training and scoring material developed during summer 2009 for field-test scoring under the direction of the DESE Communication Arts Consultant were used with no revisions to train and score for the Spring 2010 operational test.

Table 6.3 lists the training materials used for the Spring 2010 operational tests.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.3: Training Materials for Spring 2010 Operational Testing

	Algebra I	Biology	English II
Scoring Guide/Notes	Yes	Yes	Yes
Course-Level Expectation(s)	Yes	Yes	Yes
Rubrics	Yes	Yes	Yes
Supplemental Training Notes for Supervisor and Team Leaders	No	Yes	Yes
Anchors	1 set of 5	1 pt—1 set of 2 2 pt—1 set of 3 3 pt—1 set of 4 4 pt—1 set of 5	1 set of 12
Training Papers	2 sets of 10	1 pt—1 set of 5 2 pt—1 set of 6 3 pt—1 set of 8 4 pt—1 set of 10 (on average—some counts varied based on item complexity)	2 sets of 10
Qualifying Papers	2 sets of 10	2 sets of 10 per item	2 sets of 10
Definition of Condition Codes*	Yes	Yes	Yes
Definition of Alert or "Red Flag" Papers**	Yes	Yes	Yes

*Condition codes were assigned to nonscorable papers, indicating the primary reason why a paper could not be scored. The condition codes used for scoring the EOC Assessments are essentially the same for all subjects across all testing windows and are provided in Table 6.1.

**Alert or red flag papers are those that contain writing indicating that the student may be involved in, or may be the victim of, inappropriate, unethical, or criminal behavior.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

6.11 Project Staffing

Prior to Spring 2008 field testing and continuing through other seasonal testing and scoring, ARC assigned members of its content staff to manage each content area from range-finding through scoring materials development and scoring. These content area managers and content area assistants worked with scoring center management staff throughout the project. The scoring center management staff was responsible for recruiting scorer candidates.

6.11.1 Recruitment and Screening of Scoring Staff Candidates

ARC employs local temporary service agencies and the University of Missouri's Human Resources department to recruit and select highly qualified scorers. ARC has long-established relationships with area employment agencies and a pool of individuals who have worked on similar projects. In addition, ARC employs, on a temporary basis, a pool of scorers, team leaders, and supervisors to work on other scoring projects throughout the year. As a result, ARC employs experienced, qualified scorers, team leaders, and supervisors and recruits new scorers for completing the MO EOC Assessment project.

Scorer candidates for all three content areas are required to respond to a writing prompt supplied by ARC to the temporary staffing agencies. The writing assignment provides an evaluation of whether the individual is able to organize his or her thoughts and to write in idiomatically correct English. The staffing agencies also conduct a third-party verification of a candidate's baccalaureate degree before referral to ARC.

Based on ARC's evaluation of an individual's performance on the writing prompt, candidates attend a brief interview during which they are asked to sign a nondisclosure statement, answer a series of questions, and take short screening tests in either reading only or reading and mathematics. Individuals taking the reading test are considered for assignment to the English II assessment; individuals taking the mathematics test are considered for assignment to Algebra I and/or Biology assessment.

During the interview, candidates are also evaluated on other criteria, including their interest, motivation, communication skills, work history, and work ethic. ARC assigns successful scorer candidates to teams, where they receive training (see Section 6.12) and are given opportunities to qualify to score. Qualified individuals are then assigned to scoring teams and undergo further training by their team leader(s) in preparation for the operational scoring.

6.11.2 Staff Qualifications

The following are minimum requirements for EOC scorer candidates:

- A baccalaureate degree from an accredited four-year institution of higher education
- Attendance at and acceptable performance during an initial screening interview
- An acceptable writing sample
- Acceptable responses to reading and/or mathematics assessments
- Agreement to maintain security of all EOC Assessment papers and scoring materials
- Attendance at training sessions

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Additionally, scorer candidates were required to meet the qualifications for each individual test form to which they were assigned.

Team leaders are required to meet all criteria for scorers and are required to possess the following:

- Supervisory experience
- The ability to communicate effectively and lead scoring sessions
- Previous scoring experience or experience with standardized academic assessments
- The ability to discern subtle differences among papers with different score points and to be able to convey those differences to scorers
- The discretion to seek advice as needed from supervisors

In addition to scorers and team leaders, ARC employs one supervisor for each content area. At a minimum, supervisors possess the skills required for team leaders. Additionally, they are required to possess exceptional communication skills and the ability to internalize the scoring process and to foster this skill in team leaders.

6.12 Training

Training followed the general outline below for each test form in each subject for Spring 2009 field-test scoring, Summer 2009 operational test scoring, Fall 2009 operational test scoring, and Spring 2010 operational test scoring:

- Provision of complete training sets and anchors to each scorer
- Discussion of confidentiality and test security, including signing of nondisclosure form by each participant
- Review of the EOC project and specific content area
- Review of CLEs
- Review of the scoring rubric and assignment of score points
- Explanation of condition codes used for nonscorable papers
- Explanation of red flag papers
- Review of the specific PE or WP
- Explanation of the anchor papers and annotations
- An opportunity for trainees to practice scoring using training sets
- Explanation of training papers with examples
- An opportunity for each scorer to qualify to score the PE or WP
- An opportunity for additional training and a second opportunity to qualify, if needed
- Training to use ARC's Image-Based Performance Assessment Scoring System (IPASS)

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

6.12.1 Spring 2009 Field Tests

Algebra I: The scoring of the 12 Algebra I field-test items was separated into two windows. Six field-test items were scored immediately after the scoring of the Spring 2009 operational test. The scoring of the remaining six field-test items immediately followed the scoring of the Summer 2009 and Fall 2009 operational tests in February 2010. Training was conducted by the Algebra I scoring supervisor under the supervision of ARC's content manager. The trainer introduced the specific PE and reviewed the anchor set. Afterward, each scorer independently scored training set 1, and the correct responses were discussed. The scorers each then independently scored training set 2 with discussion following.

Biology: For Spring 2009 Biology field-test scoring, ARC content staff trained the scoring supervisors and team leaders, with input from the DESE Curriculum Consultant and Riverside Publishing Test Development Specialists (TDSs). Training and scoring for the first six Spring 2009 field test PEs (12 forms total) occurred in summer 2009. Participants reviewed each PE, individual item rubrics, and anchors before independently scoring the training samples. Once the scoring supervisors and team leaders reviewed the scores and annotations for the training set, they completed the qualifiers. This training routine was repeated for each individual item.

The scorers were then trained by the content staff, supervisors, and team leaders. Additional training and group discussion was conducted by ARC content staff as needed for each scoring team; these additional training sessions focused on particular samples or question types and provided clarification for scorers.

English II: The scoring of four English II WPs followed the scoring of the Spring 2009 operational tests in June–July 2010. Training was conducted by content staff and team leaders. The trainer introduced the specific WP and reviewed the anchor set. Afterward, each scorer independently scored training set 1, and the correct responses were discussed. The scorers each then independently scored training set 2 with discussion following.

An additional six English II writing prompts were scored in January–February 2010 immediately following scoring of the Summer 2009 and Fall 2009 operational tests. The same training procedures were used as in the Spring 2009 operational test scoring.

6.12.2 Summer 2009 Operational Test Scoring—Conducted January 2010

Algebra I: Algebra I training for scoring the Summer 2009 operational test was conducted by the Algebra I scoring supervisor under the supervision of ARC's content manager. The trainer introduced the specific PE and reviewed the anchor set. Afterward, each scorer independently scored training set 1, and the correct responses were discussed. The scorers each then independently scored training set 2 with discussion following.

Biology: For Summer 2009 Biology operational test scoring, ARC content staff trained the scoring supervisors and team leaders, with input from the DESE Science Consultant and Riverside Publishing TDSs. Participants reviewed the performance event, individual item rubrics, and anchors before independently scoring the training samples. Once the scoring supervisors and team leaders reviewed the scores and annotations for the training set, they completed the qualifiers. This training routine was repeated for each individual item.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

The scorers were then trained by the content staff, supervisors, and team leaders. Additional training and group discussion was conducted by ARC content staff as needed for each scoring team; these additional trainings focused on particular samples or question types and provided clarification for scorers.

English II: English II training was conducted by content staff and team leaders. The trainer introduced the specific WP and reviewed the anchor set. Afterward, each scorer independently scored training set 1, and the correct responses were discussed. The scorers each then independently scored training set 2 with discussion following. The DESE Communication Arts Consultant observed and assisted in parts of the training.

6.12.3 Fall 2009 Operational Test Scoring—Conducted January 2010

Algebra I: Algebra I training for scoring the Fall 2009 operational test was conducted by the Algebra I scoring supervisor under the supervision of ARC's content manager. The trainer introduced the specific PE and reviewed the anchor set. Afterward, each scorer independently scored training set 1, and the correct responses were discussed. The scorers each then independently scored training set 2 with discussion following.

Biology: For Fall 2009 Biology operational test scoring, ARC content staff trained the scoring supervisors and team leaders, with input from the DESE Science Consultant and Riverside Publishing TDSs. Participants reviewed the PE, individual item rubrics, and anchors before independently scoring the training samples. Once the scoring supervisors and team leaders reviewed the scores and annotations for the training set, they completed the qualifiers. This training routine was repeated for each individual item.

The scorers were then trained by the content staff, supervisors, and team leaders. Additional training and group discussion was conducted by ARC content staff as needed for each scoring team; these additional trainings focused on particular samples or question types and provided clarification for scorers.

English II: English II training was conducted by content staff and team leaders. The trainer introduced the specific WP and reviewed the anchor set. Afterward, each scorer independently scored training set 1, and the correct responses were discussed. The scorers each then independently scored training set 2 with discussion following.

6.12.4 Spring 2010 Operational Test Scoring—Conducted May–June 2010

Algebra I: Training of team leaders for spring 2010 was conducted by ARC's content manager, the scoring supervisor, and the DESE Mathematics Consultant. Additionally three RIFs attended team-leader training prior to operational test scoring. Scorer training was led by the content area manager, the scoring supervisor, and team leaders. The trainer introduced the specific PE and reviewed the anchor set. Afterward, each scorer independently scored training set 1, and the correct responses were discussed. The scorers each then independently scored training set 2 with discussion following

Biology: For Spring 2010 Biology operational test scoring, ARC content staff trained the scoring supervisors and team leaders, with input from the DESE Science Consultant and Riverside Publishing TDSs. Participants reviewed the PE, individual item rubrics, and anchors before independently scoring the training samples. Once the scoring supervisors and team leaders reviewed the scores and annotations for the training set, they completed the qualifiers. This training routine was repeated for each individual item.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

The scorers were then trained by the content staff, supervisors, and team leaders. Additional training and group discussion was conducted by ARC content staff, as needed, for each scoring team; these additional trainings focused on particular samples or question types and provided clarification for scorers.

English II: Initial training of the English II supervisor and team leaders was completed by content staff and observed by the DESE Communication Arts Consultant. Subsequently, the scorers were trained by a team leader with content area staff and the supervisor observing. The trainer introduced the specific WP and reviewed the anchor set. Afterward, each scorer independently scored training set 1, and the correct responses were discussed. The scorers each then independently scored training set 2 with discussion following.

6.13 Qualification to Score

Upon completion of training on a field-test or operational PE/WP, each trainee scored a qualifying set of student papers representing a range of score points. Each trainee worked independently to score the set of qualifying papers using anchor papers and training sets as references. To become qualified to score a particular English II WP or Algebra I PE, a candidate was required to achieve an 80% exact score match with the key, with no more than one score deviating by more than one point from the key.

Qualification of Biology scorers was based on the number of points that could be awarded to each item. Unlike Algebra I and English II, the Biology scoring guides were unique to each question, rather than being comprehensive across all the individual items within the PE. The passing percentage on scorer qualifying tests varied according to the number of rubric points possible for the given question. The requirements were 80% key match for 3- and 4-point items, 90% match for 2-point items, and 100% match for 1-point items.

6.14 Scoring Procedures

Scoring for all operational tests and field tests included the use of ARC's IPASS, which is specifically designed for hand scoring of open-ended assessment items. Using IPASS, scorers view and read digital images of student responses to the MO EOC Assessment PE/WPs, read items, and assign scores. IPASS does not allow a scorer to assign a score until all page images for a response are viewed. Scorers and team leaders do not have access to identifying information for specific school districts or individual students. All student responses are assigned randomly to scorers.

Scorers assign a score to a student response based on how well the student met the criteria described in the applicable rubric. ARC content staff members consult with scoring personnel as needed. In instances for which a scoring policy decision is required, ARC content staff members contact Riverside Publishing content leads and/or DESE Curriculum Consultants with the information necessary to make a decision. Once DESE establishes a policy, ARC creates documentation and recalibrates all staff assigned to the prompt.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.4 contains information about the scoring staff for each content area and scoring event.

Table 6.4: Scoring Staff for the MO EOC PE/WP Events

	English II			Algebra I			Biology		
	# of Teams/ Team Leaders	# of Team Leaders	# of Scorers	# of Teams	# of Team Leaders	# of Scorers	# of Teams	# of Team Leaders	# of Scorers
Spring 2008 Field Test	10	11	76	12	6	37	12	14	94
Spring 2009 Field Tests	1	1	3	1	1	4	4	4	16
Summer 2009 Operational Test	1	1	3	1	1	4	4	4	16
Fall 2009 Operational Test	1	1	3	1	1	4	4	4	16
Spring 2010 Operational Test	3	3	20	2	2	14	5	5	47

6.14.1 Algebra I

Each Algebra I PE included a number of questions requiring a student response. Scores for Algebra I were based on the overall quality of responses throughout the assignment, and a single rubric with point values of 0–4 was used. While students were expected to demonstrate achievement of individual skills to answer specific questions within the PE, score points were based on an overall assessment of these elements.

6.14.2 Biology

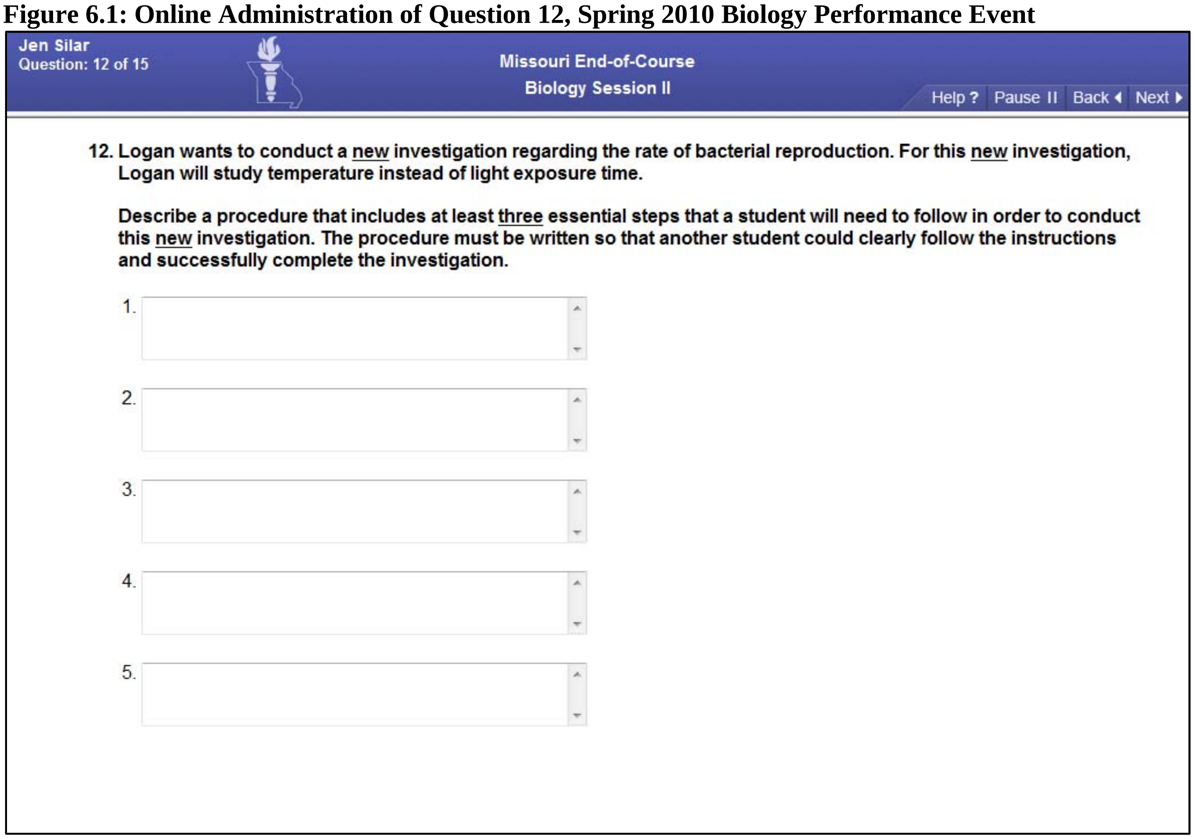
Each Biology PE included a number of discrete items, each requiring a student response. Biology PEs were made up of 10 to 16 individual questions, with maximum score-point values ranging from 1–4 for each individual question, yielding a potential total score of 20 points. Each item within the PE was scored individually and independently of all other items. The student's score for the Biology PE was the sum of the assigned scores for each individual item within the PE.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

6.14.2.1 Question 12, Spring 2010 Biology PE

In the Spring 2010 administration, an error was discovered in the images transmitted from ITS to ARC and to school districts for scoring. Question 12, one of the component items in the EOC Biology PE, was a description-of-procedure item scored with a 3-point rubric. The online assessment provided students with five independent text boxes in which to respond to the item. ARC received only the first three text boxes in the images of the online assessments transmitted. Figure 6.1 shows question 12 as it appeared on the screen for the online assessment. Figure 6.2 is an example of what ARC originally received to score without the fourth and fifth text boxes. Figure 6.3 shows an example of a corrected image including all five text boxes as completed by the student, and Figure 6.4 is question 12 as it appeared in the paper-and-pencil test book.

Figure 6.1: Online Administration of Question 12, Spring 2010 Biology Performance Event



Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 6.2: Sample of Image Originally Transmitted to ARC

Biology

12. **Logan wants to conduct a new investigation regarding the rate of bacterial reproduction. For this new investigation, Logan will study temperature instead of light exposure time.**

Describe a procedure that includes at least three essential steps that a student will need to follow in order to conduct this new investigation. The procedure must be written so that another student could clearly follow the instructions and successfully complete the investigation.

1. Be sure to have a temperature scale, a clock, and the number of containers you will need.

2. Place the bacteria into each container and be sure that each container has the same temperature level.

3. Check each container every 10-20 minutes and change the temperature if needed.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 6.3: Sample of Corrected Image Transmitted to ARC with All Five Text Boxes

Biology

12. Logan wants to conduct a new investigation regarding the rate of bacterial reproduction. For this new investigation, Logan will study temperature instead of light exposure time.

Describe a procedure that includes at least three essential steps that a student will need to follow in order to conduct this new investigation. The procedure must be written so that another student could clearly follow the instructions and successfully complete the investigation.

1. Be sure to have a temperature scale, a clock, and the number of containers you will need.

2. Place the bacteria into each container and be sure that each container has the same temperature level.

3. Check each container every 10-20 minutes and change the temperature if needed.

4. Count the bacteria every time the temperature changes and be sure to label it.

5. If you are changing the temperature from high to low be sure to label it, if you are changing the temperature from low to high be sure to identify it.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 6.4: Paper-and-Pencil Administration of Question 12, Spring 2010 Biology Performance Event

12. **Logan wants to conduct a <u>new</u> investigation regarding the rate of bacterial reproduction. For this <u>new</u> investigation, Logan will study temperature instead of light exposure time.**

Describe a procedure that includes at least <u>three</u> essential steps that a student will need to follow in order to conduct this <u>new</u> investigation. The procedure must be written so that another student could clearly follow the instructions and successfully complete the investigation.

1. ____

____

2. ____

____

3. ____

____

Immediately upon notification of the error, ARC suspended scoring of the item. As with all items in the EOC Biology PEs, question 12 was scored independently of all the other component items. In the Spring 2010 form, question 12 appeared by itself on a single page in both the paper-and-pencil and online formats. After reviewing the images, ITS and Riverside Publishing determined the student scores that could potentially have been affected by the error were limited to those that contained student response information in the fourth or fifth text boxes.

After identifying that online responses were potentially incorrect, scoring of all tests received via online transmission was halted until the problem could be fully vetted and corrected. On May 25, 2010, ARC received new copies of 13,986 test books that Riverside Publishing had identified as being impacted and confirmed that the problem was limited to a single item (question 12) on the Biology assessment. ARC was also provided with a Microsoft Excel® document that mapped the Riverside Publishing file numbers on these test books to those of the original books already been sent to ARC. In addition, ARC requested a total report from Riverside Publishing that included details on all online Biology test books that had been transmitted and were not affected by the error, giving them a complete list. These lists made it possible for ARC to maintain quality control for each test book throughout the process.

After cross-mapping the data, a total of six test books were identified as “nonmatching” between the lists provided and the materials stored in ARC systems. Conversations with Riverside Publishing on these six test books revealed that they had been removed from the stream to ARC by Riverside quality control as duplicates or partial records but had

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

been incorrectly included in the list to be reprocessed. These six items were subsequently removed from reprocessing.

ARC technical staff compared the details of the remaining items on the report against the test books already loaded into the scoring system. Items that were not yet processed by the scoring system were removed from the queue, replaced with the updated copy, and then handled normally when scoring was later resumed.

Because the item in question was contained in a single page of the multipage Biology test book, it was programmatically possible to replace just the affected page in the books containing the error. Scores for those items could be invalidated and the items returned to the pool to be scored.

Copies of the affected online test books were made and had the page for question 12 replaced. Scorers were locked out of the scoring system. A replication of the database was made, and the required database references were updated and validated. This process was then repeated on the production data and validated again, and then scoring resumed normally.

ARC proceeded to score the responses to question 12 along with the remainder of the items in the Biology PE. The question 12 rubric does not consider the position of the student response. For example, a response written completely in text box 1 would be assigned the same score as the identical response divided among two or more text boxes. After discussions between Riverside Publishing and DESE Curriculum and Assessment staff, no change to the rubric was necessary. ARC then advised team leaders and scorers that the online responses would differ in appearance from the paper-and-pencil responses, which have three response areas, but no rubric changes were implemented for scoring question 12. Scorers evaluated the responses based on their content, not their placement.

6.14.3 English II

English II papers were scored holistically using a single rubric with score points ranging from 1–4. Individual elements of student writing were described in the rubric, and the scoring notes emphasized that the score points were based on an overall assessment of these elements.

6.15 Monitoring for Quality Assurance

ARC employed a number of methods to monitor the progress and quality of scorers' work and to ensure consistent and accurate scoring with minimum score drift. Supervisors, team leaders, scoring center managers, and content area managers regularly monitored the various quality assurance reports to take corrective action when necessary. Scorers who were unable to maintain acceptable agreement rates on check sets or validation scores were required to improve by either recalibrating or retraining and requalifying.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

6.15.1 Inter-Rater Reliability Checks (Validation Scores)

One of every 10 student responses scored by each scorer was submitted automatically to the scorer's team leader for a validation score. When the scorer and team leader scored a paper differently, the team leader's score (validation score) became the score of record. At least two times each day, team leaders generated and reviewed reports indicating the number and percentage of papers that received different scores, with each score identified as belonging to the scorer or the team leader. Also, the supervisor for each content area monitored agreement rates throughout each day

Tables 6.5 through 6.11 summarize the number and percentage of a scorer's score in perfect agreement with the validation scores and the number in perfect plus adjacent agreement with the validation scores. The number of items validated indicates that approximately 10% of all scores were read by a team leader. These tables are repeated in Chapter 10: Reliability but are also included here for the reader's reference.

Inter-rater reliability for all scoring sessions (Spring 2008 field tests, Spring 2009 field Tests, Summer 2009 operational tests, Summer 2009 Large Print/Braille, Fall 2009 operational tests, Fall 2009 Large Print/Braille tests, and Spring 2010 operational tests) was monitored on a daily basis by ARC, Riverside Publishing, and DESE staff during conference calls in all subjects. Inter-rater reliability was based on the number of discrepancies between the score assigned by a scorer and the score assigned by a team leader or a more experienced scorer. Discrepancies identified training needs for each scorer. Scorers who consistently scored higher or lower than the team leaders' scores were retrained. Also, training needs for specific score-point splits were identified and retrained, such as score splits in English II between 3 and 4, 2 and 3, and 1 and 2.

A scorer who achieved less than an 80% agreement with team leaders' scores was required to attend a recalibration session led by a team leader or supervisor. If the scorer was unable to score correctly after recalibration, the team leader consulted with a supervisor and provided retraining and an opportunity for the scorer to requalify. If the scorer was unable to requalify, he or she was considered for other work assignments or was released from the project.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.5: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Spring 2008 Field-Test Administration*

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS100076683	4	5001	499	71.1%	99.0%
RPS100076622	4	5004	499	82.6%	97.8%
Biology					
RPS100076797 - 1	1	5006	476	80.9%	99.6%
RPS100076798 - 2	1	5006	476	79.8%	99.6%
RPS100076799 - 3	1	5006	476	95.6%	99.6%
RPS100076800 - 4	2	5006	476	85.5%	98.1%
RPS100076807 - 5	2	5006	476	91.2%	99.6%
RPS100076801 - 6	4	5006	476	63.2%	94.5%
RPS100076803 - 7	2	5006	476	79.4%	100.0%
RPS100076808 - 8	2	5006	476	84.5%	99.8%
RPS100076802 - 9	2	5006	476	68.9%	99.2%
RPS100076804 - 10	3	5006	476	55.5%	92.0%
RPS100077987 - 11	1	5006	476	92.9%	99.4%
RPS100076805 - 12	2	5006	476	70.8%	95.2%
RPS100076806 - 13	1	5006	476	95.4%	99.4%
RPS100075961 - 1	1	5007	483	79.1%	100.0%
RPS100075962 - 2	1	5007	483	93.2%	100.0%
RPS100075963 - 3	1	5007	483	94.8%	100.0%
RPS100075965 - 4	1	5007	483	73.1%	99.8%
RPS100075964 - 5	2	5007	483	82.0%	90.3%
RPS100075966 - 6	4	5007	483	61.3%	91.7%
RPS100075968 - 7	1	5007	483	96.1%	100.0%
RPS100075969 - 8	2	5007	483	77.6%	97.9%
RPS100075970 - 9	2	5007	483	75.8%	96.3%
RPS100075971 - 10	3	5007	483	50.1%	88.2%
RPS100075972 - 11	2	5007	483	67.7%	87.4%
English II					
RPS100076784	4	5004	492	68.3%	99.0%
RPS100076785	4	5006	494	61.1%	98.4%

*Test items reported in Table 6.5 are those that were later used in the Summer 2009 and Fall 2009 operational tests.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.6: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Spring 2009 Field-Test Administration*

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Per + Adj Agreement
Algebra I					
Forms 15 & 16	4	4945	543	83.6%	98.9%
Biology					
RPS100089159 - 12	1	10716	1178	98.6%	99.9%
RPS100089166 - 13	1	10716	1181	99.6%	100.0%
RPS100089165 - 14	4	10716	1175	82.3%	98.6%
RPS100089162 - 15	1	10716	1174	98.6%	99.9%
RPS100089167 - 16	1	10716	1163	97.0%	100.0%
RPS100089164 - 17	2	10716	1176	84.4%	99.3%
RPS100089171 - 18	1	10716	1164	96.2%	99.9%
RPS100089172 - 19	1	10716	1174	95.7%	99.8%
RPS100089160 - 12	1	2340	263	97.7%	99.2%
RPS100089161 - 13	2	2340	262	96.2%	98.9%
RPS100089169 - 14	1	2340	261	94.6%	100.0%
RPS100089163 - 15	1	2340	260	94.2%	100.0%
RPS100089168 - 16	2	2340	257	83.3%	99.2%
RPS100089173 - 17	1	2340	262	93.5%	100.0%
RPS100089174 - 18	3	2340	258	85.7%	95.0%
RPS100089175 - 19	1	2340	262	100.0%	100.0%
English II					
Form 15	4	2018	221	91.0%	100.0%

*Test items reported in Table 6.6 are those that were later used in the Spring 2010 operational test.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.7: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Summer 2009 Operational Test Administration

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS100076683	4	1589	174	90.2%	100.0%
Biology					
RPS100076797 - 1	1	653	71	95.8%	100.0%
RPS100076798 - 2	1	653	72	90.3%	100.0%
RPS100075927 - 3	1	653	72	98.6%	100.0%
RPS100076807 - 4	2	653	73	98.6%	100.0%
RPS100076801 - 5	4	653	67	89.6%	95.5%
RPS100076803 - 6	2	653	72	93.1%	100.0%
RPS100076808 - 7	2	653	71	90.1%	98.6%
RPS100076802 - 8	2	653	70	92.9%	100.0%
RPS100076804 - 9	3	653	62	83.9%	96.8%
RPS100076805 - 10	2	653	70	92.9%	100.0%
English II					
RPS 100076785	4	864	85	91.8%	100.0%

Table 6.8: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Fall 2009 Operational Test Administration

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS 100076622	4	2681	297	92.3%	98.7%
Biology					
RPS100075961 - 1	1	2259	248	95.2%	100.0%
RPS100075962 - 2	1	2259	246	98.8%	100.0%
RPS100075963 - 3	1	2259	249	98.8%	100.0%
RPS100075965 - 4	1	2259	248	95.2%	100.0%
RPS100075964 - 5	2	2259	249	98.0%	99.2%
RPS100075966 - 6	4	2259	241	85.5%	98.8%
RPS100075968 - 7	1	2259	252	98.0%	100.0%
RPS100075969 - 8	2	2259	248	91.5%	99.2%
RPS100075970 - 9	2	2259	248	94.8%	99.6%
RPS100075971 - 10	3	2259	249	91.2%	98.4%
RPS100075972 - 11	2	2259	251	91.2%	98.8%
English II					
RPS 100076785	4	1480	148	91.2%	100.0%

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.9: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Fall 2009 Large Print/Braille Test Administration

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS 100076624	4	1	1	100.0%	100.0%
Biology					
RPS100075983 - 1	2	2	1	100.0%	100.0%
RPS100075984 - 2	2	2	1	100.0%	100.0%
RPS100075985 - 3	2	2	1	100.0%	100.0%
RPS100075986 - 4	4	2	1	100.0%	100.0%
RPS100075992 - 5	3	2	1	100.0%	100.0%
RPS100075987 - 6	4	2	1	100.0%	100.0%
RPS100075989 - 7	5	2	1	100.0%	100.0%
RPS100075988 - 8	4	2	1	100.0%	100.0%
RPS100075990 - 9	2	2	1	100.0%	100.0%
RPS100075991 - 10	2	2	1	100.0%	100.0%
English II					
RPS 100076789	NA	0	0	0.0%	0.0%

Table 6.10: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders Validation Scores, Spring 2010 Operational Test Administration

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS Form 15 & 16	4	64297	7076	91.4%	99.5%
Biology					
RPS100089160 - 1	1	63662	7023	97.6%	99.9%
RPS100089166 - 2	1	63662	7054	99.3%	100.0%
RPS100089168 - 3	2	63662	6919	92.6%	99.8%
RPS100089161 - 4	2	63662	7046	98.1%	99.9%
RPS100089164 - 5	2	63662	6942	93.1%	99.9%
RPS100089165 - 6	4	63662	7037	84.9%	98.6%
RPS100089162 - 7	1	63662	7015	98.6%	99.8%
RPS100089167 - 8	1	63662	6988	96.4%	99.9%
RPS100089163 - 9	1	63662	7041	96.7%	100.0%
RPS100089172 - 10	1	63662	6948	93.9%	99.9%
RPS100089173 - 11	1	63662	7042	95.0%	100.0%
RPS100089174 - 12	3	63662	7036	90.1%	98.8%
English II					
RPS Form 15	4	63737	6916	94.2%	100.0%

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.11: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Spring 2010 Large Print/Braille Test Administration

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS 100076624	4	24	3	100.0%	100.0%
Biology					
RPS100075983 - 1	1	26	3	100.0%	100.0%
RPS100075984 - 2	1	26	3	66.7%	100.0%
RPS100075985 - 3	1	26	3	100.0%	100.0%
RPS100075986 - 4	3	26	3	100.0%	100.0%
RPS100075992 - 5	2	26	3	100.0%	100.0%
RPS100075987 - 6	3	26	3	100.0%	100.0%
RPS100075989 - 7	4	26	3	100.0%	100.0%
RPS100075988 - 8	3	26	3	100.0%	100.0%
RPS100075990 - 9	1	26	3	100.0%	100.0%
RPS100075991 - 10	1	26	3	100.0%	100.0%
English II					
RPS 100076789	4	24	3	100.0%	100.0%

6.15.2 Check Sets for Spring 2009 Field Tests, Summer 2009 Operational Tests, Fall 2009 Operational Tests, and Spring 2010 Operational Tests

For both field tests and operational tests in these time periods, a number of papers were chosen for each PE/WP during range-finding to serve as check sets. The check-set papers were representative of all score points and were prescored and approved by DESE.

After meeting qualifying requirements, the supervisor and team leaders for each subject read and independently scored each check-set paper. The content area managers, supervisors, and team leaders then discussed the papers and correct scores, identifying the applicable rubric characteristics of each.

For scorers, check-set papers were applied in two phases. Phase I occurred during the scoring of the first 100 test books. This phase included four sets of five check papers, each interspersed among 20 live papers for each individual scorer. Each scorer was required to maintain at least an 80% agreement rate with the check set. If a scorer met the first predefined check-set benchmark, he or she proceeded with scoring the next 20 live items, also interspersed with check papers, and the team leader received notification that the scorer had satisfactorily completed the check set. Tables 6.12–6.15 indicate the percentage of agreement rates for the check sets.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.12: Percentages of Exact and Adjacent Agreement for Check Sets, Spring 2009 Field-Test Administration*

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Check-Set Scores	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS Forms 15 & 16	4	4945	352	98.3%	100.0%
Biology					
RPS100089159 - 12	1	10716	403	98.5%	100.0%
RPS100089166 - 13	1	10716	422	99.8%	100.0%
RPS100089165 - 14	4	10716	370	93.5%	100.0%
RPS100089162 - 15	1	10716	407	97.8%	100.0%
RPS100089167 - 16	1	10716	443	95.3%	100.0%
RPS100089164 - 17	2	10716	404	95.8%	100.0%
RPS100089171 - 18	1	10716	405	96.5%	99.8%
RPS100089172 - 19	1	10716	441	97.3%	100.0%
RPS100089160 - 12	1	2340	236	99.6%	100.0%
RPS100089161 - 13	2	2340	204	98.5%	100.0%
RPS100089169 - 14	1	2340	198	97.5%	100.0%
RPS100089163 - 15	1	2340	236	98.3%	100.0%
RPS100089168 - 16	2	2340	184	89.7%	98.9%
RPS100089173 - 17	1	2340	235	96.6%	100.0%
RPS100089174 - 18	3	2340	217	89.9%	98.6%
RPS100089175 - 19	1	2340	201	100.0%	100.0%
English II					
RPS Form 15	4	2018	283	91.9%	100.0%

*Test items reported in Table 6.12 are those that were later used in the Spring 2010 operational test.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.13: Percentages of Exact and Adjacent Agreement for Check Sets, Summer 2009 Operational Test Administration

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Check-Set Scores	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS 100076683	4	1589	109	93.6%	100.0%
Biology					
RPS100076797 - 1	1	653	90	95.6%	100.0%
RPS100076798 - 2	1	653	86	100.0%	100.0%
RPS100075927 - 3	1	653	87	100.0%	100.0%
RPS100076807 - 4	2	653	88	97.7%	100.0%
RPS100076801 - 5	4	653	87	93.1%	100.0%
RPS100076803 - 6	2	653	90	97.8%	100.0%
RPS100076808 - 7	2	653	87	96.6%	100.0%
RPS100076802 - 8	2	653	88	95.5%	100.0%
RPS100076804 - 9	3	653	90	88.9%	98.9%
RPS100076805 -10	2	653	87	97.7%	100.0%
English II					
RPS 100076784	4	864	92	97.8%	100.0%

Table 6.14: Percentages of Exact and Adjacent Agreement for Check Sets, Fall 2009 Operational Test Administration

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Check-Set Scores	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS 100076622	4	2681	133	97.7%	100.0%
Biology					
RPS100075961 - 1	1	2259	124	97.6%	100.0%
RPS100075962 - 2	1	2259	88	98.9%	100.0%
RPS100075963 - 3	1	2259	87	100.0%	100.0%
RPS100075965 - 4	1	2259	124	96.8%	100.0%
RPS100075964 - 5	2	2259	122	98.4%	100.0%
RPS100075966 - 6	4	2259	122	98.4%	100.0%
RPS100075968 - 7	1	2259	123	100.0%	100.0%
RPS100075969 - 8	2	2259	112	98.2%	100.0%
RPS100075970 - 9	2	2259	125	96.8%	100.0%
RPS100075971 - 10	3	2259	102	97.1%	100.0%
RPS100075972 - 11	2	2259	124	97.6%	100.0%
English II					
RPS 100076785	4	1480	81	95.1%	100.0%

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 6.15: Percentages of Exact and Adjacent Agreement for Check Sets, Spring 2010 Operational Test Administration

Content Area/Code	Total # of Points Possible	Total # of Items Scored	# of Check-Set Scores	Perfect Agreement	Per + Adj Agreement
Algebra I					
RPS Forms 15 & 16	4	64297	1700	97.8%	99.8%
Biology					
RPS100089160 - 01	1	63662	1686	99.0%	100.0%
RPS100089166 - 02	1	63662	1665	99.8%	100.0%
RPS100089168 - 03	2	63662	1719	97.7%	99.9%
RPS100089161 - 04	2	63662	1667	97.8%	100.0%
RPS100089164 - 05	2	63662	1687	99.6%	100.0%
RPS100089165 - 06	4	63662	1653	96.4%	99.9%
RPS100089162 - 07	1	63662	1706	99.4%	99.9%
RPS100089167 - 08	1	63662	1612	99.5%	100.0%
RPS100089163 - 09	1	63662	1681	99.6%	99.9%
RPS100089172 - 10	1	63662	1696	98.6%	100.0%
RPS100089173 - 11	1	63662	1661	99.2%	99.9%
RPS100089174 - 12	3	63662	1117	98.5%	99.6%
English II					
RPS Forms 15	4	63737	1890	95.0%	100.0%

Check-set papers were not identified specifically to scorers, although the papers were used as additional training material by each subject's supervisor and team leaders. If a scorer did not meet the check-set benchmark, scoring automatically stopped for that scorer, and the team leader was notified. The team leader then delivered recalibration and/or retraining as necessary. Following recalibration/retraining, the scorer proceeded with scoring the subsequent items interspersed with check papers. Evaluation continued in this manner as scorers proceeded through the four initial check sets.

Phase II immediately followed Phase I and included multiple check sets of five check papers interspersed among each 250 live papers scored by an individual scorer. If a scorer met the predefined check-set benchmark, he or she proceeded with scoring the subsequent live items interspersed with check papers, and the team leader received notification that the scorer had satisfactorily completed the check set and had moved on.

If a scorer did not meet the check-set benchmark in Phase II, scoring was automatically stopped for that scorer, and the team leader was notified. As with Phase I, the team leader provided recalibration and/or retraining, and the scorer then proceeded with scoring the subsequent live items in the set.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Check papers were interspersed into live papers through IPASS at the prescribed rates for Phases I and II. The delivery scheme programmed into IPASS provided for a degree of randomness in the presentation of check papers and prevented multiple scorers from receiving identical check papers at the same time within a single check set.

6.15.3 Recalibration Procedures

At the beginning of each day's work, supervisors and team leaders led recalibration sessions with scorers. The team leaders clarified scoring issues, answered questions, reviewed the PE/WP assigned for scoring, and worked with individual scorers as needed to improve performance.

Recalibration was also required when an individual had not scored a specific form within the previous three workdays. The scorer reviewed the rubric, anchor papers, and training papers with the team leader and then was given the opportunity to clarify scoring issues and to ask questions about specific issues.

6.15.4 Retraining and Requalifying

Standard ARC policy requires that a scorer or team leader retrain and requalify on a specific test item if an individual has not scored that item in more than 30 days. Because of the multiple items on the Biology test form, this policy was applied to that content area as needed. In addition, a scorer or team leader was required to retrain and requalify on a specific item whenever an issue of reliability was raised (e.g., for a scorer's lack of adequate agreement rates on check sets or low agreement rates with the team leader's validation scores).

Retraining and requalifying activities were conducted by a team leader, supervisor, or other content staff member. The outline used for initial training and qualifying was used for retraining and requalifying.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 7: SCALING AND EQUATING

7.1 Introduction

This chapter details the scaling and equating procedures implemented by Riverside Publishing for the 2009 and 2010 Missouri End-of-Course (EOC) Assessments. While the 2009 scaling and equating procedures are described in the *2009 Missouri End-of-Course Assessments Technical Report*, the details of the procedures followed are repeated here to provide continuity for readers. The equating methods described in this chapter will serve to maintain consistency of the EOC Assessments score scales over time and ensure that the achievement levels are applied consistently from year to year.

A pre-equating model (Kolen and Brennan, 2004) was used to produce equated forms for each EOC Assessment. The chapter begins with an overview of the equating design. Then the item response theory (IRT) models used for equating are described, and the model assumptions are examined. This is followed by a description of the steps used to carry out the scaling and equating for the 2008–2009 and 2009–2010 operational assessments.

7.1.1 Equating Design

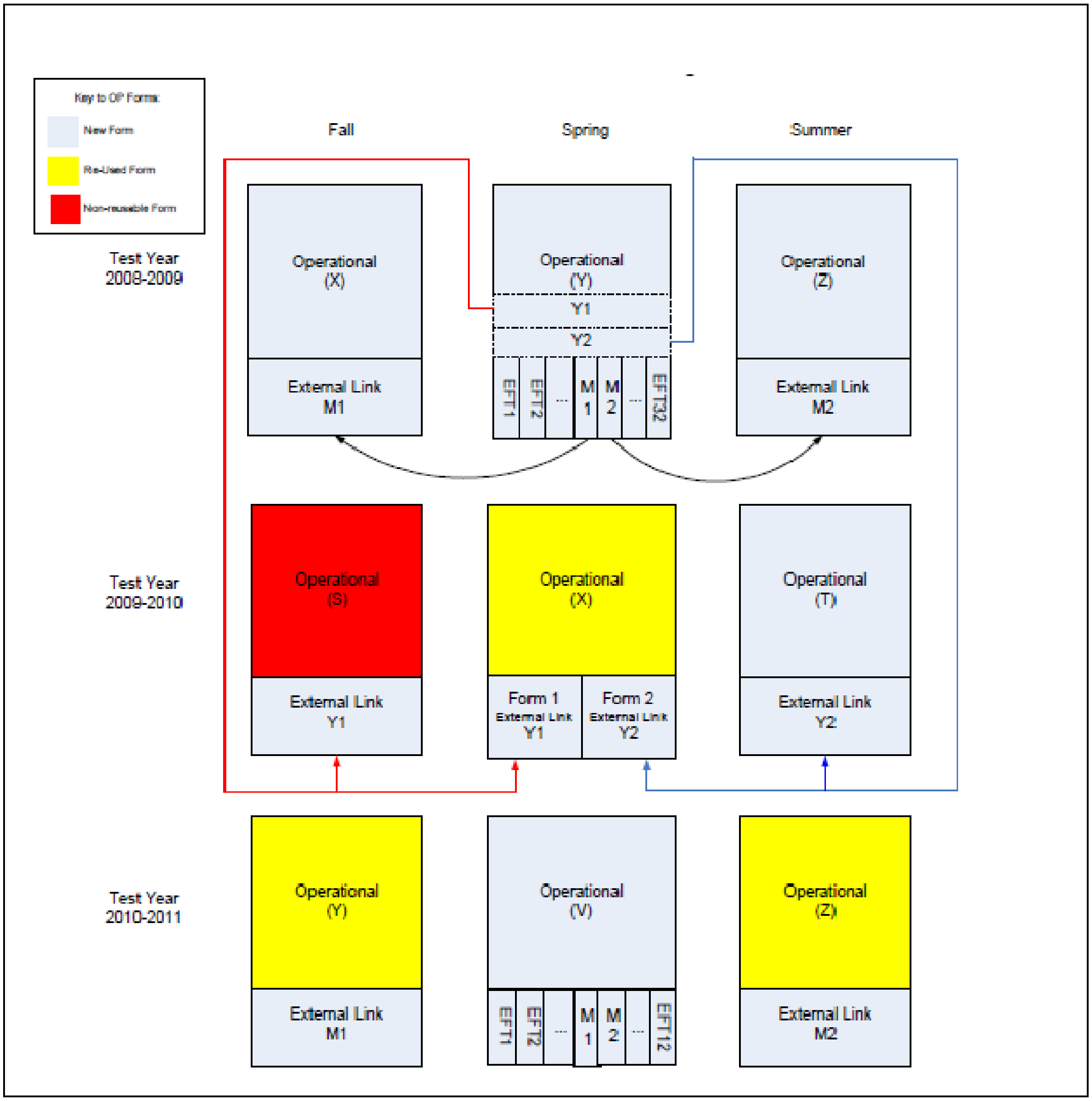
At the May 2008 meeting, Missouri's Technical Advisory Committee (TAC) recommended that post-equating be conducted as a check on the pre-equating results. To accomplish this, a common-item test design was developed in which each form was equated to a base form through a set of linking, or anchor, items.[11] Figure 7.1 shows the post-equating design for the EOC Assessments.

Each assessment contains a set of operational items as well as 12 additional (external) item slots. For the Spring 2009 assessments, these 12 external slots were used for embedded field test (FT) forms (designated as EFT1, EFT2, etc. in Figure 7.1). Two of the forms contained linking (or anchor) items embedded in the Fall (designated as M1) and Summer (designated as M2) forms. The Spring 2009 form was designated as the base form for the English II, Algebra I, and Biology Assessments. The M1 and M2 item sets were used in a post-equating check of the pre-equating results for Fall 2008 and Summer 2009 operational forms.[12] The items designated as Y1 and Y2 in Figure 7.1 are embedded in the Fall 2009, Spring 2010, and Summer 2010 forms for a post-equating check of the 2009–2010 accountability year forms.

[11] Post-equating was done only for Algebra I and Biology. Due to item development needs, there were not enough field-test slots available to use for post-equating English II.
[12] The post-equating check of the 2008–2009 forms did not reveal significant differences in the raw-score to scale-score conversions, confirming that the pre-equating results were acceptable.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 7.1: Missouri EOC Equating Design



7.2 Item Response Theory

WINSTEPS software (Linacre, 2006b) was used to accomplish the scaling and equating for the Missouri EOC Assessments. WINSTEPS is designed to produce a single scale by jointly analyzing data from students' responses to both selected response (SR) and open-ended items. SR items were calibrated using the Rasch model (Rasch, 1960; Wright and Stone, 1979), while the partial credit model (Masters 1982) was used to calibrate the Performance Event/Writing Prompt (PE/WP) items.

Rasch scaling is "a method for obtaining objective, fundamental, linear measures from stochastic observations of ordered category responses" (Linacre 2006a, p. 10). One

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

feature of the Rasch model that distinguishes it from classical test theory is the placement of estimates of a person's ability and item difficulty on the same scale. The Rasch model expresses the probability of a correct response to an item as a function of the ability of the person and the difficulty of the item. In the Rasch model, the probability of a correct response to item *i*, given θ, is

$$P_i(\theta) = \frac{e^{(\theta\text{-}b_i)}}{1 + e^{(\theta\text{-}b_i)}},$$

where θ = latent trait, or ability, level and b_i = the difficulty parameter for item *i*.

Masters (1982) developed the partial credit model as an extension of the Rasch model to handle polytomous items, or items that allow for partially correct responses (e.g., open-ended items). As noted above, all Missouri EOC item calibrations used the dichotomous Rasch model for SR items and the partial credit model for open-ended items (polytomous items).

7.3 Scaling and Equating

IRT pre-equating involves scaling item parameters and equating test forms based on field-test data before the forms are administered operationally. Note, however, that for the 2008–2009 year, the forms were pre-equated retroactively (after the Spring 2009 operational administration) to allow for a one-time recentering of the pools using Spring 2009 operational data. The follow approach was used for pre-equating the EOC Assessments:

1. Calibrate all 2008 standalone field-test forms concurrently without constraint.
2. Establish the base scale through calibration of the Spring 2009 operational forms without constraint.
3. Examine the stability of the common items from the two calibrations (i.e., the operational form items).
4. Recenter the 2008 item bank to the 2009 base scale.
5. Place the 2009 embedded field test items onto the 2009 operational scale.
6. Perform fixed calibrations on the Summer 2009, Fall 2009, and Spring 2010 operational forms.
7. Place the 2010 embedded field test items onto the 2009 operational scale.

7.3.1 Step 1: Concurrent Calibration of 2008 Field-Test Forms

Table 7.1 shows the number of field-test forms and their composition for the Spring 2008 standalone field test. For each content area, 10 forms containing SR items and 10 forms containing a PE/WP were spiraled within each classroom. Assuming randomly equivalent groups, the complete pool of items for each content area was concurrently calibrated using the WINSTEPS software program, placing all items on a common scale. Because these calibrations had to be performed before a complete set of data was available, Tables 7.2 through 7.4 provide a comparison of the calibration set and a complete set of data for

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

the Spring 2008 standalone field test. Inspection of these tables shows that the demographics for the calibration samples were similar to the census (generally less than 1% difference), or complete set, of data.

Table 7.1: 2008 Standalone Field Test, Spring 2008

Assessment	Session I Multiple Choice Number of Items	Session I Multiple Choice Number of Forms per Session	Session II Number of Performance Events/ Writing Prompts	Session II Number of Forms per Session
English II*	30	10	1	10
Algebra I	18	10	1	10
Biology	18	10	1	10

*The English II forms shared 5 unique passages across the 10 forms.

Table 7.2: Comparison of the Calibration and Census Data for the Spring 2008 Standalone Field Test, English II

	Calibration Sample		Census Data		Difference (calibration minus census)
	N	%	*N*	%	%
All Students	37,108		39,839		
Gender					
Male	18,503	49.9	19,754	49.6	0.3
Female	18,605	50.1	20,085	50.4	–0.3
Race/Ethnicity					
White	30,686	82.7	32,822	82.4	0.3
Black	4,720	12.7	5,108	12.8	–0.1
Hispanic	902	2.4	968	2.4	0.0
Asian/Pacific Islander	618	1.7	752	1.9	–0.2
American Indian	182	0.5	189	0.5	0.0

Table 7.3: Comparison of the Calibration and Census Data for the Spring 2008 Standalone Field Test, Algebra I

	Calibration Sample		Census Data		Difference (calibration minus census)
	N	%	*N*	%	%
All Students	35,449		38,823		
Gender					
Male	17,837	50.3	19,440	50.1	0.2
Female	17,612	49.7	19,383	49.9	–0.2
Race/Ethnicity					
White	27,997	79.0	30,143	77.6	1.4
Black	5,792	16.3	6,777	17.5	–1.2
Hispanic	899	2.5	1,029	2.7	–0.2
Asian/Pacific Islander	607	1.7	715	1.8	–0.1
American Indian	154	0.4	159	0.4	0.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

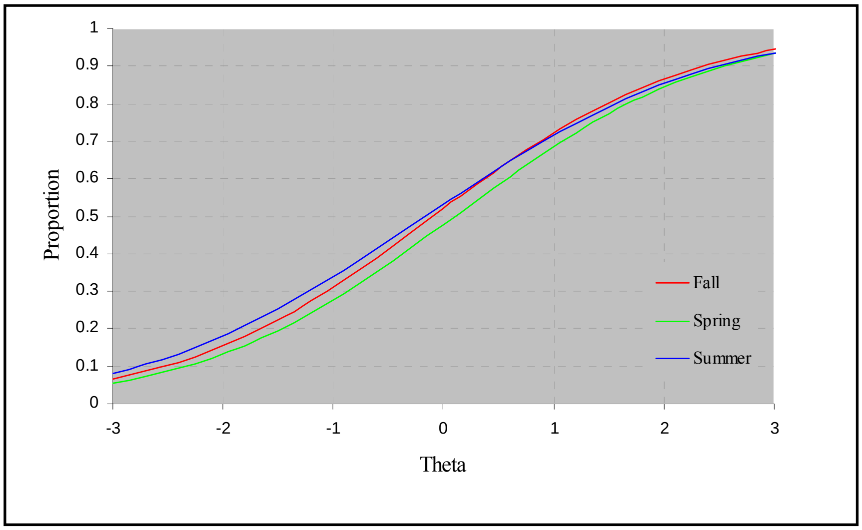
Table 7.4: Comparison of the Calibration and Census Data for the Spring 2008 Standalone Field Test, Biology

	Calibration Sample		Census Data		Difference (calibration minus census)
	N	%	*N*	%	%
All Students	27,062		39,849		
Gender					
Male	13,462	49.7	19,726	49.5	0.2
Female	13,600	50.3	20,123	50.5	–0.2
Race/Ethnicity					
White	22,539	83.3	33,274	83.5	–0.2
Black	3,286	12.1	4671	11.7	0.4
Hispanic	652	2.4	992	2.5	–0.1
Asian/Pacific Islander	476	1.8	746	1.9	–0.1
American Indian	109	0.4	166	0.4	0.0

7.3.2 Step 2: Establishing the Base Scale

Four forms (three operational forms and one form for release) were constructed for the 2008–2009 test administration. The forms were built to be consistent with the test blueprint using classical and IRT item statistics from the initial concurrent calibration. Figures 7.2 to 7.4 show the test characteristic curves (TCCs) for the three operational forms (Fall, Spring, and Summer) for each content area. The TCCs generally show the three forms to be similar (differences were within 5% of the range of test scores) across the full range of ability.

Figure 7.2: TCCs for Three Operational Forms for English II for the 2009 Test Administration



Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 7.3: TCCs for Three Operational Forms for Algebra I for the 2009 Test Administration

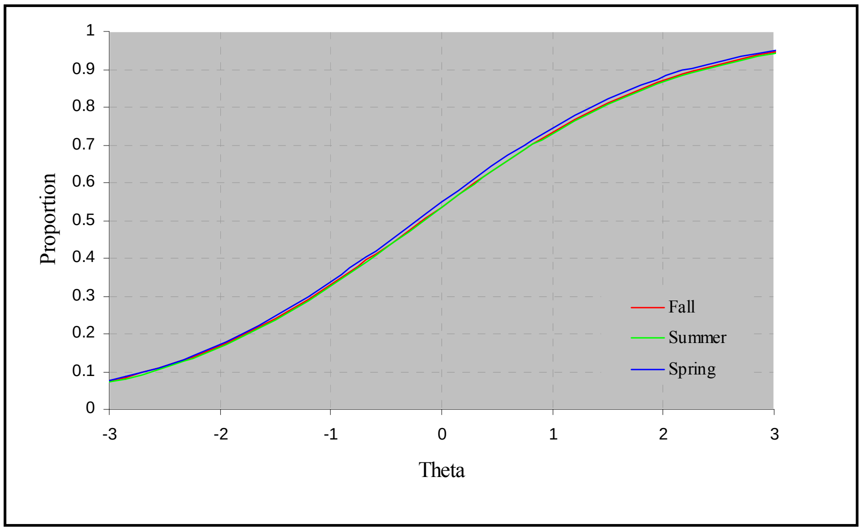
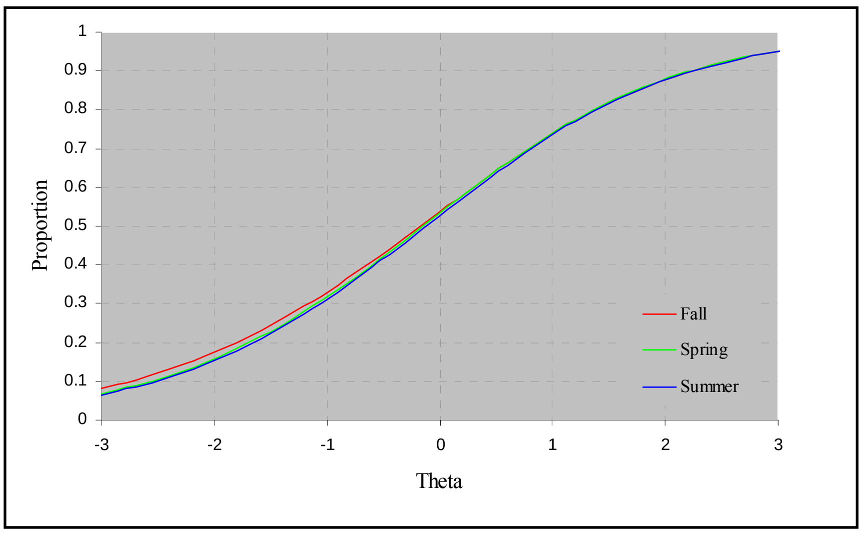


Figure 7.4: TCCs for Three Operational Forms for Biology for the 2009 Test Administration



One of the forms was chosen for release, while the other three forms were used for the Fall 2008, Spring 2009, and Summer 2009 administrations. The Spring 2009 form also contained new items embedded for field testing. No field test items were included on the Fall and Summer forms due to the small sample sizes participating in those administrations. However, the Fall and Summer forms each contained one set of 12 additional items, making the test length the same across all three administrations. These 12 items served as a means of linking the Fall and Summer forms to the Spring form for the post-equating check.[13] Table 7.5 shows the composition of the operational tests.

[13] This was done for Algebra I and Biology only. For English II, sets of filler items were included in the Fall and Summer forms.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.5: Operational Test Design for Core Assessments (Four Forms)

	Session I		Session II		
	Selected Response		Performance Event/ Writing Prompt		Writing Prompt
	OP	EFT	OP	EFT	Stand-alone
English II	35	12	1		1
Algebra I	35	12	1	1	
Biology	35	12	1	1	

OP = operational items; EFT = embedded field test items
Note: For fall and summer, EFT slots are used for linking items.

To establish the base scale for each content area test, calibrations of the Spring 2009 operational forms were executed freely, without constraint. These calibrations had to be performed before a complete set of data was available. Tables 7.6 through 7.8 provide a comparison of the calibration set and complete set of data for the Spring 2009 operational test forms. Inspection of these tables shows that the demographics for the calibration samples were similar to the census (generally less than one percent difference), or complete set, of data.

Table 7.6: Comparison of the Calibration and Census Data for the Spring 2009 Operational Test Forms, English II

	English II				
	Calibration Sample		Census Data		Difference (calibration minus census)
	N	%	*N*	%	%
All Students	49,415		59,011		
Gender					
Male	24,471	50.48	29,204	49.50	0.98
Female	24,944	49.52	29,807	50.50	–0.98
Race/Ethnicity					
White	40,306	81.57	47,277	80.10	1.47
Black	6,656	13.47	8,849	15.00	–1.53
Hispanic	1,338	2.71	1,615	2.70	0.01
Asian/Pacific Islander	870	1.76	989	1.70	0.06
American Indian	245	0.50	281	0.50	0.00

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.7: Comparison of the Calibration and Census Data for the Spring 2009 Operational Test Forms, Algebra I

	Algebra I				
	Calibration Sample		Census Data		Difference (calibration minus census)
	N	%	*N*	%	%
All Students	48,374		55,774		
Gender					
Male	23,713	49.02	27,496	49.30	–0.28
Female	24,661	50.98	28,278	50.70	0.28
Race/Ethnicity					
White	38,398	79.38	43,739	78.40	0.98
Black	7,241	14.97	8,886	15.90	–0.93
Hispanic	1,498	3.10	1,744	3.10	0.00
Asian/Pacific Islander	1,010	2.09	1,150	2.10	–0.01
American Indian	227	0.47	255	0.50	–0.03

Table 7.8: Comparison of the Calibration and Census Data for the Spring 2009 Operational Test Forms, Biology

	Biology				
	Calibration Sample		Census Data		Difference (calibration minus census)
	N	%	*N*	%	%
All Students	48,672		57,587		
Gender					
Male	23,849	49.00	28,165	48.90	0.10
Female	24,823	51.00	29,422	51.10	–0.10
Race/Ethnicity					
White	39,688	81.54	46,208	80.20	1.34
Black	6,493	13.34	8,433	14.60	–1.26
Hispanic	1,401	2.88	1,649	2.90	–0.02
Asian/Pacific Islander	849	1.74	1,026	1.80	–0.06
American Indian	241	0.50	271	0.50	0.00

Table 7.9 provides a comparison of classical item statistics for the item pool, based on 2008 field test data, and for the Fall 2008, Spring 2009, and Summer 2009 operational forms, based on the 2009 operational test administration for each content area. The comparison includes the percentage of items with *p*-values less than 0.3 and point-biserial correlations less than 0.1. Items with values below these criteria are typically considered low performing and are excluded from operational forms. However, such items may be included if the item pool is limited or if content considerations justify keeping an item. For example, an item may have poor field test statistics because of examinee motivational issues or because content is not currently being taught. Examination of the summary statistics in Table 7.9 generally supports test development efforts in selecting the highest-quality items for inclusion in each operational form. Summary statistics for the Spring 2009 operational administration are provided in Table 7.10.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.9: Comparison of 2008 Item Pool with 2008–2009 Operational Test Forms

Subject	Item Set	% *p*-Value < .3	% Point-Biserial < .1
English II	Phase I FT	6.0%	3.7%
	Fall	0.0%	0.0%
	Spring	0.0%	0.0%
	Summer	2.9%	2.9%
Algebra I	Phase I FT	17.2%	5.6%
	Fall	14.3%	2.9%
	Spring	8.6%	2.9%
	Summer	5.7%	0.0%
Biology	Phase I FT	6.7%	3.9%
	Fall	0.0%	5.7%
	Spring	2.9%	0.0%
	Summer	0.0%	0.0%

Table 7.10: Summary Statistics for the Spring 2009 Operational Administration

Content	Total Items	Total Points	Minimum	Maximum	Mean (Raw Score)	*SD* (Raw Score)
English II	36	39	3	39	27.47	6.22
Algebra I	36	39	2	39	22.43	7.19
Biology	36	55	3	55	33.09	9.66

Because the Rasch model is the basis of all scoring and scaling analyses associated with the EOC Assessments, the utility of the results from the Spring 2009 administration depends on the degree to which the assumptions of the model are met, as well as the degree to which the test data fit the model. The assumptions of the Rasch model are that (1) the data are unidimensional, and (2) the data have the quality of local independence, meaning that responses to one item do not depend on responses to another item. The sections below address these assumptions and include evaluations of the dimensionality and local independence of the data, as well as fit indices.

7.3.2.1 Assessing Unidimensionality of the Data

WINSTEPS provides a residual-based, unrotated principal components analysis (PCA) that can be used to assess the unidimensionality assumption of the Rasch model. The purpose of the analysis is to reveal contrasts between opposing factors by showing the variance explained by factors not accounted for by the Rasch model. That is, the Rasch dimension is removed first, and the residual variance is then analyzed. Consequently, with this analysis, one does not want to identify a second dimension that accounts for a practically significant amount of residual variance.

Ideally, additional factors will be at the “noise” level, implying that there are no other shared dimensions in the data. Because the WINSTEPS standardized residuals are modeled to have unit normal distributions, which are independent, a PCA of these residuals should look similar to a PCA of random normal deviates. Simulation studies

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

(such as Smith and Miao, 1994) indicate that the largest component in a set of random normal deviates would have an eigenvalue of about 1.4, which represents a small percentage of variance explained (i.e., less than 5%).

Table 7.11 shows the results of the PCA for the Spring 2009 operational form for each content area. For each analysis, the secondary dimension has an eigenvalue representing fewer than three items (less than 5% of the total variance) and, therefore, is of little practical import.

Table 7.11: Results of the PCA for the Spring 2009 Operational Tests

Content	Total Units (Items)	Second Dimension Eigenvalue	Second Dimension % of Total Variance Explained	% of Unexplained Variance	Second Dimension % of Unexplained Variance
English II	36	1.5	4.1%	56.4%	2.3%
Algebra I	36	1.5	4.1%	53.4%	2.2%
Biology	46	1.9	4.1%	56.0%	2.3%

7.3.2.2 Assessing Local Independence of the Data

Based on the PCA, WINSTEPS also provides standardized residual correlations that can be used to assess the local independence assumption of the Rasch model. The purpose of the analysis is to detect dependency between pairs of items. Figures 7.5 to 7.7 provide screen shots from WINSTEPS Table 23.99 (Linacre, 2006b) for each content area from the Spring 2009 operational test administration. Results of these analyses generally support the assumption of local independence. More specifically, values for standardized residual correlations were generally low (i.e., had absolute values below .10), indicating little dependency between pairs of items.

Figure 7.5: Standardized Residual Correlations from the Spring 2009 Administration for English II

```
TABLE 23.99 Spring 2009 Missouri EOC Local Calibr MOELAS09.OUT Oct  7 21:50 2009
INPUT: 49415 PERSONS  48 ITEMS  MEASURED: 49415 PERSONS  36 ITEMS  74 CATS  3.64.2
--------------------------------------------------------------------------------

     LARGEST STANDARDIZED RESIDUAL CORRELATIONS
     USED TO IDENTIFY DEPENDENT ITEMS
+-------------------------------------+
|RESIDUL| ENTRY        | ENTRY        |
|CORRELN|NUMBER ITEM   |NUMBER ITEM   |
|-------+--------------+--------------|
|   .13 |    45 I0045 |    46 I0046 |
|   .10 |     4 I0004 |     5 I0005 |
|   .08 |     4 I0004 |    11 I0011 |
|   .08 |    36 I0036 |    39 I0039 |
|-------+--------------+--------------|
|  -.11 |    27 I0027 |    29 I0029 |
|  -.11 |     4 I0004 |     6 I0006 |
|  -.09 |     6 I0006 |    32 I0032 |
|  -.08 |    29 I0029 |    48 I0048 |
|  -.08 |    27 I0027 |    45 I0045 |
|  -.08 |     9 I0009 |    11 I0011 |
+-------------------------------------+
```


Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 7.6: Standardized Residual Correlations from the Spring 2009 Administration for Algebra I

```
TABLE 23.99 Spring 2009 Missouri EOC Local Calibr MOMATS09.OUT Oct  7 21:51 2009
INPUT: 48374 PERSONS  48 ITEMS  MEASURED: 48374 PERSONS  36 ITEMS  75 CATS  3.64.2
--------------------------------------------------------------------------------

     LARGEST STANDARDIZED RESIDUAL CORRELATIONS
     USED TO IDENTIFY DEPENDENT ITEMS
+-----------------------------------+
|RESIDUL| ENTRY       | ENTRY       |
|CORRELN|NUMBER ITEM  |NUMBER ITEM  |
|-------+-------------+-------------|
|   .11 |    28 I0028 |    33 I0033 |
|-------+-------------+-------------|
|  -.11 |    32 I0032 |    48 I0048 |
|  -.10 |    36 I0036 |    48 I0048 |
|  -.10 |    33 I0033 |    48 I0048 |
|  -.10 |    19 I0019 |    45 I0045 |
|  -.10 |    14 I0014 |    48 I0048 |
|  -.09 |    21 I0021 |    48 I0048 |
|  -.09 |    20 I0020 |    48 I0048 |
|  -.09 |    35 I0035 |    37 I0037 |
|  -.08 |    28 I0028 |    48 I0048 |
+-----------------------------------+
```

Figure 7.7: Standardized Residual Correlations from the Spring 2009 Administration for Biology

```
TABLE 23.99 Spring 2009 Missouri EOC Local Calibr MOSCIS09.OUT Oct  7 21:51 2009
INPUT: 48672 PERSONS  58 ITEMS  MEASURED: 48672 PERSONS  46 ITEMS  101 CATS 3.64.2
--------------------------------------------------------------------------------

     LARGEST STANDARDIZED RESIDUAL CORRELATIONS
     USED TO IDENTIFY DEPENDENT ITEMS
+-----------------------------------+
|RESIDUL| ENTRY       | ENTRY       |
|CORRELN|NUMBER ITEM  |NUMBER ITEM  |
|-------+-------------+-------------|
|   .72 |    49 I0049 |    50 I0050 |
|   .18 |    13 I0013 |    33 I0033 |
|   .17 |    48 I0048 |    56 I0056 |
|   .15 |    55 I0055 |    56 I0056 |
|-------+-------------+-------------|
|  -.10 |    28 I0028 |    57 I0057 |
|  -.10 |    33 I0033 |    57 I0057 |
|  -.10 |    43 I0043 |    44 I0044 |
|  -.10 |    13 I0013 |    57 I0057 |
|  -.09 |    26 I0026 |    57 I0057 |
|  -.08 |     1 I0001 |    51 I0051 |
+-----------------------------------+
```


Copyright © 2010 by Missouri Department of Elementary and Secondary Education

7.3.2.3 Assessing Data Fit to the Model

WINSTEPS provides two statistics for indicating how well the data fit the Rasch model. Infit (inlier-sensitive or information-weighted fit) is sensitive to aberrations in item response patterns at the examinee's ability level. High infit statistics indicate unexpected responses to items that are well-targeted at the examinee's ability. Low infit statistics, while not a threat to measurement, may indicate over-fit of the data to the model (resulting in Guttman-like patterns) that may result in artificially inflated reliability statistics. Outfit (outlier-sensitive fit) is sensitive to outliers (in other words, to aberrant responses to items with difficulty far from a person's ability). High outfit values may indicate lucky guessing or careless mistakes. Relatively speaking, extremely high infit values are believed to be a greater threat to the measurement process than extreme outfit values.

Infit and outfit can be expressed as a mean square (MS) statistic or on a standardized metric (z). Both should be considered because they provide different perspectives: MS values are more oriented toward practical significance, while standardized values are more oriented toward statistical significance. Fit statistics expressed as mean squares (statistically, a chi-square statistic divided by its degrees of freedom) show the degree of practical distortion in the measurement. The expected value is 1.0, with values less than 1.0 indicating overfitting items (too predictable) and values greater than 1.0 indicating underfitting items (unpredictability, too much noise). Rules of thumb regarding "practically significant" MS fit values vary. Wright and Linacre (1994) suggest that reasonable MS fit values range from 0.8 to 1.2 for SR items. Others believe that reasonable test results can be achieved with values from 0.5 to 1.5. Riverside Publishing has typically considered values outside the range of 0.7 to 1.3 to be outside the range of acceptable fit.

Fit statistics expressed as z-scores (standardized unit normal deviates) offer a means to statistically test model fit. Standardized fit statistics show the degree of statistical improbability in the data (i.e., its significance) if the data actually do fit the model. The expected value of standardized fit statistics is 0.0, with values significantly less than 0.0 indicating too much predictability and values significantly greater than 0.0 indicating lack of predictability. Also, z-scores may be affected by sample sizes. For example, in a large sample, the test of interest might show a statistically significant difference. In practice, the difference might not be important.

Tables 7.12 to 7.14 provide summary statistics, including summary fit statistics, for the Spring 2009 operational test calibrations, which were used to establish the base scale for the EOC Assessments. The evaluation of fit values, specifically MS infit, yielded these results: Infit values for English II ranged from 0.79 to 1.18, values for Algebra I ranged from 0.85 to 1.27, and values for Biology ranged from 0.81 to 1.13. The fit values and output files are based on the local runs using WINSTEPS version 3.64.2. Tables 7.15 to 7.17 provide Rasch difficulties and item fit statistics.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.12: Summary Statistics for the Spring 2009 Operational Test Calibrations for English II

Statistic	Rasch Difficulty Estimate	*p*-value	Infit		Outfit		Point-Biserial
			MS	Standardized	MS	Statistic	
# of Items	36	36	36	36	36	36	36
Mean	0.10	0.65	0.97	–1.24	0.95	–1.11	0.39
SD	0.81	0.33	0.10	5.40	0.17	5.44	0.10
Minimum	–1.23	0.36	0.79	–9.90	0.69	–9.90	0.19
Percentiles							
10	–1.05	0.39	0.83	–9.54	0.71	–8.17	0.25
25	–0.69	0.50	0.89	–5.05	0.80	–5.52	0.33
50	0.15	0.62	0.97	–1.81	0.96	–1.42	0.39
75	0.68	0.77	1.05	3.08	1.07	2.32	0.46
90	1.22	0.82	1.12	6.60	1.19	7.06	0.54
Maximum	1.40	2.37	1.18	9.90	1.28	9.90	0.58

Table 7.13: Summary Statistics for the Spring 2009 Operational Test Calibrations for Algebra I

Statistic	Rasch Difficulty Estimate	*p*-value	Infit		Outfit		Point-Biserial
			MS	Standardized	MS	Standardized	
# of Items	36	36	36	36	36	36	36
Mean	–0.14	0.55	0.99	–0.91	1.01	–0.40	0.31
SD	0.89	0.29	0.08	4.24	0.14	4.33	0.09
Minimum	–1.86	0.12	0.85	–9.90	0.77	–8.94	0.08
Percentiles							
10	–1.19	0.29	0.90	–6.35	0.84	–5.88	0.21
25	–0.83	0.38	0.94	–3.42	0.91	–3.19	0.24
50	–0.20	0.53	0.98	–1.22	0.98	–0.86	0.31
75	0.50	0.66	1.05	2.12	1.08	3.12	0.38
90	1.06	0.74	1.08	3.98	1.21	5.23	0.43
Maximum	2.23	1.98	1.27	9.90	1.40	9.90	0.48

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.14: Summary Statistics for the Spring 2009 Operational Test Calibrations for Biology

Statistic	Rasch Difficulty Estimate	*p*-value	Infit		Outfit		Point-Biserial
			MS	Standardized	MS	Statistic	
# of Items	46	46	46	46	46	46	46
Mean	–0.16	0.65	0.99	–0.49	0.98	–0.40	0.35
SD	0.72	0.31	0.07	4.02	0.12	4.10	0.11
Minimum	–2.10	0.27	0.81	–9.90	0.74	–8.51	0.11
Percentiles							
10	–1.17	0.41	0.90	–5.34	0.84	–5.06	0.19
25	–0.58	0.50	0.93	–3.14	0.90	–3.59	0.28
50	–0.09	0.61	0.99	–0.69	0.98	–0.58	0.33
75	0.46	0.72	1.04	2.77	1.05	2.06	0.41
90	0.66	0.93	1.10	5.87	1.15	6.65	0.51
Maximum	1.32	2.26	1.13	8.68	1.33	8.52	0.5

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.15: Item Statistics for the Spring 2009 Operational Test Calibrations for English II

Item Number	Item Type	Rasch Difficulty Estimate	*N*	MS Infit	Standardized Infit	MS Outfit	Standardized Outfit
1	SR	–0.1206	49324	0.94	–9.90	0.86	–9.90
2	SR	–0.1534	49324	1.06	9.90	1.15	9.90
3	SR	–1.2769	49324	1.04	4.19	1.14	7.08
4	SR	–1.2199	49324	0.90	–9.90	0.71	–9.90
5	SR	–1.4390	49324	0.98	–2.01	0.91	–4.39
6	SR	1.7859	49324	1.16	9.90	1.34	9.90
7	SR	–0.9242	49324	0.95	–5.43	0.82	–9.90
8	SR	–0.1477	49324	0.94	–9.90	0.91	–8.89
9	SR	1.1881	49324	1.03	8.94	1.08	9.90
10	SR	–0.1864	49324	1.00	–0.15	0.99	–1.18
11	SR	–0.7860	49324	0.87	–9.90	0.69	–9.90
12	SR	0.9749	49324	1.05	9.90	1.09	9.90
24	SR	0.7396	49324	1.07	9.90	1.10	9.90
25	SR	0.9281	49324	1.04	9.90	1.05	8.41
26	SR	–1.0475	49324	0.89	–9.90	0.75	–9.90
27	SR	0.8934	49324	0.86	–9.90	0.82	–9.90
28	SR	0.3485	49324	1.08	9.90	1.13	9.90
29	SR	1.3961	49324	1.28	9.90	1.42	9.90
30	SR	1.4290	49324	0.96	–9.90	0.98	–2.75
31	SR	–0.6387	49324	0.93	–9.48	0.86	–9.90
32	SR	–0.1762	49324	0.87	–9.90	0.76	–9.90
33	SR	1.3749	49324	1.06	9.90	1.11	9.90
34	SR	0.4193	49324	1.04	8.91	1.06	8.67
35	SR	–1.3572	49324	0.90	–9.83	0.81	–9.90
36	SR	–1.6394	49324	0.86	–9.90	0.60	–9.90
37	SR	0.6507	49324	1.05	9.90	1.09	9.90
38	SR	0.7489	49324	0.98	–5.89	0.97	–5.59
39	SR	–0.8995	49324	0.93	–9.16	0.90	–6.65
40	SR	0.4086	49324	1.01	2.03	1.00	0.56
41	SR	0.0747	49324	0.98	–3.39	1.02	2.43
43	SR	–1.0969	49324	0.94	–6.47	0.84	–9.40
44	SR	0.9452	49324	1.00	0.34	1.01	2.51
45	SR	–0.5446	49324	1.04	6.21	1.17	9.90
46	SR	–0.4962	49324	1.09	9.90	1.17	9.90
47	SR	–0.0808	49324	0.99	–2.44	0.95	–5.46
48	WP	–0.0743	49324	0.94	–8.11	0.94	–8.08

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.16: Item Statistics for the Spring 2009 Operational Test Calibrations for Algebra I

Item Number	Item Type	Rasch Difficulty Estimate	*N*	MS Infit	Standardized Infit	MS Outfit	Standardized Outfit
1	SR	–2.2165	48297	0.97	–2.25	0.88	–5.89
2	SR	–1.0292	48297	1.13	9.90	1.38	9.90
3	SR	–0.7039	48297	0.93	–9.90	0.88	–9.90
4	SR	–0.8042	48297	1.04	7.03	1.07	6.66
5	SR	–0.7678	48297	0.91	–9.90	0.87	–9.90
10	SR	–1.3993	48297	0.99	–0.80	0.97	–1.96
11	SR	–0.6732	48297	1.01	2.50	1.06	6.25
12	SR	–0.8986	48297	0.94	–9.90	0.87	–9.90
13	SR	–0.1044	48297	0.94	–9.90	0.91	–9.90
14	SR	0.7574	48297	0.99	–1.54	1.01	0.91
15	SR	–0.5451	48297	0.89	–9.90	0.81	–9.90
16	SR	–0.4357	48297	0.93	–9.90	0.88	–9.90
17	SR	–1.4179	48297	0.93	–9.03	0.84	–9.90
18	SR	–1.2795	48297	0.91	–9.90	0.86	–9.90
19	SR	–0.3714	48297	0.83	–9.90	0.77	–9.90
20	SR	0.0915	48297	1.05	9.90	1.05	7.84
21	SR	0.2794	48297	0.94	–9.90	0.93	–9.90
26	SR	–0.2492	48297	0.98	–3.90	0.99	–0.90
27	SR	0.4775	48297	0.98	–4.55	0.98	–3.38
28	SR	0.1925	48297	1.01	3.02	1.00	0.76
29	SR	0.3360	48297	1.01	2.67	1.01	2.20
30	SR	0.3135	48297	1.09	9.90	1.12	9.90
31	SR	0.0591	48297	0.98	–6.00	0.96	–6.52
32	SR	0.5486	48297	0.96	–9.90	0.95	–8.96
33	SR	0.8881	48297	0.99	–3.65	1.02	2.84
34	SR	0.4535	48297	0.98	–6.12	0.98	–3.12
35	SR	–0.1195	48297	1.03	7.86	1.01	0.87
36	SR	1.0415	48297	0.98	–4.42	1.05	8.47
37	SR	0.0195	48297	0.98	–5.78	0.96	–6.20
38	SR	1.0504	48297	1.09	9.90	1.16	9.90
43	SR	1.3772	48297	1.05	9.90	1.10	9.90
44	SR	0.0171	48297	1.05	9.90	1.08	9.90
45	SR	0.8424	48297	1.19	9.90	1.27	9.90
46	SR	2.8078	48297	1.06	6.24	1.61	9.90
47	SR	1.2618	48297	1.04	9.52	1.12	9.90
48	PE	0.2005	48297	1.15	9.90	1.18	9.90

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.17: Item Statistics for the Spring 2009 Operational Test Calibrations for Biology

Item Number	Item Type	Rasch Difficulty Estimate	*N*	MS Infit	Standardized Infit	MS Outfit	Standardized Outfit
1	SR	–1.1185	48669	0.94	–8.39	0.86	–9.90
2	PE	–0.8996	48669	0.93	–9.90	0.82	–9.90
3	SR	–1.9827	48669	1.05	3.71	1.25	9.90
4	SR	0.0018	48669	1.05	9.90	1.05	8.19
5	SR	–1.3239	48669	0.98	–2.56	0.96	–2.68
10	SR	–0.2673	48669	0.96	–9.56	0.94	–8.76
11	SR	–0.6320	48669	0.95	–9.08	0.92	–9.19
12	SR	0.4324	48669	1.04	9.90	1.04	8.53
13	SR	–0.1247	48669	0.93	–9.90	0.87	–9.90
14	SR	0.8547	48669	1.05	9.90	1.07	9.90
15	SR	0.3053	48669	1.14	9.90	1.16	9.90
16	SR	–0.7564	48669	0.93	–9.90	0.88	–9.90
17	SR	–2.3836	48669	0.89	–7.52	0.59	–9.90
18	SR	–0.3570	48669	0.90	–9.90	0.85	–9.90
19	SR	0.6910	48669	0.96	–9.90	0.96	–9.07
20	SR	1.5737	48669	1.09	9.90	1.23	9.90
21	SR	0.1878	48669	1.11	9.90	1.16	9.90
26	SR	–0.6554	48669	0.90	–9.90	0.79	–9.90
27	SR	–0.1494	48669	1.00	–0.30	0.97	–4.15
28	SR	0.6145	48669	0.84	–9.90	0.82	–9.90
29	SR	1.0411	48669	0.97	–8.75	1.00	–0.56
30	SR	0.9980	48669	1.10	9.90	1.14	9.90
31	SR	0.9808	48669	0.98	–5.09	1.00	–0.03
32	SR	0.1842	48669	0.95	–9.90	0.92	–9.90
33	SR	0.1323	48669	0.94	–9.90	0.91	–9.90
34	SR	0.1575	48669	0.96	–9.90	0.93	–9.90
35	SR	–0.4869	48669	1.03	5.93	1.07	7.85
36	SR	–0.0020	48669	0.95	–9.90	0.92	–9.90
37	SR	0.7227	48669	1.03	7.49	1.04	8.21
38	SR	–0.5109	48669	0.97	–4.96	0.93	–8.03
43	SR	0.0796	48669	1.10	9.90	1.15	9.90
44	SR	0.3442	48669	1.02	5.70	1.02	4.41
45	SR	0.8969	48669	0.97	–9.81	0.98	–4.71
46	SR	0.3037	48669	0.99	–2.84	0.98	–3.09
47	SR	–1.3029	48669	0.91	–9.90	0.77	–9.90
48	PE	–0.2292	48351	0.97	–5.57	1.00	–0.11
49	SR	–0.8921	48384	0.97	–4.74	0.89	–9.90
50	PE	–0.4974	48343	0.94	–9.90	0.86	–9.90
51	PE	–0.0256	48510	1.03	4.47	1.03	4.71
52	PE	1.3149	48271	1.16	9.90	1.18	9.90
53	PE	0.7933	47900	1.15	9.90	1.17	9.90
54	PE	0.2967	48193	0.99	–1.93	0.99	–0.84
55	PE	–0.6367	48264	0.95	–9.03	0.90	–9.90
56	PE	–0.1707	47997	0.91	–9.90	0.88	–9.90
57	PE	1.4149	47800	1.26	9.90	1.36	9.90
58	PE	1.0826	47178	1.11	9.90	1.11	9.90

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

7.3.2.4 Establish Scaling Transformations

Total scores for the EOC Assessments were reported in scale scores with a range of 100–250. A scale score of 200 represents the cut point between Basic and Proficient, and a scale score of 225 represents the cut point between Proficient and Advanced. The scale score ranges are displayed in Table 7.18.

Table 7.18: Scale Score Ranges for EOC Assessment Achievement Levels

EOC Assessment	Achievement Level	Scale Score Range
English II	Below Basic	100 to 179
	Basic	180 to 199
	Proficient	200 to 224
	Advanced	225 to 250
Algebra I	Below Basic	100 to 176
	Basic	177 to 199
	Proficient	200 to 224
	Advanced	225 to 250
Biology	Below Basic	100 to 176
	Basic	177 to 199
	Proficient	200 to 224
	Advanced	225 to 250

To produce these scale score ranges, linear transformations were applied to theta estimates and scale scores. The following formula was used to obtain the slopes and intercepts for the transformation functions:

$$sc(y) = \left[\frac{sc(y_2) - sc(y_1)}{\theta_2 - \theta_1}\right] y + \left\{ (sc(y_1) - \left[\frac{sc(y_2) - sc(y_1)}{\theta_2 - \theta_1}\right]\theta_1 \right\},$$

where θ_1 and θ_2 are person parameter estimates that correspond to the cut score points, and $sc(y_1)$ and $sc(y_2)$ are scale score points. This formula was adopted from Kolen and Brennan (2004, p. 337). For the Spring 2009 base scale, $sc(y_1)$ was 200 and $sc(y_2)$ was 225. Slopes and intercepts of the transformation functions are summarized in Table 7.19. These same slopes and intercepts will be applied to all future forms for each content area.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.19: Summary of Slopes and Intercepts of Theta to Scale Score Transformation Functions by Content Area

	Basic			Proficient			Advanced				
	Raw Score	Theta	Scale Score	Raw Score	Theta	Scale Score	Raw Score	Theta	Scale Score	Slope	Intercept
English II	15	–0.71	180	24	0.51	200	33	2.04	225	16.35	191.72
Algebra I	13	–0.80	177	22	0.36	200	31	1.61	225	19.96	192.83
Biology	18	–0.69	177	32	0.51	200	45	1.79	225	19.53	189.99

In addition to the above scaling transformation, the following rules were applied for the Fall 2008 operational tests:

- The raw score cut (e.g., for Proficient) was selected as the lowest raw score associated with a rounded scale score of 200. The same strategy was also followed for a scale score of 225.
- If there was no raw score associated with a rounded scale score of 200, the raw score with the highest scale score below 200 was selected as the cut score and assigned a scale score of 200. For example, if two consecutive raw scores were associated with rounded scale scores of 198 and 201, the scale score of 198 was moved up to 200. The same strategy was also followed for a scale score of 225.
- Scale scores below 100 were rounded up to 100.
- Scale scores above 250 were rounded down to 250.
- For each test, for a perfect raw score, the scale score was set to 250.

7.3.3 Step 3: Examine Stability of the Common Items

Although the concurrent calibrations following the 2008 standalone field test were sufficient for developing a common scale for the item pools and for building alternate forms (see Step 2), the Spring 2009 operational administration of the EOC Assessments was chosen as the base form. To equate or recenter the Spring 2008 item pool to the Spring 2009 base scale, the Rasch values for the common items (i.e., the Spring 2009 operational items) were fixed to the 2009 parameter estimates. Next, using the Spring 2008 standalone field test data, the concurrent calibration with the complete pool of items was repeated, this time with the 2009 operational item parameters fixed to their Spring 2009 values. Before the concurrent calibrations were completed, the stability of the common items was assessed for each content area.

The stability of common items should be examined visually and statistically (Kolen and Brennan, 2004). For example, scatterplots can be used to check visually for outlier common items. The scatter points for items that function similarly should line up along a straight line. Outlier items will not fall on the straight line and thus can be seen visually. In addition to a visual examination, an analytical study of the stability of common items may be performed. It is recommended that a 0.30-logistic unit should be applied as a cut criterion for removing “unstable” common items (Miller, Rotou, and Twing, 2004).

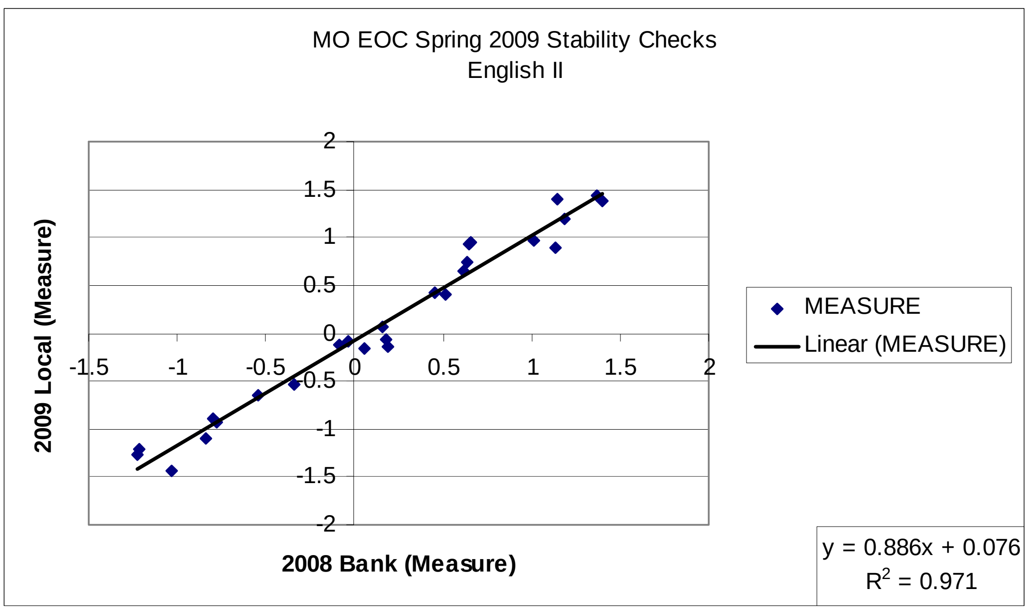
Copyright © 2010 by Missouri Department of Elementary and Secondary Education

To study the stability of the common items, the displacement value for each operational item (i.e., the common items) was evaluated after calibrating the items with the operational items fixed, or anchored, to their Spring 2009 difficulty values. Any common item with a displacement greater than 0.30 logits was removed from the common item set and treated as a new item. The fixed calibration was then performed again with the unstable common item free to be estimated. The displacement value for each of the common items was then re-evaluated. As with the previous step, any outlier items identified during this procedure were removed from the rescaling process. Table 7.20 shows the number of items dropped from the set of operationally administered items (i.e., the common set of items) for instability. Figures 7.8 to 7.10 show scatterplots for the final set of common items used to recenter each content area's item bank or pool of items.

Table 7.20: Number of Items Dropped from the Common Set of Operational Items

Subject	Number of Items Dropped
English II	10
Algebra I	5
Biology	8

Figure 7.8: Scatterplot of Stable Linking Items for English II



Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 7.9: Scatterplot of Stable Linking Items for Algebra I

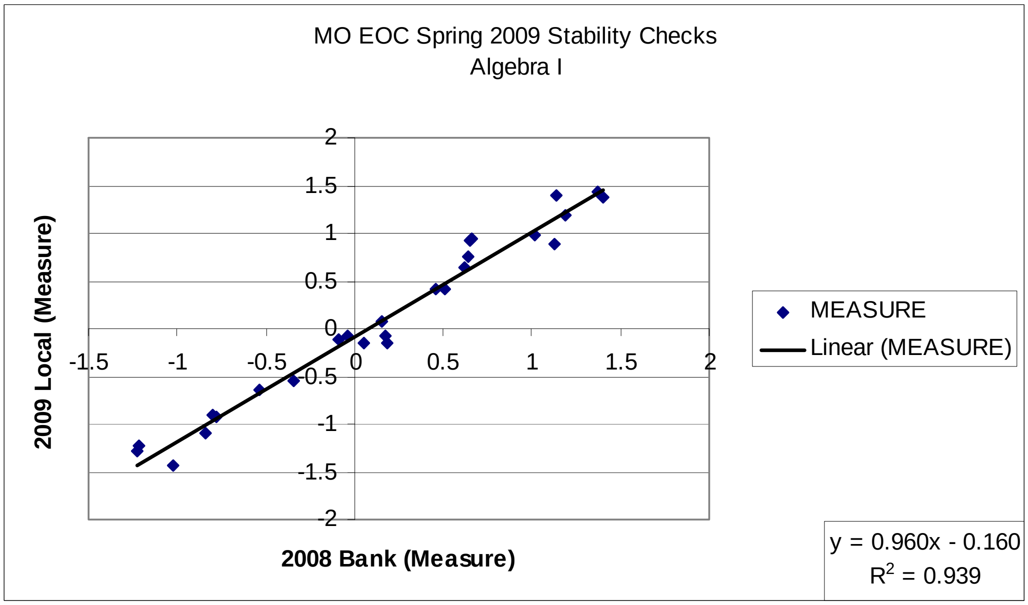


Figure 7.10: Scatterplot of Stable Linking Items for Biology

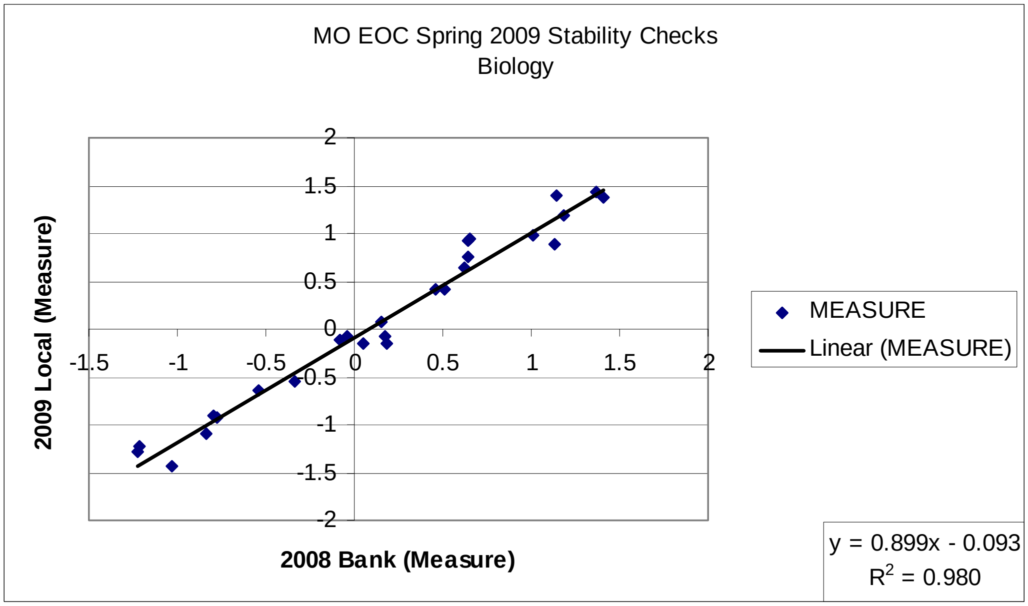


Table 7.21 summarizes displacement statistics for the common items generated with the anchored calibrations. Linacre (2006a) suggests that "random displacements of less than 0.50 logits are unlikely to have much impact in a test instrument" (p. 203). However, as discussed above, the 0.30 criteria for removing unstable items was used. Table 7.21 shows that all displacement statistics for the common items are smaller than 0.30, indicating that the anchored calibrations converged well.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.21: Displacement Statistics for the Spring 2009 Recentering of the Item Pool

English II		Algebra I		Biology	
2009 Operational Item	Displacement	2009 Operational Item	Displacement	2009 Operational Item	Displacement
1	–0.10	2	–0.21	1	0.03
2	0.07	3	–0.13	2	0.08
3	–0.09	4	0.16	3	0.01
4	–0.14	5	0.26	4	0.15
5	0.28	7	0.09	5	0.21
7	0.10	8	0.06	6	0.01
8	0.20	9	–0.16	7	0.26
9	–0.05	10	–0.10	8	–0.12
12	–0.10	11	–0.14	9	0.00
14	–0.25	12	–0.21	10	–0.29
16	0.26	13	0.08	11	–0.16
18	–0.22	14	0.26	12	0.19
19	–0.07	15	0.25	14	0.26
20	0.02	16	–0.15	15	0.19
22	0.02	17	–0.20	16	–0.15
23	–0.04	18	0.08	17	0.05
26	–0.08	19	0.18	18	0.19
27	–0.16	20	0.01	19	–0.13
28	0.05	21	–0.06	20	0.12
29	0.04	22	–0.22	21	–0.15
30	0.03	23	0.01	23	–0.24
31	0.14	24	–0.09	24	–0.02
32	–0.29	25	0.23	25	0.08
33	0.07	26	–0.06	26	0.25
35	–0.09	27	0.29	27	–0.10
36	0.16	28	–0.08	28	–0.17
		29	–0.08	29	–0.02
		30	–0.15	30	0.11
		33	–0.04	32	–0.26
		35	0.10	33	–0.16
		36	–0.19	34	–0.10
				35	0.03
				36	–0.28
				39	0.08
				42	–0.14
				43	0.21
				44	–0.26
				46	–0.30

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

7.3.4: Step 4: Recenter the 2008 Item Bank

To equate or recenter the 2008 item pool to the 2009 base scale, the Rasch values for the stable common items (i.e., the stable Spring operational items) were fixed to the 2009 parameter estimates. With the Spring 2009 operational item parameters fixed, the rest of the item pool was equated to the Spring 2009 base scale. Note that it was assumed that the latent traits measured by the 2009 operational tests and the 2008 field tests were the same. Given the common items used across the two testing events and given that the blueprint and item specifications were the same, it seems reasonable to assume that the underlying latent trait or construct measured by each assessment was the same. The above procedure was a one-time-only activity occurring after the first operational administration. With the pool recentered, the Fall 2008 and Summer 2009 forms were retroactively equated to the Spring 2009 form. Although pre-equating occurred after the administration of the Fall 2008 form, the results were not reported until after the Spring 2009 administration and the item pool recentering. Tables 7.22 to 7.27 provide the raw score to scale score conversions for Fall 2008 and Spring 2009, respectively.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.22: Raw Score to Scale Score Conversions for Fall 2008, English II

Raw Score	Scale Score	*CSEM*
0	103	30
1	124	17
2	136	12
3	143	10
4	149	9
5	153	8
6	157	8
7	160	7
8	163	7
9	166	7
10	168	6
11	171	6
12	173	6
13	175	6
14	178	6
15	180	6
16	182	6
17	184	6
18	186	6
19	188	6
20	190	6
21	192	6
22	194	6
23	196	6
24	198	6
25	200	6
26	202	6
27	205	6
28	207	6
29	210	7
30	212	7
31	215	7
32	218	7
33	225	8
34	226	8
35	231	9
36	237	11
37	245	13
38	250	17
39	250	30

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.23: Raw Score to Scale Score Conversions for Fall 2008, Algebra I

Raw Score	Scale Score	*CSEM*
0	100	37
1	112	20
2	127	15
3	136	12
4	143	11
5	149	10
6	153	9
7	157	9
8	161	8
9	165	8
10	168	8
11	171	8
12	174	8
13	177	7
14	179	7
15	182	7
16	184	7
17	187	7
18	190	7
19	192	7
20	194	7
21	197	7
22	200	7
23	202	7
24	204	7
25	207	7
26	210	7
27	213	7
28	215	8
29	218	8
30	222	8
31	225	8
32	229	9
33	233	9
34	237	10
35	243	11
36	250	12
37	250	15
38	250	20
39	250	37

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.24: Raw Score to Scale Score Conversions for Fall 2008, Biology

Raw Score	Scale Score	*CSEM*
0	100	36
1	103	20
2	118	15
3	127	12
4	134	11
5	139	10
6	144	9
7	148	9
8	151	8
9	154	8
10	157	7
11	160	7
12	162	7
13	165	7
14	167	6
15	169	6
16	171	6
17	173	6
18	175	6
19	177	6
20	178	6
21	180	6
22	181	6
23	183	6
24	185	5
25	186	5
26	188	5
27	189	5
28	191	5
29	192	5
30	194	5
31	195	5
32	197	5
33	198	5
34	200	6
35	201	6
36	203	6
37	204	6
38	206	6
39	208	6
40	210	6
41	212	6
42	213	6
43	216	6
44	218	7
45	220	7
46	223	7
47	225	7

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.24: Raw Score to Scale Score Conversions for Fall 2008, Biology (continued)

Raw Score	Scale Score	*CSEM*
48	228	8
49	232	8
50	236	9
51	240	10
52	247	12
53	250	14
54	250	20
55	250	36

Table 7.25: Raw Score to Scale Score Conversions for Spring 2009, English II

Raw Score	Scale Score	*CSEM*
0	100	30
1	105	30
2	125	17
3	137	12
4	145	10
5	150	9
6	155	8
7	158	8
8	162	7
9	165	7
10	168	7
11	171	7
12	173	6
13	176	6
14	178	6
15	180	6
16	182	6
17	185	6
18	187	6
19	189	6
20	191	6
21	193	6
22	196	6
23	198	6
24	200	6
25	202	6
26	205	6
27	207	6
28	210	7
29	212	7
30	215	7
31	218	7
32	221	7

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.25: Raw Score to Scale Score Conversions for Spring 2009, English II (continued)

Raw Score	Scale Score	*CSEM*
33	225	8
34	229	8
35	234	9
36	240	10
37	248	12
38	250	17
39	250	30

Table 7.26: Raw Score to Scale Score Conversions for Spring 2009, Algebra I

Raw Score	Scale Score	*CSEM*
0	100	37
1	111	21
2	126	15
3	136	13
4	143	11
5	148	10
6	153	9
7	157	9
8	161	9
9	165	8
10	168	8
11	171	8
12	174	8
13	177	7
14	180	7
15	182	7
16	185	7
17	187	7
18	190	7
19	192	7
20	195	7
21	197	7
22	200	7
23	203	7
24	205	7
25	208	7
26	210	7
27	213	7
28	216	8
29	219	8
30	222	8
31	225	8
32	229	9

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.26: Raw Score to Scale Score Conversions for Spring 2009, Algebra I (continued)

Raw Score	Scale Score	*CSEM*
33	232	9
34	237	10
35	242	11
36	249	12
37	250	15
38	250	21
39	250	37

Table 7.27: Raw Score to Scale Score Conversions for Spring 2009, Biology

Raw Score	Scale Score	*CSEM*
0	100	36
1	107	20
2	121	14
3	130	12
4	137	11
5	142	10
6	146	9
7	150	8
8	153	8
9	156	8
10	159	7
11	162	7
12	164	7
13	166	7
14	169	6
15	171	6
16	173	6
17	175	6
18	177	6
19	178	6
20	180	6
21	182	6
22	184	6
23	185	6
24	187	6
25	189	6
26	190	6
27	192	6
28	193	6
29	195	6
30	197	6
31	198	6
32	200	6
33	202	6
34	203	6

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.27: Raw Score to Scale Score Conversions for Spring 2009, Biology (continued)

Raw Score	Scale Score	*CSEM*
35	205	6
36	207	6
37	208	6
38	210	6
39	212	6
40	214	6
41	216	6
42	218	6
43	220	7
44	223	7
45	225	7
46	228	7
47	231	8
48	234	8
49	238	9
50	242	9
51	247	10
52	250	12
53	250	14
54	250	20
55	250	36

7.3.5 Step 5: Place the 2009 Embedded Field Test Items onto the 2009 Scale

The 2009 embedded field test items were treated separately in this process to avoid having them influence calibration of the operational items and the establishment of the base scale. To bring the field-test items onto the base scale, a second calibration of the Spring data, fixing the 2009 operational parameter estimates, was conducted.

For the 2009–2010 operational administration, three new forms were built for each content area from the calibrated and recentered item pools (one each for Fall, Spring, and Summer). These new forms are pre-equated to the base form because, after the recentering of the pool, all previously field tested items are on the operational scale. For all subsequent years, one new form will be built from the calibrated pool for the Spring administration. A form reuse plan will be implemented for the Fall and Summer administrations. Each Fall and Summer form from the 2008–2009 and 2009–2010 administrations will be used in alternating years and in alternating administrations.

For the new Spring form in 2011, new items were field tested. The field test items will be scaled to the pool, using a calibration in which the operational test item parameters are fixed and the new field test items are free to be estimated. This process will allow all new items for all Spring form administrations to be placed on the same EOC Assessment scale as the other items in the pool.

As outlined above, not only can the pre-equating model be used to annually build alternate test forms, but by using the embedded field testing approach, DESE will also be able to maintain its item pools.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

7.3.6 Step 6: Perform Fixed Calibrations on the 2010 Operational Forms

To place the 2009–2010 operational forms onto the 2009 scale, an anchored item calibration was performed by fixing the parameters with the estimates resulting from Step 5 above. Figures 7.11 to 7.13 show the TCCs for the three operational forms (Fall, Spring, and Summer) for each content area. The TCCs generally show the three forms to be similar (differences were within 5% of the range of test scores) across the full range of ability.

Figure 7.11: TCCs for Three Alternate Forms for English II for the 2010 Test Administration

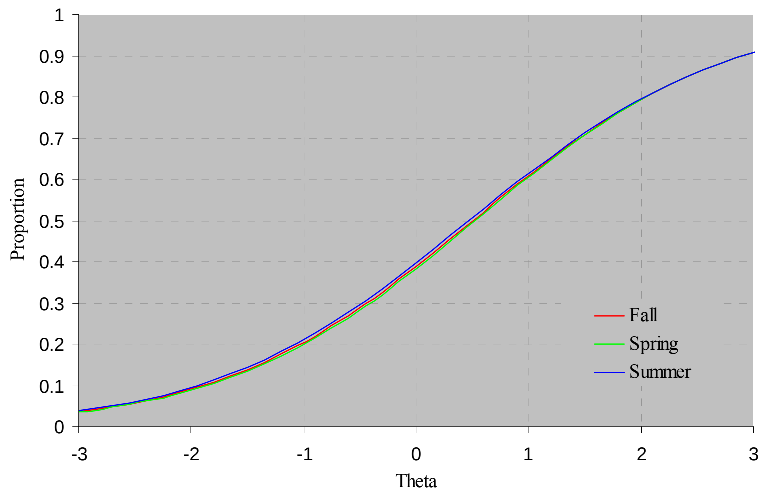
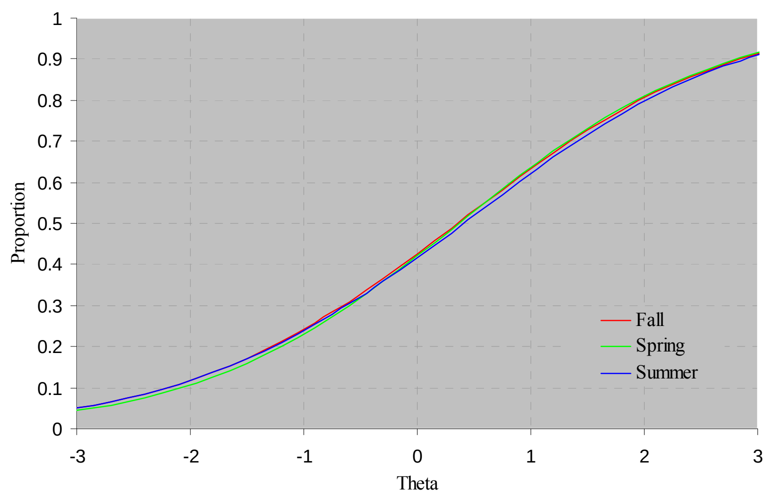


Figure 7.12: TCCs for Three Alternate Forms for Algebra I for the 2010 Test Administration



Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 7.13: TCCs for Three Alternate Forms for Biology for the 2010 Test Administration

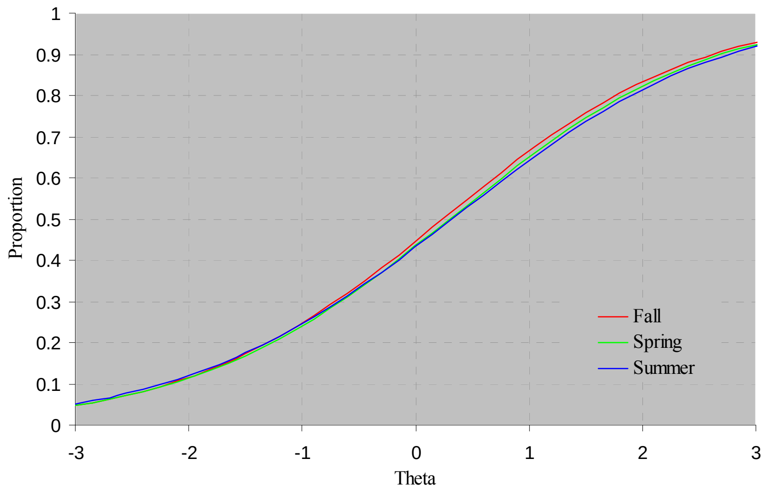


Table 7.28 provides a comparison of classical item statistics for the item pool based on 2009 embedded field test data and for the Summer 2009, Fall 2009, and Spring 2010 operational forms based on the 2009–2010 operational test administration for each content area. The comparison includes the percentage of items with *p*-values less than 0.3 and point-biserial correlations less than 0.1. Items with values below these criteria are typically considered low performing and are excluded from operational forms. However, such items may be included if the item pool is limited or if content considerations justify keeping an item. For example, an item may have poor field-test statistics because of examinee motivational issues or because content is not currently being taught. Examination of the summary statistics in Table 7.28 generally supports test development efforts in selecting the highest-quality items for inclusion in each operational form. Summary statistics for the Spring 2010 operational administration are provided in Table 7.29.

Table 7.28: Comparison of 2009 Item Pool with 2009–2010 Operational Test Forms

Subject	Item Set	% *p*-Value < .3	% Point-Biserial < .1
English II	Phase I EFT	0.00%	5.10%
	Summer	2.90%	2.90%
	Fall	0.00%	0.00%
	Spring	0.00%	0.00%
Algebra I	Phase I EFT	7.60%	3.80%
	Summer	5.70%	0.00%
	Fall	8.30%	0.00%
	Spring	11.10%	2.80%
Biology	Phase I EFT	1.70%	1.00%
	Summer	0.00%	0.00%
	Fall	0.00%	0.00%
	Spring	0.00%	2.80%

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.29: Summary Statistics for the Spring 2010 Operational Administration

Content	Total Items	Total Points	Minimum	Maximum	Mean (Raw Score)	*SD* (Raw Score)
English II	36	39	1	39	27.65	6.18
Algebra I	36	39	1	39	23.22	7.26
Biology	36	55	1	55	31.55	9.87

Because the Rasch model is the basis of all scoring and scaling analyses associated with the EOC Assessments, the utility of the results from the Spring 2010 administration depends on the degree to which the assumptions of the model are met as well as the degree to which the test data fit the model. As noted in Section 7.3.2, the assumptions of the Rasch model are that (1) the data are unidimensional, and (2) the data have the quality of local independence, meaning that responses to one item do not depend on responses to another item. The analyses below address these assumptions and include evaluations of the dimensionality and local independence of the data, as well as fit indices.

Table 7.30 shows the results of the PCA for the Spring 2010 operational form for each content area. For each analysis, the secondary dimension has an eigenvalue representing fewer than three items (less than 5% of the total variance) and, therefore, is of little practical import.

Table 7.30: Results of the PCA for the Spring 2010 Operational Tests

Content	Total Units (Items)	Second Dimension Eigenvalue	Second Dimension % of Total Variance Explained	% of Unexplained Variance	Second Dimension % of Unexplained Variance
English II	36	1.7	4.7%	54.4%	2.6%
Algebra I	36	1.8	4.9%	51.9%	2.5%
Biology	47	2.1	4.4%	51.5%	2.3%

Figures 7.14 to 7.16 provide screen shots from WINSTEPS Table 23.99 (Linacre, 2006b) for each content area from the Spring 2010 operational test administration. Results of these analyses generally support the assumption of local independence. More specifically, values for standardized residual correlations were generally low (i.e., had absolute values below .10), indicating little dependency between pairs of items.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 7.14: Standardized Residual Correlations from the Spring 2010 Administration for English II

```
TABLE 23.99 Spring 2010 Missouri EOC Local Calibra MOELA10.OUT Sep 20 12:02 2010
INPUT: 62198 PERSONS  48 ITEMS  MEASURED: 62198 PERSONS  36 ITEMS  75 CATS  3.64.2
--------------------------------------------------------------------------------

     LARGEST STANDARDIZED RESIDUAL CORRELATIONS
     USED TO IDENTIFY DEPENDENT ITEMS
+------------------------------------+
|RESIDUL| ENTRY       | ENTRY       |
|CORRELN|NUMBER ITEM  |NUMBER ITEM  |
|-------+-------------+-------------|
|   .11 |    33 I0033 |    34 I0034 |
|   .10 |     2 I0002 |     6 I0006 |
|   .08 |    28 I0028 |    33 I0033 |
|-------+-------------+-------------|
|  -.10 |     3 I0003 |    33 I0033 |
|  -.09 |    28 I0028 |    45 I0045 |
|  -.09 |    33 I0033 |    45 I0045 |
|  -.09 |     3 I0003 |    28 I0028 |
|  -.09 |    12 I0012 |    28 I0028 |
|  -.08 |     3 I0003 |    34 I0034 |
|  -.08 |     3 I0003 |    25 I0025 |
+------------------------------------+
```

Figure 7.15: Standardized Residual Correlations from the Spring 2010 Administration for Algebra I

```
TABLE 23.99 Spring 2010 Missouri EOC Local Calibra MOMAT10.OUT Sep 20 12:03 2010
INPUT: 60862 PERSONS  48 ITEMS  MEASURED: 60862 PERSONS  36 ITEMS  75 CATS  3.64.2
--------------------------------------------------------------------------------

     LARGEST STANDARDIZED RESIDUAL CORRELATIONS
     USED TO IDENTIFY DEPENDENT ITEMS
+------------------------------------+
|RESIDUL| ENTRY       | ENTRY       |
|CORRELN|NUMBER ITEM  |NUMBER ITEM  |
|-------+-------------+-------------|
|   .35 |    28 I0028 |    34 I0034 |
|   .14 |    16 I0016 |    27 I0027 |
|   .11 |    17 I0017 |    19 I0019 |
|-------+-------------+-------------|
|  -.13 |    34 I0034 |    44 I0044 |
|  -.10 |    34 I0034 |    45 I0045 |
|  -.09 |    31 I0031 |    34 I0034 |
|  -.09 |    11 I0011 |    44 I0044 |
|  -.09 |    34 I0034 |    37 I0037 |
|  -.09 |    16 I0016 |    48 I0048 |
|  -.09 |    46 I0046 |    48 I0048 |
+------------------------------------+
```


Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 7.16: Standardized Residual Correlations from the Spring 2010 Administration for Biology

```
TABLE 23.99 Spring 2010 Missouri EOC Local Calibra MOSCI10.OUT Sep 20 12:04 2010
INPUT: 60134 PERSONS  59 ITEMS  MEASURED: 60134 PERSONS  47 ITEMS  102 CATS 3.64.2
--------------------------------------------------------------------------------

     LARGEST STANDARDIZED RESIDUAL CORRELATIONS
     USED TO IDENTIFY DEPENDENT ITEMS
+------------------------------------+
|RESIDUL| ENTRY        | ENTRY        |
|CORRELN|NUMBER ITEM   |NUMBER ITEM   |
|-------+-------------+-------------|
|   .45 |    55 I0055 |    56 I0056 |
|   .19 |    57 I0057 |    58 I0058 |
|   .11 |    28 I0028 |    30 I0030 |
|   .09 |    19 I0019 |    28 I0028 |
|-------+-------------+-------------|
|  -.13 |    11 I0011 |    35 I0035 |
|  -.11 |    28 I0028 |    35 I0035 |
|  -.11 |    18 I0018 |    35 I0035 |
|  -.10 |    35 I0035 |    49 I0049 |
|  -.10 |    35 I0035 |    57 I0057 |
|  -.10 |    30 I0030 |    35 I0035 |
+------------------------------------+
```

Tables 7.31 to 7.33 provide summary statistics, including summary fit statistics, for the Spring 2010 operational test calibrations. The evaluation of fit values, specifically MS infit, yielded these results: Infit values for English II ranged from 0.81 to 1.22, values for Algebra I ranged from 0.79 to 1.33, and values for Biology ranged from 0.84 to 1.29. The fit values and output files are based on the local runs using WINSTEPS version 3.64.2. Tables 7.34 to 7.36 provide Rasch difficulties and item fit statistics.

Table 7.31: Summary Statistics for the Spring 2010 Operational Test Calibrations for English II

Statistic	Rasch Difficulty Estimate	Item Mean	Infit		Outfit		Point-Biserial
			MS	Standardized	*MS*	Statistic	
# of Items	36	36	36	36	36	36	36
Mean	0.00	0.77	0.99	–0.92	0.96	–1.30	0.34
SD	0.98	0.41	0.10	8.80	0.20	8.70	0.10
Minimum	–2.18	0.36	0.81	–9.90	0.50	–9.90	0.13
Percentiles							
10	–1.24	0.51	0.87	–9.90	0.75	–9.90	0.22
25	–0.42	0.61	0.92	–9.90	0.83	–9.90	0.27
50	–0.03	0.74	0.96	–4.13	0.96	–5.17	0.35
75	0.68	0.80	1.07	9.90	1.11	9.90	0.41
90	1.17	0.91	1.13	9.90	1.22	9.90	0.46
Maximum	1.90	2.97	1.22	9.90	1.34	9.90	0.54

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.32: Summary Statistics for the Spring 2010 Operational Test Calibrations for Algebra I

Statistic	Rasch Difficulty Estimate	Item Mean	Infit *MS*	Infit Standardized	Outfit *MS*	Outfit Standardized	Point-Biserial
# of Items	36	36	36	36	36	36	36
Mean	0.00	0.65	1.00	–1.82	1.00	–1.45	0.35
SD	0.99	0.28	0.11	8.18	0.20	7.94	0.11
Minimum	–2.09	0.24	0.79	–9.90	0.62	–9.90	0.01
Percentiles							
10	–1.23	0.35	0.89	–9.90	0.79	–9.90	0.22
25	–0.72	0.51	0.94	–9.90	0.91	–9.90	0.30
50	–0.03	0.64	0.98	–3.71	0.98	–2.73	0.36
75	0.61	0.76	1.03	6.95	1.10	9.90	0.40
90	1.38	0.84	1.15	9.90	1.22	9.90	0.47
Maximum	1.96	1.95	1.33	9.90	1.74	9.90	0.61

Table 7.33: Summary Statistics for the Spring 2010 Operational Test Calibrations for Biology

Statistic	Rasch Difficulty Estimate	Item Mean	Infit *MS*	Infit Standardized	Outfit *MS*	Outfit Statistic	Point-Biserial
# of Items	47	47	47	47	47	47	47
Mean	0.00	0.67	1.00	–1.10	0.99	–0.86	0.36
SD	0.92	0.24	0.11	8.35	0.16	8.30	0.11
Minimum	–2.45	0.37	0.84	–9.90	0.64	–9.90	0.03
Percentiles							
10	–1.21	0.43	0.87	–9.90	0.78	–9.90	0.22
25	–0.53	0.52	0.91	–9.90	0.87	–9.90	0.30
50	0.13	0.62	0.99	–3.53	0.98	–2.86	0.35
75	0.60	0.80	1.05	9.90	1.07	9.90	0.44
90	1.05	0.93	1.14	9.90	1.20	9.90	0.50
Maximum	1.73	1.76	1.29	9.90	1.41	9.90	0.58

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.34: Item Statistics for the Spring 2010 Operational Test Calibrations for English II

Item Number	Item Type	Rasch Difficulty Estimate	*N*	*MS* Infit	Standardized Infit	*MS* Outfit	Standardized Outfit
1	SR	–1.8289	62102	0.96	–3.24	0.86	–6.42
2	SR	–1.2984	62095	0.92	–8.15	0.78	–9.90
3	SR	1.2701	62098	1.22	9.90	1.34	9.90
4	SR	0.0448	62097	1.16	9.90	1.27	9.90
5	SR	–0.1727	62102	1.01	2.42	1.01	0.73
6	SR	–1.9984	62101	0.86	–9.90	0.50	–9.90
7	SR	–2.1842	62100	0.94	–4.29	0.78	–8.73
8	SR	0.6197	62101	1.08	9.90	1.12	9.90
9	SR	1.0235	62099	1.07	9.90	1.10	9.90
10	SR	1.0662	62096	1.13	9.90	1.17	9.90
11	SR	0.9704	62100	1.10	9.90	1.15	9.90
12	SR	1.8502	62100	1.08	9.90	1.24	9.90
25	SR	–0.3741	62099	0.87	–9.90	0.78	–9.90
26	SR	0.2470	62100	1.12	9.90	1.18	9.90
27	SR	1.0759	62095	1.03	9.90	1.06	9.90
28	SR	–1.1799	62092	0.82	–9.90	0.56	–9.90
29	SR	–0.4422	62096	0.92	–9.90	0.87	–9.90
30	SR	–0.3183	62100	0.94	–9.90	0.94	–6.53
31	SR	0.2825	62096	1.01	1.82	0.98	–2.21
32	SR	–0.0192	62097	0.98	–3.96	0.95	–5.64
33	SR	–0.5343	62099	0.81	–9.90	0.61	–9.90
34	SR	–0.2822	62099	0.86	–9.90	0.73	–9.90
35	SR	0.1609	62096	1.02	4.83	1.06	8.18
36	SR	–0.1852	62091	0.90	–9.90	0.86	–9.90
37	SR	1.4180	62097	0.94	–9.90	0.97	–5.30
38	SR	–0.1500	62099	0.93	–9.90	0.84	–9.90
39	SR	–0.0425	62097	0.95	–9.22	0.96	–5.03
40	SR	0.0505	62091	0.96	–9.16	0.92	–9.65
41	SR	0.3177	62095	1.12	9.90	1.20	9.90
42	SR	–0.7518	62097	0.89	–9.90	0.77	–9.90
43	SR	–0.3727	62095	0.92	–9.90	0.80	–9.90
44	SR	0.1576	62097	1.04	8.85	1.02	3.27
45	SR	1.9035	62097	1.15	9.90	1.31	9.90
46	SR	–0.7719	62096	0.96	–5.21	0.88	–9.33
47	SR	0.8659	62099	1.02	6.75	1.02	4.63
48	WP	–0.4173	61834	0.96	–4.63	0.96	–4.71

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.35: Item Statistics for the Spring 2010 Operational Test Calibrations for Algebra I

Item Number	Item Type	Rasch Difficulty Estimate	*N*	*MS* Infit	Standardized Infit	*MS* Outfit	Standardized Outfit
1	SR	–0.6825	60816	0.99	–1.60	0.96	–4.64
2	SR	–1.3229	60814	0.90	–9.90	0.79	–9.90
3	SR	–0.1609	60815	1.14	9.90	1.21	9.90
4	SR	–1.7968	60811	0.96	–4.32	0.98	–1.29
5	SR	–0.9933	60811	0.98	–3.09	0.99	–1.34
10	SR	–2.0914	60810	0.88	–9.90	0.62	–9.90
11	SR	–1.1383	60815	0.87	–9.90	0.73	–9.90
12	SR	–1.3951	60811	0.91	–9.90	0.91	–6.98
13	SR	0.5565	60807	0.89	–9.90	0.86	–9.90
14	SR	–0.8763	60808	0.95	–9.46	0.92	–7.93
15	SR	–0.9224	60808	0.96	–7.83	0.97	–2.87
16	SR	0.6633	60812	1.09	9.90	1.12	9.90
17	SR	0.1169	60809	1.00	1.16	1.00	–0.14
18	SR	0.2689	60811	1.02	4.42	1.01	1.17
19	SR	–0.3259	60810	0.94	–9.90	0.88	–9.90
20	SR	0.1576	60811	0.97	–7.81	0.94	–9.90
21	SR	–0.0997	60809	0.99	-2.56	0.98	–3.32
26	SR	0.1556	60810	1.02	6.20	1.00	–0.81
27	SR	0.5962	60810	1.00	0.78	1.00	0.03
28	SR	–0.3500	60811	0.87	–9.90	0.79	–9.90
29	SR	–0.2478	60813	1.00	0.93	0.98	–2.59
30	SR	–0.0611	60810	0.93	–9.90	0.90	–9.90
31	SR	0.2366	60807	1.16	9.90	1.22	9.90
32	SR	–0.8354	60809	0.96	–7.81	0.94	–6.48
33	SR	0.4996	60811	1.10	9.90	1.12	9.90
34	SR	–0.4071	60811	0.79	–9.90	0.68	–9.90
35	SR	–0.0066	60812	0.90	–9.90	0.87	–9.90
36	SR	–0.2378	60805	0.95	–9.90	0.94	–9.55
37	SR	1.7176	60809	1.15	9.90	1.39	9.90
38	SR	1.7149	60809	1.04	9.19	1.18	9.90
43	SR	1.7313	60809	1.01	2.24	1.12	9.90
44	SR	1.9604	60806	1.33	9.90	1.74	9.90
45	SR	1.0376	60806	1.19	9.90	1.28	9.90
46	SR	0.9396	60807	1.05	9.90	1.09	9.90
47	SR	0.9432	60814	0.96	–9.90	0.97	–5.02
48	PE	0.6555	60208	0.97	–6.49	1.00	v0.42

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.36: Item Statistics for the Spring 2010 Operational Test Calibrations for Biology

Item Number	Item Type	Rasch Difficulty Estimate	*N*	*MS* Infit	Standardized Infit	*M*S Outfit	Standardized Outfit
1	SR	–2.3495	60126	0.90	–7.70	0.64	–9.90
2	PE	–2.4533	60124	0.97	–2.58	1.05	2.22
3	SR	–1.3539	60122	1.01	1.22	1.05	3.75
4	SR	–0.9724	60125	0.91	–9.90	0.78	–9.90
5	SR	–1.3517	60124	0.97	–4.43	0.92	–6.50
10	SR	–1.1718	60121	0.95	–8.24	0.93	–6.06
11	SR	–0.7744	60123	0.85	–9.90	0.72	–9.90
12	SR	–1.0174	60124	0.88	–9.90	0.78	–9.90
13	SR	–0.5347	60123	1.11	9.90	1.11	9.90
14	SR	0.1156	60122	0.98	–5.10	0.98	–4.66
15	SR	–0.0859	60120	1.09	9.90	1.12	9.90
16	SR	–1.2723	60123	0.92	–9.90	0.97	–2.86
17	SR	–0.0012	60125	1.02	6.23	1.01	2.01
18	SR	–0.0604	60120	0.87	–9.90	0.81	–9.90
19	SR	–0.1573	60122	0.97	–7.13	0.95	–8.31
20	SR	–1.1540	60125	0.92	–9.90	0.84	–9.90
21	SR	0.2872	60121	0.99	–4.53	0.97	–5.52
26	SR	–0.1781	60122	1.00	0.55	1.01	2.20
27	SR	–0.4216	60120	1.04	9.28	1.05	7.14
28	SR	–0.3572	60119	0.86	–9.90	0.78	–9.90
29	SR	–0.0004	60122	0.95	–9.90	0.93	–9.90
30	SR	–0.7834	60124	0.86	–9.90	0.73	–9.90
31	SR	0.1284	60123	0.98	–5.38	0.97	–5.86
32	SR	0.8679	60122	1.18	9.90	1.26	9.90
33	SR	0.5627	60122	1.05	9.90	1.06	9.90
34	SR	0.2776	60124	1.05	9.90	1.05	9.90
35	SR	0.9405	60122	1.29	9.90	1.41	9.90
36	SR	0.4323	60123	1.00	0.83	1.00	0.07
37	SR	0.4548	60122	1.01	3.75	1.01	2.36
38	SR	0.9361	60120	1.17	9.90	1.25	9.90
43	SR	0.4450	60122	0.99	–3.53	0.99	–3.24
44	SR	0.5263	60123	0.99	–2.04	1.00	–0.47
45	SR	1.2068	60122	1.02	5.27	1.07	9.90
46	SR	1.0893	60121	1.11	9.90	1.19	9.90
47	SR	1.1838	60121	1.10	9.90	1.19	9.90
48	PE	0.5549	59933	0.98	–7.04	0.98	–4.23
49	SR	0.8443	59933	0.84	–9.90	0.81	–9.90
50	PE	1.7295	59933	1.12	9.90	1.16	9.90
51	PE	0.3499	59933	1.24	9.90	1.28	9.90
52	PE	0.6313	59933	1.18	9.90	1.21	9.90
53	PE	0.7974	59933	0.99	–1.46	0.98	–2.67
54	PE	1.0238	59933	0.89	–9.90	0.86	–9.90
55	PE	0.2197	59933	0.88	–9.90	0.85	–9.90
56	PE	0.1381	59933	0.89	–9.90	0.87	–9.90
57	PE	–0.5302	59933	0.86	–9.90	0.80	–9.90
58	PE	0.0955	59933	0.91	–9.90	0.90	–9.90
59	PE	1.1424	59933	1.12	9.90	1.12	9.90

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Tables 7.37 to 7.45 provide the raw score to scale score conversions for Summer 2009, Fall 2009, and Spring 2010, respectively.

Table 7.37: Raw Score to Scale Score Conversions for Summer 2009, English II

Raw Score	Scale Score	*CSEM*
0	101	30
1	121	17
2	133	12
3	140	10
4	146	9
5	150	8
6	154	8
7	158	7
8	161	7
9	164	7
10	166	7
11	169	6
12	171	6
13	174	6
14	176	6
15	180	6
16	181	6
17	183	6
18	185	6
19	187	6
20	189	6
21	192	6
22	194	6
23	196	6
24	200	6
25	201	6
26	203	6
27	206	6
28	208	7
29	211	7
30	214	7
31	217	7
32	221	8
33	225	8
34	228	9
35	233	9
36	239	11
37	248	13
38	250	17
39	250	30

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.38: Raw Score to Scale Score Conversions for Summer 2009, Algebra I

Raw Score	Scale Score	*CSEM*
0	100	37
1	113	20
2	128	15
3	137	12
4	144	11
5	149	10
6	154	9
7	158	9
8	161	8
9	165	8
10	168	8
11	171	8
12	174	7
13	177	7
14	179	7
15	182	7
16	184	7
17	187	7
18	189	7
19	192	7
20	194	7
21	196	7
22	200	7
23	201	7
24	204	7
25	206	7
26	209	7
27	212	7
28	215	8
29	218	8
30	221	8
31	225	8
32	228	9
33	232	9
34	236	10
35	242	11
36	249	12
37	250	15
38	250	20
39	250	37

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.39: Raw Score to Scale Score Conversions for Summer 2009, Biology

Raw Score	Scale Score	*CSEM*
0	100	36
1	107	20
2	121	14
3	130	12
4	136	10
5	141	10
6	146	9
7	149	8
8	153	8
9	156	7
10	158	7
11	161	7
12	163	7
13	165	6
14	167	6
15	169	6
16	171	6
17	173	6
18	175	6
19	177	6
20	178	6
21	180	6
22	181	6
23	183	6
24	185	5
25	186	5
26	188	5
27	189	5
28	191	5
29	192	5
30	194	5
31	195	5
32	197	5
33	198	6
34	200	6
35	201	6
36	203	6
37	205	6
38	206	6
39	208	6
40	210	6
41	212	6
42	214	6
43	216	7
44	218	7
45	221	7
46	225	7
47	226	8

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.39: Raw Score to Scale Score Conversions for Summer 2009, Biology (continued)

Raw Score	Scale Score	*CSEM*
48	229	8
49	233	9
50	237	9
51	242	10
52	248	12
53	250	14
54	250	20
55	250	36

Table 7.40: Raw Score to Scale Score Conversions for Fall 2009, English II

Raw Score	Scale Score	*CSEM*
0	113	30
1	134	17
2	146	12
3	154	10
4	159	9
5	164	8
6	168	8
7	171	7
8	174	7
9	177	7
10	180	6
11	182	6
12	184	6
13	186	6
14	189	6
15	191	6
16	193	6
17	195	6
18	197	6
19	200	6
20	201	6
21	203	6
22	205	6
23	206	6
24	208	6
25	211	6
26	213	6
27	215	6
28	217	6
29	219	6
30	222	7
31	225	7
32	228	7

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.40: Raw Score to Scale Score Conversions for Fall 2009, English II (continued)

Raw Score	Scale Score	*CSEM*
33	231	8
34	235	8
35	239	9
36	244	10
37	250	12
38	250	17
39	250	30

Table 7.41: Raw Score to Scale Score Conversions for Fall 2009, Algebra I

Raw Score	Scale Score	*CSEM*
0	100	37
1	119	21
2	135	15
3	144	12
4	151	11
5	156	10
6	161	9
7	165	9
8	169	9
9	173	8
10	177	8
11	179	8
12	182	8
13	185	7
14	188	7
15	190	7
16	193	7
17	195	7
18	198	7
19	200	7
20	203	7
21	205	7
22	208	7
23	210	7
24	212	7
25	215	7
26	217	7
27	220	7
28	222	7
29	225	7
30	228	8
31	231	8
32	234	8

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.41: Raw Score to Scale Score Conversions for Fall 2009, Algebra I (continued)

Raw Score	Scale Score	*CSEM*
33	238	9
34	242	10
35	247	11
36	250	12
37	250	15
38	250	20
39	250	37

Table 7.42: Raw Score to Scale Score Conversions for Fall 2009, Biology

Raw Score	Scale Score	*CSEM*
0	100	36
1	115	20
2	129	14
3	137	12
4	144	10
5	149	9
6	153	9
7	156	8
8	160	8
9	162	7
10	165	7
11	167	7
12	170	6
13	172	6
14	174	6
15	176	6
16	177	6
17	179	6
18	181	6
19	182	6
20	184	5
21	185	5
22	187	5
23	188	5
24	190	5
25	191	5
26	193	5
27	194	5
28	195	5
29	197	5
30	198	5
31	200	5
32	201	5
33	202	5
34	204	5

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.42: Raw Score to Scale Score Conversions for Fall 2009, Biology (continued)

Raw Score	Scale Score	*CSEM*
35	205	5
36	207	6
37	209	6
38	210	6
39	212	6
40	214	6
41	215	6
42	217	6
43	219	6
44	222	7
45	225	7
46	227	7
47	229	8
48	232	8
49	236	9
50	240	9
51	245	10
52	250	12
53	250	14
54	250	20
55	250	36

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.43: Raw Score to Scale Score Conversions for Spring 2010, English II

Raw Score	Scale Score	*CSEM*
0	106	30
1	126	17
2	138	12
3	145	10
4	151	9
5	155	8
6	159	8
7	162	7
8	165	7
9	168	6
10	170	6
11	173	6
12	175	6
13	177	6
14	180	6
15	181	6
16	183	5
17	184	5
18	186	5
19	188	5
20	190	5
21	192	6
22	194	6
23	195	6
24	197	6
25	200	6
26	202	6
27	204	6
28	206	6
29	209	7
30	212	7
31	214	7
32	218	7
33	225	8
34	226	9
35	231	9
36	237	11
37	245	13
38	250	18
39	250	31

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.44: Raw Score to Scale Score Conversions for Spring 2010, Algebra I

Raw Score	Scale Score	*CSEM*
0	100	37
1	114	20
2	128	15
3	138	12
4	144	11
5	150	10
6	155	9
7	159	9
8	162	8
9	166	8
10	169	8
11	172	8
12	177	7
13	178	7
14	180	7
15	183	7
16	185	7
17	188	7
18	190	7
19	192	7
20	195	7
21	197	7
22	200	7
23	202	7
24	204	7
25	207	7
26	209	7
27	212	7
28	215	7
29	217	8
30	221	8
31	225	8
32	227	9
33	231	9
34	236	10
35	241	11
36	248	12
37	250	15
38	250	20
39	250	37

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.45: Raw Score to Scale Score Conversions for Spring 2010, Biology

Raw Score	Scale Score	*CSEM*
0	100	36
1	111	20
2	125	14
3	134	12
4	140	10
5	145	10
6	149	9
7	153	8
8	156	8
9	159	8
10	162	7
11	165	7
12	167	7
13	170	7
14	172	6
15	174	6
16	177	6
17	178	6
18	180	6
19	182	6
20	183	6
21	185	6
22	187	6
23	189	6
24	190	6
25	192	6
26	193	6
27	195	6
28	197	6
29	198	6
30	200	6
31	201	6
32	203	6
33	205	6
34	206	6

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 7.45: Raw Score to Scale Score Conversions for Spring 2010, Biology (continued)

Raw Score	Scale Score	*CSEM*
35	208	6
36	209	6
37	211	6
38	212	6
39	214	6
40	216	6
41	218	6
42	219	6
43	221	6
44	225	6
45	226	7
46	228	7
47	231	7
48	233	8
49	237	8
50	241	9
51	245	10
52	250	12
53	250	14
54	250	20
55	250	36

7.3.7 Step 7: Place the 2010 Embedded Field Test Items onto the 2009 Scale

The 2010 embedded field test items were treated separately in this process to avoid having them influence calibration of the operational items. To bring the field test items onto the base scale, a second calibration of the Spring data, fixing the 2009 operational parameter estimates, was conducted.

The above embedded field test items will be used to create the new 2011 Spring form. This process will allow all new items for the Spring form administration to be placed on the same EOC Assessment scale as the other items in the pool.

As outlined above, not only can the pre-equating model be used to annually build alternate test forms, but by using the embedded field testing approach, DESE will also be able to maintain its item pools.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 8: REPORTING

8.1 Introduction

The purpose of reporting assessment data is to communicate test results to students, their parents, and their teachers. The Missouri End-of-Course (MO EOC) Assessment reports provide useful information for determining the performance of students in a particular school and classroom. These reports help describe students' knowledge of a given set of expectations, allowing educators to determine specific instructional needs, measure student mastery toward post-secondary readiness, provide evidence of accountability for Missouri and national programs, and evaluate educational programs. Additionally, districts may use locally designed assessments aligned to the Show-Me Content Standards and Course-Level Expectations (CLEs) to provide more detailed information for each student in specific test areas.

Paper reports are generated for all assessment windows following the Spring administration; therefore, for the Summer 2009, Fall 2009, and Spring 2010 assessments, the paper reports were generated and distributed following the Spring 2010 operational administration. However, teachers may access their students' raw scores for the selected response items and score their students' performance events through an online interface shortly after the district's testing materials have been received for processing in each assessment window.

For each testing event, Riverside Publishing converts each student's raw score points earned into an EOC scale score, as described in Chapter 7: Scaling and Equating. A student receives an EOC scale score when he or she has made a valid attempt in any session. EOC scale scores range in value from 100 to 250. The EOC scale score determines the student's achievement level. For all content areas, a scale score of 200 to 224 is considered Proficient, and a scale score of 225 and above is considered Advanced. The cut score for Basic varies by content area. Each achievement level represents standards of performance for each assessed content area (English II, Algebra I, and Biology). Achievement-level scores describe what students can do in terms of the content and skills assessed. These scores provide a way to compare test results with standards of academic performance. Panels drawn from Missouri's educational, business, and professional communities recommended the raw score cuts (based on the Spring 2009 test forms) to be used for each achievement level. These cuts were then reviewed and adopted by the Missouri State Board of Education. For more information on how the achievement levels were set, refer to Chapter 3: Achievement-Level Setting.

No test provides a perfect measure of a student's ability. This situation is expected because all tests have a known standard error of measurement (*SEM*). The *SEM* represents the amount of variability that can be expected in a student's test score due to the inherent imprecision of the test. For example, if the student were tested again, he or she would likely obtain a slightly different score. The range for this new score is provided as a standard error (*SE*) and gives an indication of the margin of error for the reported scale score.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

8.2 Individual Student Report

The Individual Student Report provides information about performance on the EOC Assessment, describing the results in terms of four levels of achievement in a content area. It is used for measuring and reflecting an individual student's mastery toward post-secondary readiness for the content area. It is used in instructional planning as a point of reference during a parent-teacher conference and for permanent record keeping. Teachers are informed that other sources of information should be used along with this report when determining the student's areas of strength or need.

On the report, achievement-level scores describe what students can do in terms of the CLEs for the content and skills assessed by the EOC Assessment. A student at the Proficient or Advanced level has met the standard.

A sample of the Individual Student Report appears in Figure 8.1. A brief description of selected parts of the report follows:

A. The heading of the Individual Student Report includes the content area for the results being presented. A separate report is produced for each content area tested.

B. The Student Information section contains the biographic data for the individual student taking the assessment. Identifying information, including the MOSIS ID, gender, building, and district, is listed, followed by the test period.

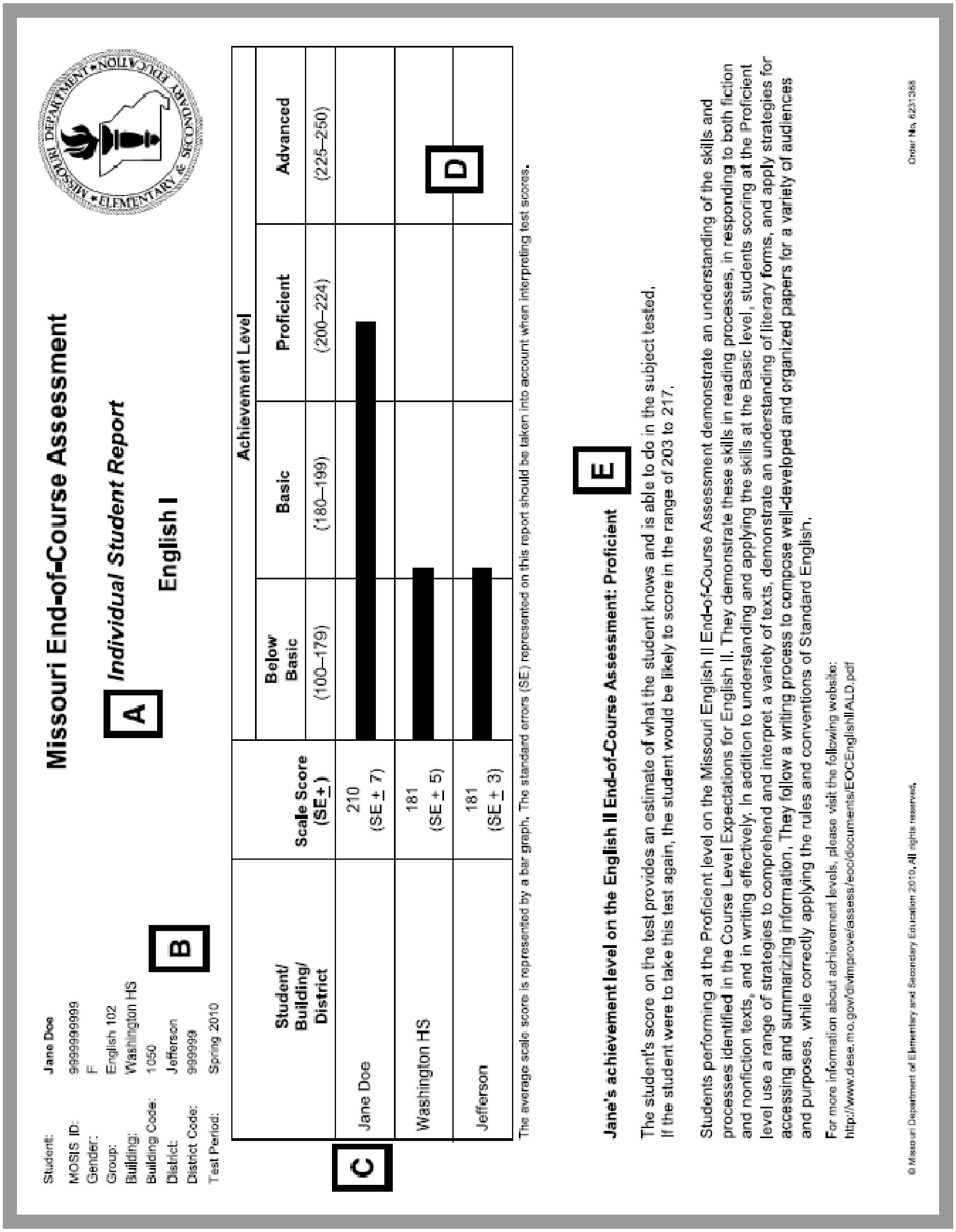
C. The individual student's results are presented numerically as a three-digit scale score with the *SE*. An accompanying bar graph to the right of the scale score illustrates the achievement level obtained by the student. Achievement levels (whether Advanced, Proficient, Basic, or Below Basic) are based on the scale score ranges listed beneath the Achievement Level heading in the table.

D. The mean scale scores for the student's building and district are displayed in the two rows below the student's individual results. The mean scale score, with an associated *SE*, and the bar graph provide a way to view the individual's results in contrast to the group's results for the content area during the same test period.

E. The narrative describes the student performance characteristics corresponding to the obtained level of achievement. The text is specific to the content area tested. At the bottom of the narrative is a URL for a website that provides additional information for all achievement levels for the content area.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Figure 8.1: Individual Student Report



Student: Jane Doe
MOSIS ID: 9999999999
Gender: F
Group: English 102
Building: Washington HS
Building Code: 1050
District: Jefferson
District Code: 999999
Test Period: Spring 2010

B

A

Missouri End-of-Course Assessment

Individual Student Report

English I

Student/ Building/ District	Scale Score (SE±)	Achievement Level: Below Basic (100–179)	Basic (180–199)	Proficient (200–224)	Advanced (225–250)
Jane Doe	210 (SE ± 7)				
Washington HS	181 (SE ± 5)				
Jefferson	181 (SE ± 3)				

C

D

The average scale score is represented by a bar graph. The standard errors (SE) represented on this report should be taken into account when interpreting test scores.

Jane's achievement level on the English II End-of-Course Assessment: Proficient E

The student's score on the test provides an estimate of what the student knows and is able to do in the subject tested. If the student were to take this test again, the student would be likely to score in the range of 203 to 217.

Students performing at the Proficient level on the Missouri English II End-of-Course Assessment demonstrate an understanding of the skills and processes identified in the Course Level Expectations for English II. They demonstrate these skills in reading processes, in responding to both fiction and nonfiction texts, and in writing effectively. In addition to understanding and applying the skills at the Basic level, students scoring at the Proficient level use a range of strategies to comprehend and interpret a variety of texts, demonstrate an understanding of literary forms, and apply strategies for accessing and summarizing information. They follow a writing process to compose well-developed and organized papers for a variety of audiences and purposes, while correctly applying the rules and conventions of Standard English.

For more information about achievement levels, please visit the following website:
http://www.dese.mo.gov/divimprove/assess/eoc/documents/EOCEnglishIIALD.pdf

© Missouri Department of Elementary and Secondary Education 2010. All rights reserved.

Order No. 6231368

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

8.3 Student Score Label

The Student Score Label provides a summary of a student's results on the EOC Assessment. A separate label is produced for each content area tested. The individual label provides the student's biographic data, raw score, scale score, and achievement level. The labels have adhesive backing so they can be easily transferred onto the student record folders. A sample label is shown below in Figure 8.2. A brief description of selected parts of the label follows:

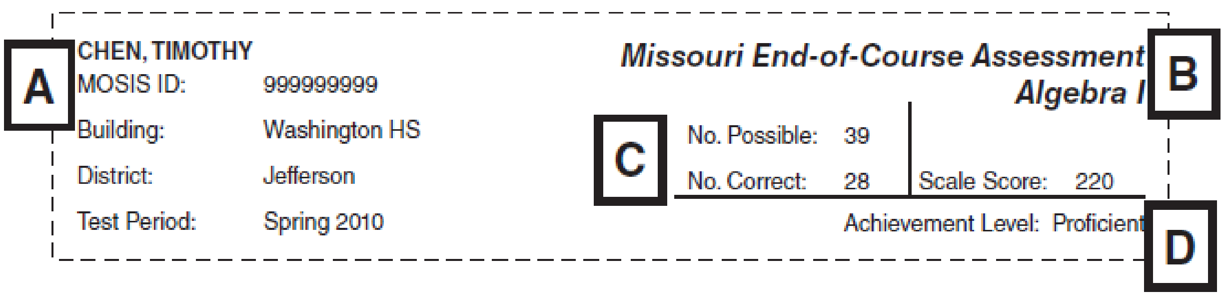
A. The student's name and identifying information are provided on the left side of the label.

B. The upper right side of the label shows the content area tested. If a student has results for more than one content area, the next label is printed below the first one.

C. The middle of the label has the Number Possible and the student's raw score (Number Correct). A corresponding column to the right of these data contains the raw score's associated Scale Score.

D. The student's achievement level is displayed in the lower right corner below the scores.

Figure 8.2: Student Score Label



8.4 Online Crystal Reports

Schools and districts are able to access summary level reports through the online Crystal Reports tool. This tool allows district and school administrators to create on-the-fly reports containing information relevant to their data needs. There are several reporting options available through the Crystal Reports tool, including administrative reports, adequate yearly progress (AYP) reports, achievement level reports, content standard reports, and item analysis reports.

For each subreport, a user selects various filters such as year, grade/content area, and level of reporting (state, district, or school) to create the desired report. For the Content Standard Reports, the user may also disaggregate results by various subgroups (e.g., race, disability).

A detailed discussion of all available reports is beyond the scope of this document. Only those reports that are first-level analyses of MO EOC data will be discussed. The Achievement Level-5 reports will not be discussed, as these are summaries of the pre-No Child Left Behind (NCLB) testing program. In addition, the AYP reports and some of the

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Administrative Reports, including the High School Career Education Student Summary and the Level Not Determined, will not be discussed.

The Crystal Reports tool is accessed through the Missouri Department of Elementary and Secondary Education (DESE) website. Each school and/or district is assigned a user name and password to access the site.

8.4.1 Administrative Reports

These reports provide student-level test data. Based on only the MO EOC Assessment results, four reports are generated: MO EOC Scale Score Summary, MO EOC Student Demographic, Student Achievement Level, and Student Report.

MO EOC Scale Score Summary: This report lists each student in the school or district along with his or her MOSIS ID, testing year, content area, grade level, MO EOC scale score, and achievement level.

MO EOC Student Demographic: This report lists all students in the school or district along with their date of birth (DOB), content area, MOSIS ID, district ID, and relevant demographic information, including if the student has been in the district for less than a year, if the student has been in the building for less than a year, if the student is limited English proficiency (LEP), the student's race, if the student qualifies for free and reduced lunch (SES), if the student has an individualized education program (IEP), if the student is an English-language learner (ELL)/LEP who has been in the school for less than one year and in the country for less than three years, if the student is an LEP/ELL Title 3, the number of months the LEP/ELL student has been in the U.S., the student's disability diagnosis, and if the student is Title 1.

Student Achievement Level: This report lists all students in a school or district along with the year of testing, content area, grade-level, achievement level, and MOSIS ID.

Student Report: For each school or district, this report contains the following information: student name, DOB, district student number, MOSIS ID, content area tested, grade level, achievement level, and scale score for each content area tested.

8.4.2 Achievement Level 4 Levels

These reports contain summary information on school or district performance in terms of the four MO EOC achievement levels. There are two types of achievement level reports: Achievement Level 4 Chart and Achievement Level 4 Report.

Achievement Level 4 Chart: This report charts the percentage of students classified as Proficient or Advanced on each MO EOC Assessment. State-level, district-level, and/or school-level performance may be displayed on the chart.

Achievement Level 4 Report: This report summarizes the number and percentage of students in each achievement level. This report is comprised of 10 columns: Total, content area, grade, year, number of accountable (ACC) students, number of reportable (REP) students, number and percentage of students classified in the Basic (B) achievement level, number and percentage of students classified in the Proficient (P) achievement level, number and percentage of students classified in the Advanced (A) achievement level, and mean MO EOC scale score. The first column, Total, shows if

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

aggregate or disaggregated information is being displayed. A key to the abbreviations is found in the bottom left corner.

8.4.3 Content Standard

The content standard reports summarize information about the content standards.

Content Standards Report: This report has 14 columns: content area, grade level, category/type, year, percentage of points earned on content standard 1 (CS-1), points possible (PP) on CS-1, percentage of points earned on CS-2, PP on CS-2, percentage of points earned on CS-3, PP on CS-3, percentage of points earned on CS-4, PP on CS-4, percentage of points earned on CS-5, and PP on CS-5. The category/type column indicates if the data are aggregated or disaggregated.

Content Standards Detail: This report shows the percentage of points each student achieved on each content standard within a particular content area.

8.4.4 Item Analysis Expanded

This set of reports provides detailed item-level results for the school or district, aggregated either by the content or the process standard.

Content Standard IBD EX: The Content Standard Item Benchmark Descriptor (IBD) Extended (EX) report contains item-level detail aggregated by content standard. The report is comprised of 11 columns: school code (SC), grade level (GR), standard number and description (desc.), code for the course-level expectation (CLE), description of the CLE, depth of knowledge (DOK) of the item, session/item number where the item was in the operational test, question type (QT), points possible for the item, the average points (avg pts) earned by students in the district on that item, and percentage of points earned by students in the district on that item.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 9: SUMMARY STATISTICS

9.1 Introduction

This chapter provides descriptive statistics for the number correct raw score and for scale scores for each of the three Missouri End-of-Course (EOC) Assessments from the Summer 2009, Fall 2009, and Spring 2010 administrations. Statistics include *N* counts, means, standard deviations (*SD*), minimum and maximum values, and a variety of data disaggregations.

9.2 Descriptive Statistics for Total Raw Score

Descriptive statistics for total raw score are summarized in Table 9.1 by test administration and content area.

Table 9.1: Descriptive Statistics for Total Raw Score

Test Period	Subject	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	757	1	38	22.21	6.180
	Algebra I	1,394	1	38	18.08	6.641
	Biology	515	5	49	26.78	9.626
Fall 2009	English II	1,465	2	38	23.02	6.404
	Algebra I	2,571	1	38	19.50	7.634
	Biology	2,159	3	55	32.33	11.461
Spring 2010	English II	62,198	1	39	27.65	6.175
	Algebra I	60,862	1	39	23.22	7.262
	Biology	60,134	1	55	31.55	9.874

9.3 Descriptive Statistics for Total Raw Score by Cluster

Table 9.2 summarizes the number correct raw score by test administration, content area, and cluster.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.2: Descriptive Statistics for Total Raw Score by Test Administration, Content Area, and Cluster

Test Period	Subject		*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	Reading	757	0	30	17.62	5.327
		Writing	757	1	9	4.59	1.507
	Algebra I	Number and Operations	1,394	0	8	4.07	1.870
		Algebraic Relationships	1,394	0	24	10.80	4.368
		Data and Probability	1,394	0	8	3.74	1.636
	Biology	Characteristics and Interactions of Living Organisms	515	2	21	10.36	3.681
		Changes in Ecosystems and Interactions of Organisms with their Environments	515	1	13	7.30	2.849
		Inquiry	515	0	18	9.12	4.564
Fall 2009	English II	Reading	1,465	2	29	17.40	5.195
		Writing	1,465	0	9	5.62	1.755
	Algebra I	Number and Operations	2,571	0	9	5.14	2.113
		Algebraic Relationships	2,571	0	21	8.95	4.509
		Data and Probability	2,571	0	9	5.42	2.052
	Biology	Characteristics and Interactions of Living Organisms	2,159	0	22	12.19	4.861
		Changes in Ecosystems and Interactions of Organisms with their Environments	2,159	0	13	9.06	2.989
		Inquiry	2,159	0	20	11.07	4.911

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.2: Descriptive Statistics for Total Raw Score by Test Administration, Content Area, and Cluster (continued)

Test Period	Subject		*N*	Minimum	Maximum	Mean	*SD*
Spring 2010	English II	Reading	62,198	0	30	21.42	5.133
		Writing	62,198	0	9	6.23	1.531
	Algebra I	Number and Operations	60,862	0	8	5.20	1.815
		Algebraic Relationships	60,862	0	23	12.45	4.808
		Data and Probability	60,862	0	8	5.56	1.614
	Biology	Characteristics and Interactions of Living Organisms	60,134	0	22	12.60	4.041
		Changes in Ecosystems and Interactions of Organisms with their Environments	60,134	0	13	9.78	2.648
		Inquiry	60,134	0	20	9.17	4.550

9.4 Descriptive Statistics for Scale Scores by Test Period and Subject

Descriptive statistics of scale scores and percentage distributions of students' achievement levels are summarized in Tables 9.3 and 9.4. Table 9.3 summarizes student scale scores by each End-of-Course Assessment for the Summer 2009, Fall 2009, and Spring 2010 administrations. Table 9.4 lists the percentage and frequency of students in each achievement level.

Table 9.3: Scale Score Distributions for Each End-of-Course Assessment

Descriptive Statistics						
Test Period	**Subject**	***N***	**Minimum**	**Maximum**	**Mean**	***SD***
Summer 2009	English II	706	133	250	195.47	15.391
	Algebra I	1,284	113	250	189.68	18.339
	Biology	491	141	233	188.92	16.504
Fall 2009	English II	1,422	146	250	207.70	14.256
	Algebra I	2,488	119	250	201.52	20.465
	Biology	2,122	144	250	203.58	19.912
Spring 2010	English II	61,911	126	250	208.60	16.215
	Algebra I	60,544	114	250	203.81	19.947
	Biology	59,904	111	250	202.58	17.450

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Scale scores range from a minimum of 100 to a maximum of 250 for the three content areas administered in Summer 2009, Fall 2009, and Spring 2010. For English II, a minimum scale score of 180 is required to earn an achievement level of Basic. For Algebra I and Biology, a minimum scale score of 177 is required to earn an achievement level of Basic. For all content areas, a scale score of 200 represents the minimum score to earn an achievement level of Proficient, and a scale score of 225 represents the minimum score to earn an achievement level of Advanced.

Table 9.4: Achievement-Level Distributions for Each End-of-Course Assessment

Test Period	Subject	Achievement	Frequency	Percentage
Summer 2009	English II	Below Basic	74	10.5
		Basic	318	45.0
		Proficient	286	40.5
		Advanced	28	4.0
		Total	706	100.0
	Algebra I	Below Basic	271	21.1
		Basic	629	49.0
		Proficient	320	24.9
		Advanced	64	5.0
		Total	1,284	100.0
	Biology	Below Basic	99	20.2
		Basic	270	55.0
		Proficient	104	21.2
		Advanced	18	3.7
		Total	491	100.0
Fall 2009	English II	Below Basic	23	1.6
		Basic	325	22.9
		Proficient	884	62.2
		Advanced	190	13.4
		Total	1,422	100.0
	Algebra I	Below Basic	208	8.4
		Basic	963	38.7
		Proficient	943	37.9
		Advanced	374	15.0
		Total	2,488	100.0
	Biology	Below Basic	187	8.8
		Basic	706	33.3
		Proficient	867	40.9
		Advanced	362	17.1
		Total	2,122	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.4: Achievement-Level Distributions for Each End-of-Course Assessment (continued)

Test Period	Subject	Achievement	Frequency	Percentage
Spring 2010	English II	Below Basic	1,830	3.0
		Basic	14,260	23.0
		Proficient	31,658	51.1
		Advanced	14,163	22.9
		Total	61,911	100.0
	Algebra I	Below Basic	3,733	6.2
		Basic	20,593	34.0
		Proficient	25,381	41.9
		Advanced	10,837	17.9
		Total	60,544	100.0
	Biology	Below Basic	3,703	6.2
		Basic	20,890	34.9
		Proficient	27,984	46.7
		Advanced	7,327	12.2
		Total	59,904	100.0

9.5 Descriptive Statistics by Demographic Group

Descriptive statistics of scale scores and percentage distributions of students' achievement levels by demographic groups are summarized in Tables 9.5 through 9.20.

The demographic variables included are gender (Tables 9.5 and 9.13), ethnicity (Tables 9.6 and 9.14), migrant status (Tables 9.7 and 9.15), free and reduced lunch (FRL) (Tables 9.8 and 9.16), limited English proficient (LEP) (Tables 9.9 and 9.17), Title I (Tables 9.10 and 9.18), individualized education program (IEP) (Tables 9.11 and 9.19), and accommodations (Tables 9.12 and 9.20). Note that for certain cells in these tables (particularly for Fall and Summer where *n*-counts are low), when the total *n*-count is less than 30, the descriptive statistics are not reported.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.5: Scale Score Distributions by Demographic Group—Gender

Test Period	Subject	Gender	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	Female	250	133	250	195.97	15.716
		Male	456	146	250	195.20	15.220
	Algebra I	Female	558	113	250	188.79	18.961
		Male	722	144	250	190.32	17.851
	Biology	Female	237	141	233	189.45	16.781
		Male	252	146	229	188.59	16.215
Fall 2009	English II	Female	671	146	250	209.71	14.572
		Male	751	164	250	205.90	13.729
	Algebra I	Female	1,253	135	250	201.09	19.983
		Male	1,233	119	250	202.00	20.929
	Biology	Female	1,045	144	250	202.11	19.712
		Male	1,075	153	250	205.03	20.017
Spring 2010	English II	Female	30,983	126	250	210.66	15.665
		Male	30,819	126	250	206.58	16.489
	Algebra I	Female	30,460	114	250	203.33	19.526
		Male	29,994	114	250	204.35	20.331
	Biology	Female	30,272	111	250	201.67	17.048
		Male	29,537	111	250	203.57	17.773

Table 9.6: Scale Score Distributions by Demographic Group—Ethnicity

Test Period	Subject	Ethnicity	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	American Indian					
		Asian/Pacific Islander					
		Black (not Hispanic)	249	146	228	191.86	13.765
		Hispanic	33	166	228	194.36	13.824
		White (not Hispanic)	404	133	250	197.97	16.194
	Algebra I	American Indian					
		Asian/Pacific Islander					
		Black (not Hispanic)	449	137	232	183.08	15.606
		Hispanic	86	137	242	183.20	17.133
		White (not Hispanic)	696	113	250	195.33	18.319
	Biology	American Indian					
		Asian/Pacific Islander					
		Black (not Hispanic)	232	141	225	182.18	13.832
		Hispanic					
		White (not Hispanic)	217	158	233	197.11	15.821

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.6: Scale Score Distributions by Demographic Group—Ethnicity (continued)

Test Period	Subject	Ethnicity	*N*	Minimum	Maximum	Mean	*SD*
Fall 2009	English II	American Indian					
		Asian/Pacific Islander					
		Black (not Hispanic)	304	146	239	199.42	12.260
		Hispanic	37	177	228	204.54	12.525
		White (not Hispanic)	1,050	171	250	210.15	13.958
	Algebra I	American Indian					
		Asian/Pacific Islander	50	156	231	202.04	17.957
		Black (not Hispanic)	453	135	250	186.21	15.789
		Hispanic	89	151	250	195.30	20.150
		White (not Hispanic)	1,887	119	250	205.52	19.760
	Biology	American Indian					
		Asian/Pacific Islander	69	172	250	215.59	20.082
		Black (not Hispanic)	374	144	245	186.11	16.053
		Hispanic	60	153	236	198.43	19.818
		White (not Hispanic)	1,605	149	250	207.40	18.343
Spring 2010	English II	American Indian	330	165	250	205.88	14.862
		Asian/Pacific Islander	1,113	145	250	213.40	17.485
		Black (not Hispanic)	9,833	126	250	199.54	14.482
		Hispanic	1,865	145	250	204.89	15.303
		White (not Hispanic)	48,602	126	250	210.53	15.888
	Algebra I	American Indian	314	144	250	199.14	18.581
		Asian/Pacific Islander	1,293	138	250	213.52	22.018
		Black (not Hispanic)	9,636	114	250	191.67	18.385
		Hispanic	2,069	144	250	198.62	18.888
		White (not Hispanic)	47,090	114	250	206.34	19.200
	Biology	American Indian	340	156	245	200.10	15.665
		Asian/Pacific Islander	1,164	140	250	209.24	19.937
		Black (not Hispanic)	9,276	111	245	189.94	16.181
		Hispanic	1,927	125	250	196.95	16.986
		White (not Hispanic)	47,069	111	250	205.20	16.433

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.7: Scale Score Distributions by Demographic Group—Migrant Status

Test Period	Subject	Migrant	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	No	705	133	250	195.48	15.400
		Yes					
	Algebra I	No	1,284	113	250	189.68	18.339
		Yes					
	Biology	No	491	141	233	188.92	16.504
		Yes					
Fall 2009	English II	No	1,420	146	250	207.70	14.265
		Yes					
	Algebra I	No	2,487	119	250	201.54	20.458
		Yes					
	Biology	No	2,122	144	250	203.58	19.912
		Yes					
Spring 2010	English II	No	61,880	126	250	208.61	16.212
		Yes	31	155	225	194.06	16.904
	Algebra I	No	60,493	114	250	203.82	19.946
		Yes	51	155	231	193.14	18.043
	Biology	No	59,864	111	250	202.59	17.447
		Yes	40	159	218	189.30	15.817

Table 9.8: Scale Score Distributions by Demographic Group—Free and Reduced Lunch

Test Period	Subject	Migrant	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	No	362	133	250	196.82	16.707
		Yes	344	146	248	194.05	13.753
	Algebra I	No	681	113	250	193.32	18.954
		Yes	603	137	242	185.58	16.706
	Biology	No	232	141	233	192.89	17.732
		Yes	259	146	221	185.37	14.456
Fall 2009	English II	No	825	168	250	211.29	14.158
		Yes	597	146	250	202.74	12.844
	Algebra I	No	1,599	119	250	206.71	20.208
		Yes	889	135	250	192.19	17.390
	Biology	No	1,408	149	250	209.38	18.467
		Yes	714	144	250	192.14	17.579
Spring 2010	English II	No	38,430	126	250	212.29	15.687
		Yes	23,481	126	250	202.57	15.228
	Algebra I	No	37,239	114	250	208.07	19.389
		Yes	23,305	114	250	197.00	18.914
	Biology	No	37,817	111	250	206.64	16.554
		Yes	22,087	111	250	195.64	16.739

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.9: Scale Score Distributions by Demographic Group—Limited English Proficient

Test Period	Subject	Migrant	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	No	696	133	250	195.53	15.330
		Yes					
	Algebra I	No	1,240	113	250	190.03	18.384
		Yes	44	149	206	180.07	14.103
	Biology	No	473	141	233	189.45	16.437
		Yes					
Fall 2009	English II	No	1,408	146	250	207.74	14.290
		Yes					
	Algebra I	No	2,460	119	250	201.64	20.455
		Yes					
	Biology	No	2,109	144	250	203.65	19.903
		Yes					
Spring 2010	English II	No	61,089	126	250	208.74	16.196
		Yes	822	138	250	198.66	14.446
	Algebra I	No	59,452	114	250	204.01	19.892
		Yes	1,092	144	250	193.13	20.061
	Biology	No	59,002	111	250	202.78	17.378
		Yes	902	134	250	189.87	17.413

Table 9.10: Scale Score Distributions by Demographic Group—Title I

Test Period	Subject	Migrant	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	No	652	133	250	195.46	15.512
		Yes	54	164	228	195.57	13.979
	Algebra I	No	1,183	113	250	190.59	18.392
		Yes	101	144	218	179.13	13.942
	Biology	No	439	141	233	190.24	16.382
		Yes	52	149	203	177.81	13.089
Fall 2009	English II	No	1,338	164	250	208.19	14.176
		Yes	84	146	228	199.85	13.269
	Algebra I	No	2,368	119	250	202.49	20.317
		Yes	120	135	212	182.48	12.566
	Biology	No	2,032	144	250	204.70	19.414
		Yes	90	144	214	178.21	13.001
Spring 2010	English II	No	59,302	126	250	209.05	16.110
		Yes	2,609	126	250	198.47	15.263
	Algebra I	No	57,548	114	250	204.37	19.785
		Yes	2,996	128	250	193.09	20.016
	Biology	No	57,584	111	250	203.22	17.183
		Yes	2,320	134	250	186.77	16.555

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.11: Scale Score Distributions by Demographic Group—Students with IEPs

Test Period	Subject	Migrant	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	No	615	133	250	196.46	15.182
		Yes	91	146	248	188.76	15.194
	Algebra I	No	1,130	113	250	190.62	18.382
		Yes	154	137	232	182.79	16.508
	Biology	No	431	146	233	189.65	16.696
		Yes	60	141	214	183.72	14.106
Fall 2009	English II	No	1,277	146	250	208.87	13.966
		Yes	145	171	228	197.43	12.598
	Algebra I	No	2,313	119	250	202.73	20.172
		Yes	175	144	250	185.62	17.483
	Biology	No	1,927	144	250	205.24	19.308
		Yes	195	144	250	187.17	18.326
Spring 2010	English II	No	55,762	126	250	210.42	15.361
		Yes	6,149	126	250	192.18	14.420
	Algebra I	No	56,307	114	250	205.06	19.442
		Yes	4,237	114	250	187.22	19.142
	Biology	No	55,392	111	250	203.85	16.862
		Yes	4,512	111	250	187.00	17.009

Table 9.12: Scale Score Distributions by Demographic Group—Students with Accommodations

Test Period	Subject	Migrant	*N*	Minimum	Maximum	Mean	*SD*
Summer 2009	English II	No	688	133	250	195.51	15.316
		Yes					
	Algebra I	No	1,260	113	250	189.83	18.374
		Yes					
	Biology	No	474	141	233	189.17	16.581
		Yes					
Fall 2009	English II	No	1,376	146	250	208.06	14.156
		Yes	46	174	228	196.91	13.099
	Algebra I	No	2,429	119	250	201.91	20.390
		Yes	59	151	250	185.76	17.165
	Biology	No	2,060	144	250	204.04	19.806
		Yes	62	156	236	188.21	17.278
Spring 2010	English II	No	58,754	126	250	209.61	15.739
		Yes	3,157	126	250	189.96	13.398
	Algebra I	No	58,996	114	250	204.35	19.732
		Yes	1,548	128	250	183.37	17.151
	Biology	No	58,333	111	250	203.04	17.249
		Yes	1,571	111	250	185.68	16.423

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.13: Achievement-Level Distributions by Gender

Test Period	Subject	Gender	Achievement Level	Frequency	Percentage
Summer 2009	English II	Female	Below Basic	27	3.8
			Basic	104	14.7
			Proficient	111	15.7
			Advanced	8	1.1
			Total	250	100.0
		Male	Below Basic	47	6.7
			Basic	214	30.3
			Proficient	175	24.8
			Advanced	20	2.8
			Total	456	100.0
	Algebra I	Female	Below Basic	126	9.8
			Basic	279	21.8
			Proficient	121	9.5
			Advanced	32	2.5
			Total	558	100.0
		Male	Below Basic	145	11.3
			Basic	349	27.3
			Proficient	196	15.3
			Advanced	32	2.5
			Total	722	100.0
	Biology	Female	Below Basic	46	9.4
			Basic	133	27.2
			Proficient	50	10.2
			Advanced	8	1.6
			Total	237	100.0
		Male	Below Basic	52	10.6
			Basic	136	27.8
			Proficient	54	11.0
			Advanced	10	2.0
			Total	252	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.13: Achievement-Level Distributions by Gender (continued)

Test Period	Subject	Gender	Achievement Level	Frequency	Percentage
Fall 2009	English II	Female	Below Basic	8	0.6
			Basic	126	8.9
			Proficient	423	29.7
			Advanced	114	8.0
			Total	671	100.0
		Male	Below Basic	15	1.1
			Basic	199	14.0
			Proficient	461	32.4
			Advanced	76	5.3
			Total	751	100.0
	Algebra I	Female	Below Basic	110	4.4
			Basic	471	18.9
			Proficient	508	20.4
			Advanced	164	6.6
			Total	1,253	100.0
		Male	Below Basic	97	3.9
			Basic	491	19.8
			Proficient	435	17.5
			Advanced	210	8.4
			Total	1,233	100.0
	Biology	Female	Below Basic	100	4.7
			Basic	376	17.7
			Proficient	405	19.1
			Advanced	164	7.7
			Total	1,045	100.0
		Male	Below Basic	87	4.1
			Basic	329	15.5
			Proficient	461	21.7
			Advanced	198	9.3
			Total	1,075	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.13: Achievement-Level Distributions by Gender (continued)

Test Period	Subject	Gender	Achievement Level	Frequency	Percentage
Spring 2010	English II	Female	Below Basic	533	0.9
			Basic	6,063	9.8
			Proficient	16,320	26.4
			Advanced	8,067	13.1
			Total	30,983	100.0
		Male	Below Basic	1,287	2.1
			Basic	8,144	13.2
			Proficient	15,296	24.8
			Advanced	6,092	9.9
			Total	30,819	100.0
	Algebra I	Female	Below Basic	1,813	3.0
			Basic	10,663	17.6
			Proficient	12,899	21.3
			Advanced	5,085	8.4
			Total	30,460	100.0
		Male	Below Basic	1,894	3.1
			Basic	9,886	16.4
			Proficient	12,468	20.6
			Advanced	5,746	9.5
			Total	29,994	100.0
	Biology	Female	Below Basic	1,841	3.1
			Basic	11,295	18.9
			Proficient	13,918	23.3
			Advanced	3,218	5.4
			Total	30,272	100.0
		Male	Below Basic	1,832	3.1
			Basic	9,552	16.0
			Proficient	14,051	23.5
			Advanced	4,102	6.9
			Total	29,537	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.14: Achievement-Level Distribution by Ethnicity

Test Period	Subject	Ethnicity	Achievement Level	Frequency	Percentage
Summer 2009	English II	American Indian	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Asian/Pacific Islander	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Black (not Hispanic)	Below Basic	38	5.4
			Basic	124	17.8
			Proficient	85	12.2
			Advanced	2	0.3
			Total	249	100.0
		Hispanic	Below Basic	2	0.3
			Basic	20	2.9
			Proficient	9	1.3
			Advanced	2	0.3
			Total	33	100.0
		White (not Hispanic)	Below Basic	33	4.7
			Basic	162	23.2
			Proficient	185	26.5
			Advanced	24	3.4
			Total	404	100.0
	Algebra I	American Indian	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Asian/Pacific Islander	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Black (not Hispanic)	Below Basic	145	11.6
			Basic	233	18.6
			Proficient	68	5.4
			Advanced	3	0.2
			Total	449	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.14: Achievement-Level Distribution by Ethnicity (continued)

Test Period	Subject	Ethnicity	Achievement Level	Frequency	Percentage
Summer 2009	Algebra I	Hispanic	Below Basic	28	2.2
			Basic	40	3.2
			Proficient	16	1.3
			Advanced	2	0.2
			Total	86	100.0
		White (not Hispanic)	Below Basic	79	6.3
			Basic	331	26.4
			Proficient	228	18.2
			Advanced	58	4.6
			Total	696	100.0
	Biology	American Indian	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Asian/Pacific Islander	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Black (not Hispanic)	Below Basic	72	15.0
			Basic	140	29.2
			Proficient	19	4.0
			Advanced	1	0.2
			Total	232	100.0
		Hispanic	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		White (not Hispanic)	Below Basic	17	3.5
			Basic	103	21.5
			Proficient	81	16.9
			Advanced	16	3.3
			Total	217	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.14: Achievement-Level Distribution by Ethnicity (continued)

Test Period	Subject	Ethnicity	Achievement Level	Frequency	Percentage
Fall 2009	English II	American Indian	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Asian/Pacific Islander	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Black (not Hispanic)	Below Basic	13	0.9
			Basic	127	8.9
			Proficient	156	11.0
			Advanced	8	0.6
			Total	304	100.0
		Hispanic	Below Basic	1	0.1
			Basic	9	0.6
			Proficient	24	1.7
			Advanced	3	0.2
			Total	37	100.0
		White (not Hispanic)	Below Basic	9	0.6
			Basic	182	12.8
			Proficient	685	48.2
			Advanced	174	12.3
			Total	1,050	100.0
	Algebra I	American Indian	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Asian/Pacific Islander	Below Basic	4	0.2
			Basic	19	0.8
			Proficient	21	0.8
			Advanced	6	0.2
			Total	50	100.0
		Black (not Hispanic)	Below Basic	90	3.6
			Basic	286	11.5
			Proficient	65	2.6
			Advanced	12	0.5
			Total	453	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.14: Achievement-Level Distribution by Ethnicity (continued)

Test Period	Subject	Ethnicity	Achievement Level	Frequency	Percentage
Fall 2009	Algebra I	Hispanic	Below Basic	13	0.5
			Basic	43	1.7
			Proficient	24	1.0
			Advanced	9	0.4
			Total	89	100.0
		White (not Hispanic)	Below Basic	101	4.1
			Basic	608	24.4
			Proficient	831	33.4
			Advanced	347	14.0
			Total	1,887	100.0
	Biology	American Indian	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
		Asian/Pacific Islander	Below Basic	2	0.1
			Basic	15	0.7
			Proficient	23	1.1
			Advanced	29	1.4
			Total	69	100.0
		Black (not Hispanic)	Below Basic	100	4.7
			Basic	208	9.8
			Proficient	56	2.6
			Advanced	10	0.5
			Total	374	100.0
		Hispanic	Below Basic	7	0.3
			Basic	27	1.3
			Proficient	18	0.9
			Advanced	8	0.4
			Total	60	100.0
		White (not Hispanic)	Below Basic	75	3.5
			Basic	452	21.4
			Proficient	764	36.1
			Advanced	314	14.8
			Total	1,605	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.14: Achievement-Level Distribution by Ethnicity (continued)

Test Period	Subject	Ethnicity	Achievement Level	Frequency	Percentage
Spring 2010	English II	American Indian	Below Basic	11	0.0
			Basic	93	0.2
			Proficient	178	0.3
			Advanced	48	0.1
			Total	330	100.0
		Asian/Pacific Islander	Below Basic	21	0.0
			Basic	199	0.3
			Proficient	508	0.8
			Advanced	385	0.6
			Total	1,113	100.0
		Black (not Hispanic)	Below Basic	617	1.0
			Basic	4,009	6.5
			Proficient	4,497	7.3
			Advanced	710	1.1
			Total	9,833	100.0
		Hispanic	Below Basic	62	0.1
			Basic	546	0.9
			Proficient	970	1.6
			Advanced	287	0.5
			Total	1,865	100.0
		White (not Hispanic)	Below Basic	1,106	1.8
			Basic	9,328	15.1
			Proficient	25,441	41.2
			Advanced	12,727	20.6
			Total	48,602	100.0
	Algebra I	American Indian	Below Basic	31	0.1
			Basic	121	0.2
			Proficient	132	0.2
			Advanced	30	0.0
			Total	314	100.0
		Asian/Pacific Islander	Below Basic	53	0.1
			Basic	268	0.4
			Proficient	505	0.8
			Advanced	467	0.8
			Total	1,293	100.0
		Black (not Hispanic)	Below Basic	1,566	2.6
			Basic	4,820	8.0
			Proficient	2,712	4.5
			Advanced	538	0.9
			Total	9,636	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.14: Achievement-Level Distribution by Ethnicity (continued)

Test Period	Subject	Ethnicity	Achievement Level	Frequency	Percentage
Spring 2010	Algebra I	Hispanic	Below Basic	181	0.3
			Basic	861	1.4
			Proficient	800	1.3
			Advanced	227	0.4
			Total	2,069	100.0
		White (not Hispanic)	Below Basic	1,863	3.1
			Basic	14,454	23.9
			Proficient	21,203	35.1
			Advanced	9,570	15.8
			Total	47,090	100.0
	Biology	American Indian	Below Basic	23	0.0
			Basic	127	0.2
			Proficient	169	0.3
			Advanced	21	0.0
			Total	340	100.0
		Asian/Pacific Islander	Below Basic	60	0.1
			Basic	291	0.5
			Proficient	508	0.8
			Advanced	305	0.5
			Total	1,164	100.0
		Black (not Hispanic)	Below Basic	1,699	2.8
			Basic	4,892	8.2
			Proficient	2,485	4.2
			Advanced	200	0.3
			Total	9,276	100.0
		Hispanic	Below Basic	177	0.3
			Basic	909	1.5
			Proficient	717	1.2
			Advanced	124	0.2
			Total	1,927	100.0
		White (not Hispanic)	Below Basic	1,710	2.9
			Basic	14,603	24.4
			Proficient	24,087	40.3
			Advanced	6,669	11.2
			Total	47,069	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.15: Achievement-Level Distribution—Migrant

Test Period	Subject	Migrant	Achievement Level	Frequency	Percentage
Summer 2009	English II	No	Below Basic	74	10.5
			Basic	317	44.9
			Proficient	286	40.5
			Advanced	28	4.0
			Total	705	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Algebra I	No	Below Basic	271	21.1
			Basic	629	49.0
			Proficient	320	24.9
			Advanced	64	5.0
			Total	1,284	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Biology	No	Below Basic	99	20.2
			Basic	270	55.0
			Proficient	104	21.2
			Advanced	18	3.7
			Total	491	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.15: Achievement-Level Distribution—Migrant (continued)

Test Period	Subject	Migrant	Achievement Level	Frequency	Percentage
Fall 2009	English II	No	Below Basic	23	1.6
			Basic	325	22.9
			Proficient	882	62.0
			Advanced	190	13.4
			Total	1,420	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Algebra I	No	Below Basic	207	8.3
			Basic	963	38.7
			Proficient	943	37.9
			Advanced	374	15.0
			Total	2,487	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Biology	No	Below Basic	187	8.8
			Basic	706	33.3
			Proficient	867	40.9
			Advanced	362	17.1
			Total	2,122	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.15: Achievement-Level Distribution—Migrant (continued)

Test Period	Subject	Migrant	Achievement Level	Frequency	Percentage
Spring 2010	English II	No	Below Basic	1,824	2.9
			Basic	14,248	23.0
			Proficient	31,647	51.1
			Advanced	14,161	22.9
			Total	61,880	100.0
		Yes	Below Basic	6	0.0
			Basic	12	0.0
			Proficient	11	0.0
			Advanced	2	0.0
			Total	31	100.0
	Algebra I	No	Below Basic	3,730	6.2
			Basic	20,562	34.0
			Proficient	25,371	41.9
			Advanced	10,830	17.9
			Total	60,493	100.0
		Yes	Below Basic	3	0.0
			Basic	31	0.1
			Proficient	10	0.0
			Advanced	7	0.0
			Total	51	100.0
	Biology	No	Below Basic	3,695	6.2
			Basic	20,870	34.8
			Proficient	27,972	46.7
			Advanced	7,327	12.2
			Total	59,864	100.0
		Yes	Below Basic	8	0.0
			Basic	20	0.0
			Proficient	12	0.0
			Advanced	0	0.0
			Total	40	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.16: Achievement-Level Distribution—FRL

Test Period	Subject	FRL	Achievement Level	Frequency	Percentage
Summer 2009	English II	No	Below Basic	45	6.4
			Basic	134	19.0
			Proficient	165	23.4
			Advanced	18	2.5
			Total	362	100.0
		Yes	Below Basic	29	4.1
			Basic	184	26.1
			Proficient	121	17.1
			Advanced	10	1.4
			Total	344	100.0
	Algebra I	No	Below Basic	109	8.5
			Basic	312	24.3
			Proficient	211	16.4
			Advanced	49	3.8
			Total	681	100.0
		Yes	Below Basic	162	12.6
			Basic	317	24.7
			Proficient	109	8.5
			Advanced	15	1.2
			Total	603	100.0
	Biology	No	Below Basic	36	7.3
			Basic	121	24.6
			Proficient	57	11.6
			Advanced	18	3.7
			Total	232	100.0
		Yes	Below Basic	63	12.8
			Basic	149	30.3
			Proficient	47	9.6
			Advanced	0	0.0
			Total	259	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.16: Achievement-Level Distribution—FRL (continued)

Test Period	Subject	FRL	Achievement Level	Frequency	Percentage
Fall 2009	English II	No	Below Basic	8	0.6
			Basic	126	8.9
			Proficient	535	37.6
			Advanced	156	11.0
			Total	825	100.0
		Yes	Below Basic	15	1.1
			Basic	199	14.0
			Proficient	349	24.5
			Advanced	34	2.4
			Total	597	100.0
	Algebra I	No	Below Basic	86	3.5
			Basic	479	19.3
			Proficient	704	28.3
			Advanced	330	13.3
			Total	1,599	100.0
		Yes	Below Basic	122	4.9
			Basic	484	19.5
			Proficient	239	9.6
			Advanced	44	1.8
			Total	889	100.0
	Biology	No	Below Basic	61	2.9
			Basic	343	16.2
			Proficient	682	32.1
			Advanced	322	15.2
			Total	1,408	100.0
		Yes	Below Basic	126	5.9
			Basic	363	17.1
			Proficient	185	8.7
			Advanced	40	1.9
			Total	714	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.16: Achievement-Level Distribution—FRL (continued)

Test Period	Subject	FRL	Achievement Level	Frequency	Percentage
Spring 2010	English II	No	Below Basic	652	1.1
			Basic	6,269	10.1
			Proficient	20,123	32.5
			Advanced	11,386	18.4
			Total	38,430	100.0
		Yes	Below Basic	1,178	1.9
			Basic	7,991	12.9
			Proficient	11,535	18.6
			Advanced	2,777	4.5
			Total	23,481	100.0
	Algebra I	No	Below Basic	1,342	2.2
			Basic	10,317	17.0
			Proficient	16,931	28.0
			Advanced	8,649	14.3
			Total	37,239	100.0
		Yes	Below Basic	2,391	3.9
			Basic	10,276	17.0
			Proficient	8,450	14.0
			Advanced	2,188	3.6
			Total	23,305	100.0
	Biology	No	Below Basic	1,209	2.0
			Basic	10,693	17.9
			Proficient	19,693	32.9
			Advanced	6,222	10.4
			Total	37,817	100.0
		Yes	Below Basic	2,494	4.2
			Basic	10,197	17.0
			Proficient	8,291	13.8
			Advanced	1,105	1.8
			Total	22,087	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.17: Achievement-Level Distribution—LEP

Test Period	Subject	LEP	Achievement Level	Frequency	Percentage
Summer 2009	English II	No	Below Basic	71	10.1
			Basic	315	44.6
			Proficient	283	40.1
			Advanced	27	3.8
			Total	696	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Algebra I	No	Below Basic	255	19.9
			Basic	606	47.2
			Proficient	315	24.5
			Advanced	64	5.0
			Total	1,240	100.0
		Yes	Below Basic	16	1.2
			Basic	23	1.8
			Proficient	5	0.4
			Advanced	0	0.0
			Total	44	100.0
	Biology	No	Below Basic	91	18.5
			Basic	260	53.0
			Proficient	104	21.2
			Advanced	18	3.7
			Total	473	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.17: Achievement-Level Distribution—LEP (continued)

Test Period	Subject	LEP	Achievement Level	Frequency	Percentage
Fall 2009	English II	No	Below Basic	23	1.6
			Basic	321	22.6
			Proficient	874	61.5
			Advanced	190	13.4
			Total	1,408	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Algebra I	No	Below Basic	203	8.2
			Basic	947	38.1
			Proficient	938	37.7
			Advanced	372	15.0
			Total	2,460	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Biology	No	Below Basic	184	8.7
			Basic	700	33.0
			Proficient	864	40.7
			Advanced	361	17.0
			Total	2,109	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.17: Achievement-Level Distribution—LEP (continued)

Test Period	Subject	LEP	Achievement Level	Frequency	Percentage
Spring 2010	English II	No	Below Basic	1,774	2.9
			Basic	13,912	22.5
			Proficient	31,291	50.5
			Advanced	14,112	22.8
			Total	61,089	100.0
		Yes	Below Basic	56	0.1
			Basic	348	0.6
			Proficient	367	0.6
			Advanced	51	0.1
			Total	822	100.0
	Algebra I	No	Below Basic	3,571	5.9
			Basic	20,052	33.1
			Proficient	25,089	41.4
			Advanced	10,740	17.7
			Total	59,452	100.0
		Yes	Below Basic	162	0.3
			Basic	541	0.9
			Proficient	292	0.5
			Advanced	97	0.2
			Total	1,092	100.0
	Biology	No	Below Basic	3,537	5.9
			Basic	20,391	34.0
			Proficient	27,785	46.4
			Advanced	7,289	12.2
			Total	59,002	100.0
		Yes	Below Basic	166	0.3
			Basic	499	0.8
			Proficient	199	0.3
			Advanced	38	0.1
			Total	902	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.18: Achievement-Level Distribution—Title I

Test Period	Subject	Title I	Achievement Level	Frequency	Percentage
Summer 2009	English II	No	Below Basic	70	9.9
			Basic	291	41.2
			Proficient	265	37.5
			Advanced	26	3.7
			Total	652	100.0
		Yes	Below Basic	4	0.6
			Basic	27	3.8
			Proficient	21	3.0
			Advanced	2	0.3
			Total	54	100.0
	Algebra I	No	Below Basic	229	17.8
			Basic	580	45.2
			Proficient	310	24.1
			Advanced	64	5.0
			Total	1,183	100.0
		Yes	Below Basic	42	3.3
			Basic	49	3.8
			Proficient	10	0.8
			Advanced	0	0.0
			Total	101	100.0
	Biology	No	Below Basic	77	15.7
			Basic	243	49.5
			Proficient	101	20.6
			Advanced	18	3.7
			Total	439	100.0
		Yes	Below Basic	22	4.5
			Basic	27	5.5
			Proficient	3	0.6
			Advanced	0	0.0
			Total	52	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.18: Achievement-Level Distribution—Title I (continued)

Test Period	Subject	Title I	Achievement Level	Frequency	Percentage
Fall 2009	English II	No	Below Basic	18	1.3
			Basic	296	20.8
			Proficient	836	58.8
			Advanced	188	13.2
			Total	1,338	100.0
		Yes	Below Basic	5	0.4
			Basic	29	2.0
			Proficient	48	3.4
			Advanced	2	0.1
			Total	84	100.0
	Algebra I	No	Below Basic	176	7.1
			Basic	886	35.6
			Proficient	932	37.5
			Advanced	374	15.0
			Total	2,368	100.0
		Yes	Below Basic	32	1.3
			Basic	77	3.1
			Proficient	11	0.4
			Advanced	0	0.0
			Total	120	100.0
	Biology	No	Below Basic	150	7.1
			Basic	658	31.0
			Proficient	862	40.6
			Advanced	362	17.1
			Total	2,032	100.0
		Yes	Below Basic	37	1.7
			Basic	48	2.3
			Proficient	5	0.2
			Advanced	0	0.0
			Total	90	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.18: Achievement-Level Distribution—Title I (continued)

Test Period	Subject	Title I	Achievement Level	Frequency	Percentage
Spring 2010	English II	No	Below Basic	1,601	2.6
			Basic	13,199	21.3
			Proficient	30,527	49.3
			Advanced	13,975	22.6
			Total	59,302	100.0
		Yes	Below Basic	229	0.4
			Basic	1,061	1.7
			Proficient	1,131	1.8
			Advanced	188	0.3
			Total	2,609	100.0
	Algebra I	No	Below Basic	3,242	5.4
			Basic	19,205	31.7
			Proficient	24,506	40.5
			Advanced	10,595	17.5
			Total	57,548	100.0
		Yes	Below Basic	491	0.8
			Basic	1,388	2.3
			Proficient	875	1.4
			Advanced	242	0.4
			Total	2,996	100.0
	Biology	No	Below Basic	3,108	5.2
			Basic	19,696	32.9
			Proficient	27,490	45.9
			Advanced	7,290	12.2
			Total	57,584	100.0
		Yes	Below Basic	595	1.0
			Basic	1,194	2.0
			Proficient	494	0.8
			Advanced	37	0.1
			Total	2,320	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.19: Achievement-Level Distribution—IEP

Test Period	Subject	IEP	Achievement Level	Frequency	Percentage
Summer 2009	English II	No	Below Basic	57	8.1
			Basic	263	37.3
			Proficient	269	38.1
			Advanced	26	3.7
			Total	615	100.0
		Yes	Below Basic	17	2.4
			Basic	55	7.8
			Proficient	17	2.4
			Advanced	2	0.3
			Total	91	100.0
	Algebra I	No	Below Basic	218	17.0
			Basic	548	42.7
			Proficient	302	23.5
			Advanced	62	4.8
			Total	1,130	100.0
		Yes	Below Basic	53	4.1
			Basic	81	6.3
			Proficient	18	1.4
			Advanced	2	0.2
			Total	154	100.0
	Biology	No	Below Basic	84	17.1
			Basic	233	47.5
			Proficient	96	19.6
			Advanced	18	3.7
			Total	431	100.0
		Yes	Below Basic	15	3.1
			Basic	37	7.5
			Proficient	8	1.6
			Advanced	0	0.0
			Total	60	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.19: Achievement-Level Distribution—IEP (continued)

Test Period	Subject	IEP	Achievement Level	Frequency	Percentage
Fall 2009	English II	No	Below Basic	15	1.1
			Basic	251	17.7
			Proficient	825	58.0
			Advanced	186	13.1
			Total	1,277	100.0
		Yes	Below Basic	8	0.6
			Basic	74	5.2
			Proficient	59	4.1
			Advanced	4	0.3
			Total	145	100.0
	Algebra I	No	Below Basic	163	6.6
			Basic	866	34.8
			Proficient	914	36.7
			Advanced	370	14.9
			Total	2,313	100.0
		Yes	Below Basic	45	1.8
			Basic	97	3.9
			Proficient	29	1.2
			Advanced	4	0.2
			Total	175	100.0
	Biology	No	Below Basic	128	6.0
			Basic	613	28.9
			Proficient	833	39.3
			Advanced	353	16.6
			Total	1,927	100.0
		Yes	Below Basic	59	2.8
			Basic	93	4.4
			Proficient	34	1.6
			Advanced	9	0.4
			Total	195	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.19: Achievement-Level Distribution—IEP (continued)

Test Period	Subject	IEP	Achievement Level	Frequency	Percentage
Spring 2010	English II	No	Below Basic	855	1.4
			Basic	11,013	17.8
			Proficient	29,937	48.4
			Advanced	13,957	22.5
			Total	55,762	100.0
		Yes	Below Basic	975	1.6
			Basic	3,247	5.2
			Proficient	1,721	2.8
			Advanced	206	0.3
			Total	6,149	100.0
	Algebra I	No	Below Basic	2,708	4.5
			Basic	18,466	30.5
			Proficient	24,490	40.4
			Advanced	10,643	17.6
			Total	56,307	100.0
		Yes	Below Basic	1,025	1.7
			Basic	2,127	3.5
			Proficient	891	1.5
			Advanced	194	0.3
			Total	4,237	100.0
	Biology	No	Below Basic	2,541	4.2
			Basic	18,630	31.1
			Proficient	26,995	45.1
			Advanced	7,226	12.1
			Total	55,392	100.0
		Yes	Below Basic	1,162	1.9
			Basic	2,260	3.8
			Proficient	989	1.7
			Advanced	101	0.2
			Total	4,512	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.20: Achievement-Level Distribution—Accommodations

Test Period	Subject	Accommodations	Achievement Level	Frequency	Percentage
Summer 2009	English II	No	Below Basic	71	10.1
			Basic	309	43.8
			Proficient	281	39.8
			Advanced	27	3.8
			Total	688	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Algebra I	No	Below Basic	263	20.5
			Basic	615	47.9
			Proficient	319	24.8
			Advanced	63	4.9
			Total	1,260	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		
	Biology	No	Below Basic	93	18.9
			Basic	261	53.2
			Proficient	102	20.8
			Advanced	18	3.7
			Total	474	100.0
		Yes	Below Basic		
			Basic		
			Proficient		
			Advanced		
			Total		

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.20: Achievement-Level Distribution—Accommodations (continued)

Test Period	Subject	Accommodations	Achievement Level	Frequency	Percentage
Fall 2009	English II	No	Below Basic	21	1.5
			Basic	300	21.1
			Proficient	866	60.9
			Advanced	189	13.3
			Total	1,376	100.0
		Yes	Below Basic	2	0.1
			Basic	25	1.8
			Proficient	18	1.3
			Advanced	1	0.1
			Total	46	100.0
	Algebra I	No	Below Basic	193	7.8
			Basic	930	37.4
			Proficient	934	37.5
			Advanced	372	15.0
			Total	2,429	100.0
		Yes	Below Basic	15	0.6
			Basic	33	1.3
			Proficient	9	0.4
			Advanced	2	0.1
			Total	59	100.0
	Biology	No	Below Basic	173	8.2
			Basic	671	31.6
			Proficient	857	40.4
			Advanced	359	16.9
			Total	2,060	100.0
		Yes	Below Basic	14	0.7
			Basic	35	1.6
			Proficient	10	0.5
			Advanced	3	0.1
			Total	62	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 9.20: Achievement-Level Distribution—Accommodations (continued)

Test Period	Subject	Accommodations	Achievement Level	Frequency	Percentage
Spring 2010	English II	No	Below Basic	1,234	2.0
			Basic	12,492	20.2
			Proficient	30,915	49.9
			Advanced	14,113	22.8
			Total	58,754	100.0
		Yes	Below Basic	596	1.0
			Basic	1,768	2.9
			Proficient	743	1.2
			Advanced	50	0.1
			Total	3,157	100.0
	Algebra I	No	Below Basic	3,274	5.4
			Basic	19,777	32.7
			Proficient	25,136	41.5
			Advanced	10,809	17.9
			Total	58,996	100.0
		Yes	Below Basic	459	0.8
			Basic	816	1.3
			Proficient	245	0.4
			Advanced	28	0.0
			Total	1,548	100.0
	Biology	No	Below Basic	3,272	5.5
			Basic	20,089	33.5
			Proficient	27,668	46.2
			Advanced	7,304	12.2
			Total	58,333	100.0
		Yes	Below Basic	431	0.7
			Basic	801	1.3
			Proficient	316	0.5
			Advanced	23	0.0
			Total	1,571	100.0

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Chapter 10: Reliability

10.1 Introduction

The Missouri Department of Elementary and Secondary Education (DESE) is required by federal law to ensure that the instruments used to measure student achievement for school accountability provide reliable results. This chapter provides evidence that scores from the Missouri End-of-Course (EOC) Assessments measure student achievement in a reliable manner and that the size of the measurement error associated with reported test scores is reasonable, especially at the Proficient cut score.

10.2 Reliability and Measurement Error

10.2.1 Defining Reliability

Reliability refers to the consistency of student test scores. *Measurement error* refers to the random variability in the test scores. Both are indicators of the degree of precision in a test score. In general, measurement error and reliability are inversely related. When measurement error is large, reliability is small. Increasing reliability by minimizing measurement error is an important goal in the construction of any test.

Estimating the size of the measurement error associated with a true score is the key to estimating reliability. Errors in measurement can result from any of a multitude of factors, including environmental factors (e.g., testing conditions) and examinee factors (e.g., fatigue, stress). Feldt and Brennan (1989) note that “Quantification of the consistency and inconsistency in examinee performance constitutes the essence of reliability analysis” (p. 105). Classical test theory (CTT) provides a means for this quantification of examinee inconsistency (i.e., measurement error). This approach builds on the notion of an ideal error-free, or true, measurement score. Any observed measurement, such as test score *X*, is defined as a composite of true score, *T*, and its associated error:

$$X = T + \text{error}.$$

The definitions or assumptions in CTT lead to several important properties. For example, it can be demonstrated that observed score variance equals the sum of true score variance plus error variance:

$$\sigma_X^2 = \sigma_t^2 + \sigma_e^2.$$

The relationship among variance terms (i.e., $\sigma_X^2, \sigma_t^2, \sigma_e^2$) is critical to a more thorough understanding of important CTT concepts, including reliability and the standard error of measurement (*SEM*). For example, CTT equivalence reliability is defined as the correlation between observed scores on parallel forms, which is equal to

$$\rho_{x_1 x_2} = \sigma_t^2 \big/ \sigma_X^2.$$

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Reliability in CTT is thus conceptualized as true score variance divided by observed score variance. With just a few algebraic steps, the CTT definition of the *SEM* can be derived:

$$\sigma_e = \sigma_x \sqrt{1 - \rho_{x_1 x_2}}.$$

Although the conceptualizations of reliability and *SEM* are relatively straightforward, issues underlying the estimation of reliability are not.

10.2.2 Estimating Reliability

Reliability can be estimated via the correlation of scores on parallel forms (equivalence reliability) or from test-retest data (stability reliability), or it can be estimated from a single test administration (internal consistency reliability) using any one of a variety of techniques (e.g., Brown, 1910; Cronbach, 1951; Kuder and Richardson, 1937). A very popular index for describing internal consistency reliability based on a single test administration is Cronbach's coefficient alpha, which provides an estimate of reliability that is mathematically equivalent to the average of all possible split-half reliability estimates.

10.2.3 Sources of Measurement Error

As noted above, errors in measurement can result from environmental factors and examinee factors. To reduce other sources of measurement error, the scoring of student responses to selected response (SR) items was done electronically. Scoring error may result from improper coding or extraneous marks on scannable response sheets. The size of this sort of error is usually small and is controlled through standardized test administration procedures (including detailed instructions on how to fill out response sheets and how to erase extraneous markings) and quality control measures implemented during the scanning process.

The Performance Event (PE) and Writing Prompt (WP) items are susceptible to scoring error due to ambiguity in the scoring rubric and differences among raters. Rubrics were written to balance generality and specificity and to cover the range of student responses, while at the same time allowing raters to easily identify the response characteristics distinguishing each score category. To minimize rater error, the Assessment Resource Center (ARC) at the University of Missouri—the organization that conducted the hand scoring of the PE/WP items—thoroughly trained raters and monitored the scoring process. Only raters who met ARC's criteria for consistent scoring during training were retained as scorers.

10.3 Evidence of Raw-Score Internal Consistency

Consistency of individual student performance was estimated using Cronbach's coefficient alpha. As previously noted, coefficient alpha provides an estimate of reliability that is mathematically equivalent to the average of all possible split-half reliability estimates. Alpha is an appropriate index of internal consistency for use on untimed tests such as the MO EOC Assessments.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Separate analyses were performed for each EOC content area. Both SR and PE items were used in the computations. Cronbach's alpha can be interpreted as a lower bound to reliability and can be estimated using the following formula:

$$\alpha = \frac{n}{n-1}[1 - \frac{\sum_{i=1}^{n} \sigma_{Y_i}^2}{\sigma_X^2}],$$

where *n* is the number of items, $\sigma_{Y_i}^2$ is the variance of item *i*, and σ_X^2 is the variance of the total score. Following this, *SEM* can be interpreted as "the square root of the average of the person-specific error variances of all examinees who participated in the reliability estimation experiment" (Traub, 1994, p. 114). *SEM*s were calculated using the following formula:

$$SEM = S_X \sqrt{1-\alpha},$$

where S_X is the standard deviation of observed total scores. Tables 10.1 to 10.9 show the reliability coefficients (Cronbach's alpha) and *SEM*s based on the raw-score metric for the total population and for select student subgroups.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.1: Alpha Coefficients and Standard Errors of Measurement, English II, Summer 2009

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	22.31	6.16	706	0.81	2.67
Gender					
Female	22.54	6.20	250	0.82	2.64
Male	22.19	6.14	456	0.81	2.68
Ethnicity					
White	23.31	6.30	404	0.83	2.61
Black	20.88	5.80	249	0.78	2.73
Hispanic	21.85	5.65	33	0.77	2.70
Asian	N/A	N/A	N/A	N/A	N/A
American Indian	N/A	N/A	N/A	N/A	N/A
LEP					
Yes	N/A	N/A	N/A	N/A	N/A
No	22.34	6.13	696	0.81	2.67
IEP					
Yes	19.47	5.95	91	0.79	2.74
No	22.73	6.08	615	0.81	2.65
Migrant					
Yes	N/A	N/A	N/A	N/A	N/A
No	22.32	6.16	705	0.81	2.67
FRL					
Yes	21.75	5.63	344	0.77	2.70
No	22.85	6.58	362	0.84	2.63
Title I					
Yes	22.37	5.77	54	0.80	2.60
No	22.31	6.19	652	0.81	2.67
Accommodations					
Yes	22.33	6.15	688	0.81	2.66
No	22.31	6.16	706	0.81	2.67

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.2: Alpha Coefficients and Standard Errors of Measurement, English II, Fall 2009

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	23.11	6.40	1,422	0.82	2.75
Gender					
Female	23.99	6.39	671	0.82	2.71
Male	22.32	6.31	751	0.81	2.77
Ethnicity					
White	24.21	6.19	1,050	0.81	2.71
Black	19.34	5.70	304	0.75	2.85
Hispanic	21.78	5.99	37	0.79	2.73
Asian	N/A	N/A	N/A	N/A	N/A
American Indian	N/A	N/A	N/A	N/A	N/A
LEP					
Yes	N/A	N/A	N/A	N/A	N/A
No	23.12	6.41	1,408	0.82	2.74
IEP					
Yes	18.32	6.00	145	0.78	2.84
No	23.65	6.22	1,277	0.81	2.73
Migrant					
Yes	N/A	N/A	N/A	N/A	N/A
No	23.11	6.41	1,420	0.82	2.74
FRL					
Yes	20.86	5.94	597	0.77	2.82
No	24.73	6.24	825	0.82	2.68
Title I					
Yes	19.64	5.99	84	0.78	2.81
No	23.32	6.37	1,338	0.82	2.74
Accommodations					
Yes	18.11	6.29	46	0.80	2.83
No	23.27	6.34	1,376	0.81	2.74

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.3: Alpha Coefficients and Standard Errors of Measurement, English II, Spring 2010

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	27.67	6.16	61,911	0.84	2.44
Gender					
Female	28.49	5.71	30,983	0.82	2.39
Male	26.87	6.47	30,819	0.85	2.49
Ethnicity					
White	28.41	5.86	48,602	0.83	2.39
Black	24.19	6.32	9,833	0.83	2.64
Hispanic	26.34	6.18	1,865	0.83	2.52
Asian	29.21	6.08	1,113	0.85	2.33
American Indian	26.79	5.85	330	0.81	2.52
LEP					
Yes	23.77	6.36	822	0.83	2.65
No	27.72	6.14	61,089	0.84	2.44
IEP					
Yes	20.74	6.70	6,149	0.83	2.77
No	28.44	5.59	55,762	0.82	2.40
Migrant					
Yes	21.58	7.66	31	0.87	2.76
No	27.67	6.15	61,880	0.84	2.44
FRL					
Yes	25.40	6.37	23,481	0.84	2.58
No	29.06	5.58	38,430	0.82	2.35
Title I					
Yes	23.65	6.72	2,609	0.84	2.65
No	27.85	6.07	59,302	0.84	2.43
Accommodations					
Yes	19.72	6.42	3,157	0.81	2.80
No	28.10	5.84	58,754	0.83	2.42

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.4: Alpha Coefficients and Standard Errors of Measurement, Algebra I, Summer 2009

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	18.21	6.72	1,284	0.83	2.79
Gender					
Female	17.87	6.86	558	0.84	2.78
Male	18.46	6.61	722	0.82	2.80
Ethnicity					
White	20.30	6.65	696	0.83	2.76
Black	15.77	5.84	449	0.77	2.82
Hispanic	15.85	6.21	86	0.80	2.81
Asian	N/A	N/A	N/A	N/A	N/A
American Indian	N/A	N/A	N/A	N/A	N/A
LEP					
Yes	14.68	5.17	44	0.71	2.81
No	18.34	6.74	1,240	0.83	2.79
IEP					
Yes	15.70	6.08	154	0.79	2.80
No	18.56	6.74	1,130	0.83	2.79
Migrant					
Yes	N/A	N/A	N/A	N/A	N/A
No	18.21	6.72	1,284	0.83	2.79
FRL					
Yes	16.70	6.21	603	0.80	2.79
No	19.56	6.88	681	0.84	2.79
Title I					
Yes	14.30	5.11	101	0.70	2.79
No	18.55	6.74	1,183	0.83	2.79
Accommodations					
Yes	N/A	N/A	N/A	N/A	N/A
No	18.27	6.74	1,260	0.83	2.79

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.5: Alpha Coefficients and Standard Errors of Measurement, Algebra I, Fall 2009

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	19.64	7.61	2,488	0.86	2.89
Gender					
Female	19.49	7.44	1,253	0.85	2.88
Male	19.81	7.78	1,233	0.86	2.90
Ethnicity					
White	21.15	7.37	1,887	0.85	2.90
Black	13.88	5.67	453	0.76	2.75
Hispanic	17.28	7.43	89	0.85	2.83
Asian	19.92	6.90	50	0.83	2.81
American Indian	N/A	N/A	N/A	N/A	N/A
LEP					
Yes	N/A	N/A	N/A	N/A	N/A
No	19.68	7.61	2,460	0.86	2.89
IEP					
Yes	13.74	6.10	175	0.80	2.75
No	20.09	7.53	2,313	0.85	2.90
Migrant					
Yes	N/A	N/A	N/A	N/A	N/A
No	19.65	7.61	2,487	0.86	2.89
FRL					
Yes	16.10	6.43	889	0.81	2.83
No	21.61	7.51	1,599	0.85	2.90
Title I					
Yes	12.50	4.21	120	0.58	2.73
No	20.00	7.57	2,368	0.85	2.89
Accommodations					
Yes	13.69	6.09	59	0.79	2.78
No	19.79	7.59	2,429	0.85	2.89

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.6: Alpha Coefficients and Standard Errors of Measurement, Algebra I, Spring 2010

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	23.24	7.25	60,544	0.85	2.80
Gender					
Female	23.08	7.14	30,460	0.85	2.80
Male	23.42	7.36	29,994	0.86	2.79
Ethnicity					
White	24.19	6.92	47,090	0.84	2.76
Black	18.71	6.93	9,636	0.83	2.89
Hispanic	21.34	7.05	2,069	0.84	2.86
Asian	26.55	7.57	1,293	0.87	2.68
American Indian	21.59	6.98	314	0.84	2.83
LEP					
Yes	19.16	7.42	1,092	0.85	2.91
No	23.31	7.23	59,452	0.85	2.80
IEP					
Yes	17.00	7.11	4,237	0.84	2.88
No	23.71	7.04	56,307	0.84	2.78
Migrant					
Yes	19.16	6.88	51	0.83	2.86
No	23.24	7.25	60,493	0.85	2.80
FRL					
Yes	20.73	7.08	23,305	0.84	2.86
No	24.80	6.91	37,239	0.84	2.74
Title I					
Yes	19.21	7.47	2,996	0.85	2.89
No	23.45	7.18	57,548	0.85	2.79
Accommodations					
Yes	15.56	6.38	1,548	0.80	2.87
No	23.44	7.16	58,996	0.85	2.79

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.7: Alpha Coefficients and Standard Errors of Measurement, Biology, Summer 2009

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	26.89	9.62	491	0.87	3.49
Gender					
Female	27.20	9.78	237	0.87	3.49
Male	26.70	9.45	252	0.86	3.48
Ethnicity					
White	31.72	9.23	217	0.86	3.46
Black	22.93	8.00	232	0.81	3.48
Hispanic	N/A	N/A	N/A	N/A	N/A
Asian	N/A	N/A	N/A	N/A	N/A
American Indian	N/A	N/A	N/A	N/A	N/A
LEP					
Yes	N/A	N/A	N/A	N/A	N/A
No	27.20	9.60	473	0.87	3.49
IEP					
Yes	23.95	8.14	60	0.81	3.51
No	27.30	9.75	431	0.87	3.48
Migrant					
Yes	N/A	N/A	N/A	N/A	N/A
No	26.89	9.62	491	0.87	3.49
FRL					
Yes	24.86	8.61	259	0.84	3.49
No	29.16	10.19	232	0.88	3.47
Title I					
Yes	20.40	7.47	52	0.79	3.38
No	27.66	9.56	439	0.87	3.49
Accommodations					
Yes	N/A	N/A	N/A	N/A	N/A
No	27.04	9.66	474	0.87	3.48

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.8: Alpha Coefficients and Standard Errors of Measurement, Biology, Fall 2009

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	32.55	11.36	2,122	0.91	3.45
Gender					
Female	31.75	11.39	1,045	0.91	3.45
Male	33.34	11.29	1,075	0.91	3.44
Ethnicity					
White	34.85	10.28	1,605	0.89	3.39
Black	22.10	9.50	374	0.86	3.50
Hispanic	29.75	11.79	60	0.91	3.50
Asian	38.91	10.61	69	0.92	3.09
American Indian	N/A	N/A	N/A	N/A	N/A
LEP					
Yes	N/A	N/A	N/A	N/A	N/A
No	32.59	11.35	2,109	0.91	3.45
IEP					
Yes	22.76	10.67	195	0.89	3.54
No	33.54	10.95	1,927	0.90	3.43
Migrant					
Yes	N/A	N/A	N/A	N/A	N/A
No	32.55	11.36	2,122	0.91	3.45
FRL					
Yes	25.80	10.57	714	0.89	3.58
No	35.97	10.15	1,408	0.89	3.33
Title I					
Yes	17.43	7.32	90	0.79	3.39
No	33.22	11.04	2,032	0.90	3.44
Accommodations					
Yes	23.37	10.33	62	0.88	3.60
No	32.83	11.28	2,060	0.91	3.44

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.9: Alpha Coefficients and Standard Errors of Measurement, Biology, Spring 2010

Group	Mean Raw Score	*SD* Raw Score	*N* Count	Reliability	*SEM*
All Students	31.58	9.86	59,904	0.88	3.42
Gender					
Female	31.06	9.70	30,272	0.87	3.43
Male	32.14	9.99	29,537	0.88	3.40
Ethnicity					
White	33.08	9.29	47,069	0.87	3.41
Black	24.36	9.09	9,276	0.86	3.37
Hispanic	28.34	9.73	1,927	0.88	3.43
Asian	35.15	10.82	1,164	0.91	3.33
American Indian	30.26	9.05	340	0.85	3.45
LEP					
Yes	24.26	9.74	902	0.88	3.42
No	31.69	9.82	59,002	0.88	3.42
IEP					
Yes	22.72	9.42	4,512	0.88	3.33
No	32.30	9.54	55,392	0.87	3.41
Migrant					
Yes	24.03	9.02	40	0.85	3.45
No	31.58	9.86	59,864	0.88	3.42
FRL					
Yes	27.62	9.54	22,087	0.87	3.42
No	33.89	9.30	37,817	0.87	3.39
Title I					
Yes	22.59	9.18	2,320	0.87	3.32
No	31.94	9.72	57,584	0.88	3.42
Accommodations					
Yes	21.98	9.09	1,571	0.87	3.30
No	31.84	9.76	58,333	0.88	3.42

10.4 Conditional Standard Error Estimates for Scale Scores

The overall *SEM* in Tables 10.1 to 10.9 represents the standard deviation of projected replications of the testing procedure averaged over all students. In contrast, conditional standard errors of measurement (*CSEM*s) are conditioned on the ability of the student. Rasch-based *CSEM*s ($CSEM(\theta)$) for each scale score are defined as the reciprocal of the square root of the test information function ($I(\theta)$) at the point on the ability continuum that corresponds to each scale score (Hambleton and Swaminathan, 1985):

$$CSEM(\theta) = \frac{1}{\sqrt{I(\theta)}}.$$

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

*CSEM*s are especially useful for characterizing measurement precision in the neighborhood of score levels used for decision making, such as cut scores at various achievement levels. The *CSEM*s for the Proficient cut scores for the MO EOC Assessments are presented in Table 10.10. *CSEM*s for other scale scores are reported in Chapter 7 of this technical report. Note that *CSEM*s are smaller in the middle of the score distribution than at the extremes. This pattern is expected for item response theory (IRT) based *CSEM*s. The value for all *CSEM*s was either 6 or 7 scale-score points. These values reflect a reasonable amount of measurement error at the Proficient cut for making adequate yearly progress (AYP) determinations for federal accountability.

Table 10.10: *CSEM*s at the Proficient Cut Score

Test Event	Subject	SS Cut*	*CSEM*
Summer 2009	English II	200	6
	Algebra I	200	7
	Biology	200	6
Fall 2009	English II	200	6
	Algebra I	200	7
	Biology	200	5
Spring 2010	English II	200	6
	Algebra I	200	7
	Biology	200	6

*See Tables 7.22 through 7.27 in Chapter 7 for the *CSEM* at each scale score.

10.5 Evidence Supporting Scorer Reliability

Ten percent of the PE/WP items were read and scored by the table leader and a scorer during the hand-scoring process. The purpose of the 10% validation for the PE/WP items was to monitor the consistency of scorers. Tables 10.11 to 10.17 show the percentages of PE/WP items scored with exact agreement and adjacent agreement for each assessment for the Spring 2008 field test, Spring 2009 embedded field test, Summer 2009 operational administration, Fall 2009 operational administration, Fall 2009 Large Print/Braille operational administration, Spring 2010 operational administration, and Spring 2010 Large Print/Braille operational administration. (Note that these tables also appear in Chapter 6.) The scoring rubrics used for raters had a score range of 0 to 4 for English II and Algebra I. For Biology, the rubrics had score points that ranged from 0 to 1, 0 to 2, 0 to 3, and 0 to 4. There were no half points assigned for the PE/WP items.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.11: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Spring 2008 Field-Test Administration*

Content Area/ Item Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Perfect + Adjacent
English II					
RPS100076784	4	5004	492	68.3%	99.0%
RPS100076785	4	5006	494	61.1%	98.4%
Algebra I					
RPS100076683	4	5001	499	71.1%	99.0%
RPS100076622	4	5004	499	82.6%	97.8%
Biology					
RPS100076797 - 1	1	5006	476	80.9%	99.6%
RPS100076798 - 2	1	5006	476	79.8%	99.6%
RPS100076799 - 3	1	5006	476	95.6%	99.6%
RPS100076800 - 4	2	5006	476	85.5%	98.1%
RPS100076807 - 5	2	5006	476	91.2%	99.6%
RPS100076801 - 6	4	5006	476	63.2%	94.5%
RPS100076803 - 7	2	5006	476	79.4%	100.0%
RPS100076808 - 8	2	5006	476	84.5%	99.8%
RPS100076802 - 9	2	5006	476	68.9%	99.2%
RPS100076804 - 10	3	5006	476	55.5%	92.0%
RPS100077987 - 11	1	5006	476	92.9%	99.4%
RPS100076805 - 12	2	5006	476	70.8%	95.2%
RPS100076806 - 13	1	5006	476	95.4%	99.4%
RPS100075961 - 1	1	5007	483	79.1%	100.0%
RPS100075962 - 2	1	5007	483	93.2%	100.0%
RPS100075963 - 3	1	5007	483	94.8%	100.0%
RPS100075965 - 4	1	5007	483	73.1%	99.8%
RPS100075964 - 5	2	5007	483	82.0%	90.3%
RPS100075966 - 6	4	5007	483	61.3%	91.7%
RPS100075968 - 7	1	5007	483	96.1%	100.0%
RPS100075969 - 8	2	5007	483	77.6%	97.9%
RPS100075970 - 9	2	5007	483	75.8%	96.3%
RPS100075971 - 10	3	5007	483	50.1%	88.2%
RPS100075972 - 11	2	5007	483	67.7%	87.4%

*Test items reported in Table 10.11 are those that were later used in the Summer 2009 and Fall 2009 operational tests.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.12: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Spring 2009 Embedded Field Test Administration*

Content Area/ Item Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Perfect + Adjacent
English II					
Form 15	4	2018	221	91.0%	100.0%
Algebra I					
Forms 15 & 16	4	4945	543	83.6%	98.9%
Biology					
RPS100089159 - 12	1	10716	1178	98.6%	99.9%
RPS100089166 - 13	1	10716	1181	99.6%	100.0%
RPS100089165 - 14	4	10716	1175	82.3%	98.6%
RPS100089162 - 15	1	10716	1174	98.6%	99.9%
RPS100089167 - 16	1	10716	1163	97.0%	100.0%
RPS100089164 - 17	2	10716	1176	84.4%	99.3%
RPS100089171 - 18	1	10716	1164	96.2%	99.9%
RPS100089172 - 19	1	10716	1174	95.7%	99.8%
RPS100089160 - 12	1	2340	263	97.7%	99.2%
RPS100089161 - 13	2	2340	262	96.2%	98.9%
RPS100089169 - 14	1	2340	261	94.6%	100.0%
RPS100089163 - 15	1	2340	260	94.2%	100.0%
RPS100089168 - 16	2	2340	257	83.3%	99.2%
RPS100089173 - 17	1	2340	262	93.5%	100.0%
RPS100089174 - 18	3	2340	258	85.7%	95.0%
RPS100089175 - 19	1	2340	262	100.0%	100.0%

*Test items reported in Table 10.12 are those that were later used in the Summer 2009 and Fall 2009 operational tests.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.13: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Summer 2009 Operational Administration

Content Area/ Item Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Perfect + Adjacent
English II					
RPS100076784	4	864	85	91.8%	100.0%
Algebra I					
RPS100076683	4	1589	174	90.2%	100.0%
Biology					
RPS100076797 - 1	1	653	71	95.8%	100.0%
RPS100076798 - 2	1	653	72	90.3%	100.0%
RPS100075927 - 3	1	653	72	98.6%	100.0%
RPS100076807 - 4	2	653	73	98.6%	100.0%
RPS100076801 - 5	4	653	67	89.6%	95.5%
RPS100076803 - 6	2	653	72	93.1%	100.0%
RPS100076808 - 7	2	653	71	90.1%	98.6%
RPS100076802 - 8	2	653	70	92.9%	100.0%
RPS100076804 - 9	3	653	62	83.9%	96.8%
RPS100076805 -10	2	653	70	92.9%	100.0%

Table 10.14: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Fall 2009 Operational Administration

Content Area/ Item Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Perfect + Adjacent
English II					
RPS 100076785	4	1480	148	91.2%	100.0%
Algebra I					
RPS 100076622	4	2681	297	92.3%	98.7%
Biology					
RPS100075961 - 1	1	2259	248	95.2%	100.0%
RPS100075962 - 2	1	2259	246	98.8%	100.0%
RPS100075963 - 3	1	2259	249	98.8%	100.0%
RPS100075965 - 4	1	2259	248	95.2%	100.0%
RPS100075964 - 5	2	2259	249	98.0%	99.2%
RPS100075966 - 6	4	2259	241	85.5%	98.8%
RPS100075968 - 7	1	2259	252	98.0%	100.0%
RPS100075969 - 8	2	2259	248	91.5%	99.2%
RPS100075970 - 9	2	2259	248	94.8%	99.6%
RPS100075971 - 10	3	2259	249	91.2%	98.4%
RPS100075972 - 11	2	2259	251	91.2%	98.8%

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.15: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Fall 2009 Large Print/Braille Operational Administration

Content Area/ Item Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Perfect + Adjacent
English II					
RPS 100076789	N/A	N/A	N/A	N/A	N/A
Algebra I					
RPS 100076624	4	1	1	100.0%	100.0%
Biology					
RPS100075983 - 1	1	2	1	100.0%	100.0%
RPS100075984 - 2	1	2	1	100.0%	100.0%
RPS100075985 - 3	1	2	1	100.0%	100.0%
RPS100075986 - 4	3	2	1	100.0%	100.0%
RPS100075992 - 5	2	2	1	100.0%	100.0%
RPS100075987 - 6	3	2	1	100.0%	100.0%
RPS100075989 - 7	4	2	1	100.0%	100.0%
RPS100075988 - 8	3	2	1	100.0%	100.0%
RPS100075990 - 9	1	2	1	100.0%	100.0%
RPS100075991 - 10	1	2	1	100.0%	100.0%

Table 10.16: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Spring 2010 Operational Administration

Content Area/ Item Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Perfect + Adjacent
English II					
Form 15 - Spring 2010	4	63737	6916	94.2%	100.0%
Algebra I					
Forms 15 & 16	4	64297	7076	91.4%	99.5%
Biology					
RPS100089160 - 01	1	63662	7023	97.6%	99.9%
RPS100089166 - 02	1	63662	7054	99.3%	100.0%
RPS100089168 - 03	2	63662	6919	92.6%	99.8%
RPS100089161 - 04	2	63662	7046	98.1%	99.9%
RPS100089164 - 05	2	63662	6942	93.1%	99.9%
RPS100089165 - 06	4	63662	7037	84.9%	98.6%
RPS100089162 - 07	1	63662	7015	98.6%	99.8%
RPS100089167 - 08	1	63662	6988	96.4%	99.9%
RPS100089163 - 09	1	63662	7041	96.7%	100.0%
RPS100089172 - 10	1	63662	6948	93.9%	99.9%
RPS100089173 - 11	1	63662	7042	95.0%	100.0%
RPS100089174 - 12	3	63662	7036	90.1%	98.8%

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.17: Percentages of Exact and Adjacent Agreement Between Scorers' and Team Leaders' Validation Scores, Spring 2010 Large Print/Braille Operational Administration

Content Area/ Item Code	Total # of Points Possible	Total # of Items Scored	# of Items Validated	Perfect Agreement	Perfect + Adjacent
English II					
RPS 100076789	4	24	3	100.0%	100.0%
Algebra I					
RPS 100076624	4	24	3	100.0%	100.0%
Biology					
RPS100075983 - 01	1	26	3	100.0%	100.0%
RPS100075984 - 02	1	26	3	66.7%	100.0%
RPS100075985 - 03	1	26	3	100.0%	100.0%
RPS100075986 - 04	3	26	3	100.0%	100.0%
RPS100075992 - 05	2	26	3	100.0%	100.0%
RPS100075987 - 06	3	26	3	100.0%	100.0%
RPS100075989 - 07	4	26	3	100.0%	100.0%
RPS100075988 - 08	3	26	3	100.0%	100.0%
RPS100075990 - 09	1	26	3	100.0%	100.0%
RPS100075991 - 10	1	26	3	100.0%	100.0%

10.6 Reliability of Classifications

The reliability of student achievement-level classifications (i.e., Below Basic, Basic, Proficient, and Advanced) was evaluated using a computer program developed by Huynh (1979). This FORTRAN program is based on the beta-binomial model that also provides SEs for the consistency estimates. Classification consistency refers to the degree to which each student's achievement level can be replicated and is similar to the traditional test-retest or equivalent forms reliability. Using the maximum possible score, mean, standard deviation, and KR-21 reliability estimate, the program computes parameters (α, β) for the beta-binomial distribution. Kappa indices, which estimate the level of improvement in decision consistency beyond chance when test data are used, are then computed (Huynh, 1979).

Tables 10.18 and 10.19 show the results of the classification and decision consistency analyses for Summer 2009, Fall 2009, and Spring 2010 administrations for the three MO EOC Assessments. As noted above, the raw agreement index is a classification consistency index that estimates the percentage of examinees who would (hypothetically) be assigned to the same achievement level if the same test was administered a second time or an equivalent test was administered under the same conditions. The agreement consistency indices (p) for the EOC Assessments were generally in the mid 60s to lower 70s. These values reflect classification agreement consistency for the four performance categories: Below Basic, Basic, Proficient, and Advanced. Had a student been regarded as "pass" if his or her achievement level was Proficient or Advanced and as "fail" if his or her achievement level was Below Basic or Basic, the agreement consistency indices would have been 10 to 15 percent higher, as indicated in Tables 10.20 and 10.21. The latter classification accuracy is directly related to determining the accuracy of proficiency classifications for the No Child Left Behind Act (NCLB).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.18: Classification Consistency Coefficients

Year	*N* (Items)	Raw Cut Scores			Mean	*SD*	*Kappa*	*SE* (*κ*)
		Basic	Proficient	Advanced				
Summer 2009								
English II	39	15	24	33	22.21	6.18	0.43	0.0088
Algebra I	39	13	22	31	18.08	6.64	0.46	0.0070
Biology	55	18	32	45	26.78	9.63	0.55	0.0092
Fall 2009								
English II	39	15	24	33	23.02	6.40	0.44	0.0061
Algebra I	39	13	22	31	19.50	7.63	0.50	0.0047
Biology	55	18	32	45	32.33	11.46	0.61	0.0041
Spring 2010								
English II	39	15	24	33	27.65	6.18	0.47	0.0011
Algebra I	39	13	22	31	23.22	7.26	0.49	0.0010
Biology	55	18	32	45	31.55	9.87	0.56	0.0008

Table 10.19: Raw Agreement Consistency Coefficients

Year	*N* (Items)	Raw Cut Scores			Mean	*SD*	*p*	*SE* (*p*)
		Basic	Proficient	Advanced				
Summer 2009								
English II	39	15	24	33	22.21	6.18	0.64	0.0014
Algebra I	39	13	22	31	18.08	6.64	0.65	0.0021
Biology	55	18	32	45	26.78	9.63	0.72	0.0021
Fall 2009								
English II	39	15	24	33	23.02	6.40	0.64	0.0010
Algebra I	39	13	22	31	19.50	7.63	0.66	0.0018
Biology	55	18	32	45	32.33	11.46	0.73	0.0015
Spring 2010								
English II	39	15	24	33	27.65	6.18	0.67	0.0003
Algebra I	39	13	22	31	23.22	7.26	0.65	0.0003
Biology	55	18	32	45	31.55	9.87	0.71	0.0002

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 10.20: Classification Consistency Coefficients (Two Classification Categories)

Year	*N* (Items)	Raw Cut Scores	Mean	*SD*	*Kappa*	*SE* (*κ*)
		Proficient/ Advanced				
Summer 2009						
English II	39	24	22.21	6.18	0.57	0.0122
Algebra I	39	22	18.08	6.64	0.59	0.0089
Biology	55	32	26.78	9.63	0.67	0.0115
Fall 2009						
English II	39	24	23.02	6.40	0.59	0.0082
Algebra I	39	22	19.50	7.63	0.66	0.0051
Biology	55	32	32.33	11.46	0.75	0.0041
Spring 2010						
English II	39	24	27.65	6.18	0.60	0.0014
Algebra I	39	22	23.22	7.26	0.65	0.0011
Biology	55	32	31.55	9.87	0.69	0.0009

Table 10.21: Raw Agreement Consistency Coefficients (Two Classification Categories)

Year	*N* (Items)	Raw Cut Scores	Mean	*SD*	*p*	*SE* (*p*)
		Proficient/ Advanced				
Summer 2009						
English II	39	24	22.21	6.18	0.79	0.0059
Algebra I	39	22	18.08	6.64	0.82	0.0036
Biology	55	32	26.78	9.63	085	0.0048
Fall 2009						
English II	39	24	23.02	6.40	0.80	0.0041
Algebra I	39	22	19.50	7.63	0.84	0.0024
Biology	55	32	32.33	11.46	0.88	0.0020
Spring 2010						
English II	39	24	27.65	6.18	0.85	0.0005
Algebra I	39	22	23.22	7.26	0.83	0.0005
Biology	55	32	31.55	9.87	0.85	0.0005

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

CHAPTER 11: VALIDITY

11.1 Introduction

According to the *Standards* (AERA, APA, and NCME, 1999), "Ultimately, the validity of an intended interpretation of test scores relies on all the available evidence relevant to the technical quality of a testing program. This includes evidence of careful test construction; adequate score reliability; appropriate test administration and scoring; accurate score scaling, equating, and standard setting; and careful attention to fairness for all examinees" (p. 17). While this chapter summarizes evidence that supports claims about the validity of Missouri End-of-Course (MO EOC) Assessment scores, many other parts of this technical report also provide appropriate evidence for validity. Some of this evidence is cross-referenced below for added convenience. The procedural and empirical evidence available, along with the rationale presented below, provides support for the standards-based interpretations of the MO EOC Assessments.

This chapter begins with a brief review of important federal statutes that require the MO EOC Assessments and explains the purposes and intended uses of test scores, suggesting the value implications of these assessments for schools, teachers, students, and parents. Validity evidence related to test content is presented in terms of the adequacy and appropriateness of the EOC Assessments for measuring progress on the Missouri content standards. Then, validity evidence based on the internal structure of the MO EOC Assessments is provided through a correlational analysis of MO EOC Assessment content clusters. References to specific standards are provided where appropriate.

11.2 Federal Authority for School Accountability

The United States Department of Education bases accountability on a school's achievement of adequate yearly progress (AYP) in reading/language arts and mathematics. AYP determinations refer to the minimum improvement required of each school and district during the course of one year. For Missouri high schools and school districts, AYP is set in terms of the percentage of all students, and all student groups of sufficient size, scoring Proficient or above on the MO EOC tests in English II and Algebra I.

11.3 Purpose and Intended Uses of Test Scores

The *Standards* state that "Validation logically begins with an explicit statement of the proposed interpretation of the test scores, along with a rationale for the relevance of the interpretation to the proposed use" (AERA, APA, and NCME 1999).[14] The MO EOC Assessments were developed for the following purposes and uses:

- Measuring and reflecting students' mastery toward post-secondary readiness
- Identifying students' strengths and weaknesses
- Communicating expectations for all students

[14] **Standard 1.2:** The test developer should set forth clearly how test scores are intended to be interpreted and used. The population(s) for which a test is appropriate should be clearly delimited, and the construct that the test is intended to assess should be clearly described (p. 17).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Serving as the basis for state and national accountability plans
- Evaluating programs

The valid interpretation and appropriate use of MO EOC Assessment scores are supported in a variety of ways, including the training and consultation provided by personnel of the Missouri Department of Elementary and Secondary Education (DESE) and publications such as the *Test Examiner's Manual*, *Guide to Interpreting Results*, and this technical report. The training and documentation provided to test users help them better administer, understand, and use test score results.

11.4 MO EOC Assessment Scores

The MO EOC Assessment scores are scaled in several ways: raw-score points, item response theory (IRT) derived scale scores, and achievement level (based on scale-score cuts). Missouri actively promotes the use of achievement-level results, reporting them annually on each assessment at the student, school, district, and state levels. Individual student and average scale scores are also used, but they play a secondary role and are generally interpreted with reference to their distance from achievement-level cut points. Test results are reported for students as a whole as well as by student group, including gender, ethnicity, migrant status, free and reduced lunch (FRL) status, English language proficiency, Title I, Individualized Education Program (IEP) status, and accommodations used during testing. Scores are reported to schools and districts in annually published reports (for more information, see Chapter 8: Reporting).

The MO EOC Assessment score indicates that an individual student performs at the Below Basic, Basic, Proficient, or Advanced level in a given content area. Achievement-level descriptors provide details about the content expectations that students at each level meet or exceed. No stakes for teachers are attached to student-level scores by the state. Teachers are counseled to interpret individual student scores only in the context of other assessment results and their own experience.

11.5 Content-Related Evidence of Validity

Baker and Linn (2002) suggest that "Two questions are central in the evaluation of content aspects of validity. Is the definition of the content domain to be assessed adequate and appropriate? Does the test provide an adequate representation of the content domain the test is intended to measure?" (p. 6). The following sections help answer these two very important questions and also address Standard 1.6[15] of the *Standards* (AERA, APA, NCME, 1999), which specifically relates to the definition and development of test content.

[15] **Standard 1.6:** When the validation rests in part on the appropriateness of test content, the procedures followed in specifying and generating test content should be described and justified in reference to the construct the test is intended to measure or the domain it is intended to represent. If the definition of the content sampled incorporates criteria such as importance, frequency, or criticality, these criteria should also be clearly explained and justified (p. 18).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

11.5.1 Appropriateness of Content Definition

In 1993, the Missouri legislature passed the Outstanding Schools Act (Senate Bill 380), requiring the State Board of Education to adopt challenging academic performance standards that define the skills and competencies necessary for students to successfully advance through the public school system, prepare for post-secondary education and the workplace, and participate as citizens in a democratic society. The Missouri State Board of Education formally adopted the academic standards known as the Show-Me Standards in January 1996.

In addition to mandating the development of rigorous academic standards, the Outstanding Schools Act of 1993 required the development and implementation of a comprehensive, primarily performance-based assessment program to measure student proficiency in the knowledge, skills, and competencies identified in the standards. Upon adoption of the standards in 1996, Missouri began developing the Missouri Assessment Program (MAP).

In January 2007, the Missouri State Board of Education approved a plan to replace the MAP for high school students, beginning in August of the 2008–2009 school year, with EOC Assessments in English II, Algebra I, and Biology. The intent was to provide MO EOC Assessments that are an integral part of the statewide assessment system and, as such, are a logical extension of MAP tests at the elementary and middle grade levels.

11.5.2 Adequacy of Content Representation

Adequacy of the content representation of the MO EOC Assessments is critically important because the tests must provide an indication of student progress toward achieving the knowledge and skills identified in the Missouri Course-Level Expectations (CLEs), and they must fulfill the requirements of the No Child Left Behind Act.

Adequate representation of the content domains defined in the CLEs is assured through the use of a test blueprint and a carefully documented test construction process. CLEs and the Show-Me Standards are taken into consideration in the writing of selected response and Performance Event/Writing Prompt (PE/WP) items and in PE/WP rubric development. Each assessment must align with and proportionally represent the subdomains of the test blueprint. Evidence to support the content validity of the MO EOC Assessments was provided in Chapter 2: Test Development through the documentation of the test specifications and blueprints, item-writing processes, and item-review processes.

Additional evidence to support the content validity of the MO EOC Assessments was provided in Tables 2.1 through 2.3 in Chapter 2: Test Development and also in Chapter 4: Item Analysis. Chapter 2 outlined the target strand and CLE point distributions on the English II, Algebra I, and Biology operational forms. All forms administered in 2009–2010 met the point ranges specified in the blueprints. In addition, Riverside Publishing strove to equitably represent the strands on each assessment by balancing CLE and sub-CLE coverage according to the targets outlined in the test specifications and by matching item format to the requirements of the content and standards descriptions.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

11.6 Validity Evidence Based on the Internal Structure of the MO EOC Assessments

Standard 1.11[16] pertains to the relationships between the parts of the test. Because the MO EOC Assessments measure student performance in several content areas using a variety of item types, it is important to study the pattern of relationships among the content areas and item types (i.e., testing methods). One way to study patterns of relationships to provide evidence supporting the inferences made from test scores is the multitrait, multimethod matrix. Tables 11.1 through 11.3 summarize Pearson correlation coefficients among test domains and clusters for English II, Algebra I, and Biology. The correlations between clusters within each assessment are in the moderate to moderately high range, suggesting strong relationships between the clusters. Note that the high correlations between cluster scores and total assessment scores are inflated due to the overlap of items.

Table 11.1: Pearson Correlation Coefficients Between Domains and Clusters for English II

		Reading	Writing
Fall 2008	English II	0.97	0.80
	Reading	1.00	0.63
	Writing	0.63	1.00
Spring 2009	English II	0.98	0.73
	Reading	1.00	0.57
	Writing	0.57	1.00
Summer 2009	English II	0.98	0.65
	Reading	1.00	0.47
	Writing	0.47	1.00
Fall 2009	English II	0.98	0.76
	Reading	1.00	0.60
	Writing	0.60	1.00
Spring 2010	English II	0.98	0.75
	Reading	1.00	0.60
	Writing	0.60	1.00

[16] **Standard 1.11:** If the rationale for a test use or interpretation depends on premises about the relationships among parts of the test, evidence concerning the internal structure of the test should be provided (p. 20).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 11.2: Pearson Correlation Coefficients Between Domains and Clusters for Algebra I

		Number and Operations	Algebraic Relationships	Data and Probability
Fall 2008	Algebra I	0.81	0.95	0.80
	Number and Operations	1.00	0.66	0.56
	Algebraic Relationships	0.66	1.00	0.66
	Data and Probability	0.56	0.66	1.00
Spring 2009	Algebra I	0.82	0.95	0.79
	Number and Operations	1.00	0.66	0.58
	Algebraic Relationships	0.66	1.00	0.64
	Data and Probability	0.58	0.64	1.00
Summer 2009	Algebra I	0.79	0.95	0.73
	Number and Operations	1.00	0.64	0.45
	Algebraic Relationships	0.64	1.00	0.58
	Data and Probability	0.45	0.58	1.00
Fall 2009	Algebra I	0.82	0.94	0.80
	Number and Operations	1.00	0.66	0.58
	Algebraic Relationships	0.66	1.00	0.63
	Data and Probability	0.58	0.63	1.00
Spring 2010	Algebra I	0.79	0.96	0.76
	Number and Operations	1.00	0.64	0.52
	Algebraic Relationships	0.64	1.00	0.62
	Data and Probability	0.52	0.62	1.00

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Table 11.3: Pearson Correlation Coefficients Between Domains and Clusters for Biology

		Characteristics and Interactions	Changes in Ecosystems	Inquiry
Fall 2008	Biology	0.91	0.83	0.87
	Characteristics and Interactions	1.00	0.72	0.63
	Changes in Ecosystems	0.72	1.00	0.57
	Inquiry	0.63	0.57	1.00
Spring 2009	Biology	0.90	0.83	0.86
	Characteristics and Interactions	1.00	0.67	0.61
	Changes in Ecosystems	0.67	1.00	0.58
	Inquiry	0.61	0.58	1.00
Summer 2009	Biology	0.86	0.85	0.89
	Characteristics and Interactions	1.00	0.65	0.59
	Changes in Ecosystems	0.65	1.00	0.64
	Inquiry	0.59	0.64	1.00
Fall 2009	Biology	0.91	0.86	0.91
	Characteristics and Interactions	1.00	0.72	0.70
	Changes in Ecosystems	0.72	1.00	0.69
	Inquiry	0.70	0.69	1.00
Spring 2010	Biology	0.89	0.83	0.90
	Characteristics and Interactions	1.00	0.65	0.66
	Changes in Ecosystems	0.65	1.00	0.63
	Inquiry	0.66	0.63	1.00

11.7 Additional Validity Evidence for the MO EOC Assessments

Validity evidence related to other standards is described below.

Standard 1.5[17] relates to the characteristics of the sample of examinees from which validity evidence is inferred. The sample of examinees from which the validity evidence for the MO EOC Assessments was obtained is described in detail in Chapter 9: Summary Statistics, which includes tables with descriptive statistics for raw score, scale score, and achievement-level distributions. Statistics include *n*-counts, means, standard deviations, minimum and maximum values, and a variety of data disaggregations.

[17] **Standard 1.5:** The composition of any sample of examinees from which validity evidence is obtained should be described in as much detail as is practical, including major relevant sociodemographic and developmental characteristics (p. 18).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Standard 1.7[18] relates to human judgment at various points in the test development, scoring, and reporting process. For the MO EOC Assessments, human judgment was especially prevalent during the hand scoring of the PE/WP items and during the standard-setting process. Chapter 6: Scanning, Scoring, and Quality Control Procedures contains detailed information about the processes involved with Assessment Resource Center's hand scoring of the PE/WP items, including scorer selection, training, qualifications, and quality-control measures. Chapter 3: Achievement-Level Setting contains detailed information about the standard-setting procedures used for the MO EOC Assessments, including the selection process for and characteristics of the standard-setting participants.

Standard 1.13[19] relates to the conditions under which the data used to support validity claims were collected. Chapter 5: Test Administration contains information about how data were gathered in both the online and paper-and-pencil administrations, including the testing environment, materials distribution and security, Test Examiner training, student preparation, and allowable accommodations.

11.8 Summary

Validity is not an all-or-nothing property of a test; rather, validity evidence must be documented for a specific purpose and in the context of how the test scores will be interpreted and used. Much of the information contained in this technical report is, in and of itself, documentation of the validity of the MO EOC Assessments for their stated purpose. This chapter provides a summary of the evidence presented elsewhere in the manual and provides some additional types of validity evidence relevant to the content and internal structure of the assessments.

The overall technical quality of the EOC Assessments was sound. The Spring 2008 standalone field test and the Spring 2009 embedded field test produced pools of technically sound items, with a 91% retention rate after psychometric and content criteria were applied. From those pools, Riverside Publishing was able to assemble forms that were psychometrically very similar, and that similarity helped support the pre-equating model that is in place. Application of item response theory (IRT) pre-equating resulted in perfect or nearly congruent raw-to-scale score conversions between the Summer, Fall, and Spring forms at the proficiency level cuts.

[18] **Standard 1.7:** When a validation rests in part on the opinions or decisions of expert judges, observers, or raters, procedures for selecting such experts and for eliciting judgments or ratings should be fully described. The qualifications, and experience, of the judges should be presented. The description of procedures should include any training and instructions provided, should indicate whether participants reached their decisions independently, and should report the level of agreement reached. If participants interacted with one another or exchanged information, the procedures through which they may have influenced one another should be set forth (p. 19).

[19] **Standard 1.13:** When validity evidence includes statistical analyses of test results, either alone or together with data on other variables, the conditions under which the data were collected should be described in enough detail that users can judge the relevance of the statistical findings to local conditions. Attention should be drawn to any features of a validation data collection that are likely to differ from typical operational testing conditions and that could plausibly influence test performance (p. 20).

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Post-administration test analyses supported the technical quality of the MO EOC Assessments. Evaluations of IRT model assumptions supported the use of the Rasch model for all tests. Test reliabilities ranged from .81 to .91 across the content areas for the 2009–2010 test forms. Conditional standard errors of measurement (*CSEM*s) were between 5 and 7 scale score points at the cut scores. The item analyses also showed that the MO EOC Assessments have sound psychometric properties. The *p*-value ranges were sufficiently broad, indicating that the items do measure achievement across a broad range of difficulty. Most of the items had discrimination values > .15, and only nine items had a value < .10. Speededness was not a factor in students' test performance. Item bias analyses conducted on the pools further indicated that items were functioning equivalently for different gender and ethnic groups.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

References

American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME). 1999. *Standards for educational and psychological testing*. Washington, D.C.: American Psychological Association.

Baker, E. L., and R. L. Linn. 2002. *Validity issues for accountability systems*. Technical report 585. Los Angeles: Center for the Study of Evaluation.

Brown, W. 1910. Some experimental results in the correlation of mental abilities. *British Journal of Psychology* 3:296–322.

Council of Chief State School Officers. 2003. *Quality control checklist for processing, scoring, and reporting.* http://www.ccsso.org/Documents/2003/Quality_Control_Checklist_Scoring_2003.pdf (accessed November 5, 2009).

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

Dorans, N. J., and P. W. Holland. 1993. DIF detection and description: Mantel-Haenszel and standardization. In *Differential item functioning*, ed. P. W. Holland and H. Wainer, 35–66. Hillsdale, NJ: Erlbaum.

Feldt, L. S., and R. L. Brennan. 1989. Reliability. In *Educational measurement*, 3rd ed., ed. R. L. Linn, 105–146. New York: Macmillan.

Hambleton, R. K., and H. Swaminathan. 1985. *Item response theory: Principles and applications*. Hingham, MA: Kluwer-Nijhoff.

Huynh, H. 1979. Computational and statistical inference for two reliability indices based on the beta-binomial model. *Journal of Educational Statistics* 4:231–246.

Kolen, M. J., and R. L. Brennan. 2004. *Test equating, scaling, and linking: Methods and practices*, 2d ed. New York: Springer-Verlag.

Kuder, G. F., and M. W. Richardson. 1937. The theory of the estimation of test reliability. *Psychometrika* 2:151–160.

Linacre, J. M. 2006a. *A user's guide to WINSTEPS Rasch-model computer programs*. Chicago: Winsteps.

———. 2006b. WINSTEPS v3.64.

Linacre, J. M., and B. D. Wright. 1989. The equivalence of Rasch, PROX, and Mantel-Haenszel. *Transactions of the Rasch Measurement Special Interest Group* 3:1–3.

Liu, Y., and D. Mix. 2006. Categorizations of WINSTEPS DIF output. Paper presented at the American Educational Research Association conference, San Francisco, California.

Mantel, N., and W. Haenszel. 1959. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of National Cancer Institute* 22:719–748.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Masters, G. N. 1982. A Rasch model for partial credit scoring. *Psychometrika* 47:149–174.

Miller, G. E., O. Rotou, and J. S. Twing. 2004. Evaluation of the 0.3 logits screening criterion in common item equating. *Journal of Applied Measurement* 5(2):172–177.

Rasch, G. 1960. *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Schulz, E. M., C. Perlman, W. K. Rice, and B. D. Wright. 1996. An empirical comparison of Rasch and Mantel-Haenszel procedures for assessing differential item functioning. In *Objective measurement: Theory into practice*, Vol. 3, ed. G. Engelhard and M. Wilson, 65–84. Norwood, NJ: Ablex Publishing.

Smith R. M., and C. Y. Miao. 1994. Assessing unidimensionality for Rasch measurement. In *Objective measurement: Theory into practice*, Vol. 2, ed. M. Wilson, 316–328. Norwood, NJ: Ablex Publishing.

Thompson, S. J., C. J. Johnstone, and M.L. Thurlow. 2002. *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. http://education.umn.edu/NCEO/OnlinePubs/Synthesis44.html (accessed December 21, 2009).

Traub, R. E. 1994. *Reliability for the social sciences.* London: Sage Publications.

Wright, B. D., and Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8, 370.

Wright, B. D., and M. H. Stone. 1979. *Best test design*. Chicago: MESA Press.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX A: DEMOGRAPHIC CHARACTERISTICS OF STANDARD-SETTING PARTICIPANTS

Appendix Table A.1: English II

Region	District	Gender	Position	Ethnicity	% Minority*	% Free and Reduced Lunch*
Heart of Missouri	Jefferson City	F	Classroom Teacher	White	43%	25%
Heart of Missouri	Lincoln University	M	Non-Teacher Educator	White	NA	NA
Heart of Missouri	Not Available	M	Non-School	White	NA	NA
Kansas City	Blue Springs	F	Classroom Teacher	White	21%	19%
Kansas City	Lee's Summit	F	Classroom Teacher	White	13%	19%
Kansas City	North Kansas City	F	Classroom Teacher	White	41%	30%
Kansas City	Park Hill	M	Non-Teacher Educator	White	22%	22%
Northwest	St. Joseph	F	Classroom Teacher	White	56%	17%
Southeast	Jackson	F	Classroom Teacher	White	29%	4%
Southwest	Neosho	F	Classroom Teacher	White	56%	17%
St. Louis	Affton	F	Classroom Teacher	White	32%	16%
St. Louis	Rockwood	F	Classroom Teacher	White	13%	17%
West Central	Raymore-Peculiar	F	Classroom Teacher	White	21%	14%

*Percent minority and percent free and reduced lunch refers to the population of the district represented by the panelist. NA = Not available.

Appendix Table A.2: Algebra I

Region	District	Gender	Position	Ethnicity	% Minority*	% Free and Reduced Lunch*
Heart of Missouri	Jefferson City	M	Classroom Teacher	Asian/PI	43%	25%
Heart of Missouri	Keytesville	M	Non-Teacher Educator	White	46%	1%
Heart of Missouri	Moberly	F	Classroom Teacher	White	56%	13%
Kansas City	Center 58	F	Non-Teacher Educator	White	67%	86%
Kansas City	Kearney	M	Non-Teacher Educator	White	11%	4%
Kansas City	Lee's Summit	M	Classroom Teacher	White	13%	19%
Northwest	Hamilton	F	Classroom Teacher	NA	42%	4%
South Central	Saint Clair	F	Classroom Teacher	White	45%	3%
Southeast	North St. Francis County	F	Classroom Teacher	White	51%	2%
Southwest	Neosho	M	Classroom Teacher	White	56%	17%
Southwest	Nixa	F	Classroom Teacher	White	32%	7%
Southwest	Springfield	F	Non-Teacher Educator	White	45%	14%
St. Louis	Northwest	F	Classroom Teacher	White	34%	2%
St. Louis	Rockwood	F	Classroom Teacher	NA	13%	17%
West Central	Sherwood Cass	M	Classroom Teacher	White	47%	3%

*Percent minority and percent free and reduced lunch refers to the population of the district represented by the panelist. NA = Not available.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Appendix Table A.3: Biology

Region	District	Gender	Position	Ethnicity	% Minority*	% Free and Reduced Lunch*
Heart of Missouri	Fayette	F	Classroom Teacher	White	44%	13%
Heart of Missouri	Lincoln University	M	Non-Teacher Educator	White	NA	NA
Kansas City	Independence	M	Non-Teacher Educator	White	55%	25%
Kansas City	Kansas City	M	Classroom Teacher	Black	81%	85%
Northeast	North Shelby	F	Classroom Teacher	White	37%	0%
Northwest	Maryville	F	Classroom Teacher	White	29%	6%
Northwest	St. Joseph	M	Classroom Teacher	White	56%	17%
South Central	Maries County	M	Classroom Teacher	White	42%	2%
South Central	Waynesville	F	Classroom Teacher	Black	39%	39%
Southeast	Jackson	M	Classroom Teacher	White	29%	4%
Southwest	Branson	M	Classroom Teacher	White	48%	13%
Southwest	Carl Junction	F	Classroom Teacher	White	35%	6%
Southwest	Mansfield	F	Classroom Teacher	White	60%	6%
St. Louis	Clayton	M	Classroom Teacher	White	16%	24%
St. Louis	Ferguson-Florissant	M	Classroom Teacher	Asian/PI	60%	81%

*Percent minority and percent free and reduced lunch refers to the population of the district represented by the panelist. NA = Not available.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX B: STANDARD-SETTING SESSION AGENDA

Missouri EOC Achievement-Level Setting Agenda
Capitol Plaza Hotel and Convention Center
Jefferson City, Missouri—November 3–5, 2008

(*NOTE: Times are approximate.*)

Monday, November 3

Morning

7:30–8:30	Registration and Breakfast
8:30–9:15	Welcome, Introductions, Logistics, and Overview of Missouri's EOC Assessments (DESE)
9:15–9:35	Overview of the Standard-Setting Sessions (Questar Assessment)
9:35–10:00	Introduction to Achievement-Level Descriptors (ALDs) (Questar Assessment)
10:00–10:15	Break
10:15–11:15	Setting Performance Standards—General Process
11:15–12:15	"Experience" the Assessments
12:15–1:30	Lunch

Afternoon

1:30–3:15	Definitions and Description of Performance Standards
3:15–3:30	Break
3:30–4:30	Orientation to the Specific Standard-Setting Methodology
4:30–4:45	Questions and Dismissal for the Day

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Tuesday, November 4

Morning

7:30–8:30	Breakfast
8:30–9:15	Review of Day 1 Activities and Discussions
9:15–10:15	Preparation for Round 1 of Judgments
10:15–10:30	Break
10:30–12:00	Round 1 Judgments
12:00–1:15	Lunch

Afternoon

1:15–1:45	Review of Round 1 Issues and Problems
1:45–3:15	Feedback and Discussion of Round 1 Judgments
3:15–3:30	Break
3:30–3:45	Preparation for Round 2 Judgments
3:45–5:00	Round 2 Judgments

Wednesday, November 5

Morning

7:45–8:45	Breakfast
8:45–9:45	Review of Round 2 Judgments
9:45–10:00	Break
10:00–10:45	Preparation for Final Judgments
10:45–12:30	Final Round of Judgments and Evaluation
12:30–1:15	Lunch

Afternoon

1:15–2:15	Final review of ALDs and Session Wrap-up

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX C: OPENING SESSION POWERPOINT PRESENTATION

Standard Setting Overview

Missouri
Algebra I, Biology & English II
EOC Assessments

November, 2008

Setting Performance Standards

- *Who's Involved?* State and contractor roles
- *Why Questar?* Who's facilitating? Our role
- *Why you?* **Individually & collectively:**

  You are the *experts*.
  You *represent* various audiences.

Session Outline - Day 1

I. What is "standard setting" - in general and for the EOC Assessments?

II. Describe the performance "categories"; refine achievement level descriptors (ALDs)

III. Review & discuss the actual EOC test;

IV. The "Angoff procedure" – how it works

Setting Performance Standards

- *Who's Involved?* State and contractor roles
- *Why Questar?* Who's facilitating? Our role
- *Why you?* **Individually & collectively:**

  You are the *experts*.
  You *represent* various audiences.
  You are *judges*, not psychometricians.
  You are *advisors*, not policy makers

Setting Performance Standards

- *Who's Involved?* State and contractor roles
- *Why Questar?* Who's facilitating? Our role:
  Not content experts, but facilitators

Groundrules

NO DISCUSSIONS about the ***EOC*** *program or its underlying content standards*

OR

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Groundrules

NO DISCUSSIONS about the *EOC program*
OR

- why to set standards
- the philosophy of educational assessment
- why these particular tasks/assessments
- why a particular procedure is being used

What *IS* Standard Setting?

- another frame of reference to interpret test scores ("how good is *good*"?)
- a routine, daily activity

Groundrules

NO DISCUSSIONS about the *EOC program*
OR

- why to set standards
- the philosophy of educational assessment
- why these particular tasks/assessments
- why a particular procedure is being used

Confidentiality of all materials & discussions.

What *IS* Standard Setting?

- another frame of reference to interpret test scores ("how good is *good*"?)
- a routine, daily activity
- true "criterion-referencing"

Groundrules

NO DISCUSSIONS about the *EOC program*
OR

- why to set standards
- the philosophy of educational assessment
- why these particular tasks/assessments
- the fairness of assessing special students
- why a particular procedure is being used

Confidentiality of all materials & discussions.

All discussions should be *as a group*.

What *IS* Standard Setting?

- another frame of reference to interpret test scores ("how good is *good*"?)
- a routine, daily activity
- true "criterion-referencing"
- a *semi*-quantitative, *semi*-standardized, socio-political judgment process

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

What *IS* Standard Setting?

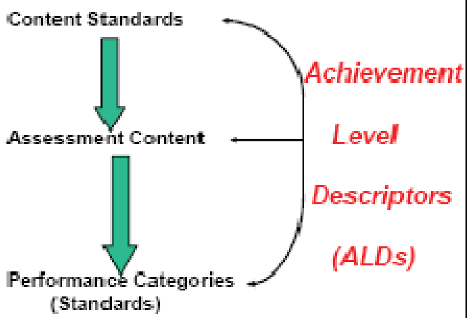
- just a frame of reference for test scores
- a routine, daily activity
- true "criterion-referencing"
- essentially, a judgment process

- **NOT** "science" !

Advice on Setting Standards

- **Set demanding, but *attainable* standards**

Critical Elements of An Assessment System



Advice on Setting Standards

- Set demanding, but *attainable* standards
- **What "*should be*" probably shouldn't disregard what "*is*"**

4 Keys to Being a Great Judge:

1. **Judgments** vs. **Data**
2. **"Should"** *vs.* **"Will"**
3. **Consider *ALL Missouri* students who took this EOC assessment**
4. **Think of *threshold* students, not *all* who are Proficient**

Advice on Setting Standards

- Set demanding, but *attainable* standards
- What "*should be*" probably shouldn't disregard what "*is*"
- **Focus on *concrete* behaviors, skills, responses**

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Advice on Setting Standards

- Set demanding, but *attainable* standards
- What "*should be*" probably shouldn't disregard what "*is*"
- Focus on *concrete* behaviors, skills, responses
- (*for M-C items*) Item difficulty resides in the answer choices, not the item "stem"

Missouri EOC Achievement-Level Setting

English II

End-of-Course Assessment

November 3—5, 2008

Advice on Setting Standards

- Set demanding, but *attainable* standards
- Don't disregard what "*is*"
- Focus on the *concrete*
- (*for M-C items*) Item difficulty resides in the answer choices, not the item "stem"
- (for constructed-response items) Judge the *response quality*, not the task difficulty.

"Housekeeping"

- Security Forms
- Judges' Numbers
- Break and lunch locations
- General agenda for the day

Advice on Setting Standards

- Set demanding, but *attainable* standards
- What "*should be*" shouldn't disregard what "*is*"
- Focus on the *concrete*
- Remember the type of item you're judging

Use your best judgment !!

Session Outline - Day 1

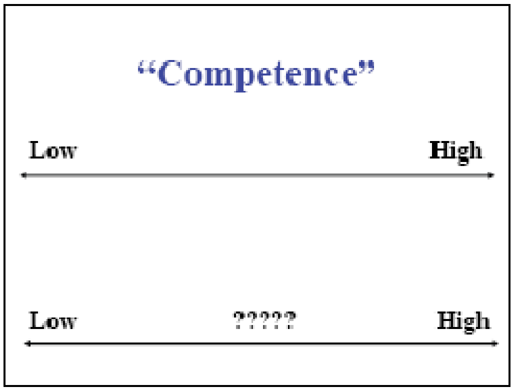
I. What is "standard setting" - in general, and for the EOC Assessments?

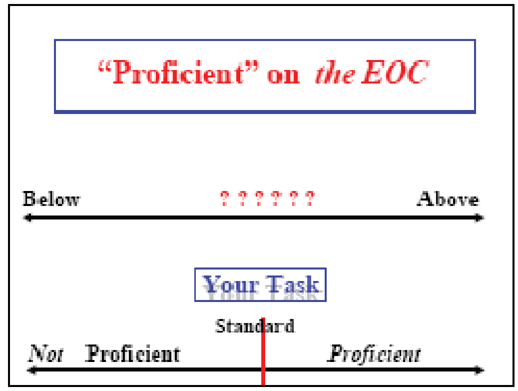
II. Describe the performance "categories"; refine Achievement Level Descriptors (ALDs)

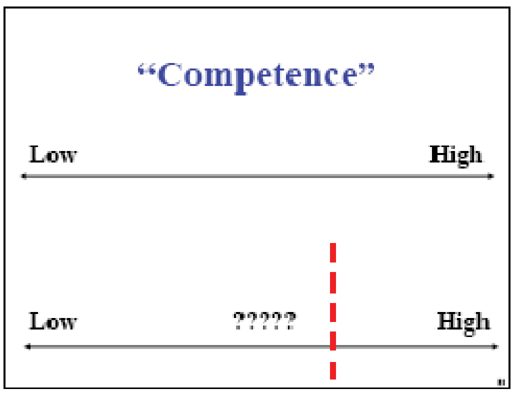
III. Review & discuss the actual EOC test

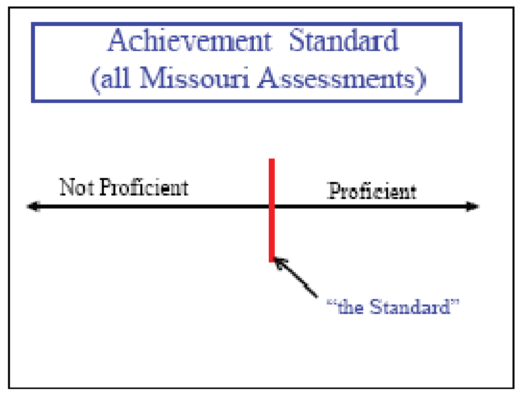
IV. The "Angoff procedure" – how it works

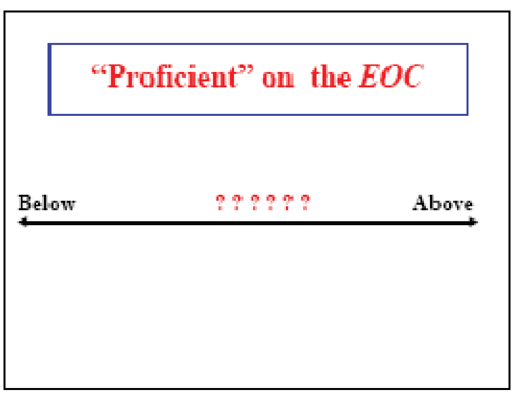
Copyright © 2010 by Missouri Department of Elementary and Secondary Education

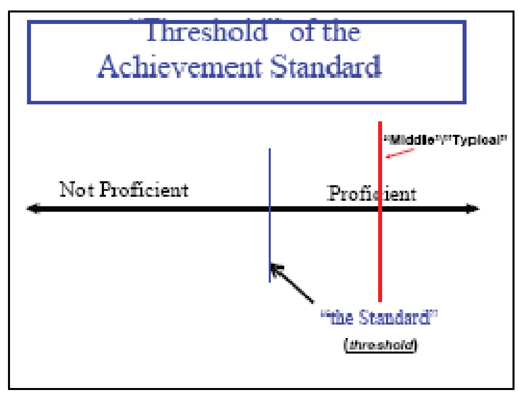






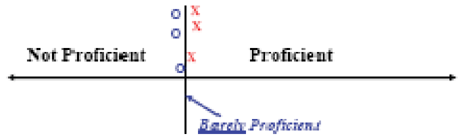






Copyright © 2010 by Missouri Department of Elementary and Secondary Education

“Proficient”



Are the Xs *really* better than the Os??

Achievement Level Descriptors ALDs

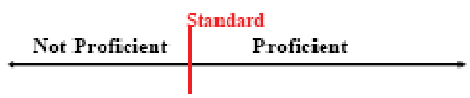
Start with the “labels”:

1. Below Basic
2. Basic
3. Proficient
4. Advanced

“Proficient” on the EOC

From the abstract:

To the concrete:



Problem:

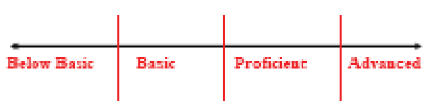
What do these *general* descriptions of achievement levels mean *concretely* for *EOC* students in each content area?

But that would be far too easy!

- So, we added *two more* cuts –

  Now we have:

  Below Basic   Proficient
  Basic   Advanced



Key Elements of the ALDs

- **Advanced**
  - Demonstrate thorough understanding
  - Demonstrate higher-level skills
  - Consistently apply a variety of strategies
- **Proficient**
  - Demonstrate understanding of skills and processes
  - Use a range of strategies
- **Basic**
  - Demonstrate incomplete understanding
  - Demonstrate skills inconsistently
  - Use some strategies
- **Below Basic**
  - Demonstrate little understanding
  - Demonstrate skills inconsistently or incorrectly
  - Use few strategies

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Don't Forget the Assessment !

Why? Standards are set on the actual EOC assessments, not *in general*

What to do? "Be" a student
Think about each item / task

Think about: Skill(s) / behaviors / expectations tapped
Basic, Proficient, Advanced
"Threshold" students

ASK: How well SHOULD a student who is JUST barely Proficient be able to do this?

Achievement Level Descriptors ALDs – Your Task

- Action verbs, e.g., determine, analyze, evaluate, utilize, identify, compare, describe, etc.
- Qualifiers, e.g., adjectives and adverbs that describe:
  - Differences in amount (most, some, few, etc.)
  - Degree of
    - understanding (thorough, partial, etc.)
    - frequency (consistently, rarely, etc.)
    - effectiveness (highly, moderately, somewhat)

Now that you've seen the "tasks"… let's debrief.

- Return to the general descriptors.
- Think about the tasks and items on the assessment.
- Which activities seemed to be hard (Advanced?) and easy (Basic?)?

  These are the grounding of your work to establish standards. The descriptions "define" the categories and should *anchor* your judgments.

Key Elements of the ALDs

- Advanced
  - Demonstrate thorough understanding
  - Demonstrate higher-level skills
  - Consistently apply a variety of strategies
- Proficient
  - Demonstrate understanding of skills and processes
  - Use a range of strategies
- Basic
  - Demonstrate incomplete understanding
  - Demonstrate skills inconsistently
  - Use some strategies
- Below Basic
  - Demonstrate little understanding
  - Demonstrate skills inconsistently or incorrectly
  - Use few strategies

Achievement Level Descriptors ALDs – Your Task

- ALDs probably should be *broader* than any specific assessment.
- ALDs should be *descriptive*, not definitional.
- ALDs "anchor" the standards, as they describe the behaviors of students whose performances "fit" each category

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Describe the assessed students *concretely*

Beginning with the *Proficient* category, describe the assessed students *concretely*.

What do they know? What can they do?
What skills do they possess in order to demonstrate this behavior?
What does the skill look like?
What are examples?
What behaviors/actions "fit" a certain category?

"Angoff" – What to *Do* ?

- Read each (*MC*) item in the test. Think about what is assessed/required.
- Conceptualize 100 "just barely" *Proficient* students all across the state who took this EOC.
- For each item, decide what percent of "barely Proficient" students <u>should</u> answer correctly.

"Angoff Procedure" for Setting Performance Standards

- *A* way, not *the* way to establish performance standards
- Recommended by the state's TAC
- Preferred procedure when statewide data are not available
- Requires judgments about each item on the assessment

"Angoff" – What to *Do* ?

- For each item, decide what percent of "barely **Proficient**" students should answer correctly.
- Repeat the decision for "barely **Advanced**" and "barely **Basic.**"
- After making the 3 judgments about an item, move to the next item.

"Angoff Procedure" for Setting Performance Standards

- For each test item, simply judge the percent of students in each performance category who should answer correctly.
- You can expect NO students to answer correctly, all students, or somewhere in between.
- In general – *maybe without exception?* – you should expect Basic students to perform less well than Proficient students, and less well yet than Advanced students.

"Angoff" – What to *Do* ?

- Read each (MC) item in the test. Think about what is assessed/required.
- Conceptualize 100 "just barely" **Proficient** students all across the state who took this EOC
- For each item, decide what percent of the "barely Proficient" students should answer correctly.
- Repeat the decision for "barely Basic" & "barely Advanced." Move to the next item.

- The 100 kids in each group aren't identical in skill/background/instruction and don't all know the same things, so the decision can't be all-or-none.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

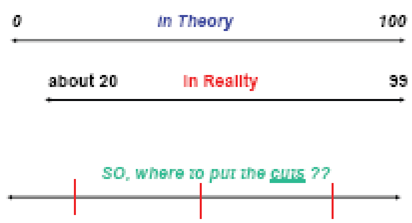
"Item Difficulty" Values

- The values you're going to work with are often termed "*p*-values."
- They're the proportion of students answering the item correctly.
- Let's look at what these values "mean."

Judgments for the Constructed-Response Item

- The *mechanics* differ, but the *intent* is the same – how should *borderline* students do?
- Think of the same 100 "threshold" students at each performance level.
- Decide what their AVERAGE SCORE should be on the constructed-response item rubric.

"Item Difficulty" Values



What About the Constructed-Response Item?

- Think of the same 100 "threshold" students at each performance level.
- Decide what their average score should be on the writing prompt (0-4).
- For example, perhaps you expect the following averages:

Basic	Proficient	Advanced
1.0	2.5	3.5

Judgments for the Constructed-Response Item

- The English II EOC assessment includes one constructed-response (CR) item – a writing prompt.
- Making your judgments for this item differs from the process used for the MC items.
- How to "judge" this item:

What About the Constructed-Response Item?

- Think of the same 100 "threshold" students at each performance level. Decide what their average score should be on the constructed-response item.
- E.g. perhaps you expect the following averages:

Basic	*Proficient*	*Advanced*
1.0	2.5	3.5

- Enter these averages on your Rating Sheet.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Session Overview - Day 2

I. Review Day 1 activities and outcomes.
II. Practice the Angoff procedure.
III. Round 1 of independent judgments
IV. Feedback & discussion of Round 1
V. Round 2 ratings -- *reconsider* Round 1

ISSUES:

Should / Ought

What *just* separates "Below Basic" from "Basic"?

Basic from "Proficient"?

Proficient from "Advanced"?

Threshold Students

All Assessed EOC Students in Missouri

Practice Activity: What to *Do*?

Think about:

The *item* – what's measured, intentionally or not?

The *curriculum* – Is this taught? Will it be?

The *performance category* – what does it mean?

Threshold students

How students *should* perform. What % should answer this item correctly

Record *three* judgments about each item.

Jot down any notes, questions, reactions as you work.

"Rules" for Ratings

- Anonymity
- Independence
- Don't perseverate -- Make a best guess.
- Find the "neighborhoods"; then refine.

Reminders for Round 1

- In a group of 100 students, all of whom are just *barely* Proficient, how many *should* answer this item correctly?
- Don't think of a "clearly Proficient" student. Focus on the cut score, and someone who *barely* makes it into the category. *Above* the cut, but *just* above – at the threshold.
- Remember the ALDs – they're your anchors.

Marking Your Judgments

- For MC items –
  - What percent of *barely* Proficient students should answer this item correctly?
- For CR items –
  - What should be the average score of *barely* Proficient students on this scale?
- Then, ask the same questions of *barely* Advanced and *barely* Basic.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Marking Your Judgments

- Record your Judge Number.
- For *MC Items*, bubble in 3 numbers per row/item.
  Just Proficient, just Advanced, just Basic.
- For *CR Item*, remember - judge *average* performance by students in each category.
- If you change your mind, erase completely.
- Double check your form before turning it in.

Are these data helpful?

- *Sorta*. What's the issue??
- Small samples. *More importantly*, from a field test. However, it's all we have.
  - Data tell how students *DID* perform.
  - Data CANNOT tell how students SHOULD perform, NOR how those who demonstrate a particular level of competence perform.

ISSUES:

Should / Ought
What *just* separates "Below Basic" from "Basic"?
Basic from "Proficient"?
Proficient from "Advanced"?
Threshold Students
All Assessed EOC Students in Missouri

Why Reratings?

- You are now a *different* judge.
- Consider the judgments & views of your peers.
- Goal: NOT "consensus," but *reflection*

YOU ARE NOW a better judge, because you are a *better-informed* judge.

Discuss the Preliminary Ratings

- *WHY* ????
- Hearing from your peers helps you to:
  - become more comfortable with your judgments -- both the *how* and *where*.
  - reconsider your earlier judgments.

Reratings: What to Do?

1. Reflect on earlier ratings – yours & peers.
2. Reflect on the discussions we have had.
3. Consider expanding the "zones" around your earlier judgments.
4. Reconsider each judgment –how well *should* a barely Basic/Proficient/Advanced student *do* on this item?

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Session Overview -- Day 3

I. Round 2 feedback & discussion

II. *Real* ratings & session evaluation

III. Final review of the ALDs

Thanks for your all your hard work, your patience, your many contributions, and your generous gift of time!

"How do I know if I'm *right*?"

- *There is no "right."*
- Did you keep in mind:
  1. "*Should*"?
  2. The *threshold* student?
  3. What "Below Basic," "Basic," "Proficient" & "Advanced" mean?
  4. *All* assessed students for this EOC?
  5. The discussions you've had?

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX D: DRAFT ALDS

Missouri End-of-Course Assessment Achievement-Level Descriptors—DRAFT

English II

Advanced: Students performing at the Advanced level on the Missouri-End-of-Course Assessment consistently demonstrate a thorough understanding of the skills and processes identified in the Course-Level Expectations for English II. They demonstrate higher level skills in reading processes, in responding to both fiction and nonfiction texts, and in writing effectively. In addition to understanding and applying the skills at the Proficient level, students scoring at the Advanced level use a range of strategies to comprehend and interpret a variety of texts, demonstrate a thorough understanding of literary forms, and consistently apply different strategies for accessing and summarizing information. They follow a writing process to compose well developed and organized papers for a variety of audiences and purposes while consistently and correctly applying the rules and conventions of Standard English.

Reading—In fiction and nonfiction, a student can

- Determine vocabulary meaning;
- Analyze the main idea and evaluate supporting details;
- Make sophisticated connections—compare, contrast, evaluate;
- Evaluate text features;
- Analyze complex figurative language and literary techniques;
- Draw insightful conclusions;
- Summarize and paraphrase ideas and information;
- Analyze story components and theme;
- Analyze literary elements;
- Evaluate reasoning, inferences, and sources;
- Evaluate proposed solutions;
- Evaluate accuracy and adequacy of evidence;
- Utilize organizational patterns;
- Evaluate the author's point of view, viewpoint/perspective, and/ or purpose;
- Evaluate the author's tone.

Writing

A student is able to write across genres a paper that

- Contains a strong controlling idea, along with an effective beginning, middle, and end;
- Uses paragraphing effectively;
- Progresses in a logical order and uses cohesive devices effectively;
- Addresses the topic clearly and provides specific and relevant details, reasons, and examples;
- Uses precise, vivid language in sentences that are clear and varied in structure;
- Effectively uses writing techniques;

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Shows complexity, freshness of thought, and individual perspective;
- Shows an awareness of audience and purpose;
- Contains few errors in Standard English and spelling.

A student is able to consistently and correctly apply the conventions of capitalization, punctuation, and standard usage.

Proficient: Students performing at the Proficient level on the Missouri-End-of-Course Assessment demonstrate an understanding of the skills and processes identified in the Course-Level Expectations for English II. They demonstrate these skills in reading processes, in responding to both fiction and nonfiction texts, and in writing effectively. In addition to understanding and applying the skills at the Basic level, students scoring at the Proficient level use a range of strategies to comprehend and interpret a variety of texts, demonstrate an understanding of literary forms, and apply strategies for accessing and summarizing information. They follow a writing process to compose well developed and organized papers for a variety of audiences and purposes while correctly applying the rules and conventions of Standard English.

Reading—In fiction and nonfiction, a student can

- Determine vocabulary meaning;
- Identify the main idea and supporting details;
- Make connections—compare, contrast, evaluate;
- Analyze text features;
- Analyze figurative language and literary techniques;
- Draw accurate conclusions;
- Summarize and paraphrase ideas and information;
- Analyze story components and theme;
- Analyze literary elements;
- Analyze reasoning, inferences, and sources;
- Analyze proposed solutions;
- Analyze evidence and use of information;
- Utilize organizational patterns;
- Analyze author's point of view, viewpoint/perspective, and/or purpose;
- Analyze the author's tone.

Writing

A student is able to write across genres a paper that

- Contains a controlling idea, along with a clear beginning, middle, and end;
- Uses paragraphing appropriately;
- Progresses in a generally logical order and uses cohesive devices;
- Addresses the topic and provides details, reasons, and examples;
- Uses precise language in sentences that are clear in structure;
- Uses writing techniques;
- Shows some complexity, freshness of thought, and/or individual perspective;
- Shows awareness of audience and purpose;
- Contains some errors in Standard English and spelling.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

A student is able to apply the conventions of capitalization, punctuation, and standard usage correctly.

Basic: Students performing at the Basic level on the Missouri-End-of-Course Assessment demonstrate an incomplete understanding of the skills and processes identified in the Course-Level Expectations for English II. They demonstrate these skills inconsistently in reading processes, in responding to both fiction and nonfiction texts, and in writing. In addition to understanding and applying the skills at the Below Basic level, students scoring at the Basic level use some strategies to comprehend and interpret a variety of texts, demonstrate a partial understanding of literary forms, and inconsistently apply few strategies for accessing and summarizing information. They may follow a writing process to compose papers while inconsistently applying the rules of Standard English.

Reading—In fiction and nonfiction, a student can

- Determine vocabulary meaning;
- Identify the main idea and major details;
- Make simple connections—compare, contrast;
- Identify text features;
- Identify figurative language and literary techniques;
- Draw simple conclusions;
- Summarize and paraphrase basic ideas and information;
- Identify characters, plot, setting, and basic theme;
- Identify basic literary elements;
- Make simple inferences;
- Identify proposed solutions;
- Determine reliability of information;
- Identify organizational patterns;
- Identify author's purpose; and point of view.

Writing

A student is able to write across genres a paper that

- Contains an idea, though it may lack focus, along with a beginning, middle, and end;
- Shows evidence of paragraphing;
- Progresses generally in a somewhat logical order and may use cohesive devices;
- Addresses the topic but relies on generalities rather than specifics;
- May use imprecise language in sentences that are generally clear in structure;
- May lack writing techniques;
- May lack complexity, freshness of thought, and individual perspective;
- Shows some awareness of audience and purpose;
- Contains errors in Standard English and spelling that may be distracting.

A student inconsistently applies the conventions of capitalization, punctuation, and standard usage.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Below Basic: Students performing at the Below Basic level on the Missouri-End-of-Course Assessment demonstrate little understanding of the skills and processes identified in the Course-Level Expectations for English II. They demonstrate these skills inconsistently and/or incorrectly in reading processes, in responding to both fiction and nonfiction texts, and in writing. Students scoring at the Below Basic level use few strategies to comprehend and interpret texts, demonstrate little understanding of literary forms, and apply few strategies for accessing information. They may not follow a writing process to compose papers and/or incorrectly apply the rules and conventions of Standard English.

Reading—In fiction and nonfiction, a student can

- Determine vocabulary meaning;
- Identify the main idea and some details;
- Make simple connections;
- Identify simple text features;
- Identify figurative language;
- Identify characters, plot and setting;
- Determine literal meaning;
- Identify point of view.

Writing

A student is able to write across genres a paper that

- May contain an unfocused idea and may lack a beginning, middle, and/or end;
- May lack evidence of paragraphing;
- Does not progress in a logical order and lacks cohesion;
- May address the topic but lacks details;
- May use imprecise language in sentences that may be unclear in structure;
- Shows little evidence of writing techniques;
- Lacks complexity, freshness of thought, and individual perspective;
- Shows little or no awareness of audience or purpose;
- Contains repeated errors in Standard English and spelling that are distracting.

A student incorrectly applies the conventions of capitalization, punctuation, and standard usage.

Algebra I

Advanced: Students performing at the Advanced level on the Missouri Algebra I End-of-Course Assessment demonstrate a thorough understanding of the Course-Level Expectations for Algebra I. They demonstrate these skills in algebraic relationships. In addition to understanding and applying the skills at the Proficient level, students scoring at the Advanced level use a wide range of strategies to solve problems and demonstrate a thorough understanding of important mathematical content and concepts.

Algebraic Relationships—Using algebraic relationships, a student can

- Generalize patterns using explicitly or recursively defined functions

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Describe the effects of parameter changes on exponential growth/decay and quadratic functions including intercepts
- Use symbolic algebra to represent and solve problems that involve quadratic relationships including equations and inequalities
- Describe and use algebraic manipulations, including factoring and apply properties of exponents to simplify expressions
- Use and solve equivalent forms of quadratic equations
- Use and solve systems of linear inequalities with 2 variables
- Analyze quadratic functions by investigating rates of change, intercepts, and zeros

Proficient: Students performing at the Proficient level on the Missouri Algebra I End-of-Course Assessment demonstrate an understanding of most Course-Level Expectations for Algebra I. They demonstrate these skills in number and operations, algebraic relationships, and data and probability. In addition to understanding and applying the skills at the Basic level, students scoring at the Proficient level use a range of strategies to solve problems and demonstrate understanding of important mathematical content and concepts.

Number and Operations—Using numbers and operations, a student can

- Compare and order rational and irrational numbers, including finding their approximate locations on a number line
- Use real numbers and various models, drawings, etc. to solve problems

Algebraic Relationships—Using algebraic relationships, a student can

- Generalize patterns using explicitly or recursively defined linear functions
- Compare and contrast various forms of representations of patterns
- Compare the properties of linear and nonlinear functions
- Describe the effects of parameter changes on linear functions including intercepts
- Use symbolic algebra to represent problems that involve linear relationships including equations and inequalities
- Describe and use algebraic manipulations, including rules of integer exponents to simplify expressions
- Use and solve equivalent forms of absolute value and linear equations
- Use and solve systems of linear equations with 2 variables
- Identify quantitative relationships and determine type(s) of functions that might model the situation to solve the problem
- Analyze linear functions by investigating rates of change, intercepts, and zeros

Data and Probability—Using data and probability, a student can

- Determine the distributions of the outcome of an experiment
- Use appropriate graphical representations of data
- Given one-variable quantitative data, display the distribution and describe its shape
- Apply statistical methods to measures of center to solve problems
- Given a scatterplot, determine an equation for a line of best fit

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Make conjectures about possible relationships between 2 characteristics of a sample on the basis of scatterplots of the data

Basic: Students performing at the Basic level on the Missouri Algebra I End-of-Course Assessment demonstrate an incomplete understanding of the Course-Level Expectations for Algebra I. They demonstrate these skills in number and operations, algebraic relationships, and data and probability. In addition to understanding and applying the skills at the Below Basic level, students scoring at the Basic level use some strategies to solve problems and demonstrate some understanding of important mathematical content and concepts.

Number and Operations—Using numbers and operations, a student can

- Compare and order rational numbers, including finding their approximate locations on a number line

Algebraic Relationships—Using algebraic relationships, a student can

- Generalize patterns using recursively defined single operation functions
- Compare the properties of linear functions
- Use symbolic algebra to solve problems that involve linear relationships including equations and inequalities
- Describe and use algebraic manipulations, including order of operations to simplify expressions
- Use equivalent forms of linear equations
- Use and solve systems of linear equations with 2 variables

Data and Probability—Using data and probability, a student can

- Formulate questions and collect data about a characteristic

Below Basic: Students performing at the Below Basic level on the Missouri Algebra I End-of-Course Assessment demonstrate a limited understanding of the Course-Level Expectations for Algebra I. They demonstrate these skills in number and operations, algebraic relationships, and data and probability. Students scoring at the Below Basic level use very few strategies to solve problems and demonstrate a limited understanding of important mathematical content and concepts.

Number and Operations—Using numbers and operations, a student can

- Compare and order rational numbers

Algebraic Relationships—Using algebraic relationships, a student can

- Identify a function as linear or nonlinear
- Use symbolic algebra to solve problems that involve 2 step linear equations

Data and Probability—Using data and probability, a student can

- Identify the sample space of an experiment
- Select appropriate graphical representation of data

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Determine measures of center

Biology

Advanced: Students performing at the Advanced level on the Missouri End-of-Course Assessment consistently demonstrate a thorough understanding of the Course-Level Expectations for Biology. They demonstrate these skills in …

In addition to understanding and applying the skills at the Proficient level, students scoring at the Advanced level use a range of strategies to …

Characteristics and Interactions of Living Organisms—A student can

- List of skills here
- Apply the law of conservation of mass and energy to a biochemical process
- Classify different ways to store energy and describe the transfer of energy in a food web
- Relate structure of organic compounds to their role in living systems
- Predict the movement of molecules across a selectively permeable membrane needed for a cell to maintain homeostasis
- Compare and contrast process used in movement of molecules across a semipermeable membrane—taking energy use into consideration
- Predict patterns of inheritance using Mendelian genetics, including sex-linked, in a monohybrid cross
- Relate the expression of genetic diseases in offspring to the genetic makeup of the parents

Changes in Ecosystems and Interactions of Organisms with Their Environments—A student can

- List of skills here
- Predict how populations within an ecosystem may change in response to changes in abiotic or biotic factors
- Predict the impact of changes within a food chain on energy use and flow
- Explain how natural selection is related to environmental changes or species adaptations
- Predict local and global effects on environmental resources when given a scenario describing natural phenomena

Scientific Inquiry—A student can

- List of skills here
- Use quantitative data to calculate results
- Communicate information from investigations in data tables and appropriate graphical forms
- Identify and justify constants and variables in a repeatable scientific investigation
- Design a repeatable multi-step scientific investigation
- Gather evidence in qualitative and quantitative forms
- Determine how technological advances can affect real-world situations

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Proficient: Students performing at the Proficient level on the Missouri End-of-Course Assessment demonstrate an understanding of the Course-Level Expectations for Biology. They demonstrate these skills in …

In addition to understanding and applying the skills at the Basic level, students scoring at the Proficient level use a range of strategies to…

Characteristics and Interactions of Living Organisms—A student can

- List of skills here
- Explain cell differentiation
- Explain the chemical and physical interactions between organelles as they carry out life processes
- Explain interrelationships between photosynthesis and respiration
- Determine factors that affect the processes of photosynthesis and respiration
- Explain how enzymes affect chemical reactions
- Explain homeostasis and its effect on cellular activities
- Identify the causes of mutations in DNA and explain the possible effects on the organism
- Describe transcription and translation in DNA and identify steps in the processes of mitosis and meiosis
- Explain the advantages and disadvantages of sexual and asexual reproduction within a population
- Describe diploid and haploid chromosome number
- Explain how daughter cells compare to the original parent cell (heredity information and number)
- Describe how new genetic combinations result in new heritable characteristics
- Explain how genotypes contribute to phenotypic variation within a species

Changes in Ecosystems and Interactions of Organisms with Their Environments—A student can

- List of skills here
- Identify and explain limiting factors (abiotic and biotic) that may affect carrying capacity
- Explain the impact a natural environmental event may have on the diversity of different species in an ecosystem
- Explain the impact human activity may have on the diversity of different species in an ecosystem
- Describe energy flow in a food web
- Explain the natural and/or human factors that may lead to the extinction of a species
- Identify the evidence found in the fossil records to support relationship among species over time

Scientific Inquiry—A student can

- List of skills here

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Formulate a testable hypothesis
- Identify constants and variables in an investigation
- Determine scientific conclusions based on observations
- Use patterns to extrapolate data to form conclusions
- Identify factors required to make investigative results reliable
- Analyze quantitative data
- Design scientific investigations consisting of at least three steps
- Identify technology used to collect data to increase scientific knowledge
- Explain why accurate records and replications are essential for experimental creditability
- Calculate percent and ratios from sets of data
- Communicate procedures and results of investigations
- Explain the importance of peer review of scientific findings

Basic: Students performing at the Basic level on the Missouri End-of-Course Assessment demonstrate an incomplete understanding of the Course-Level Expectations for Biology. They demonstrate these skills inconsistently in …

In addition to understanding and applying the skills at the Below Basic level, students scoring at the Basic level use some strategies to …

Characteristics and Interactions of Living Organisms—A student can

- List of skills here
- Identify and describe cell structures and functions
- Define organelles by their functions
- Describe the equation for photosynthesis and respiration
- Identify that the carbon that organisms use for growth comes from the carbon dioxide in the air (this probably needs a better word than growth, but this is a huge misconception that needs to be addressed)
- Explain how water is important to cells
- Use a Punnett square to show a simple monohybrid cross

Changes in Ecosystems and Interactions of Organisms with Their Environments—A student can

- List of skills here
- Describe interactions between organisms in a predator/prey relationship
- Explain how interactions within an ecosystem maintain balance
- Define carrying capacity of a population within an ecosystem
- Describe how a natural environmental event impacts diversity in an ecosystem
- Describe how human caused change impacts the diversity in an ecosystem
- Construct a simple food web
- Define species in terms of the ability to mate and reproduce
- Describe similarities in DNA between species
- Describe how adaptations may have provided a population an advantage for survival

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Explain how environmental factors can be agents of natural selection

Scientific Inquiry—A student can

- List of skills here
- Select appropriate investigation methods
- Use data to formulate an explanation
- Calculate average/mean for sets of data
- Identify possible effects of errors in data collection and calculations
- Identify and describe how scientific explanations have changed over time or as a result of new evidence (strand 8?)

Below Basic: Students performing at the Below Basic level on the Missouri End-of-Course Assessment demonstrate little understanding of the Course-Level Expectations for Biology. They demonstrate these skills inconsistently and/or incorrectly in …

Students scoring at the Below Basic level inconsistently use some strategies to …

Characteristics and Interactions of Living Organisms—A student can

- List of skills here
- Identify that all organisms progress through life cycles
- Identify that all organisms are made of cells
- Identify that water is important to cells (life?)
- Identify that all living organisms have DNA
- Identify that DNA carries inherited information

Changes in Ecosystems and Interactions of Organisms with Their Environments—A student can

- List of skills here
- Use a model to show that populations interact in an ecosystem
- Identify examples of adaptations resulting from natural selection

Scientific Inquiry—A student can

- List of skills here
- Identify a valid conclusion in an experiment
- Use simple tools to measure length, mass, and volume
- Communicate basic information from an experiment
- Construct a simple graph of independent variable versus dependent variable from given data
- Identify how humans impact the environment (strand 8)
- Identify one impact of technology on an environmental factor (also strand 8)

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Appendix E: Final ALDs

Missouri End-of-Course Assessment Achievement-Level Descriptors—FINAL

English II

Achievement Levels

Advanced: Students performing at the Advanced level on the Missouri English II End-of-Course Assessment consistently demonstrate a thorough understanding of the skills and processes identified in the Course-Level Expectations for English II. They demonstrate higher-level skills in reading processes, in responding to both fiction and nonfiction texts, and in writing effectively. In addition to understanding and applying the skills at the Proficient level, students scoring at the Advanced level use a wide range of strategies to comprehend and interpret a variety of texts, demonstrate a thorough understanding of literary forms, and consistently apply different strategies for accessing and summarizing information. They follow a writing process to compose well-developed and organized papers for a variety of audiences and purposes, while consistently and correctly applying the rules and conventions of Standard English. Raw Score Cut: 33–39; Scale Score Cut: 225–250.

Proficient: Students performing at the Proficient level on the Missouri English II End-of-Course Assessment demonstrate an understanding of the skills and processes identified in the Course-Level Expectations for English II. They demonstrate these skills in reading processes, in responding to both fiction and nonfiction texts, and in writing effectively. In addition to understanding and applying the skills at the Basic level, students scoring at the Proficient level use a range of strategies to comprehend and interpret a variety of texts, demonstrate an understanding of literary forms, and apply strategies for accessing and summarizing information. They follow a writing process to compose well-developed and organized papers for a variety of audiences and purposes, while correctly applying the rules and conventions of Standard English. Raw Score Cut: 24–32; Scale Score Cut: 200–224.

Basic: Students performing at the Basic level on the Missouri English II End-of-Course Assessment demonstrate an incomplete understanding of the skills and processes identified in the Course-Level Expectations for English II. They demonstrate these skills inconsistently in reading processes, in responding to both fiction and nonfiction texts, and in writing. In addition to understanding and applying the skills at the Below Basic level, students scoring at the Basic level use some strategies to comprehend and interpret a variety of texts, demonstrate a partial understanding of literary forms, and inconsistently apply few strategies for accessing and summarizing information. They may follow a writing process to compose papers while inconsistently applying the rules and conventions of Standard English. Raw Score Cut: 15–23; Scale Score Cut: To be determined after operational data are complete.

Below Basic: Students performing at the Below Basic level on the Missouri English II End-of-Course Assessment demonstrate little understanding of the skills and processes identified in the Course-Level Expectations for English II. They demonstrate these skills inconsistently and/or incorrectly in reading processes, in responding to both fiction and

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

nonfiction texts, and in writing. Students scoring at the Below Basic level use few strategies to comprehend and interpret texts, demonstrate little understanding of literary forms, and apply few strategies for accessing information. They may not follow a writing process to compose papers and/or incorrectly apply the rules and conventions of Standard English. Raw Score Cut: 0–14; Scale Score Cut: To be determined after operational data are complete.

Achievement Descriptors

Advanced

Raw Score Cut: 33–39; Scale Score Cut: 225–250

Reading—In both fiction and nonfiction, a student can

- Determine vocabulary meaning
- Analyze the main idea and evaluate supporting details
- Make sophisticated connections—compare, contrast, evaluate
- Evaluate text features
- Analyze complex figurative language and literary techniques
- Draw insightful conclusions
- Summarize and paraphrase complex ideas and information
- Analyze literary elements
- Evaluate reasoning, inferences, and sources
- Evaluate proposed solutions
- Evaluate accuracy and adequacy of evidence
- Evaluate organizational patterns
- Evaluate the author's point of view, viewpoint/perspective, and purpose
- Evaluate the author's tone

Writing—A student is able to write across genres a paper that

- Contains a strong controlling idea, along with an effective beginning, middle, and end
- Uses paragraphing effectively
- Progresses in a logical order and uses cohesive devices effectively
- Addresses the topic clearly and provides specific and relevant details, reasons, and examples
- Uses precise, vivid language in sentences that are clear and varied in structure
- Effectively uses writing techniques
- Shows complexity, freshness of thought, and individual perspective
- Shows a clear awareness of audience and purpose
- Contains few errors in Standard English and spelling

A student is able to consistently and correctly apply the conventions of capitalization, punctuation, and standard usage.

Proficient

Raw Score Cut: 24–32; Scale Score Cut: 200–224

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Reading—In both fiction and nonfiction, a student can

- Determine vocabulary meaning
- Identify the main idea and supporting details
- Make connections—compare, contrast, analyze
- Analyze text features
- Analyze figurative language and literary techniques
- Draw accurate conclusions
- Summarize and paraphrase ideas and information
- Analyze literary elements
- Analyze reasoning, inferences, and sources
- Analyze proposed solutions
- Analyze evidence and use of information
- Analyze organizational patterns
- Analyze the author's point of view, viewpoint/perspective, and purpose
- Analyze the author's tone

***Writing*—**A student is able to write across genres a paper that

- Contains a controlling idea, along with a clear beginning, middle, and end
- Uses paragraphing appropriately
- Progresses in a generally logical order and uses cohesive devices
- Addresses the topic and provides details, reasons, and examples
- Uses precise language in sentences that are clear and show some variety in structure
- Uses writing techniques
- Shows some complexity, freshness of thought, and/or individual perspective
- Shows awareness of audience and purpose
- Contains some errors in Standard English and spelling

A student is able to apply the conventions of capitalization, punctuation, and standard usage correctly.

Basic

Raw Score Cut: 15–23; Scale Score Cut: To be determined after operational data are complete.

Reading—In fiction and nonfiction, a student can

- Determine vocabulary meaning
- Identify the main idea and major details
- Make simple connections—compare, contrast
- Identify text features
- Identify figurative language and literary techniques
- Draw basic/simple conclusions
- Summarize and paraphrase basic ideas and information
- Identify basic literary elements

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Make simple inferences
- Identify proposed solutions
- Determine reliability of information
- Identify organizational patterns
- Identify author's purpose and point of view
- Identify author's tone

***Writing*—**A student is able to write across genres a paper that

- Contains an idea, though it may lack focus, along with a beginning, middle, and end
- Shows evidence of paragraphing
- Progresses generally in a somewhat logical order and may use cohesive devices
- Addresses the topic but relies on generalities rather than specifics
- May use imprecise language in sentences that are generally clear in structure
- May lack writing techniques
- May lack complexity, freshness of thought, and individual perspective
- Shows some awareness of audience and purpose
- Contains errors in Standard English and spelling that may be distracting

A student inconsistently applies the conventions of capitalization, punctuation, and standard usage.

Below Basic

Raw Score Cut: 0–14; Scale Score Cut: To be determined after operational data are complete.

***Readin*g—**In fiction and nonfiction, a student can

- Determine vocabulary meaning
- Identify the main idea and some details
- Make simple connections
- Identify simple text features
- Identify figurative language
- Identify characters, plot, and setting
- Determine literal meaning
- Identify point of view

***Writing*—**A student is able to write across genres a paper that

- May contain an unfocused idea and may lack a beginning, middle, and/or end
- May lack evidence of paragraphing
- Does not progress in a logical order and lacks cohesion
- May address the topic but lacks details
- May use imprecise language in sentences that may be unclear in structure
- Shows little evidence of writing techniques
- Lacks complexity, freshness of thought, and individual perspective
- Shows little or no awareness of audience or purpose
- Contains repeated errors in Standard English and spelling that are distracting

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

A student incorrectly applies the conventions of capitalization, punctuation, and standard usage.

Algebra I

Achievement Levels

Advanced: Students performing at the Advanced level on the Missouri Algebra I End-of-Course Assessment demonstrate a thorough understanding of the Course-Level Expectations for Algebra I. They demonstrate these skills in number and operations, algebraic relationships, and data and probability. In addition to understanding and applying the skills at the Proficient level, students scoring at the Advanced level use a wide range of strategies to solve problems and demonstrate a thorough understanding of important mathematical content and concepts. Raw Score Cut: 31–39; Scale Score Cut: 225–250

Proficient: Students performing at the Proficient level on the Missouri Algebra I End-of-Course Assessment demonstrate an understanding of most Course-Level Expectations for Algebra I. They demonstrate these skills in number and operations, algebraic relationships, and data and probability. In addition to understanding and applying the skills at the Basic level, students scoring at the Proficient level use a range of strategies to solve problems and demonstrate an understanding of important mathematical content and concepts. Raw Score Cut: 22–30; Scale Score Cut: 200–224

Basic: Students performing at the Basic level on the Missouri Algebra I End-of-Course Assessment demonstrate some understanding of the Course-Level Expectations for Algebra I. They demonstrate these skills in number and operations, algebraic relationships, and data and probability. In addition to understanding and applying the skills at the Below Basic level, students scoring at the Basic level use some strategies to solve problems and demonstrate some understanding of important mathematical content and concepts. Raw Score Cut: 13–21; Scale Score Cut: To be determined after operational data are complete.

Below Basic: Students performing at the Below Basic level on the Missouri Algebra I End-of-Course Assessment demonstrate a limited understanding of the Course-Level Expectations for Algebra I. They demonstrate these skills in number and operations, algebraic relationships, and data and probability. In addition, students scoring at the Below Basic level use very few strategies to solve problems and demonstrate a limited understanding of important mathematical content and concepts. Raw Score Cut: 0–12; Scale Score Cut: To be determined after operational data are complete.

Achievement Descriptors

Advanced

Raw Score Cut: 31–39; Scale Score Cut: 225–250

Algebraic Relationships—Using algebraic relationships, a student can

- Generalize patterns using explicitly or recursively defined functions
- Describe the effects of parameter changes on exponential growth/decay and quadratic functions, including intercepts

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Use symbolic algebra to represent and solve problems that involve quadratic relationships, including equations and inequalities
- Describe and use algebraic manipulations, including factoring, and apply properties of exponents to simplify expressions
- Use and solve equivalent forms of quadratic and absolute value equations
- Identify quantitative relationships and determine type(s) of functions that might model the situation to solve a problem, including quadratic and exponential growth/decay
- Use and solve systems of linear inequalities with two variables
- Analyze quadratic functions by investigating rates of change, intercepts, and zeros

Proficient

Raw Score Cut: 22–30; Scale Score Cut: 200–224

Number and Operations—Using numbers and operations, a student can

- Compare and order rational and irrational numbers, including finding their approximate locations on a number line
- Use real numbers and various models, drawings, etc. to solve problems

Algebraic Relationships—Using algebraic relationships, a student can

- Generalize patterns using explicitly or recursively defined linear functions
- Compare and contrast various forms of representations of patterns
- Compare and contrast the properties of linear and nonlinear functions
- Describe the effects of parameter changes on linear functions, including intercepts
- Use symbolic algebra to represent problems that involve linear relationships, including equations and inequalities
- Describe and use algebraic manipulations, including rules of integer exponents, to simplify expressions
- Use and solve equivalent forms of absolute value and linear equations
- Use and solve systems of linear equations with two variables
- Identify quantitative relationships that can be modeled by linear functions to solve a problem
- Analyze linear functions by investigating rates of change, intercepts, and zeros

Data and Probability—Using data and probability, a student can

- Use appropriate graphical representations of data
- Given one-variable quantitative data, display the distribution and describe its shape
- Apply statistical methods to measures of center to solve problems
- Given a scatterplot, determine an equation for a line of best fit
- Make conjectures about possible relationships between two characteristics of a sample on the basis of scatterplots of the data

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Basic

Raw Score Cut: 13–21; Scale Score Cut: To be determined after operational data are complete.

Number and Operations—Using numbers and operations, a student can

- Compare and order rational numbers, including finding their approximate locations on a number line

***Algebraic Relationships*—**Using algebraic relationships, a student can

- Generalize patterns using recursively defined single-operation functions
- Compare the properties of linear functions
- Use symbolic algebra to solve problems that involve linear relationships, including equations and inequalities
- Describe and use algebraic manipulations, including order of operations, to simplify expressions
- Use equivalent forms of linear equations

***Data and Probability*—**Using data and probability, a student can

- Determine the sample space of an experiment
- Formulate questions about a characteristic which include sample spaces and distributions

Below Basic

Raw Score Cut: 0–12; Scale Score Cut: To be determined after operational data are complete.

Number and Operations—Using numbers and operations, a student can

- Compare and order rational numbers

Algebraic Relationships—Using algebraic relationships, a student can

- Identify a function as linear or nonlinear
- Use symbolic algebra to solve problems that involve two-step linear equations

***Data and Probability*—**Using data and probability, a student can

- Identify the sample space of an experiment
- Select appropriate graphical representations of data
- Determine measures of center

Biology

Achievement Levels

Advanced: Students performing at the Advanced level on the Missouri End-of-Course Assessment demonstrate a thorough understanding of the Course-Level Expectations for Biology. They demonstrate these skills in addition to understanding and applying the

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

skills at the Proficient level; students scoring at the Advanced level use a range of strategies. Raw Score Cut: 45–55; Scale Score Cut: 225–250

Proficient: Students performing at the Proficient level on the Missouri End-of-Course Assessment demonstrate an understanding of the Course-Level Expectations for Biology. They demonstrate these skills in addition to understanding and applying the skills at the Basic level; students scoring at the Proficient level use a range of strategies. Raw Score Cut: 32–44; Scale Score Cut: 200–224

Basic: Students performing at the Basic level on the Missouri End-of-Course Assessment demonstrate a partial understanding of the Course-Level Expectations for Biology. They demonstrate these skills in addition to understanding and applying the skills at the Below Basic level; students scoring at the Basic level use some strategies. Raw Score Cut: 18–31; Scale Score Cut: To be determined after operational data are complete.

Below Basic: Students performing at the Below Basic level on the Missouri End-of-Course Assessment demonstrate a limited understanding of the Course-Level Expectations for Biology. Students scoring at the Below Basic level use very few strategies and demonstrate a limited understanding of important Biological content and concepts. Raw Score Cut: 0–17; Scale Score Cut: To be determined after operational data are complete.

Achievement Descriptors

Advanced

Raw Score Cut: 45–55; Scale Score Cut: 225–250

Characteristics and Interactions of Living Organisms—A student can

- Predict the movement of molecules across a selectively permeable membrane needed for a cell to maintain homeostasis
- Compare and contrast process used in movement of molecules across a semipermeable membrane, taking energy use into consideration
- Predict patterns of inheritance, using Mendelian genetics, in a monohybrid cross

Changes in Ecosystems and Interactions of Organisms with Their Environments—A student can

- Predict how populations within an ecosystem may change in response to changes in abiotic or biotic factors
- Predict the impact of changes within in a food chain based on energy use and flow
- Explain how natural selection is related to environmental changes or species adaptations

Scientific Inquiry—A student can

- Use quantitative data to calculate results
- Communicate information from investigations in data tables and appropriate graphical forms
- Identify and justify constants and variables in a repeatable scientific investigation
- Design a repeatable multistep scientific investigation

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Recognize it is not always possible, for practical or ethical reasons, to control some conditions (e.g., when sampling or testing humans, when observing animal behaviors in nature)

Proficient

Raw Score Cut: 32–44; Scale Score Cut: 200–224

Characteristics and Interactions of Living Organisms—A student can

- Identify cell differentiation
- Explain the chemical and physical interactions between organelles as they carry out life processes
- Explain interrelationships between photosynthesis and respiration (reactant and product only)
- Determine factors that affect the processes of photosynthesis and respiration (excludes light intensity)
- Identify homeostasis and its effect on cellular activities
- Identify the causes of mutations in DNA and explain the possible effects on the organism
- Describe the chemical and structural properties of DNA
- Recognize that DNA codes for proteins, which are expressed as the heritable characteristics of an organism
- Compare the processes of mitosis and meiosis (excludes identification of steps)
- Explain the advantages and disadvantages of sexual and asexual reproduction within a population
- Identify diploid and haploid chromosome number
- Explain how daughter cells compare to the original parent cell
- Explain how genotypes contribute to phenotypic variation within a species

Changes in Ecosystems and Interactions of Organisms with Their Environments—A student can

- Identify and explain limiting factors (abiotic and biotic) that may affect carrying capacity
- Describe how a natural environmental event impacts diversity in an ecosystem
- Explain the impact human activity may have on the diversity of different species in an ecosystem
- Predict the energy flow in a food web
- Explain the natural and/or human factors that may lead to the extinction of a species
- Given a scenario describing an environmental change, hypothesize why a given species was unable to survive

Scientific Inquiry—A student can

- Formulate a testable hypothesis
- Identify constants and variables in an investigation

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Determine the appropriate tools and techniques to collect, analyze, and interpret data
- Determine scientific conclusion based on observations
- Identify factors required to make investigative results reliable
- Analyze quantitative data
- Design scientific investigations consisting of at least three steps
- Explain why accurate records and replications are essential for experimental creditability (includes peer review)
- Communicate procedures and results of investigations

Basic

Raw Score Cut: 18–31; Scale Score Cut: To be determined after operational data are complete.

Characteristics and Interactions of Living Organisms—A student can

- Identify and describe cell structures and functions
- Define organelles by their functions
- Explain how water is important to cells
- Use a Punnett square to show a simple monohybrid cross

Changes in Ecosystems and Interactions of Organisms with Their Environments—A student can

- Explain how interactions within an ecosystem maintain balance
- Explain the nature of interactions between organisms in predator/prey relationships and different symbiotic relationships (i.e., mutualism, commensalism, parasitism)
- Define carrying capacity of a population within an ecosystem
- Identify how adaptations may have provided a population an advantage for survival
- Identify the impact a natural environmental event may have on the diversity of different species in an ecosystem
- Explain how environmental factors can be agents of natural selection
- Explain the importance of reproduction to the survival of a species

Scientific Inquiry—A student can

- Select appropriate investigation methods (techniques only)
- Use data to formulate an explanation
- Calculate average/mean for sets of data
- Identify possible effects of errors in data collection and calculations

Below Basic

Raw Score Cut: 0–17; Scale Score Cut: To be determined after operational data are complete.

Characteristics and Interactions of Living Organisms—A student can

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

- Identify that all organisms progress through life cycles
- Identify that all organisms are made of cells
- Identify that water is important to cells
- Identify that all living organisms have DNA
- Identify that DNA carries inherited information

Changes in Ecosystems and Interactions of Organisms with Their Environments—A student can

- Describe interactions between organisms in a predator/prey relationship
- Use a model to show that populations interact in an ecosystem
- Identify examples of adaptations resulting from natural selection

Scientific Inquiry—A student can

- Identify a valid conclusion in an experiment
- Use simple tools to measure length, mass, and volume
- Communicate basic information from an experiment
- Construct a simple graph of independent variable versus dependent variable from given data

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX F: QUALIFYING TEST

EOC Assessment: E A B **Judge # _______**

Pre-Standard-Setting Self-Evaluation Assessment for Judges of the Missouri EOC Assessments (PSSSEAJMEOCA)

Directions: Circle the letter next to your answer for each item. Don't copy from your neighbor; he/she hasn't been listening very closely.

1. Why are the Achievement-Level Descriptors such an integral part of the standard-setting process?
   A. They provide an anchor that gives concrete meaning to the terms Basic, Proficient, and Advanced.
   B. All students at a given performance level should possess all critical behaviors and understandings listed in the ALDs.
   C. They define all of the items that are contained on the EOC.
   D. They summarize all of the elements of the Course-Level Expectations for the course.
2. Which of these statements about standard setting is TRUE?
   A. Panelists should use their best judgment to make their recommendations, but should rely more on various data to be provided during the sessions.
   B. While the EOC assessments are given statewide, judges should make recommendations based on the unique characteristics of *their* districts since other panelists will focus on other district types.
   C. A judge who concludes that the "proper" cut score for Proficient is 24 should make a final recommendation of 22 or 23 to account for errors that are present in any assessment.
   D. Judges must consider both the "stem" *and* answer options in multiple-choice items in deciding the percent of students who should answer correctly.
3. Joe the Judge decided that about 50% of the typical Proficient children in Missouri taking the EOC assessment should answer Item 32 correctly. He coded 50% under Proficient on his Rating Form. What error did he make?
   A. He should have coded 45% since some percent of special-needs students will take the assessment.
   B. He should have considered *barely* Proficient, not *typical* Proficient, students.
   C. He should reconsider his judgment, as 50% correct couldn't possibly be considered Proficient.
   D. He made no error here. This was the correct procedure.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

4. Judge Jan reviewed the performance event for her EOC and decided that the average score of borderline Proficient students should be 2 out of 4. What should she enter on her Rating Form?

   A. 50%, since 2 out of 4 is 50%

   B. 2.5, since she decided that 2 was the minimum acceptable score

   C. 1.5, since the minimum expected score should be somewhat lower than the average score

   D. 2, since her judgment is that 2 should be the average score of the target group

5. Which of these sets of "Angoff" judgments for a multiple-choice item appears to be improper and why?

	Below Basic/Basic	Basic/ Proficient	Proficient/ Advanced
A.	25%	35%	40%
B.	80%	90%	100%
C.	50%	50%	55%
D.	40%	75%	95%

   A. A, because these are very low expectations for a multiple-choice item

   B. B, because it is unrealistic to expect students to score this well on a multiple-choice item

   C. C, because the judge doesn't expect higher-classified students to perform any better on the item than lower-classified students

   D. D, because the increase in percentages across the three groups is probably unrealistically large

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX G: PARTICIPANT EVALUATION

Missouri End-of-Course Standard Setting

EVALUATION FORM

This form contains six sections, five of which ask for feedback on specific aspects of this standard-setting meeting. The last section asks for general reactions to the standard-setting meeting. Please fill out each of these sections as completely as possible in order to provide information that will help in the improvement of similar meetings in the future. Your identification number is used for analysis purposes only. Your responses to these questions will be held in strict confidence and will be analyzed in conjunction with those of the other judges who participated in this meeting.

Judge's I.D. (optional) ________________________

Section I: Opening Training Sessions

The following statements seek your judgments about the Opening Sessions for the Missouri End-of-Course standard-setting meeting. Please circle the value on the scale under each statement that best characterizes your judgment.

1. The Opening Sessions provided adequate background information about the Missouri End-of-Course Assessments.

5	4	3	2	1
Completely		Somewhat		Not at all

2. The topics covered in the Opening Sessions were appropriate to providing a context for my role in this meeting.

5	4	3	2	1
Completely		Somewhat		Not at all

3. The content of the Opening Sessions was:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

4. The organization of the Opening Sessions was:

5	4	3	2	1
Very good		Acceptable		Very poor

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

The following statements also seek your judgments about the Opening Sessions for the Missouri End-of-Course standard-setting meeting. Please write your responses to each prompt on the lines provided.

5. Were there questions or concerns that were not answered or addressed in the Opening Sessions? Please indicate these below. (Use reverse side for additional space.)

6. What was most helpful about the Opening Sessions?

7. Please use the space below to provide additional comments concerning the adequacy, appropriateness, usefulness, or organization of the Opening Sessions.

Section II: Discussing Proficient Performance

The following statements seek your judgments about the discussions of Proficient performance as they relate to Missouri's End-of-Course Assessments. Please circle the value on the scale under each statement that best characterizes your judgment.

8. The activities used to help operationalize Proficient performance were:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

9. By the end of the activity, my conception of Proficient performance was:

5	4	3	2	1
Very well formed		Moderately well formed		Not well formed

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

The following statement also seeks your judgments about the discussions of Proficient performance as they relate to Missouri's End-of-Course Assessments. Please write your responses to each prompt on the lines provided.

10. Please use the space below to provide additional comments concerning the activities around operationalizing Proficient performance for Missouri's End-of-Course Assessments.

____

____

____

Section III: Discussing Basic Performance

The following statements seek your judgments about the discussions of Basic performance as they relate to Missouri's End-of-Course Assessments. Please circle the value on the scale under each statement that best represents your judgment.

11. The activities used to help operationalize Basic performance were:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

12. By the end of this activity my conception of Basic performance was:

5	4	3	2	1
Very well formed		Moderately well formed		Not well formed

The following statement also seeks your judgments about the discussions of Basic performance as they relate to Missouri's End-of-Course Assessments. Please write your responses to each prompt on the lines provided.

13. Please use the space below to provide additional comments concerning the activities around operationalizing Basic performance for Missouri's End-of-Course Assessments.

____

____

____

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Section IV: Discussing Advanced Performance

The following statements seek your judgments about the discussions of Advanced performance as they relate to Missouri's End-of-Course Assessments. Please circle the value on the scale under each statement that best represents your judgment.

14. The activities used to help operationalize Advanced performance were:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

15. By the end of this activity my conception of Advanced performance was:

5	4	3	2	1
Very well formed		Moderately well formed		Not well formed

The following statement also seeks your judgments about the discussions of Advanced performance as they relate to Missouri's End-of-Course Assessments. Please write your responses to each prompt on the lines provided.

16. Please use the space below to provide additional comments concerning the activities around operationalizing Advanced performance for Missouri's End-of-Course Assessments.

____

____

____

Section V: Item Rating Activities

The following statements seek your judgments about the item rating activities as they relate to the Missouri End-of-Course standard-setting meeting. Please circle the value on the scale under each statement that best represents your judgment.

17. Using the sample items to prepare for the actual item rating was:

5	4	3	2	1
Very helpful		Somewhat helpful		Not helpful

18. The explanation of the item data during the sample item portion of the training was:

5	4	3	2	1
Very helpful		Somewhat helpful		Not helpful

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

19. The Item Rating Form was:

5	4	3	2	1
Very easy to use		Somewhat easy to use		Not at all easy to use

20. The information provided prior to each round of rating was:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

21. My level of understanding of the tasks I was to accomplish for each round was:

5	4	3	2	1
Very good		Acceptable		Very poor

22. The amount of time I had to complete the tasks during each round was:

5	4	3	2	1
Far too long		About right		Far too short

The following statement seeks your judgments about the item rating activities as they relate to the Missouri End-of-Course standard-setting meeting. Please write your responses to each prompt on the lines provided.

23. Please use the space below to provide additional comments concerning the instructions and explanations you received, the adequacy of the time available, your levels of understanding of the process, or any other aspects of the estimates for the multiple-choice items. (Use reverse side for additional space.)

____

____

____

Section VI: The Overall Missouri End-of-Course Standard-Setting Meeting

The following statements seek your judgments about the overall processes and procedures used during the Missouri End-of-Course standard-setting meeting in which you participated as a panelist and the resulting recommended standards. Please circle the value on the scale under each statement that best represents your judgment.

24. I feel that this standard-setting meeting provided me an opportunity to use my best judgment in selecting and revising estimates for a recommended standard of Proficient performance.

5	4	3	2	1
To a great extent		To some extent		Not at all

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

25. I feel that this standard-setting meeting provided me an opportunity to use my best judgment in selecting and revising estimates for a recommended standard of Basic performance.

5	4	3	2	1
To a great extent		To some extent		Not at all

26. I feel that this standard-setting meeting provided me an opportunity to use my best judgment in selecting and revising estimates for a recommended standard of Advanced performance.

5	4	3	2	1
To a great extent		To some extent		Not at all

27. I believe that this standard-setting meeting has produced recommended cut scores that are defensible.

5	4	3	2	1
To a great extent		To some extent		Not at all

28. I feel that this standard-setting meeting has produced recommended cut scores that would generally be considered as reasonable.

5	4	3	2	1
To a great extent		To some extent		Not at all

The following statements seek your judgments about the overall processes and procedures used during the Missouri End-of-Course standard-setting meeting. Please write your responses to each prompt on the lines provided.

29. Please provide any comments you wish to share regarding the quality of assistance provided by the standard-setting staff.

____

____

____

30. Please provide any additional comments you wish to share regarding the overall meeting.

____

____

____

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX H: RESULTS FOR ENGLISH II

Standard Setting for the Missouri EOC Assessment
English II

Round 1 Ratings Summary

Rater	Basic	Proficient	Advanced
	Individual Rater Cut Scores		
E211	16	27	34
E231	18	24	32
E232	18	24	30
E121	17	25	33
E223	18	23	30
E233	18	25	33
E331	13	26	31
E123	14	29	33
E311	18	29	35
E221	16	22	26
E313	10	23	32
E113	24	30	33
E222	13	20	27
E332	16	22	31

	Basic	Proficient	Advanced
Median Rating:	16.5	24.5	32.0
Average Rating:	16.36	24.93	31.43
Standard Deviation:	3.18	2.87	2.44
Lowest Rating:	10	20	26
Highest Rating:	24	30	35
Number of Items:	36	36	36
Points Possible:	39	39	39
Number of Raters:	14	14	14

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Standard Setting for the Missouri EOC Assessment

English II

Round 2 Ratings Summary

	Individual Rater Cut Scores		
Rater	Basic	Proficient	Advanced
E331	12	26	31
E113	16	30	33
E121	17	26	33
E123	14	28	33
E211	15	25	33
E221	17	25	30
E222	13	21	28
E223	18	24	30
E232	17	23	29
E233	19	26	33
E311	16	27	34
E313	13	24	32
E332	18	24	33
E231	19	27	33

Median Rating:	16.5	25.5	33.0
Average Rating:	16.00	25.43	31.79
Standard Deviation:	2.20	2.16	1.78
Lowest Rating:	12	21	28
Highest Rating:	19	30	34
Number of Items:	36	36	36
Points Possible:	39	39	39
Number of Raters:	14	14	14

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

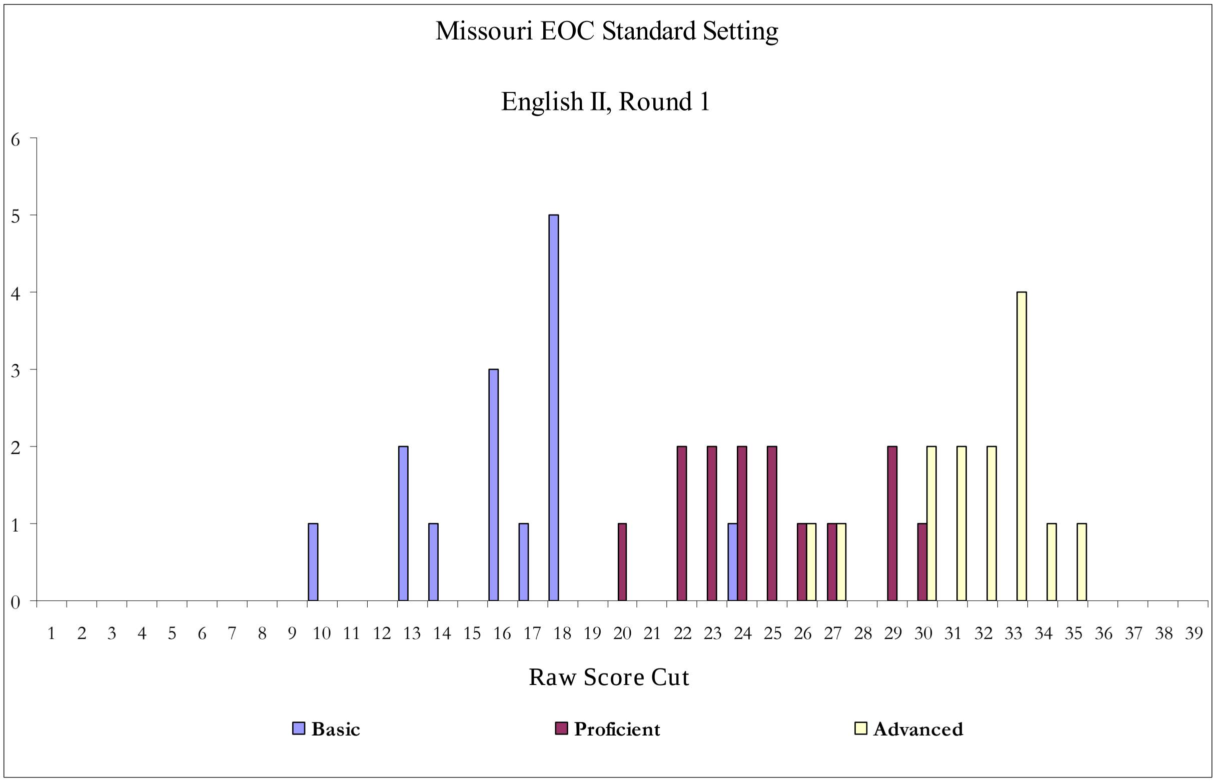
Standard Setting for the Missouri EOC Assessment

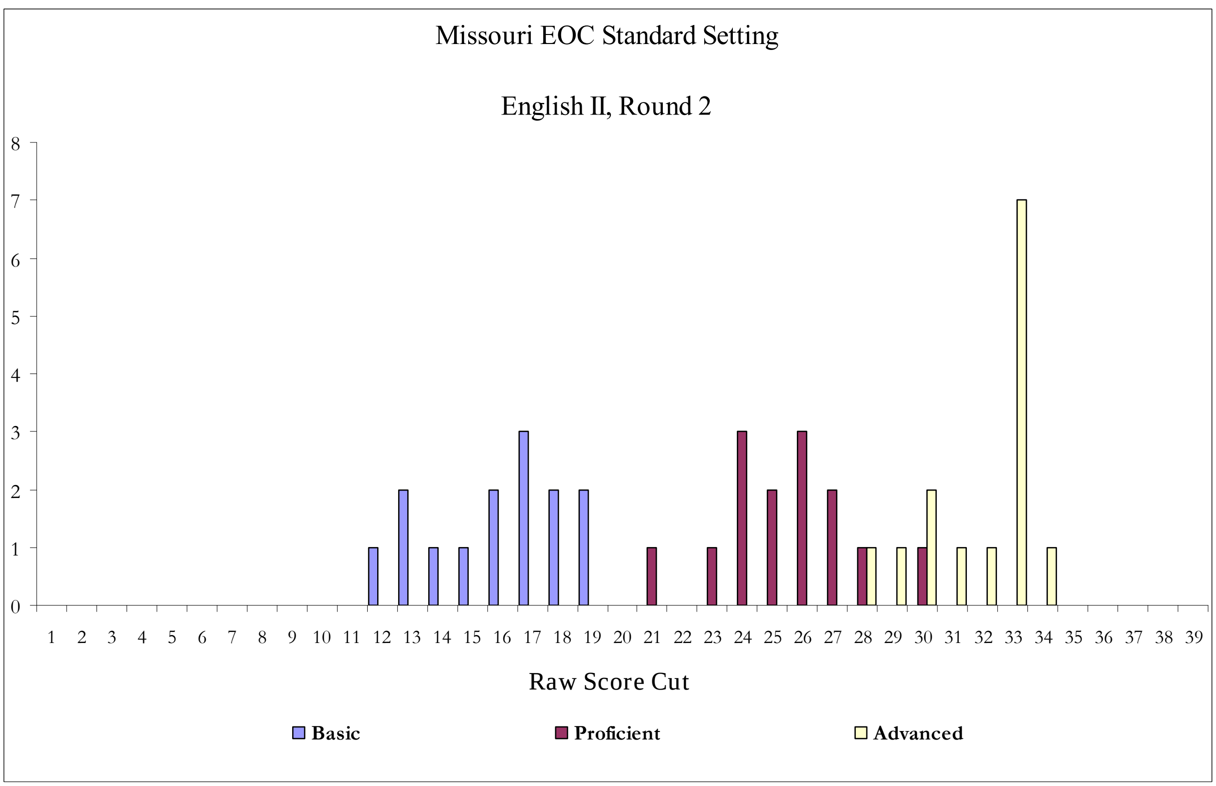
English II

Round 3 Ratings Summary

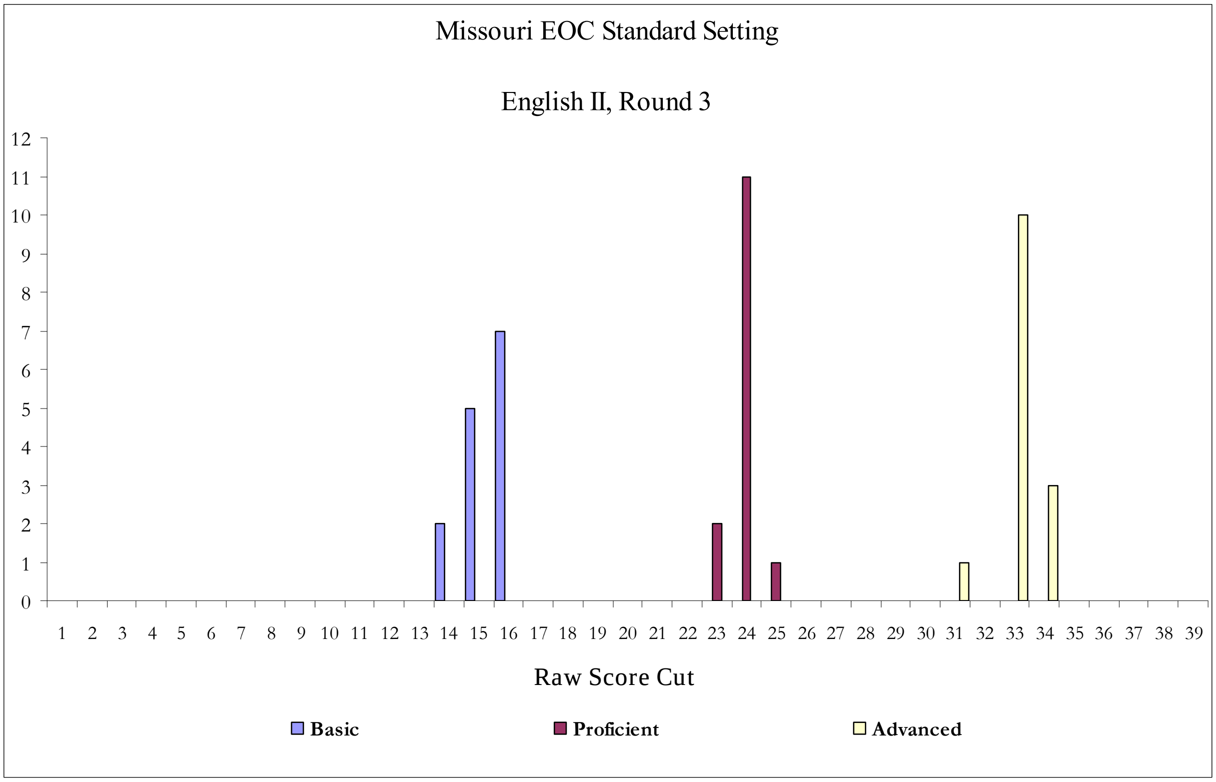
	Individual Rater Cut Scores		
Rater	Basic	Proficient	Advanced
E311	15	23	34
E232	15	24	33
E233	16	24	33
E222	14	23	31
E331	14	24	33
E223	16	24	33
E211	15	24	33
E121	15	24	34
E123	15	24	33
E221	16	24	33
E231	16	24	33
E113	16	24	33
E313	16	25	34
E332	16	24	33
Median Rating:	15.5	24.0	33.0
Average Rating:	15.36	23.93	33.07
Standard Deviation:	0.72	0.46	0.70
Lowest Rating:	14	23	31
Highest Rating:	16	25	34
Number of Items:	36	36	36
Points Possible:	39	39	39
Number of Raters:	14	14	14

Copyright © 2010 by Missouri Department of Elementary and Secondary Education





Copyright © 2010 by Missouri Department of Elementary and Secondary Education



Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX I: RESULTS FOR ALGEBRA I

Standard Setting for the Missouri EOC Assessment
Algebra I

Round 1 Ratings Summary

	Individual Rater Cut Scores		
Rater	Basic	Proficient	Advanced
A311	16	29	35
A232	10	26	36
A321	13	25	34
A112	13	23	31
A233	17	22	26
A313	12	19	24
A211	12	20	32
A122	13	23	33
A312	13	23	34
A222	16	25	32
A333	18	25	32
A123	19	28	33
A121	13	24	33
A322	12	21	27
A213	11	21	31

	Basic	Proficient	Advanced
Median Rating:	13.0	23.0	32.0
Average Rating:	13.87	23.60	31.53
Standard Deviation:	2.58	2.73	3.26
Lowest Rating:	10	19	24
Highest Rating:	19	29	36
Number of Items:	36	36	36
Points Possible:	39	39	39
Number of Raters:	15	15	15

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Standard Setting for the Missouri EOC Assessment

Algebra I

Round 2 Ratings Summary

	Individual Rater Cut Scores		
Rater	Basic	Proficient	Advanced
A122	13	21	33
A321	14	25	33
A121	12	23	32
A222	14	23	30
A322	13	21	27
A211	13	20	33
A333	16	25	32
A213	11	21	31
A312	12	23	31
A112	14	21	29
A123	16	25	31
A233	9	21	30
A311	14	27	34
A313	18	22	27
A232	11	23	35

	Basic	Proficient	Advanced
Median Rating:	13.0	23.0	31.0
Average Rating:	13.33	22.73	31.20
Standard Deviation:	2.18	1.95	2.26
Lowest Rating:	9	20	27
Highest Rating:	18	27	35
Number of Items:	36	36	36
Points Possible:	39	39	39
Number of Raters:	15	15	15

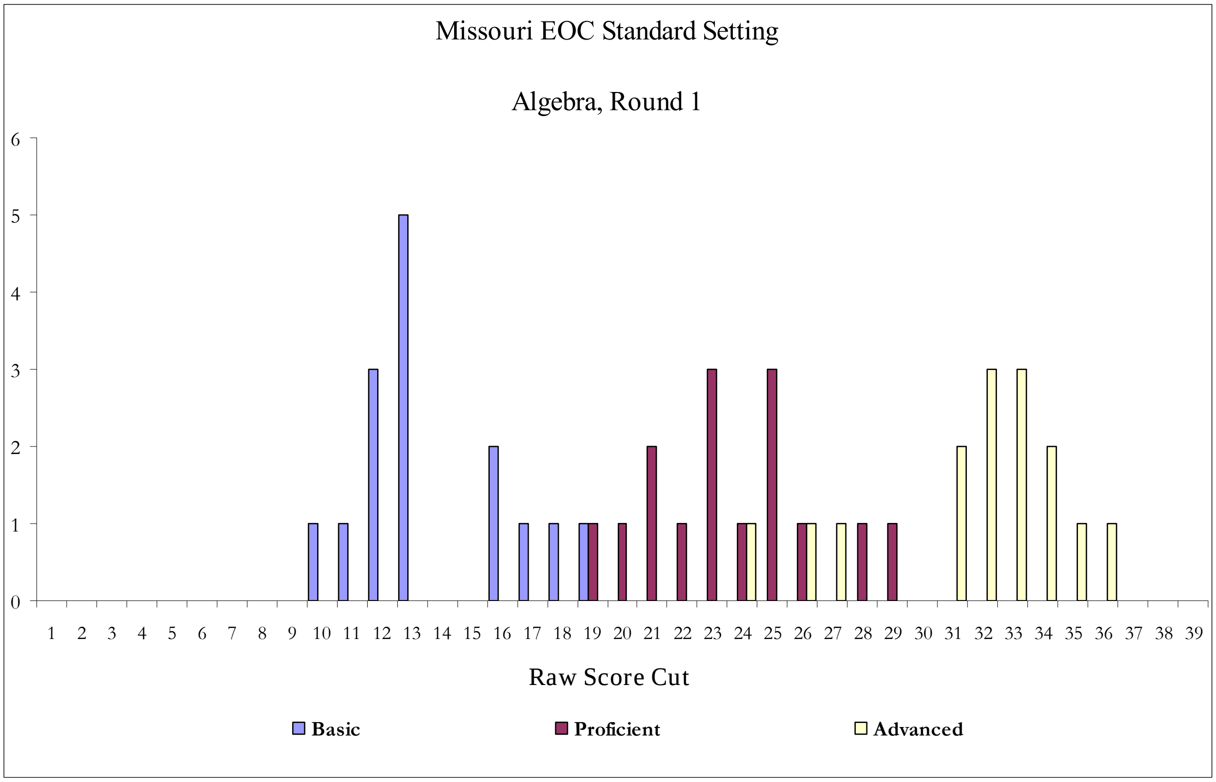
Copyright © 2010 by Missouri Department of Elementary and Secondary Education

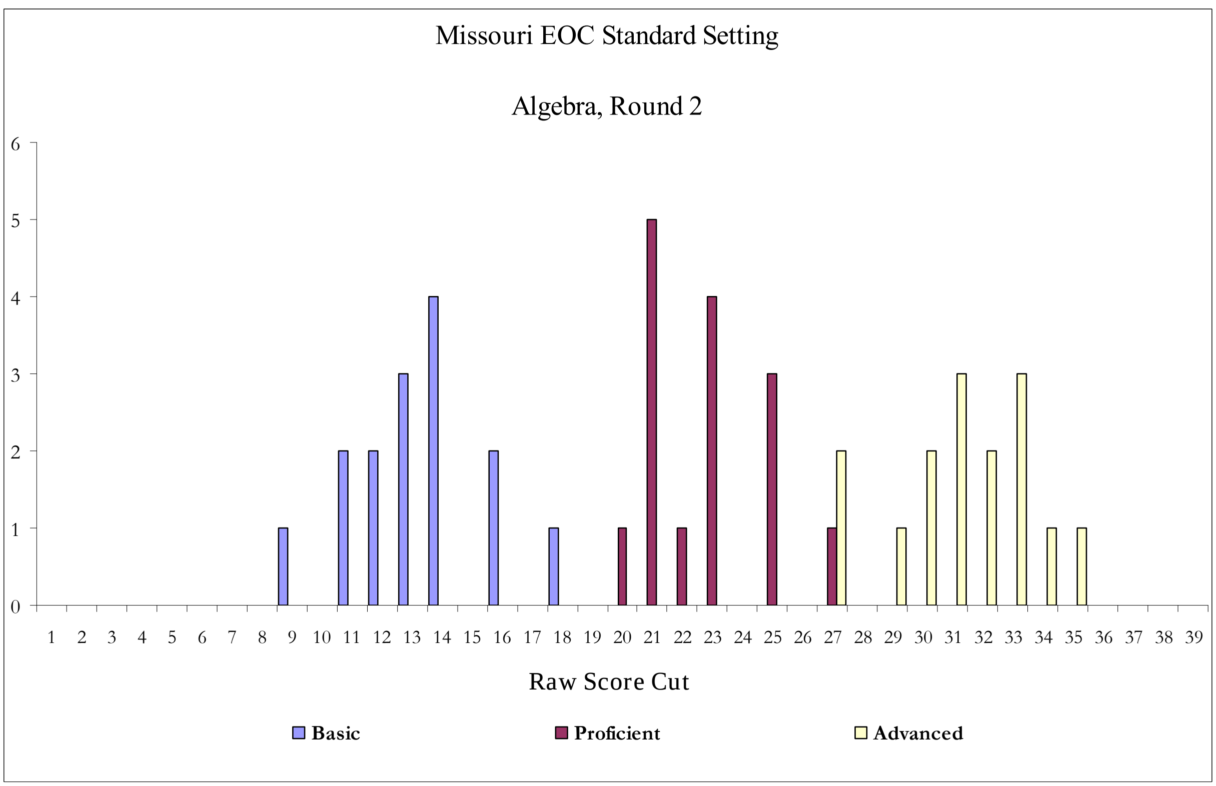
Standard Setting for the Missouri EOC Assessment
Algebra I

Round 3 Ratings Summary

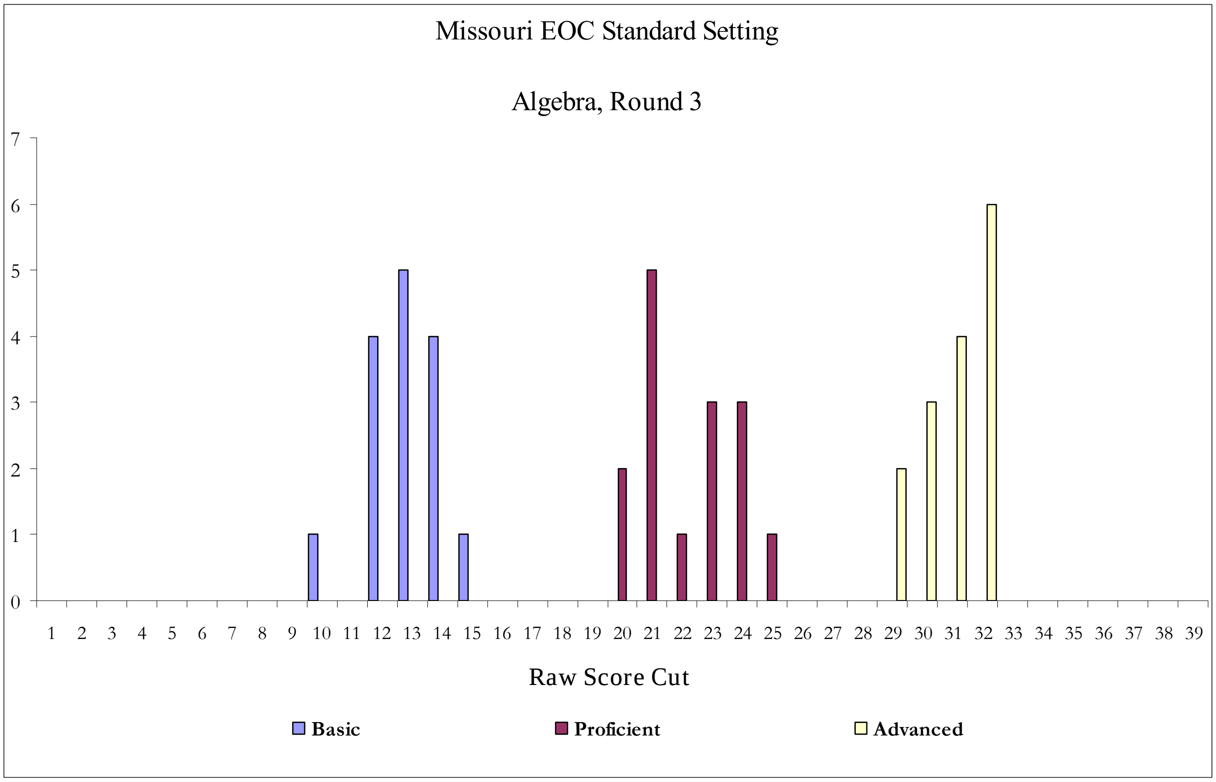
	Individual Rater Cut Scores		
Rater	Basic	Proficient	Advanced
A312	13	23	31
A122	12	20	31
A211	13	21	32
A232	10	20	31
A112	13	21	29
A121	12	21	32
A322	13	22	30
A313	13	23	29
A321	14	24	32
A311	14	25	32
A233	12	21	30
A333	15	24	32
A123	14	24	31
A213	12	21	32
A222	14	23	30
Median Rating:	13.0	22.0	31.0
Average Rating:	12.93	22.20	30.93
Standard Deviation:	1.18	1.56	1.06
Lowest Rating:	10	20	29
Highest Rating:	15	25	32
Number of Items:	36	36	36
Points Possible:	39	39	39
Number of Raters:	15	15	15

Copyright © 2010 by Missouri Department of Elementary and Secondary Education





Copyright © 2010 by Missouri Department of Elementary and Secondary Education



Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Standard Setting for the Missouri EOC Assessment
Biology

Round 1 Ratings Summary

Rater	Basic	Proficient	Advanced
	Individual Rater Cut Scores		
B213	18	32	49
B312	18	33	50
B112	14	29	42
B231	20	37	48
B311	20	31	42
B321	13	29	41
B223	20	37	48
B232	23	36	45
B322	23	36	48
B233	19	32	42
B131	23	35	46
B211	30	43	51
B212	21	30	41
B122	17	37	50
B332	21	40	54
B113	18	32	44
B111	27	38	51

	Basic	Proficient	Advanced
Median Rating:	20.0	35.0	48.0
Average Rating:	20.29	34.53	46.59
Standard Deviation:	4.08	3.85	3.94
Lowest Rating:	13	29	41
Highest Rating:	30	43	54
Number of Items:	46	46	46
Points Possible:	55	55	55
Number of Raters:	17	17	17

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Standard Setting for the Missouri EOC Assessment

Biology

Round 2 Ratings Summary

Rater	Basic	Proficient	Advanced
	Individual Rater Cut Scores		
B232	22	34	43
B212	21	30	42
B321	14	27	39
B211	28	42	51
B322	22	36	47
B233	16	30	42
B332	18	34	47
B112	12	27	45
B131	25	35	46
B223	16	34	47
B122	18	39	50
B311	22	32	42
B213	18	31	46
B231	19	35	46
B111	26	38	51
B312	19	35	49
B113	15	28	43

	Basic	Proficient	Advanced
Median Rating:	19.0	34.0	46.0
Average Rating:	19.47	33.35	45.65
Standard Deviation:	4.23	4.09	3.36
Lowest Rating:	12	27	39
Highest Rating:	28	42	51
Number of Items:	46	46	46
Points Possible:	55	55	55
Number of Raters:	17	17	17

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

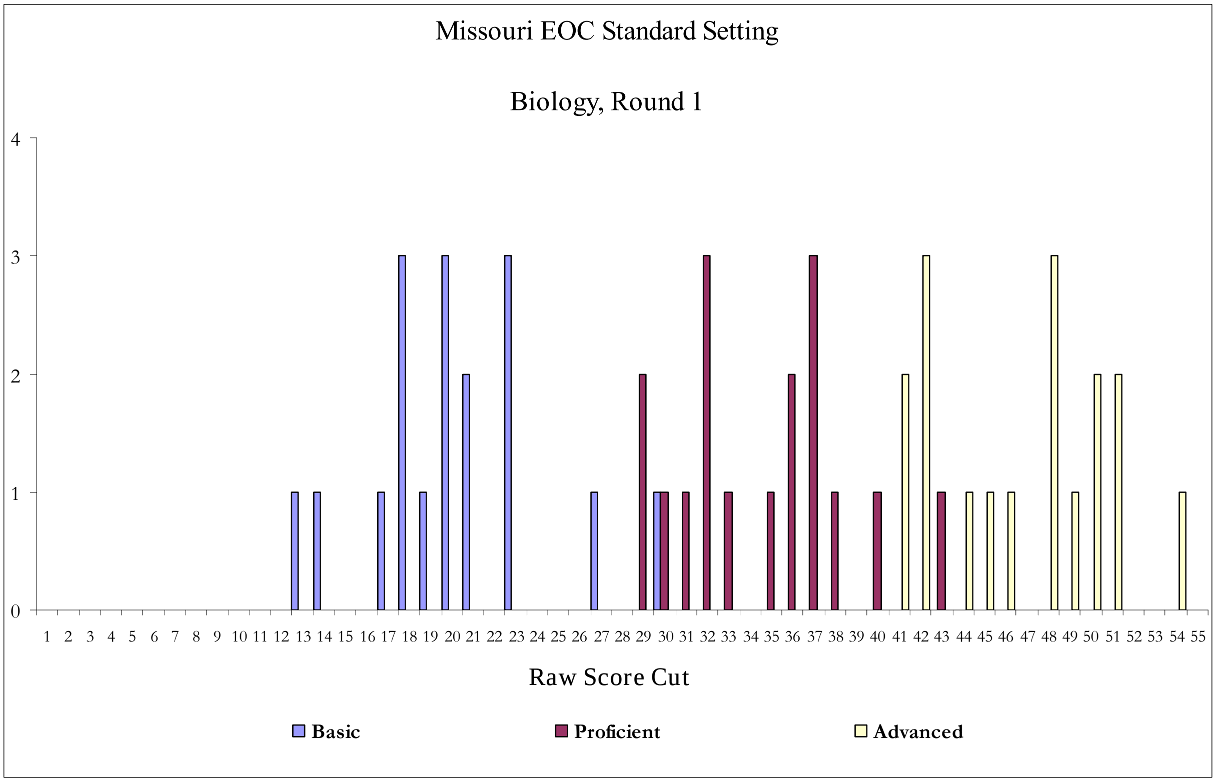
Standard Setting for the Missouri EOC Assessment
Biology

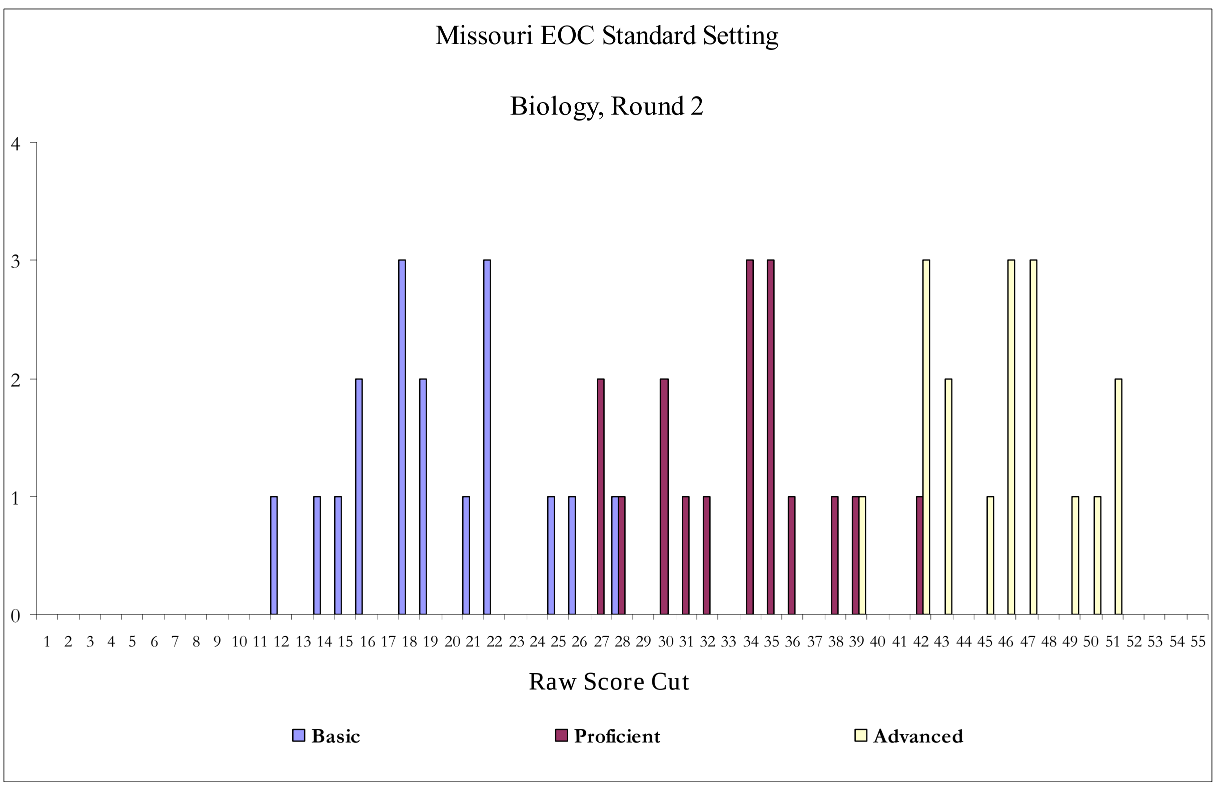
Round 3 Ratings Summary

Rater	Basic	Proficient	Advanced
	Individual Rater Cut Scores		
B223	18	35	46
B232	19	32	46
B322	20	35	46
B111	20	34	48
B213	20	32	45
B332	15	30	46
B211	11	28	40
B212	15	30	40
B321	14	27	40
B131	20	35	45
B312	18	33	47
B231	19	32	46
B112	12	27	46
B311	20	32	45
B233	15	30	40
B113	15	30	44
B122	14	29	43

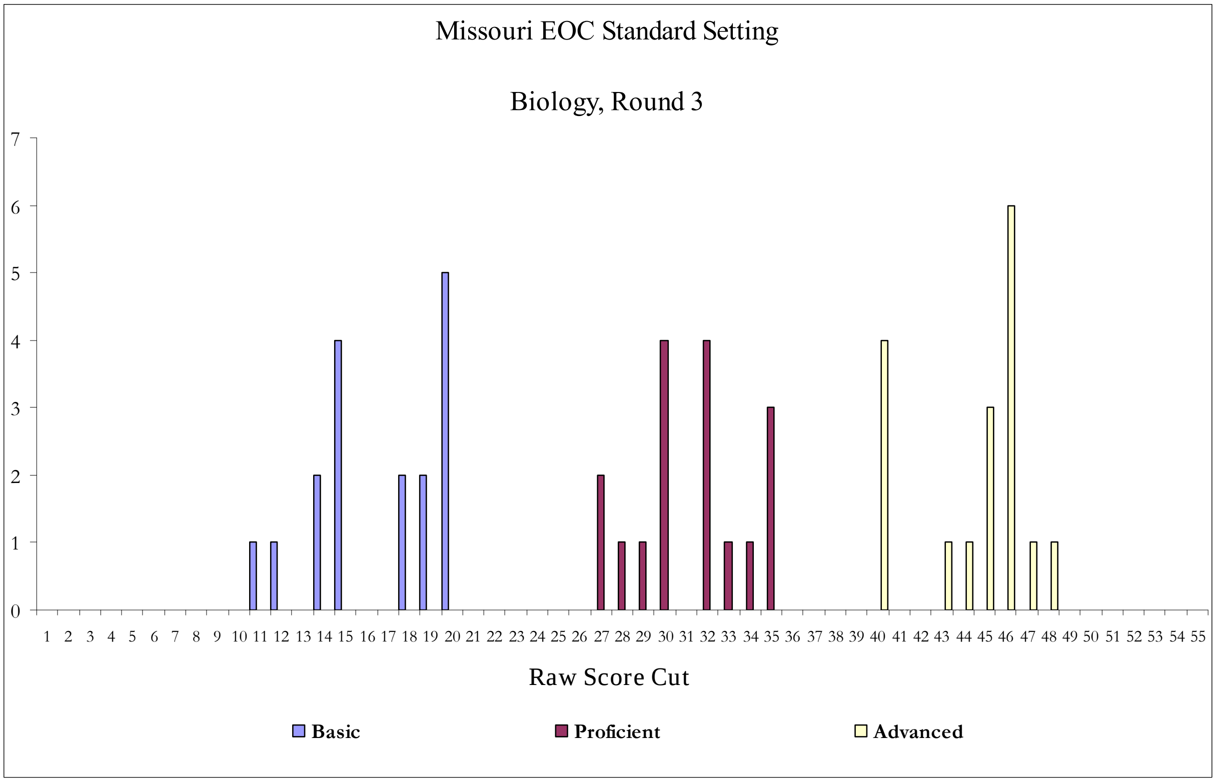
	Basic	Proficient	Advanced
Median Rating:	18.0	32.0	45.0
Average Rating:	16.76	31.24	44.29
Standard Deviation:	2.96	2.58	2.61
Lowest Rating:	11	27	40
Highest Rating:	20	35	48
Number of Items:	46	46	46
Points Possible:	55	55	55
Number of Raters:	17	17	17

Copyright © 2010 by Missouri Department of Elementary and Secondary Education





Copyright © 2010 by Missouri Department of Elementary and Secondary Education



Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX K: RESULTS OF PARTICIPANT EVALUATION

Missouri End-of-Course Standard Setting

EVALUATION FORM

This form contains six sections, five of which ask for feedback on specific aspects of this standard-setting meeting. The last section asks for general reactions to the standard-setting meeting. Please fill out each of these sections as completely as possible in order to provide information that will help in the improvement of similar meetings in the future. Your identification number is used for analysis purposes only. Your responses to these questions will be held in strict confidence and will be analyzed in conjunction with those of the other judges who participated in this meeting.

Judge's I.D. (optional) ________________________

Section I: Opening Training Sessions

The following statements seek your judgments about the Opening Sessions for the Missouri End-of-Course standard-setting meeting. Please circle the value on the scale under each statement that best characterizes your judgment.

1. The Opening Sessions provided adequate background information about the Missouri End-of-Course Assessments.

5	4	3	2	1
Completely		Somewhat		Not at all

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (17)**	**%**
5	10	67%	5	36%	6	35%
4	3	20%	8	57%	11	65%
3	2	13%	1	7%	0	0%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.5	.74	4.3	.61	4.4	.49

*Percentages may not sum to 100% due to rounding.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

2. The topics covered in the Opening Sessions were appropriate to providing a context for my role in this meeting.

5	4	3	2	1
Completely		Somewhat		Not at all

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**		**Count (15)**	**%**
5	9	60%	8	57%	7	41%
4	4	27%	4	29%	10	59%
3	2	13%	2	14%	0	0%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.5	.74	4.2	.76	4.4	.51

*Percentages may not sum to 100% due to rounding.

3. The content of the Opening Sessions was:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (17)**	**%**
5	8	53%	2	14%	3	18%
4	4	27%	11	79%	11	65%
3	3	20%	1	7%	3	18%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.3	.82	4.1	.47	4.0	.61

*Percentages may not sum to 100% due to rounding.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

4. The organization of the Opening Sessions was:

5	4	3	2	1
Very good		Acceptable		Very poor

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (17)**	**%**
5	8	53%	3	21%	2	12%
4	4	27%	9	64%	9	53%
3	1	7%	1	7%	6	35%
2	2	13%	1	7%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.2	1.08	4.0	.78	3.8	.66

*Percentages may not sum to 100% due to rounding.

The following statements also seek your judgments about the Opening Sessions for the Missouri End-of-Course standard-setting meeting. Please write your responses to each prompt on the lines provided.

5. Were there questions or concerns that were not answered or addressed in the Opening Sessions? Please indicate these below. (Use reverse side for additional space.)

6. What was most helpful about the Opening Sessions?

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

7. Please use the space below to provide additional comments concerning the adequacy, appropriateness, usefulness, or organization of the Opening Sessions.

____

____

____

Section II: Discussing Proficient Performance

The following statements seek your judgments about the discussions of Proficient performance as they relate to Missouri's End-of-Course Assessments. Please circle the value on the scale under each statement that best characterizes your judgment.

8. The activities used to help operationalize Proficient performance were:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (17)**	**%**
5	8	53%	5	36%	4	24%
4	4	27%	7	50%	9	53%
3	3	20%	2	14%	4	24%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.3	.82	4.2	.70	4.0	.71

*Percentages may not sum to 100% due to rounding.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

9. By the end of the activity, my conception of Proficient performance was:

5	4	3	2	1
Very well formed		Moderately well formed		Not well formed

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**		**Count (15)**	**%**
5	9	60%	8	57%	7	41%
4	5	33%	3	21%	8	47%
3	1	7%	3	21%	2	12%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.5	.64	4.4	.84	4.3	.69

*Percentages may not sum to 100% due to rounding.

The following statement also seeks your judgments about the discussions of Proficient performance as they relate to Missouri's End-of-Course Assessments. Please write your responses to each prompt on the lines provided.

10. Please use the space below to provide additional comments concerning the activities around operationalizing Proficient performance for Missouri's End-of-Course Assessments.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Section III: Discussing Basic Performance

The following statements seek your judgments about the discussions of Basic performance as they relate to Missouri's End-of-Course Assessments. Please circle the value on the scale under each statement that best represents your judgment.

11. The activities used to help operationalize Basic performance were:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

	Algebra I*		English II*		Biology*	
	Count (15)	%	Count (14)	%	Count (17)	%
5	7	47%	5	36%	2	12%
4	6	40%	5	36%	10	59%
3	2	13%	4	29%	5	29%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.3	.72	4.1	.83	3.8	.64

*Percentages may not sum to 100% due to rounding.

12. By the end of this activity my conception of Basic performance was:

5	4	3	2	1
Very well formed		Moderately well formed		Not well formed

	Algebra I*		English II*		Biology*	
	Count (15)	%	Count (14)	%	Count (17)	%
5	7	47%	6	43%	5	29%
4	7	47%	6	43%	9	53%
3	1	7%	2	14%	3	18%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.4	.63	4.3	.73	4.1	.70

*Percentages may not sum to 100% due to rounding.

The following statement also seeks your judgments about the discussions of Basic performance as they relate to Missouri's End-of-Course Assessments. Please write your responses to each prompt on the lines provided.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

13. Please use the space below to provide additional comments concerning the activities around operationalizing Basic performance for Missouri's End-of-Course Assessments.

____

____

____

Section IV: Discussing Advanced Performance

The following statements seek your judgments about the discussions of Advanced performance as they relate to Missouri's End-of-Course Assessments. Please circle the value on the scale under each statement that best represents your judgment.

14. The activities used to help operationalize Advanced performance were:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

	Algebra I*		**English II***		**Biology***	
	Count (15)	**%**	**Count (14)**	**%**	**Count (16)****	**%**
5	6	40%	6	43%	3	19%
4	6	40%	7	50%	9	56%
3	3	20%	1	7%	4	25%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.2	.77	4.4	.63	3.9	.68

*Percentages may not sum to 100% due to rounding.
**One panelist did not respond to this question.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

15. By the end of this activity my conception of Advanced performance was:

5	4	3	2	1
Very well formed		Moderately well formed		Not well formed

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (16)****	**%**
5	8	53%	7	50%	4	25%
4	5	33%	7	50%	10	63%
3	2	13%	0	0%	2	13%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.4	.74	4.5	.52	4.1	.62

*Percentages may not sum to 100% due to rounding.
**One panelist did not respond to this question.

The following statement also seeks your judgments about the discussions of Advanced performance as they relate to Missouri's End-of-Course Assessments. Please write your responses to each prompt on the lines provided.

16. Please use the space below to provide additional comments concerning the activities around operationalizing Advanced performance for Missouri's End-of-Course Assessments.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Section V: Item Rating Activities

The following statements seek your judgments about the item rating activities as they relate to the Missouri End-of-Course standard-setting meeting. Please circle the value on the scale under each statement that best represents your judgment.

17. Using the sample items to prepare for the actual item rating was:

5	4	3	2	1
Very helpful		Somewhat helpful		Not helpful

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (17)**	**%**
5	9	60%	6	43%	9	53%
4	4	27%	5	36%	4	24%
3	1	7%	1	7%	4	24%
2	1	7%	1	7%	0	0%
1	0	0%	1	7%	0	0%
Mean, *SD*	4.4	.91	4.0	1.24	4.3	.85

*Percentages may not sum to 100% due to rounding.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

18. The explanation of the item data during the sample item portion of the training was:

5	4	3	2	1
Very helpful		Somewhat helpful		Not helpful

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (17)**	**%**
5	6	40%	8	57%	8	47%
4	7	47%	5	36%	3	18%
3	2	13%	1	7%	6	35%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.3	.70	4.5	.65	4.1	.93

*Percentages may not sum to 100% due to rounding.

19. The Item Rating Form was:

5	4	3	2	1
Very easy to use		Somewhat easy to use		Not at all easy to use

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (17)**	**%**
5	8	53%	9	64%	9	53%
4	6	40%	5	36%	7	41%
3	0	0%	0	14%	1	6%
2	1	7%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.4	.83	4.6	.50	4.5	.62

*Percentages may not sum to 100% due to rounding.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

20. The information provided prior to each round of rating was:

5	4	3	2	1
Very useful		Somewhat useful		Not useful

	Algebra I*		English II*		Biology*	
	Count (15)	%	Count (14)	%	Count (17)	%
5	10	67%	8	57%	8	47%
4	4	27%	5	36%	9	53%
3	1	7%	1	7%	0	0%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.6	.63	4.5	.65	4.5	.51

*Percentages may not sum to 100% due to rounding.

21. My level of understanding of the tasks I was to accomplish for each round was:

5	4	3	2	1
Very good		Acceptable		Very poor

	Algebra I*		English II*		Biology*	
	Count (15)	%	Count (14)	%	Count (17)	%
5	14	93%	7	50%	13	76%
4	0	0%	6	43%	2	12%
3	1	7%	1	7%	2	12%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.9	.52	4.4	.65	4.6	.70

*Percentages may not sum to 100% due to rounding.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

22. The amount of time I had to complete the tasks during each round was:

5	4	3	2	1
Far too long		About right		Far too short

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (14)**	**%**	**Count (17)**	**%**
5	2	13%	0	57%	3	18%
4	1	7%	5	36%	4	24%
3	12	80%	9	64%	10	59%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	3.3	.72	3.4	.50	3.6	.80

*Percentages may not sum to 100% due to rounding.

The following statement seeks your judgments about the item rating activities as they relate to the Missouri End-of-Course standard-setting meeting. Please write your responses to each prompt on the lines provided.

23. Please use the space below to provide additional comments concerning the instructions and explanations you received, the adequacy of the time available, your levels of understanding of the process, or any other aspects of the estimates for the multiple-choice items. (Use reverse side for additional space.)

____

____

____

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

Section VI: The Overall Missouri End-of-Course Standard-Setting Meeting

The following statements seek your judgments about the overall processes and procedures used during the Missouri End-of-Course standard-setting meeting in which you participated as a panelist and the resulting recommended standards. Please circle the value on the scale under each statement that best represents your judgment.

24. I feel that this standard-setting meeting provided me an opportunity to use my best judgment in selecting and revising estimates for a recommended standard of Proficient performance.

5	4	3	2	1
To a great extent		To some extent		Not at all

	Algebra I*		**English II***		**Biology***	
	Count (15)	**%**	**Count (13)****	**%**	**Count (17)**	**%**
5	11	73%	10	77%	7	41%
4	3	20%	3	23%	9	53%
3	1	7%	0	14%	1	6%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.7	.62	4.8	.44	4.4	.61

*Percentages may not sum to 100% due to rounding.
**One panelist did not respond to this question.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

25. I feel that this standard-setting meeting provided me an opportunity to use my best judgment in selecting and revising estimates for a recommended standard of Basic performance.

5	4	3	2	1
To a great extent		To some extent		Not at all

	Algebra I*		English II*		Biology*	
	Count (15)	%	Count (13)**	%	Count (17)	%
5	10	67%	9	69%	7	41%
4	4	27%	3	23%	10	59%
3	1	7%	1	8%	0	0%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.6	.63	4.6	.65	4.4	.51

*Percentages may not sum to 100% due to rounding.
**One panelist did not respond to this question.

26. I feel that this standard-setting meeting provided me an opportunity to use my best judgment in selecting and revising estimates for a recommended standard of Advanced performance.

5	4	3	2	1
To a great extent		To some extent		Not at all

	Algebra I*		English II*		Biology*	
	Count (15)	%	Count (13)**	%	Count (17)	%
5	11	73%	9	69%	7	41%
4	2	13%	4	31%	9	53%
3	2	13%	0	14%	1	6%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.6	.74	4.7	.48	4.4	.61

*Percentages may not sum to 100% due to rounding.
**One panelist did not respond to this question.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

27. I believe that this standard-setting meeting has produced recommended cut scores that are defensible.

5	4	3	2	1
To a great extent		To some extent		Not at all

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (13)****	**%**	**Count (17)**	**%**
5	9	60%	10	77%	10	59%
4	5	33%	1	8%	6	35%
3	1	7%	2	15%	1	6%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.5	.64	4.6	.77	4.5	.62

*Percentages may not sum to 100% due to rounding.
**One panelist did not respond to this question.

28. I feel that this standard-setting meeting has produced recommended cut scores that would generally be considered as reasonable.

5	4	3	2	1
To a great extent		To some extent		Not at all

	Algebra I*		English II*		Biology*	
	Count (15)	**%**	**Count (13)****	**%**	**Count (17)**	**%**
5	10	67%	8	62%	10	59%
4	4	27%	3	23%	6	35%
3	1	7%	2	15%	1	6%
2	0	0%	0	0%	0	0%
1	0	0%	0	0%	0	0%
Mean, *SD*	4.6	.63	4.5	.78	4.5	.62

*Percentages may not sum to 100% due to rounding.
**One panelist did not respond to this question.

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

The following statements seek your judgments about the overall processes and procedures used during the Missouri End-of-Course standard-setting meeting. Please write your responses to each prompt on the lines provided.

29. Please provide any comments you wish to share regarding the quality of assistance provided by the standard-setting staff.

____

____

____

30. Please provide any additional comments you wish to share regarding the overall meeting.

____

____

____

Copyright © 2010 by Missouri Department of Elementary and Secondary Education

APPENDIX L: DATE RANGES FOR HAND SCORING OF PERFORMANCE EVENT/WRITING PROMPT ITEMS

Appendix Table 6.1: Hand Scoring Dates for the Spring 2008 Field Test

	English II	Algebra I	Biology
Team Leader Training	May 29–30, 2008	May 29, 2008	May 29, 2008
Scorer Training	June 9, 2008	June 2, 2008	June 2, 2008
Scoring Window	June 9–25, 2008	June 2–24, 2008	June 2–24, 2008

Training schedules for the Spring 2008 field test varied because they were scheduled based on the availability of the RIFs and the completion of the preparation of the original training materials.

Appendix Table 6.2: Hand Scoring Dates for the Fall 2008 Operational Test

	English II	Algebra I	Biology
Team Leader Training	February 2, 2009	February 2, 2009	February 2, 2009
Scorer Training	February 3, 2009	February 3, 2009	February 6, 2009
Scoring Window	February 3–17, 2009	February 3–13, 2009	February 3–13, 2009

ARC used this scoring process as an opportunity to identify potential team leaders for the Spring 2009 operational scoring. In addition to scoring the Fall 2008 booklets, these candidates learned to conduct training, use reports, handle personnel issues, and oversee other administrative duties for which they would be responsible.

Appendix Table 6.3: Hand Scoring Dates for the Summer 2009 Operational Test

	English II	Algebra I	Biology
Team Leader Training	April 21–23, 2009	April 15, 2009	April 22–23, 2009
Scorer Training	April 27, 2009	April 28, 2009	April 28, 2009
Scoring Window	April 28–June 5, 2009	April 28–June 8, 2009	April 28–June 4, 2009

Appendix Table 6.4: Hand Scoring Dates for the Fall 2009 Operational Test

	English II	Algebra I	Biology
Team Leader Training	April 21–23, 2009	April 15, 2009	April 22–23, 2009
Scorer Training	April 27, 2009	April 28, 2009	April 28, 2009
Scoring Window	April 28–June 5, 2009	April 28–June 8, 2009	April 28–June 4, 2009

Appendix Table 6.5: Hand Scoring Dates for the Spring 2010 Operational Test

	English II	Algebra I	Biology
Team Leader Training	April 21–23, 2009	April 15, 2009	April 22–23, 2009
Scorer Training	April 27, 2009	April 28, 2009	April 28, 2009
Scoring Window	April 28–June 5, 2009	April 28–June 8, 2009	April 28–June 4, 2009

Copyright © 2010 by Missouri Department of Elementary and Secondary Education