



# Independent Alignment Review of the Science Missouri Assessment Program (MAP) Final

**Prepared for:** Missouri Department of Elementary and Secondary Education  
205 Jefferson Street  
P.O. Box 480  
Jefferson City, Missouri 65102

**Authors:** Yvette M. Nemeth  
Richard Deatz  
Erin Smith  
Jessica Johnston-Fisher

**Prepared under:** Missouri Department of Elementary and Secondary Education  
205 Jefferson Street  
P.O. Box 480  
Jefferson City, Missouri 65102

**Date:** December 30, 2014



# Independent Alignment Review of the Science Missouri Assessment Program (MAP)

## Table of Contents

Executive Summary.....	v
Scope of Work.....	v
Methodology.....	v
Review of Content Alignment.....	vi
Review of Test Quality.....	vi
Summary of Results .....	vii
Key Findings and Conclusions.....	vii
Alignment of Science MAP to Missouri Learning Standards .....	vii
Quality of Science MAP Tests.....	viii
Recommendations.....	ix
Chapter 1: Introduction .....	1
Organization and Contents of the Report .....	1
Chapter 2: Alignment Study Design and Methodology.....	3
Alignment of Assessments and Standards on Content .....	3
Content Alignment and Accessibility .....	3
Webb Alignment Method.....	3
Scope of Alignment Evaluations for Science MAP .....	4
Review of Content Alignment and Accessibility .....	4
Panelists .....	5
Training.....	5
Materials .....	5
Procedures.....	6
Chapter 3: Results: Content Alignment .....	9
Reliability Results .....	9
Panelist-Test Developer Analyses .....	9
Webb Alignment Results.....	11
Categorical Concurrence .....	11
Depth-of-Knowledge Consistency .....	13
Range of Knowledge Correspondence.....	15
Balance-of-Knowledge Representation .....	16
Summary and Discussion on Webb Alignment Indicators.....	18
Chapter 4: Results: Test Quality of Science MAP .....	21
Written Content .....	21
Figures and Graphics .....	22

Overall Item Quality .....	22
Summary and Discussion of Test Quality Results .....	23
Chapter 5: Summary and Recommendations .....	25
References .....	27
Appendix A. Content Alignment Results per Grade Level Assessment .....	A-1
Categorical Concurrence .....	A-1
Depth-of-Knowledge Consistency .....	A-4
Range-of-Knowledge Correspondence .....	A-8
Balance-of-Knowledge Representation .....	A-12
GLEs Matched to Items by Panelists .....	A-16
Appendix B. Summary of Panelist Comments on Items .....	B-1
Appendix C. Sample Alignment Review Materials .....	C-1

### List of Tables

Table 1. Summary Alignment Outcomes on Each Webb Criterion by Grade Level for Science MAP .....	viii
Table 2. Item Quality Ratings for Science MAP by Grade .....	ix
Table 2.1. Professional and Demographic Characteristics of Science MAP Panelists .....	5
Table 2.2. Characteristics of 2015 Science MAP Test Forms .....	6
Table 3.1. Missouri Learning Standards Example .....	10
Table 3.2. Percent Agreement between Panelists and Item Bank on Target Content – Grade 5 .....	10
Table 3.3. Summary of Categorical Concurrence Results for Science MAP – Grade 5 .....	11
Table 3.4. Summary of Categorical Concurrence Results for Science MAP – Grade 8 .....	12
Table 3.5. Panelist Ratings on Overall Item Alignment Grade per Forms – Grade 5 .....	13
Table 3.6. Panelist Ratings on Overall Item Alignment Grade per Forms – Grade 8 .....	13
Table 3.7. Summary of Depth-of-Knowledge Results for Science MAP – Grade 5 .....	14
Table 3.8. Summary of Depth-of-Knowledge Results for Science MAP – Grade 8 .....	14
Table 3.9. Number of GLEs per Grade Level .....	15
Table 3.10. Summary of Range-of-Knowledge Results for the Science MAP – Grade 5 .....	16
Table 3.11. Summary of Range-of-Knowledge Results for the Science MAP – Grade 8 .....	16
Table 3.12. Summary of Balance-of-Knowledge Representation Results Science MAP – Grade 5 .....	17
Table 3.13. Summary of Balance-of-Knowledge Representation Results Science MAP – Grade 8 .....	18
Table 3.14. Summary Alignment Outcomes on Each Webb Criterion by Grade Level for Science MAP .....	19
Table 4.1. Mean Number of Items Rated As Accessible in Content to Range of Students per Grade Assessment .....	21
Table 4.2. Mean Ratings on Accessibility of Figures or Graphics to Range of Students per Grade Assessment .....	22
Table 4.3. Panelist Ratings on Overall Item Quality – Grade 5 .....	23
Table 4.4. Panelist Ratings on Overall Item Quality – Grade 8 .....	23

Table 4.5. Item Quality Ratings for Science MAP by Grade .....	24
Table A-1. Categorical Concurrence for Science MAP, Grade 5: Mean Number of Items per Strand Form 1.....	A-1
Table A-2. Categorical Concurrence for Science MAP, Grade 5: Mean Number of Items per Strand Form 2.....	A-2
Table A-3. Categorical Concurrence for Science MAP, Grade 8: Mean Number of Items per Strand Form 1.....	A-2
Table A-4. Categorical Concurrence for Science MAP, Grade 8: Mean Number of Items per Strand Form 2.....	A-3
Table A-5. DOK Consistency for Science MAP, Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of GLEs Form 1.....	A-4
Table A-6. DOK Consistency for Science MAP, Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of GLEs Form 2.....	A-5
Table A-7. DOK Consistency for Science MAP, Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of GLEs Form 1.....	A-6
Table A-8. DOK Consistency for Science MAP, Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of GLEs Form 2.....	A-7
Table A-9. Range-of-Knowledge for Science MAP, Grade 5: Mean Percent of GLEs per Strand Linked with Items Form 1 .....	A-8
Table A-10. Range-of-Knowledge for Science MAP, Grade 5: Mean Percent of GLEs per Strand Linked with Items Form 2.....	A-9
Table A-11. Range-of-Knowledge for Science MAP, Grade 8: Mean Percent of GLEs per Strand Linked with Items Form 1 .....	A-10
Table A-12. Range-of-Knowledge for Science MAP, Grade 8: Mean Percent of GLEs per Strand Linked with Items Form 2.....	A-11
Table A-13. Balance-of-Knowledge Representation for Science MAP, Grade 5: Mean Balance Index per Strand Form 1 .....	A-12
Table A-14. Balance-of-Knowledge Representation for Science MAP, Grade 5: Mean Balance Index per Strand Form 2.....	A-13
Table A-15. Balance-of-Knowledge Representation for Science MAP, Grade 8: Mean Balance Index per Strand Form 1 .....	A-14
Table A-16. Balance-of-Knowledge Representation for Science MAP, Grade 8: Mean Balance Index per Strand Form 2.....	A-15
Table A-17. Grade 5 MAP: Grade Span GLEs Matched to Items by Panelists .....	A-16
Table A-18. Grade 8 MAP: Grade Span GLEs Matched to Items by Panelists .....	A-21
Table B-1. Grade 5 Science MAP: Summary of Panelists' (N=8) Comments on Items by Topic.....	B-1
Table B-2. Grade 8 Science MAP: Summary of Panelists' (N=8) Comments on Items by Topic.....	B-1



# Independent Alignment Review of the Science Missouri Assessment Program (MAP)

## Executive Summary

### *Scope of Work*

The Human Resources Research Organization (HumRRO) was contracted by the Missouri Department of Elementary and Secondary Education (DESE) to conduct an external, independent alignment study of the Science Missouri Assessment Program (MAP). The evaluation included a review and analysis of the alignment of the 2015 MAP for grades 5 and 8 to the 2009 Revision of Missouri Learning Standards for Science<sup>1</sup>. The Science MAP is a grade-span assessment. This means that the Grade Level Expectations (GLEs) from three grade levels are included in each evaluation. Thus, the grade 5 assessment contains items that may assess GLEs from grades 3, 4, or 5 while the grade 8 assessment contains items that may assess GLEs from grades 6, 7, or 8. Students will receive one of two forms with either 40 or 41 items in grade 5 and either 38 or 39 items in grade 8, totaling 60 points in either grade.

DESE requested the alignment study in order to meet both state and federal accountability requirements related to its use of the Science MAP. The federal requirement of the U.S. Department of Education (USDE) stems from the Elementary and Secondary Education Act (ESEA). ESEA challenges each state to establish a coherent assessment system based on solid academic standards. This law calls for states to provide independent evidence of the validity of their assessments used to calculate Adequate Yearly Progress (AYP). All states receiving Title I funds must present evidence that their assessment system is consistent and fair, that it is based on rigorous standards with sufficient alignment between standards and assessments, and that it generates high-quality educational results.

An alignment review can provide one form of evidence supporting the validity of the state assessment system. Alignment results should demonstrate that the assessments represent the full range of the content standards and that the assessments measure student knowledge in the same manner and at the same level of complexity as specified in the content standards. All aspects of the state assessment system must coincide and be coherent. Aspects of the system include its academic content standards, achievement standards (as linked to cut scores), performance level descriptors, and each individual assessment within the system.

### *Methodology*

To conduct the study, HumRRO facilitated a review of the alignment between the 2015 Science MAP items and the 2009 Revision of Missouri Learning Standards for Science by two panels of current Missouri educators. The panelists also evaluated items on universal design characteristics, namely written content and figures/graphics as well as overall item quality. Following the reviews and examination of the alignment, HumRRO analyzed the results and presents them in this report.

---

<sup>1</sup> Missouri Learning Standards can be found at <http://dese.mo.gov/college-career-readiness/curriculum/missouri-learning-standards>

## **Review of Content Alignment**

HumRRO convened panels of Missouri educators to review the extent of the alignment between the grades 5 and 8 Science MAP and the standards it is intended to assess. The review involved two major tasks for panelists to complete: (a) providing depth of knowledge (DOK) ratings for the Missouri Learning Standards for Science, and (b) evaluating the Science items by matching them to grade level Missouri Learning Standards for Science, providing an item DOK rating, selecting an overall alignment rating, determining presence of universal design characteristics for written content and figures/graphics, and rating overall item quality. To maintain the independent and external nature of the study, DESE did not take part in this process. This process was conducted and directed solely by HumRRO.

HumRRO developed two review panels with the administrative assistance of DESE. Panelists were recruited by HumRRO from a database of Missouri educators provided by DESE. Every effort was made to produce panels consisting of teachers reflecting the population of students who take the assessments. Once selected, the panels were convened at the University of Missouri Assessment Resource Center (ARC) in Columbia, MO on September 25–26, 2014. Each panel included eight reviewers, referred to as panelists.

To conduct the content alignment review, HumRRO applied the Webb (2005) alignment method as recommended by the Missouri Technical Advisory Committee (TAC). This procedure, developed by Dr. Norman Webb, is based on four indicators (or statistics) using the data gathered from the two tasks mentioned above. These statistics provide data and measurement in determining how well the items on the assessment, regardless of item type and point value, cover the content standards in terms of content breadth and depth. The alignment indicators include:

- Categorical concurrence – determines the degree of overall content coverage by the assessment for each content strand. Webb recommends a minimum of six test questions to adequately assess each content strand.
- Range-of-knowledge correspondence – indicates the specific content expectations (e.g., standard, GLE) assessed within each strand. Webb recommends at least 50% of the GLEs per strand are linked with items.
- Balance-of-knowledge representation – provides a statistical index reflecting the distribution of assessed content within each strand (i.e., how evenly the content is assessed). Webb recommends a minimum index of 70 for a single content strand.
- Depth-of-knowledge consistency – compares the cognitive complexity ratings of the items with the complexity ratings of each content standard. Webb recommends that at least 50% of the items should have complexity ratings at or above the level corresponding to GLEs as determined by panelists.

## **Review of Test Quality**

Alignment of assessments to the state content standards serves as one form of test validity evidence. Other areas of validity are critical as well, such as whether the assessment enables students to demonstrate what they know. To address this, panelists evaluated the Science MAP items on several dimensions. During the second task listed above, panelists evaluated items on universal design characteristics—namely written content and figures/graphics. Item ratings for written content and figures/graphics were based on a simple yes–no evaluation. For written

content, panelists rated the language used in the items for the extent to which students of various backgrounds and ability levels could access the Science content being assessed. The same evaluation was made for those Science items accompanied by pictures, figures, or graphs. One important note is that the Science MAP is administered online; however, panelists only viewed the paper version of items. Finally, panelists provided a general rating reflecting their judgment of overall item quality that included aspects such as clarity and appropriateness.

## Summary of Results

### Key Findings and Conclusions

Overall, the results of the alignment review provide positive support for the content validity of the Science MAP for each grade (5 and 8) based on two outcomes. First, panelists found that the test items assessed a level of cognitive complexity that was at or above the cognitive complexity level of the GLE associated with each item for both forms in each grade. Second, items were distributed rather evenly across the content expectations. However, there were issues identified with the categorical concurrence and the range-of-knowledge results. The categorical concurrence results indicate there are an insufficient number of items on a form to cover the content strands, and the range-of-knowledge results imply a restricted range of content assessed by items. These findings stem from more GLEs than items being available for the assessment. In turn, this is a direct result of the Science MAP assessment being a grade-span assessment, and the way in which the Webb indicators do not account for the state's intentions/emphasis of content particularly for a grade span assessment.

### Alignment of Science MAP to Missouri Learning Standards

Table 1 provides summary conclusions on the alignment of the Science MAP to the Missouri Learning Standards per grade tested. The conclusions are based on the following decision criteria (Webb, 2005):

- Fully aligned – assessments align to all content strands (91%–100%);
- Highly aligned – assessments align to the majority of strands (70%–90%);
- Partially aligned – assessments align well to some strands (50%–69%);
- Weakly aligned – assessments align to less than half the strands (below 50%).

Webb's alignment method does not allow for a *single* judgment of overall alignment across the four alignment indicators. However, one can get a sense of overall alignment between the assessments and standards by looking at all of the alignment indicators together.

**Table 1. Summary Alignment Outcomes on Each Webb Criterion by Grade Level for Science MAP**

Grade Assessment	Forms	Percentage of Strands that Met Webb Criteria			
		Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Knowledge Representation
5	Form 1	Weakly aligned (25%)	Highly aligned (88%)	Weakly aligned (0%)	Fully aligned (100%)
	Form 2	Weakly aligned (25%)	Highly aligned (88%)	Weakly aligned (0%)	Fully aligned (100%)
8	Form 1	Weakly aligned (25%)	Fully aligned (100%)	Weakly aligned (0%)	Fully aligned (100%)
	Form 2	Weakly aligned (13%)	Highly aligned (88%)	Weakly aligned (0%)	Fully aligned (100%)

Overall, panelists’ findings, based on the minimum criteria of Webb’s indicators, show that the MAP Science assessments in grades 5 and 8 were not aligned in overall content breadth. In particular, the range of knowledge correspondence indicates that none of the strands met the minimum criteria of having 50% of the GLEs within a strand matched to an item. The highest percentage of GLEs within a strand matched to an item occur in the ‘Changes in Ecosystems and Interactions of Organisms with their Environments’ (EC) strand with 34–38% in grade 5 and 18–28% in grade 8. However, the weak categorical concurrence and the restricted range of content is not a large concern as the lower percentages in both grades appear to be a result of a large number of GLEs available to be assessed compared to the number of items on the test, 171 GLEs to 40/41 items in grade 5 and 264 GLEs to 38/39 items in grade 8.

Panelists found that the majority of the items assessed student knowledge at the same or higher level of cognitive complexity than expected in the content standards, as evidenced by the depth-of-knowledge consistency results. In addition, the balance-of-knowledge results suggest that items seem to be distributed reasonably, at least across GLEs matched by panelists.

### **Quality of Science MAP Tests**

Table 2 presents the summary outcomes on the item quality ratings. The table includes conclusions regarding the quality of the items on each assessment, along with the percentage of items that received favorable ratings. The conclusions are based on the following decision criteria (adapted from Thompson, Johnstone, Anderson, & Miller, 2005):

- Excellent – all items are acceptable;
- Good – most items are acceptable (at least 90%);
- Acceptable – many items are acceptable (70%–90%);
- Questionable – few items are acceptable (less than 70%).

**Table 2. Item Quality Ratings for Science MAP by Grade**

Grade	Forms	Percentage of Items with Acceptable Ratings		
		Written Content	Graphics	Overall Item Quality
5	Form 1	Good (94%)	Acceptable (89%)	Acceptable (87%)
	Form 2	Good (95%)	Good (92%)	Good (92%)
8	Form 1	Good (96%)	Good (95%)	Good (92%)
	Form 2	Good (93%)	Good (90%)	Good (92%)

As a whole, the independent item ratings suggest that the Science MAP items operate as they should for the majority of students who take these assessments. A number of items on one of the grade 5 forms may require review to enhance clarity in accompanying graphics and reduce potential bias against student subgroups.

### Recommendations

HumRRO makes the following recommendations to strengthen the alignment between the components of the Missouri assessment system:

- **Review content coverage (categorical concurrence).** The panelists reviewing these assessments found that in grade 5 only two of the eight strands in the Science Missouri Learning Standards are assessed with at least six items. In grade 8, the same can be seen for form 1, while on form 2 only one of the eight strands is assessed with at least six items. Thus, the assessments may not adequately reflect the content that students are expected to know based solely on the number of items on the assessment (not the item type or point value as these are not factors in Webb’s categorical concurrence indicator). From strictly an item count perspective, there are several ways DESE can choose to mitigate this situation such as increase the number of items on the assessment, reduce the number of strands in the state standards, or designate some of the strands for local assessment only.
- **Review whether the Science MAP is a grade-level or grade-span assessment.** With the large number of GLEs across the grade-span, being able to link each GLE with at least one item is impractical. Identifying the Science MAP as a grade-level assessment would reduce the number of GLEs substantially when evaluating the range-of-knowledge correspondence if that is the state’s intention of the Science MAP test in grade 5 and 8.
- **Review those items that received the lowest ratings on test quality for possible revision.** Panelists’ review suggests that there are issues with items aligning to multiple GLEs and a few items not aligning to any GLE. They also identified some items that could be improved by changes in word clarity, language, and specificity level. In particular, some items were considered too broad to appropriately assess a single GLE.



# Independent Alignment Review of the Science Missouri Assessment Program (MAP)

## Chapter 1: Introduction

The Human Resources Research Organization (HumRRO) was contracted by the Missouri Department of Elementary and Secondary Education (DESE) to conduct an external, independent alignment study of the Science Missouri Assessment Program (MAP), which occurred September 25–26, 2014. The evaluation included a review and analysis of the alignment of the 2015 MAP for grades 5 and 8 to the 2009 Revision of Missouri Learning Standards for Science<sup>2</sup>. The Science MAP is a grade-span assessment. This means that the Grade Level Expectations (GLEs) from three grade levels are included in the evaluation of each assessment. Thus, the grade 5 assessment contains items that may assess GLEs from grades 3, 4, or 5 while the grade 8 assessment contains items that may assess GLEs from grades 6, 7, or 8. Students will receive one of two forms with either 40 or 41 items in grade 5 and either 38 or 39 items in grade 8, totaling 60 points in either grade.

DESE requested the alignment study in order to meet both state and federal accountability requirements related to its use of the Science MAP. The federal requirement of the U.S. Department of Education (USDE) stems from the Elementary and Secondary Education Act (ESEA). ESEA challenges each state to establish a coherent assessment system based on solid academic standards. This law calls for states to provide independent evidence of the validity of their assessments used to calculate Adequate Yearly Progress (AYP). All states receiving Title I funds must present evidence that their assessment system is consistent and fair, that it is based on rigorous standards with sufficient alignment between standards and assessments, and that it generates high-quality educational results.

An alignment review can provide one form of evidence supporting the validity of the state assessment system. Alignment results should demonstrate that the assessments represent the full range of the content standards and that the assessments measure student knowledge in the same manner and at the same level of complexity as specified in the content standards. All aspects of the state assessment system must coincide and be coherent. Aspects of the system include its academic content standards, achievement standards (as linked to cut scores), performance level descriptors, and each individual assessment within it.

### *Organization and Contents of the Report*

This report contains five chapters. Chapter 2 explains the alignment methodologies used in the study and chapters 3 and 4 provide alignment results for comparisons between the components of the assessment system. Chapter 3 presents results of the alignment comparison between the Science assessments and the Missouri Learning Standards; Chapter 4 presents results on the accessibility of the assessments to all students; and Chapter 5 provides recommendations for DESE to strengthen the alignment of the Science MAP over time.

Additional information is provided in the appendices of this report. Appendix A contains tables with additional details for each Webb indicator regarding the content alignment results for each grade-span test form in both grades, Appendix B contains a synthesis of panelists' comments on items, and Appendix C provides examples of rating forms and training materials used in the alignment workshops.

---

<sup>2</sup> Missouri Learning Standards can be found at <http://dese.mo.gov/college-career-readiness/curriculum/missouri-learning-standards>



## Chapter 2: Alignment Study Design and Methodology

In this section, we discuss key concepts related to standards to assessment alignment research. This discussion is followed by a description of the alignment evaluations and methods used for this study.

### *Alignment of Assessments and Standards on Content*

Alignment studies, at their heart, answer one vital question related to the validity of an assessment, “Does the assessment content adequately reflect the content that students are expected to learn as provided in the state standards?” School curriculum must include appropriate content also meet the goals specified by the state standards and consequently assessments must measure the same content.

In general, alignment evaluations for any Kindergarten to grade 12 educational assessments in the United States reveal (a) the breadth, or scope, of knowledge and (b) the depth of knowledge, or cognitive processing, expected of students by the state’s content standards. In addition to the question related to assessment validity, alignment analyses help to answer questions such as the following:

- How much and what type of content is covered by the assessment?
- Are students asked to demonstrate this knowledge at the same level of rigor as expected in the content standards?

### *Content Alignment and Accessibility*

Several methods of alignment are in current use (e.g., Porter, 2002; Webb, 1997, 1999, 2005). These methods involve panelists subjectively evaluating several aspects of the assessment items relative to the content standards. The data from the evaluations are analyzed statistically to determine the extent of alignment. HumRRO used the alignment method developed by Norman Webb (1997; 1999; 2005) to evaluate the Science MAP. Webb’s alignment methodology is the most widely used.

### *Webb Alignment Method*

The Webb alignment method was originally designed for use with standard large-scale assessments. Dr. Webb has researched and refined this method over time (e.g., Webb, 1997; 1999; 2005), and his approach is supported by the Council of Chief State School Officers (CCSSO).

The Webb method includes four major indicators to evaluate alignment. These indicators link with statistical procedures used to assess how well items on the assessment, regardless of item type and point value, and the state’s standards document actually match. The four alignment indicators are: categorical concurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance-of-knowledge representation.

**Categorical concurrence** is a basic measure of alignment between content standards and test items. This term refers to the proportion of overlap between the content stated in the standards document and that assessed by items on the test.

**Depth of Knowledge (DOK)** measures the type of cognitive processing required by items and content standards. For example, is a student expected to simply identify or recall basic facts or use reason to manipulate information, or to strategize how to best solve a complex problem? Using Science as an example, a student may be asked to identify the planets of our solar system among several answer choices. This task should be less complex than trying to compare and contrast the composition of the planets in preparation of planning the landing of unmanned probes.

The purpose of using DOK as a measure of alignment is to determine whether a test item and its corresponding standard are written at the same level of cognitive complexity. Panelists make two separate judgments about cognitive complexity, one for the standard and one for the item. These two judgments are compared to determine whether the item is written at the same level as the standard to which it is linked. Webb refers to this comparison as *Depth-of-Knowledge consistency*.

**Range-of-knowledge correspondence** examines the range-of-knowledge correspondence between the assessment and content standards. The range-of-knowledge correspondence measure looks in greater detail at the breadth of knowledge represented by test items. Categorical concurrence simply notes whether a sufficient number of items on the test covers each general content topic (individual strands). However, states usually lay out more specific *content objectives*, or standards, under each strand. The range-of-knowledge correspondence indicates the number of content objectives assessed by items.

**Balance-of-knowledge representation** focuses on content coverage in yet more detail. In this case, the number of items matched to the content objective does matter. The balance of representation determines whether the assessment measures the content objectives equitably within each standard using only the content objectives identified by panelists and not all content objectives eligible to be assessed. Based on Webb's method, items should be distributed evenly across the objectives per standard for good balance. The balance-of-knowledge representation is determined by calculating an index, or score, for each standard. Each standard should meet or surpass a minimum index level to demonstrate adequate balance.

### ***Scope of Alignment Evaluations for Science MAP***

The alignment evaluation performed for this study involved a comparison of the Science MAP to the Missouri Learning Standards. Missouri educators highly familiar with the content standards and the assessment provided alignment ratings for the evaluation. To maintain the independent and external nature of the study, DESE did not take part in this process. This process was conducted and directed solely by HumRRO.

### ***Review of Content Alignment and Accessibility***

For the content alignment review, HumRRO convened panels of Missouri educators to review the grades 5 and 8 Science MAP forms. The review involved two major tasks for panelists to complete: (a) providing depth of knowledge (DOK) ratings for the Missouri Learning Standards for Science, and (b) evaluating the Science items by matching them to grade level Missouri Learning Standards for Science, providing an item DOK rating, selecting an overall alignment rating, determining presence of universal design characteristics for written content and figures/graphics, and rating overall item quality.

## Panelists

HumRRO developed two review panels with the administrative assistance of DESE. Panelists were recruited by HumRRO from a database of Missouri educators provided by DESE. Every effort was made to produce panels consisting of teachers reflecting the population of students who take the assessments. Panels were convened at the University of Missouri Assessment Resource Center (ARC) in Columbia, MO on September 25–26, 2014. Each panel included eight reviewers, referred to as panelists. Table 2.1 presents the characteristics of the panels per grade-level of the Science MAP.

**Table 2.1. Professional and Demographic Characteristics of Science MAP Panelists**

Professional Position	Number of Panelists	Average Years of Experience	Special Certifications	Region of Origin in Missouri							Gender	
				South-west	St Louis	North-west	Central	South-east	Columbia	Kansas City	M	F
Grade 5												
Teacher	8	16.4	8	2	1	1	1	1	1	1	1	7
Administrator	0	0	0	0	0	0	0	0	0	0	0	0
Grade 8												
Teacher	7	15.6	7	0	2	1	0	2	0	2	1	6
Administrator	1	7	1	0	0	0	0	0	0	1	0	1

## Training

One week prior to the alignment study workshop, panelists were required to take part in a one-hour training session via a web teleconference service that allowed panelists to view the screen of a HumRRO trainer through an internet browser. The training covered general alignment study information, roles and responsibilities, key alignment concepts, security and confidentiality concerns, and the alignment workshop procedures. This initial training allowed HumRRO facilitators to immediately begin specific alignment task training in the workshop grade span panels prior to panelists beginning their evaluations.

## Materials

During the alignment workshop, panelists evaluated the alignment of the MAP items with the Missouri Learning Standards using printed test booklets and electronic rating forms adapted from Webb (2005). All rating forms were completed electronically in Excel®. The test booklets and rating forms are discussed in further detail below.

**Test Forms.** Panelists evaluated the 2015 Science MAP operational items per form. Table 2.2 lists the characteristics of the forms for the 2015 administration for each grade-span test. The Science MAP tests are administered as online assessments. Many items include dynamic graphics that demonstrate concepts or require student interaction to formulate a response. The 2015 Science MAP test contains Missouri-owned items and licensed items. Even though the test is administered online, panelists viewed the items on paper, as the online versions of the items were not yet available. Because the test form is a secure document, this report does not include any examples of items or references to specific item content.

**Table 2.2. Characteristics of 2015 Science MAP Test Forms**

Grade Level	Form	Total Items per Form	Number of Missouri Items	Number of Licensed Items
5	1	41	23	18
	2	40	22	
8	1	39	23	16
	2	38		15

**Rating Forms and Instructions.** Panelists were given instruction sheets describing the rating tasks, the codes to be used, and the excel documents used during their review (see Appendix C). Panelists completed two rating forms, the first was completed as a group (by consensus) to provide depth of knowledge (DOK) ratings for the content GLEs and the second form, an item rating form, captures individual ratings for the items (see Appendix C for samples of each).

### Procedures

HumRRO conducted the alignment study at the University of Missouri ARC, in Columbia, MO. The workshop began with a general session that included introductions of staff and observers followed by a brief review of the agenda for the two-day workshop. Panelists then moved to their grade span groups to receive specific alignment task training before starting to work. The Science MAP panelists were split into two groups, one for grades 3–5 and another for grades 6-8. Both groups contained eight panelists and were facilitated by HumRRO staff members. Prior to beginning their review, panelists read and signed affidavits of nondisclosure for the secure materials they would be reviewing during the workshop.

Before each of the rating tasks, a HumRRO staff member trained panelists on the procedures to complete the task, answered questions on the rating criteria, and conducted a short calibration activity to ensure panelists were comfortable applying ratings. HumRRO staff provided general suggestions and comments when appropriate; however, they emphasized to panelists that staff would not give explicit direction on how to rate standards or items because panelists were valued as content experts. Each panelist was assigned a workstation with rating forms already uploaded. HumRRO staff provided instructions as needed for working with the electronic rating forms.

Panelists began with DOK evaluations of the content GLEs. Panelists independently assigned DOK levels to the first few GLEs and then the group reviewed and discussed each of the initial ratings. This was done as an initial calibration to familiarize panelists with assigning DOK ratings. Panelists then proceeded to rate the remaining GLEs from the Missouri Learning Standards relevant to each grade span test individually. For example, panelists reviewing the grade 5 test rated the GLEs for grades 3, 4, and 5. As panelists completed their DOK ratings for one grade, they discussed their ratings to achieve consensus for each GLE before working on their next grade level. A volunteer scribe within each group recorded these consensus ratings.

Panelists then received specific instructions for rating the items. As a calibration activity, HumRRO staff asked panelists to rate the first few items individually and then discuss the ratings as a group. Once panelists were comfortable using the ratings, they continued the item rating activity on their own. Panelists rated the individual items on the test forms on several dimensions: (a) depth of knowledge required by the item, (b) content match to the GLEs in the

Missouri Learning Standards, (c) degree of alignment (i.e., how well the item links to the GLE), (d) content clarity (i.e., readability), (e) quality of accompanying graphics (if applicable), and the overall item quality. Panelists also assigned a *primary GLE* to an item based on a judgment that an item clearly measured this content. Panelists could assign an *additional GLE* if the item seemed to assess another GLE as well (or nearly as well) as the primary GLE. Again, these were individual ratings, not consensus.

All panelists finished their rating tasks within the 2 days allotted for the workshop. Once panelists finished the review, their session ended.



## Chapter 3: Results: Content Alignment

The content alignment evaluation analyses are based on panelists' ratings of the Science MAP items. The grade 5 and 8 forms consist of Missouri-owned items and licensed items. At the beginning of September, a group of licensed items were selected for inclusion on the Science MAP and evaluated during the alignment workshop. The licensed items were selected to match missing content in the test blueprint that could not be accounted for by the Missouri-owned items. It is important to note that the licensed items were not developed to align with the Missouri Learning Standards in Science. Panelists' evaluations of the licensed items in relation to the Missouri Learning Standards were an important piece of information to obtain to assist DESE in constructing the 2015 Science MAP test. After the alignment workshop, panelists' review of the licensed items was provided to DESE. With this information, DESE worked with the vendor in November to identify new items, which better matched content than the first items chosen, to include on the assessment.

With the inclusion of items that were not developed to match the Missouri Learning Standards in Science, the alignment study before the administration is even more important. This allowed DESE to make adjustments by removing items that were potentially not evaluating content in the Missouri Learning Standards. Alignment studies are best conducted before an assessment is delivered, in general, to ensure that the content that is intended to be covered by the assessment is indeed being covered. Thus, time is allotted to make changes to the items before administration with the goal of strengthening the match between the content intended to be covered and the content that is actually being assessed.

As a result of the inclusion of the new licensed items, some of which are common across forms, a total of 9 items in grade 5 across the two forms and a total of 3 items in grade 8 across the two forms were not evaluated during the alignment workshop. The alignment results presented in this chapter and the next are based on the subset of items on each assessment that panelists evaluated during the alignment workshop.

### *Reliability Results*

In this section, we report on the comparison of panelists' ratings of content match to the item bank's documented content match. In other words, do panelists assign the same GLE to an item as the item writer during item development?

#### *Panelist-Test Developer Analyses*

This analysis examined the agreement outcomes between the GLE assigned to an item by panelists, and the GLE assigned to an item as noted in the item bank on the content assessed by each item on each form for grade 5 and 8. For the licensed items, a GLE was assigned by the vendor and DESE that was deemed a best fit to the item content. Table 3.1 shows a breakdown of an item bank code that may be assigned to an item. The item bank codes are direct reflections of the Missouri Learning Standards in Science.

**Table 3.1. Missouri Learning Standards Example**

Code	Standards Grade	Strand	Big Idea	Concept	Grade Level Expectation (GLE)
ME.1.C.5.a	5	Properties and Principles of Matter and Energy (ME)	1. Changes in properties and states of matter provide evidence of the atomic theory of matter.	C. Properties of matter can be explained in terms of moving particles too small to be seen without tremendous magnification.	a. Describe how changes in state (i.e., freezing/melting, condensation/evaporation/boiling) provide evidence that matter is made of particles too small to be seen.

Table 3.2 presents the agreement outcomes between panelists and the item bank on the content assessed by items. Agreement was analyzed at several levels of specificity. All of the items were analyzed first for ‘Exact Match’, which indicates that panelists chose the same standard grade, Strand, Big Idea, Concept, and GLE for the item as documented in the item bank. If panelists did not show an exact match with the item bank, we determined the percent agreement at the *Concept* level (panelists selected the same standard grade, Strand, Big Idea, and Concept as in the item bank). For each following step, a match between the next highest level (Big Idea, Strand, standards grade) was determined. The last column in Table 3.2 shows the percentage of ratings by panelists that did not match the item bank coding at all on items. Because panelists could assign more than one content code to a single item if they felt that another GLE measured an item as well (or nearly as well), we counted all content codes assigned to items. For this reason, the value in the ‘Total Number of Panelists Ratings across Items’ does not equal the number of items evaluated on a form multiplied by the 8 panelists. Additionally, a panelist could choose not to assign a content code to an item if they felt that none of the GLEs adequately measured the item.

**Table 3.2. Percent Agreement between Panelists and Item Bank on Target Content – Grade 5**

Grade	Form	Number of Items per Form	Total Number of Panelist Ratings across Items	Percent Agreement with Item Bank Codes					
				Exact Match	Concept Match	Big Idea Match	Strand Match	Standards Grade Match	No Match
5	1	35 of 41	318	53%	9%	5%	9%	3%	20%
	2	35 of 40	318	57%	10%	5%	8%	<1%	20%
8	1	36 of 39	335	67%	11%	9%	2%	<1%	12%
	2	36 of 38	337	65%	15%	2%	1%	4%	12%

As Table 3.2 indicates, panelists were moderately consistent with the item bank in identifying the assessment content codes of items. Panelists identified an exact match for 53–67% of the items and a content match at the strand level or below for 78–89% of the items. Panelists differed completely from the item bank on content match for 12–20% of the items. Several factors may have contributed to panelists’ difficulty in matching some of the items. First, items needed to cover content from three grades (3, 4, 5 and 6, 7, 8) instead of just one (i.e. there may be similar standards across grades causing potential mismatches). Second, a subset of items on each form is licensed items. This means the items were not developed with the Missouri GLEs as reference points but instead were selected for use and a GLE was assigned to them. Overall however, these findings suggest that the majority of Science items do, in fact, measure the intended content.

## *Webb Alignment Results*

In this section, we review the general outcomes of item analyses on the four Webb alignment indicators.

All of Webb’s measures begin with calculations for each panelist and build up to a summary of results across panelists per content strand. First, we calculated the mean ratings across items for each panelist, and then we determined the mean rating across panelists per strand. Results are presented at the strand level. In Missouri, content strand refers to the highest level of distinction in the Missouri Learning Standards.

### *Categorical Concurrence*

Categorical concurrence describes the extent to which the MAP items, regardless of item type and point value, cover the content strands in the Missouri Learning Standards for Science. Webb recommends a minimum of six test questions to adequately assess each content strand. This criterion serves as a guideline for reasonable content coverage based on earlier research on the reliability of tests compared to the number of items (Subkoviak, 1988). Tables 3.3 and 3.4 summarize the MAP alignment results for categorical concurrence for each grade span. The strands that meet Webb’s indicator criterion are in bold. Tables A-1 through A-4 in Appendix A also contain the standard deviations for each strand.

**Table 3.3. Summary of Categorical Concurrence Results for Science MAP – Grade 5**

Strands	Mean Number of Items per Form	
	Form 1	Form 2
Properties and Principles of Matter and Energy (ME)	3.00	3.00
Properties and Principles of Force and Motion (FM)	2.38	3.00
Characteristics and Interactions of Living Organisms (LO)	3.38	4.00
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	<b>7.38</b>	<b>7.13</b>
Processes and Interactions of the Earth’s Systems (ES)	3.38	4.88
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	3.63	3.00
Scientific Inquiry (IN)	<b>9.50</b>	<b>8.00</b>
Science, Technology, and Human Activity (ST)	2.00	2.00
<b>Strands with at Least Six Items</b>	<b>2 of 8</b>	<b>2 of 8</b>

**Table 3.4. Summary of Categorical Concurrence Results for Science MAP – Grade 8**

Strands	Mean Number of Items per Form	
	Form 1	Form 2
Properties and Principles of Matter and Energy (ME)	4.25	5.00
Properties and Principles of Force and Motion (FM)	1.75	1.75
Characteristics and Interactions of Living Organisms (LO)	4.00	4.50
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	4.25	4.50
Processes and Interactions of the Earth's Systems (ES)	<b>6.00</b>	4.13
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	3.75	4.25
Scientific Inquiry (IN)	<b>10.25</b>	<b>9.88</b>
Science, Technology, and Human Activity (ST)	1.75	2.00
<b>Strands with at Least Six Items</b>	<b>2 of 8</b>	<b>1 of 8</b>

As Tables 3.3 and 3.4 indicate, none of the assessments include a sufficient number of items to meet the minimum requirements for categorical concurrence on all Science content strands. The grade 5 forms have two content strands covered by the minimum number of items, according to Webb's indicator criterion. A similar outcome is seen in grade 8 where one form has two content strands covered by the minimum number of items while the other form only has one content strand. These results indicate that the Science MAP does not adequately cover the Science content students are expected to know at either grade level. However, this result is not surprising given the number of items on the test, 35 items out of 40 or 41 items in grade 5 and 36 items out of 38 or 39 items in grade 8. In order for Webb's indicator criterion to be met in either grade, a minimum of 48 items would need to be assessed with 6 items from each strand. This does not occur in either grade assessment.

The one strand that consistently met the Webb criterion in both grades was 'Scientific Inquiry' which is assessed by the performance event items. According to the test blueprint, 7–9 performance event items should be on each form. This test blueprint criterion for the performance event items is met in both grades. Unfortunately, Webb's criterion for categorical concurrence does not take into account the test blueprint for an assessment. Even though the test blueprint is met in this one particular instance, overall the majority of strands do not meet Webb's indicator criterion.

In addition to identifying the GLE assessed by each item, we asked panelists to indicate *how well* the item assessed the GLEs. Panelists subjectively rated the extent of item alignment to the GLEs on a 4-point scale ranging from 'Not aligned to any GLE' to 'Fully aligned'. Tables 3.5 and 3.6 present the mean number of items (across panelists) at each level of alignment. For each grade assessment, panelists rated items as well aligned to the GLEs matched to that item.

**Table 3.5. Panelist Ratings on Overall Item Alignment Grade per Forms – Grade 5**

Forms	Degree of Alignment	Mean Number of Items (N=35) per Level	SD	Percent of Items per Level
1	Not at all aligned	1.50	0.71	1.07
	Weakly aligned	3.67	2.50	7.86
	Highly aligned	22.00	7.91	62.86
	Fully aligned	9.88	6.83	28.21
2	Not at all aligned	0.00	0.00	0.00
	Weakly aligned	3.00	1.58	5.36
	Highly aligned	20.13	6.75	57.50
	Fully aligned	13.00	6.00	37.14

**Table 3.6. Panelist Ratings on Overall Item Alignment Grade per Forms – Grade 8**

Forms	Degree of Alignment	Mean Number of Items (N=36) per Level	SD	Percent of Items per Level
1	Not at all aligned	1.00	n/a	0.35
	Weakly aligned	2.25	0.96	3.13
	Highly aligned	16.33	11.55	34.03
	Fully aligned	22.50	13.14	62.50
2	Not at all aligned	1.00	0.00	0.70
	Weakly aligned	2.40	1.14	4.20
	Highly aligned	18.00	13.55	37.76
	Fully aligned	23.43	14.09	57.34

In general, panelists across all grades rated at least 96% of the items as being ‘Highly aligned’ or ‘Fully aligned’. The grade 5 assessment had the most items rated by panelists as being ‘Weakly aligned’ or ‘Not at all aligned’ at 5–9%.

### **Depth-of-Knowledge Consistency**

Analyses of depth-of-knowledge (DOK) measure the type of cognitive processing required of students by content standards. The DOK requirements implied by the GLEs should be matched by assessment items. To confirm this match, panelists were asked to rate the GLEs and the Science items separately. Webb includes an alignment indicator that directly compares panelists’ DOK ratings of content standards and test items, which he refers to as *depth-of-knowledge consistency*.

To make their ratings, panelists used a rating scale (adapted from Webb, 2005) with four levels of cognitive complexity.

- Level 1 Recognition – simple recall of information (i.e., facts, terms); sequencing; more automatic.
- Level 2 Skills/Concepts – beyond habitual response; applying concepts; problem-solving.
- Level 3 Strategic Thinking – requires basic reasoning, planning, or use of evidence; generating hypotheses.

- Level 4 Extended Thinking – complex reasoning; evaluation of multiple sources or independent pieces of evidence; often over an extended period of time.

Tables 3.7 and 3.8 summarize the depth-of-knowledge consistency results for each grade level of the Science MAP. Because panelists evaluated depth of knowledge at the most specific level of the standards document (GLEs), the table refers to consistency between the items and the GLEs to which they were matched. Results are summarized in terms of the percentage of items with cognitive complexity ratings at or above (more complex than) the rating for the corresponding GLE. Tables A-5 through A-8 in Appendix A contain the means and standard deviations for DOK ratings at all levels.

Webb’s suggested criterion for this alignment indicator is that at least 50% of the items should have complexity ratings at or above the level of the corresponding GLE. The percentages on strands that reach the 50% criterion are bolded.

**Table 3.7. Summary of Depth-of-Knowledge Results for Science MAP – Grade 5**

Strand	Percent of Items with DOK At or Above the Level of the GLEs per Form	
	Form 1	Form 2
Properties and Principles of Matter and Energy (ME)	<b>87.5</b>	<b>87.5</b>
Properties and Principles of Force and Motion (FM)	<b>83.4</b>	<b>91.7</b>
Characteristics and Interactions of Living Organisms (LO)	<b>79.2</b>	<b>80.2</b>
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	<b>67.9</b>	<b>77.1</b>
Processes and Interactions of the Earth’s Systems (ES)	<b>100.0</b>	<b>56.3</b>
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	<b>93.3</b>	<b>91.7</b>
Scientific Inquiry (IN)	<b>56.2</b>	<b>59.7</b>
Science, Technology, and Human Activity (ST)	43.8	37.6
<b>Number of Strands with item DOK at or above GLE DOK</b>	<b>7 of 8</b>	<b>7 of 8</b>

**Table 3.8. Summary of Depth-of-Knowledge Results for Science MAP – Grade 8**

Strand	Percent of Items with DOK At or Above the Level of the GLEs per Form	
	Form 1	Form 2
Properties and Principles of Matter and Energy (ME)	<b>59.0</b>	<b>76.3</b>
Properties and Principles of Force and Motion (FM)	<b>87.6</b>	<b>87.6</b>
Characteristics and Interactions of Living Organisms (LO)	<b>70.8</b>	<b>90.6</b>
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	<b>71.5</b>	46.2
Processes and Interactions of the Earth’s Systems (ES)	<b>80.5</b>	<b>67.3</b>
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	<b>80.2</b>	<b>93.8</b>
Scientific Inquiry (IN)	<b>87.5</b>	<b>91.5</b>
Science, Technology, and Human Activity (ST)	<b>50.0</b>	<b>73.0</b>
<b>Number of Strands with item DOK at or above GLE DOK</b>	<b>8 of 8</b>	<b>7 of 8</b>

In both grades, panelists' ratings using Webb DOK levels indicate that items on almost all of the strands assess students at the appropriate cognitive complexity. In grade 5, items assessing Science, Technology, and Human Activity did not meet the 50% criterion. On only one of the forms in grade 8, panelists' DOK ratings, for items assessing Changes in Ecosystems and Interactions of Organisms with their Environments, were just below the criterion.

### Range of Knowledge Correspondence

The *range-of-knowledge correspondence* measure examines in greater detail the breadth of knowledge covered by the assessment. In addition to evaluating which content strands are assessed, we must look at how many of the GLEs within a strand are represented by items. The GLEs should be linked with at least one item. Webb's minimum level of acceptability for range-of-knowledge correspondence is that at least 50% of GLEs per strand link with items. Table 3.9 provides a breakdown of the number of GLEs per grade level on each assessment.

**Table 3.9. Number of GLEs per Grade Level**

Strand	Number of GLEs							
	Grade 3	Grade 4	Grade 5	Grade 5 Assessment	Grade 6	Grade 7	Grade 8	Grade 8 Assessment
Properties and Principles of Matter and Energy (ME)	12	13	7	32	26	20	12	58
Properties and Principles of Force and Motion (FM)	0	11	6	17	0	18	0	18
Characteristics and Interactions of Living Organisms (LO)	5	0	7	12	7	0	27	34
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	4	11	0	15	12	0	3	15
Processes and Interactions of the Earth's Systems (ES)	4	9	9	22	12	15	12	39
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	8	0	8	16	0	27	0	27
Scientific Inquiry (IN)	12	13	15	40	17	19	19	55
Science, Technology, and Human Activity (ST)	5	6	6	17	6	6	6	18
<b>Total</b>	<b>50</b>	<b>63</b>	<b>58</b>	<b>171</b>	<b>80</b>	<b>105</b>	<b>79</b>	<b>264</b>

Tables 3.10 and 3.11 summarize the range-of-knowledge results for each grade level assessment of the Science MAP per content strand. The strands that meet Webb's indicator criterion are in bold.

**Table 3.10. Summary of Range-of-Knowledge Results for the Science MAP – Grade 5**

Strand	Percent of GLEs per Strand Matched to at Least One Item	
	Form 1	Form 2
Properties and Principles of Matter and Energy (ME)	9.4	9.4
Properties and Principles of Force and Motion (FM)	12.5	16.2
Characteristics and Interactions of Living Organisms (LO)	21.9	26.0
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	34.1	37.5
Processes and Interactions of the Earth's Systems (ES)	12.5	11.5
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	18.0	18.8
Scientific Inquiry (IN)	25.9	23.6
Science, Technology, and Human Activity (ST)	13.6	15.9
<b>Number of Strands Assessed Adequately</b>	<b>0 of 8</b>	<b>0 of 8</b>

**Table 3.11. Summary of Range-of-Knowledge Results for the Science MAP – Grade 8**

Strand	Percent of GLEs per Strand Matched to at Least One Item	
	Form 1	Form 2
Properties and Principles of Matter and Energy (ME)	7.1	7.8
Properties and Principles of Force and Motion (FM)	9.7	9.7
Characteristics and Interactions of Living Organisms (LO)	11.4	11.8
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	18.3	27.5
Processes and Interactions of the Earth's Systems (ES)	14.4	8.7
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	12.0	13.4
Scientific Inquiry (IN)	11.1	10.7
Science, Technology, and Human Activity (ST)	9.0	9.7
<b>Number of Strands Assessed Adequately</b>	<b>0 of 8</b>	<b>0 of 8</b>

None of the assessments in either grade met the minimum range-of-knowledge criterion for any of the strands. This finding is a direct result of the large number of GLEs available to be assessed in comparison to the number of items on the assessment. Tables A-9 through A-12 in Appendix A contain the means and standard deviations for each strand and the number of assessable GLEs per strand.

### **Balance-of-Knowledge Representation**

The fourth measure of alignment included in the Webb method is *balance-of-knowledge representation*. This measure describes the distribution of items linked to each GLE within each strand. The number of items should be distributed rather evenly between the GLEs to achieve good balance.

The content balance is determined by calculating an index, or score, for each strand<sup>3</sup>. According to Webb, the minimum acceptable index for a single strand is 70 (on a scale of 0 to 100 with 100 representing perfect balance). An index of 70 or higher suggests that items broadly assess the GLEs for a strand instead of clustering around one or two GLEs.

Two cautions should be noted regarding the balance index when interpreting the results. First, only those GLEs actually matched to items by the panelists are included in calculations of the balance index. A given strand may include more GLEs than are actually linked to items by panelists. For example, if a particular strand includes eight GLEs in the state content standards document but panelists found items matching to just three GLEs, only these three GLEs are evaluated for item distribution. Recognizing this feature of the balance index is important in cases when the range measure and balance measure produce seemingly contrasting results. And second, when states choose to emphasize particular content strands over others, the balance statistic becomes uninterpretable. Missouri does not emphasize any particular content strands on the Science MAP.

Tables 3.12 and 3.13 summarize the results on balance-of-content representation per grade for the Science MAP. All of the grades assessed surpassed the minimum level of acceptability (index of 70) for demonstrating good content balance among those GLEs matched to items for each strand. The strands that meet Webb’s indicator criterion are in bold. Tables A-13 through A-16 contain means associated with the calculation of the balance index.

**Table 3.12. Summary of Balance-of-Knowledge Representation Results Science MAP – Grade 5**

Strand	Balance Index per Form	
	Form 1	Form 2
Properties and Principles of Matter and Energy (ME)	<b>100</b>	<b>100</b>
Properties and Principles of Force and Motion (FM)	<b>100</b>	<b>96</b>
Characteristics and Interactions of Living Organisms (LO)	<b>92</b>	<b>86</b>
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	<b>100</b>	<b>89</b>
Processes and Interactions of the Earth’s Systems (ES)	<b>96</b>	<b>100</b>
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	<b>93</b>	<b>98</b>
Scientific Inquiry (IN)	<b>82</b>	<b>86</b>
Science, Technology, and Human Activity (ST)	<b>100</b>	<b>98</b>
<b>Number of Strands Assessed Adequately</b>	<b>8 of 8</b>	<b>8 of 8</b>

<sup>3</sup> The exact formula for calculating the balance index is explained in detail in Webb’s (2005) alignment training manual: <http://www.wcer.wisc.edu/WAT/index.aspx> .

**Table 3.13. Summary of Balance-of-Knowledge Representation Results Science MAP – Grade 8**

Strand	Balance Index per Form	
	Form 1	Form 2
Properties and Principles of Matter and Energy (ME)	<b>98</b>	<b>90</b>
Properties and Principles of Force and Motion (FM)	<b>100</b>	<b>100</b>
Characteristics and Interactions of Living Organisms (LO)	<b>98</b>	<b>96</b>
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	<b>80</b>	<b>95</b>
Processes and Interactions of the Earth’s Systems (ES)	<b>98</b>	<b>86</b>
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	<b>92</b>	<b>90</b>
Scientific Inquiry (IN)	<b>74</b>	<b>81</b>
Science, Technology, and Human Activity (ST)	<b>100</b>	<b>100</b>
<b>Number of Strands Assessed Adequately</b>	<b>8 of 8</b>	<b>8 of 8</b>

### **Summary and Discussion on Webb Alignment Indicators**

The overall alignment results provide positive support for the content validity of the Science MAP for each grade (5 and 8) based on two outcomes. At each grade level, the assessments met, to the full extent, the minimum requirements for two of the Webb indicators. First, panelists found that the test items assessed a level of cognitive complexity that was at or above the cognitive complexity level of the GLE associated with each item for both forms in each grade. Second, items were distributed rather evenly across the content expectations as indicated by the Balance Index for both forms in each grade. Results on the alignment indicator categorical concurrence suggest that the Science MAP does not adequately cover the Science content while results on range-of-knowledge imply a restricted range of content assessed by items resulting from there being more GLEs than items available for the assessment. This finding is a direct result of the Science MAP assessment being a grade-span assessment, and the way in which the Webb indicators do not account for the state’s intentions/emphasis of content particularly for a grade span assessment. We present summary alignment judgments for the Science MAP in this section based on the statistical outcomes.

Summary alignment judgments are based on Webb (2005). These summary judgments focus on the percentage of content strands represented well by the assessment. Webb outlined a scale with a range of potential alignment outcomes applied to each of the four indicators:

- Fully aligned – assessments align to all content strands (91%–100%);
- Highly aligned – assessments align to the majority of strands (70%–90%)
- Partially aligned – assessments align well to some strands (50%–69%)
- Weakly aligned – assessments align to less than half the strands (below 50%).

Webb’s alignment method does not allow for a *single* judgment of overall alignment across the four alignment indicators. However, one can get a sense of overall alignment between the assessments and standards by looking at all of the alignment indicators together.

Table 3.14 presents the summary alignment outcomes for the Science MAP based on the above scale. The table includes a summary judgment for each Webb alignment indicator per grade assessment based on the percentage of strands that met the minimum alignment criteria. This summary table is linked to the bottom row of Tables A-1 through A-16 in Appendix A. Thus, these summary judgments reflect a final evaluation of each grade assessment per Webb indicator criteria *across* the strands.

As shown in Table 3.14 with green highlighting, half of the results indicate strong content alignment of the Science MAP to the Missouri Learning Standards. Each grade assessment clearly includes a sufficient percent of operational items assessing DOK levels at or above the DOK assigned to the GLEs, as demonstrated by the outcomes on depth-of-knowledge consistency. Furthermore, across the grade assessments, balance-of-knowledge representation results suggest that items seem to be distributed reasonably, at least across GLEs matched by panelists.

Categorical Concurrence and Range-of-Knowledge Correspondence demonstrated weak alignment to the content standards, as indicated by the red highlighting.

**Table 3.14. Summary Alignment Outcomes on Each Webb Criterion by Grade Level for Science MAP**

Grade Assessment	Forms	Percentage of Strands that Met Webb Criteria			
		Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Knowledge Representation
5	Form 1	Weakly aligned (25%)	Fully aligned (88%)	Weakly aligned (0%)	Fully aligned (100%)
	Form 2	Weakly aligned (25%)	Fully aligned (88%)	Weakly aligned (0%)	Fully aligned (100%)
8	Form 1	Weakly aligned (25%)	Fully aligned (100%)	Weakly aligned (0%)	Fully aligned (100%)
	Form 2	Weakly aligned (13%)	Fully aligned (88%)	Weakly aligned (0%)	Fully aligned (100%)

The Missouri Learning Standards include eight strands, which means in order to meet Webb’s Categorical Concurrence criterion of at least 6 items per strand, each Science MAP assessment should contain a minimum of 48 items. The grade 5 assessment contains 40 or 41 items and the grade 8 assessment contains 38 or 39 items. As a result, neither of the forms in grade 5 or 8 includes a sufficient number of items to meet the categorical concurrence for all Science content strands. According to Webb’s criterion for categorical concurrence, the Science MAP in grades 5 and 8 do not adequately cover the Science content students are expected to know.

Finally, there was a restricted range of content assessed at grades 5 and 8 across all strands. The range of knowledge correspondence results indicate that none of the strands met the minimum criteria of having 50% of the GLEs within a strand matched to an item. The highest percentage of GLEs within a strand matched to an item occur in the ‘Changes in Ecosystems and Interactions of Organisms with their Environments’ (EC) strand with 34–38% in grade 5 and 18–28% in grade 8. This restricted range of content assessed is a result of a much larger number of GLEs available to be assessed than items on the assessment. For the grade 5 assessment, there are 171 GLEs across grades 3, 4, and 5 that are assessable but only 40 or 41 items are administered on the two forms. An even greater discrepancy is seen for the

grade 8 assessment where a total of 264 GLEs across grades 6, 7, and 8 are assessable but only 38 or 39 items are administered on the two forms.

Tables A-17 and A-18 in Appendix A present the mean number of items matched to each GLE and the number of panelists represented.

Suggestions for improving the alignment between the Science assessments and Missouri Learning Standards are discussed in Chapter 5, Summary and Recommendations.

## Chapter 4: Results: Test Quality of Science MAP

In this chapter, we report the results of panelists’ evaluations of test quality. Alignment of assessments to the state content standards serves as one form of test validity evidence. Other areas of validity are critical as well, such as whether the assessment enables students to demonstrate what they know. For example, are test items free of biases, clear in language, and appropriate for the grade level?

All assessments should “be designed from the beginning to be accessible and valid with respect to the widest possible range of students, including students with disabilities and students with limited English proficiency” (NCLB, 2001, Section 200.2(b)(2)). The Science MAP items underwent bias reviews as part of the item development process; however, review of quality and accessibility by an independent evaluator provides further evidence of a fair process and assessment. This evaluation of test quality for the Science MAP items represented a broad review of student access to test content.

Panelists evaluated the Science MAP items on several dimensions at the item level. Item ratings included review of written content and figures or graphics, and were based on simple yes–no evaluations of item quality. Panelists also made “overall item quality” ratings with annotations to report the rationale for their ratings. Results reported in this section include those for operational items from the 2015 Science MAP forms. Panelists made their content alignment ratings based on the paper version of the online test forms.

### Written Content

Panelists rated the language used in the items for the extent to which students of various backgrounds and ability levels could access the Science content. Ratings consisted of ‘yes’ or ‘no’ responses. Table 4.1 below indicates the mean number of items per grade test form rated as accessible or not. As the table demonstrates, the majority of items were rated favorably on accessibility.

**Table 4.1. Mean Number of Items Rated As Accessible in Content to Range of Students per Grade Assessment**

Grade	Forms	Is item content accessible to the range of students who take the assessment?					
		Yes			No		
		Mean number of items	SD	Percent of Items per Level	Mean number of items	SD	Percent of Items per Level
5	Form 1	27.88	3.23	94.49	3.25	3.30	5.51
	Form 2	28.38	2.07	94.98	2.00	2.00	5.02
8	Form 1	35.13	1.55	96.23	1.83	0.75	3.77
	Form 2	33.88	1.89	92.81	2.63	1.06	7.19

If panelists responded ‘no’, we asked them to provide an explanation of their responses. Comments pertained to confusing or misleading language and keeping terminology consistent

within the question. Appendix B contains a summary of panelists’ comments on items by common topics.

### *Figures and Graphics*

For those Science items accompanied by pictures, figures, or graphs, panelists evaluated whether these graphics would be understandable to a wide range of students from different backgrounds and ability levels. Panelists were instructed to evaluate any figures and graphics on accuracy of representation. One caveat to this evaluation is that panelists saw the paper version of any figures and graphics while students will be administered the assessment on computers. Table 4.2 indicates that panelists’ ratings were mostly positive.

**Table 4.2. Mean Ratings on Accessibility of Figures or Graphics to Range of Students per Grade Assessment.**

Grade	Forms	Is item content accessible to the range of students who take the assessment?					
		Yes			No		
		Mean number of items	SD	Percent of Items per Level	Mean number of items	SD	Percent of Items per Level
5	Form 1	11.75	2.66	88.68	3.00	2.71	11.32
	Form 2	11.38	2.20	91.92	4.00	2.83	8.08
8	Form 1	15.63	4.75	95.42	2.00	0.00	4.58
	Form 2	7.13	2.30	90.48	1.00	0.00	9.52

For those items with graphics rated as not accessible, panelists’ comments focused on adding clarity to graphics through labels and using more graphics to support the text, such as illustrating the location of the mirror on a microscope. Appendix B contains a summary of panelists’ comments on items by common topics.

### *Overall Item Quality*

In addition to rating items on accessibility, panelists had the opportunity to give items a general rating reflecting their judgments of quality. This rating encompassed aspects such as clarity (e.g., wording or item scene, prompt, or response options) and appropriateness (e.g., off-grade, exceeds GLE).

- Poor quality – item exhibits serious flaw; recommend replacement.
- Fair quality – item exhibits minor but repairable flaw.
- Good quality – item exhibits no real flaws and is typical for this type of assessment.
- Exceptional quality – item is exemplary for this type of assessment.

Tables 4.3 and 4.4 display the mean ratings on overall item quality per grade assessment. As the table illustrates, panelists considered the majority of items to be ‘good’ to ‘exceptional’ in quality.

**Table 4.3. Panelist Ratings on Overall Item Quality – Grade 5**

Form	Item Quality	Mean Number of Items per Level	SD <sup>a</sup>	Percent of Items per Level
Form 1	Poor	0.00	n/a	0.00
	Fair	6.40	4.34	13.50
	Good	18.75	8.31	63.29
	Exceptional	9.17	5.53	23.21
Form 2	Poor	0.00	n/a	0.00
	Fair	3.17	2.23	7.95
	Good	18.13	7.45	60.67
	Exceptional	12.50	4.93	31.38

<sup>a</sup> Not all panelists rated an item at each item quality level.

**Table 4.4. Panelist Ratings on Overall Item Quality – Grade 8**

Form	Item Quality	Mean Number of Items per Level	SD <sup>a</sup>	Percent of Items per Level
Form 1	Poor	0.00	n/a	0.00
	Fair	3.67	1.86	7.61
	Good	13.57	11.79	32.87
	Exceptional	21.50	13.26	59.52
Form 2	Poor	1.00	0.00	1.05
	Fair	2.53	1.51	7.37
	Good	15.43	13.16	37.89
	Exceptional	25.50	11.38	53.68

<sup>a</sup> Not all panelists rated an item at each item quality level.

For those items rated as ‘fair’ or ‘poor’ in quality, we asked panelists to provide comments to identify the issue and suggest improvements. Comments on the items focus generally on two issues. First, item content was either not part of any GLE or it was covered by multiple GLEs. Secondly, some questions were not clearly written or were too broad to understand precisely what is being asked of the students. Appendix B contains a summary of panelists’ comments on items by common topics.

### **Summary and Discussion of Test Quality Results**

The results of the test quality review by panelists suggest that the Science MAP allow a wide range of students the opportunity to demonstrate their knowledge of Science. The majority of items received positive ratings by panelists, and global judgments about test quality also emphasized this point.

Table 4.5 presents the summary outcomes on the item quality ratings. The table includes conclusions regarding the quality of the items on each assessment, along with the percentage of items that received favorable ratings. These conclusions are based on the following decision criteria (adapted from Thompson, Johnstone, Anderson, & Miller, 2005).

- Excellent – all items are acceptable;

- Good – most items are acceptable (at least 90%);
- Acceptable – many items are acceptable (70%–90%);
- Questionable – few items are acceptable (less than 70%).

**Table 4.5. Item Quality Ratings for Science MAP by Grade**

Grade	Forms	Percentage of Items with Acceptable Ratings		
		Written Content	Graphics	Overall Item Quality
5	Form 1	Good (94%)	Acceptable (89%)	Acceptable (87%)
	Form 2	Good (95%)	Good (92%)	Good (92%)
8	Form 1	Good (96%)	Good (95%)	Good (92%)
	Form 2	Good (93%)	Good (90%)	Good (92%)

Table 4.5 shows that none of the grade assessments included enough items with low ratings on any dimension to warrant a conclusion of questionable quality. However, one grade 5 form included *some* items with lower ratings (and corresponding annotations highlighting possible issues), as demonstrated by findings of ‘acceptable’ quality (70%–90% of items). Panelists commented on a number of items with graphics that were either unclear or unnecessary (not adding to the item) as well as items with lower overall item quality. In both of these instances, the percentages were just below the criterion level for ‘good.’ For all assessments, panelists commented on the written content by noting that the majority of the items were ‘good’, but there were a percentage of items that were rated as ‘fair’ showing minor but repairable flaws.

As a whole, the independent item ratings suggest that the Science MAP assessments function well for the majority of students who take them. A handful of items on one of the grade 5 forms may require review to enhance clarity in accompanying graphics and reduce potential bias against student subgroups.

## Chapter 5: Summary and Recommendations

HumRRO conducted a review of the Science MAP to examine the following: (a) content alignment to the Missouri Learning Standards for Science and (b) item quality and accessibility for all students who take these assessments. Alignment of assessments and achievement standards to the state academic content standards is a requirement of the No Child Left Behind Act of 2001.

The cumulative results provide validity evidence to support that the content of the Science MAP assessment items match the intended content as specified in the standards. Each assessment covers the content strands specified in the Missouri Learning Standards for Science. Panelists also determined that the majority of items are appropriate for a wide range of students.

Even though none of the forms in grade 5 or 8 met Webb's minimum criteria for categorical concurrence or range-of-knowledge correspondence, the Science MAP assessment is not deficient on these indicators when the structure of the Missouri Learning Standards and the Science MAP are taken into consideration. The Science MAP assessment is a grade-span test which means that the number of GLEs available to be assessed is far greater than the number of items on the test. This in turn means that the results of the range-of-knowledge correspondence cannot be favorable for the Science MAP. Additionally, the Science Missouri Learning Standards contain eight strands. In order to meet Webb's minimum criterion of six items per content strand for categorical concurrence, a minimum of 48 items would need to be assessed on each test. The Science MAP in grade 5 and 8 has 38–41 items.

As with most reviews of state assessment systems, these findings point to areas where Missouri could strengthen the alignment between the assessments and the content standards. For this reason, HumRRO makes the following recommendations to Missouri on ways in which alignment might be improved. These recommendations focus on the more critical findings:

- **Review content coverage (categorical concurrence).** The panelists reviewing these assessments found that in grade 5 only two of the eight strands in the Science Missouri Learning Standards are assessed with at least six items. In grade 8, the same can be seen for form 1 while form 2 only one of the eight strands are assessed with at least six items. Thus, the assessments may not adequately reflect the content that students are expected to know based solely on the number of items on the assessment (not the item type or point value as these are not factors in Webb's categorical concurrence indicator). From strictly an item count perspective, there are several ways DESE can choose to mitigate this situation such as increase the number of items on the assessment, reduce the number of strands in the state standards, or designate some of the strands for local assessment only.
- **Review whether the Science MAP is a grade-level or grade-span assessment.** With the large number of GLEs across the grade-span, being able to link each GLE with at least one item is impractical. Identifying the Science MAP as a grade-level assessment would reduce the number of GLEs substantially when evaluating the range-of-knowledge correspondence if that is the state's intention of the Science MAP test in grade 5 and 8.
- **Review those items that received the lowest ratings on test quality for possible revision.** Panelists' review suggests that there are issues with items aligning to multiple GLEs and a few items not aligning to any GLE. They also identified some items that could be improved by changes in word clarity, language, and specificity level. In particular, some items were considered too broad to appropriately assess a single GLE.



## References

- No Child Left Behind Act of 2001. Public Law 107-110.
- Porter, A. C. (2002, October). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, 31(7), 3–14.
- Subkoviak, M. J. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *Journal of Educational Measurement*, 25(1), 47–55.
- Thompson, S. J., Johnstone, C. J., Anderson, M. E., & Miller, N. A. (2005). *Considerations for the development and review of universally designed assessments* (Technical Report 42). Minneapolis, MN: University of Missouri, National Center on Educational Outcomes.
- U.S. Department of Education. (April, 2004). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education. Retrieved from <http://www.ed.gov/policy/elsec/guid/saaprguidance.doc>
- Webb, N. L. (1997). *Research Monograph No. 6: Criteria for alignment of expectations and assessments in Science and Science education*. Washington, D.C.: Council of Chief State Schools Officers.
- Webb, N. L. (1999). *Alignment of Science and Science standards and assessments in four states. (Research Monograph 18)*. Madison, WI: National Institute for Science Education and Council of Chief State School Officers. (ERIC Document Reproduction Service No. ED440852).
- Webb, N. L. (2005). *Webb alignment tool: Training manual*. Madison, WI: Wisconsin Center for Education Research. Available: <http://www.wcer.wisc.edu/WAT/index.aspx>



## Appendix A. Content Alignment Results per Grade Level Assessment

The following tables include complete statistical results on the Webb alignment indicators, including means and standard deviations per strand for each grade Science MAP test.

### Categorical Concurrence

The categorical concurrence results for grades 5 and 8 of the Science MAP are presented below. Each table includes: the percent of items emphasized in each strand from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row indicates the percentage of strands that met the minimum alignment indicator criterion. Note that the total mean items matched may exceed the number of items on the assessment, as panelists were able to match items to more than one strand.

**Table A-1. Categorical Concurrence for Science MAP, Grade 5: Mean Number of Items per Strand Form 1**

Title of Strand	Target # Items from Blueprint	Number of Items per Strand		At Least Six Items per Strand
		Mean Items Matched	SD	
Properties and Principles of Matter and Energy (ME)	5 – 6	3.00	0.00	No
Properties and Principles of Force and Motion (FM)	3 – 4	2.38	0.52	No
Characteristics and Interactions of Living Organisms (LO)	3 – 4	3.38	1.19	No
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	4 – 5	7.38	1.06	Yes
Processes and Interactions of the Earth's Systems (ES)	5 – 7	3.38	0.52	No
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	4 – 5	3.63	0.52	No
Scientific Inquiry (IN)	9 – 11	9.50	0.76	Yes
Science, Technology, and Human Activity (ST)	3 – 4	2.00	0.00	No
Total Number of Items		35 of 41		
<b>Percentage of strands with at least six items: 25%</b>				

**Table A-2. Categorical Concurrence for Science MAP, Grade 5: Mean Number of Items per Strand Form 2**

Title of Strand	Target # Items from Blueprint	Number of Items per Strand		At Least Six Items per Strand
		Mean Items Matched	SD	
Properties and Principles of Matter and Energy (ME)	5 – 6	3.00	0.00	No
Properties and Principles of Force and Motion (FM)	3 – 4	3.00	0.00	No
Characteristics and Interactions of Living Organisms (LO)	3 – 4	4.00	0.76	No
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	4 – 5	7.13	0.99	Yes
Processes and Interactions of the Earth’s Systems (ES)	4 – 6	4.88	0.35	No
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	4 – 5	3.00	0.00	No
Scientific Inquiry (IN)	9 – 11	8.00	0.00	Yes
Science, Technology, and Human Activity (ST)	3 – 4	2.00	0.00	No
Total Number of Items		35 of 40		
<b>Percentage of strands with at least six items: 25%</b>				

**Table A-3. Categorical Concurrence for Science MAP, Grade 8: Mean Number of Items per Strand Form 1**

Title of Strand	Target # Items from Blueprint	Number of Items per Strand		At Least Six Items per Strand
		Mean Items Matched	SD	
Properties and Principles of Matter and Energy (ME)	4 – 5	4.25	0.71	No
Properties and Principles of Force and Motion (FM)	3 – 4	1.75	0.46	No
Characteristics and Interactions of Living Organisms (LO)	3 – 4	4.00	0.00	No
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	3 – 4	4.25	0.46	No
Processes and Interactions of the Earth’s Systems (ES)	4 – 6	6.00	0.00	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	4 – 5	3.75	0.46	No
Scientific Inquiry (IN)	9 – 11	10.25	0.46	Yes
Science, Technology, and Human Activity (ST)	3 – 4	1.75	0.46	No
Total		36 of 39		
<b>Percentage of strands with at least six items: 25%</b>				

**Table A-4. Categorical Concurrence for Science MAP, Grade 8: Mean Number of Items per Strand Form 2**

Title of Strand	Target # Items from Blueprint	Number of Items per Strand		At Least Six Items per Strand
		Mean Items Matched	SD	
Properties and Principles of Matter and Energy (ME)	4 – 5	5.00	0.76	No
Properties and Principles of Force and Motion (FM)	3 – 4	1.75	0.46	No
Characteristics and Interactions of Living Organisms (LO)	3 – 4	4.50	0.53	No
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	3 – 4	4.50	0.93	No
Processes and Interactions of the Earth’s Systems (ES)	4 – 6	4.13	0.64	No
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	4 – 5	4.25	0.46	No
Scientific Inquiry (IN)	9 – 11	9.88	0.83	Yes
Science, Technology, and Human Activity (ST)	3 – 4	2.00	0.53	No
Total 36 of 38				
<b>Percentage of strands with at least six items: 13%</b>				

### *Depth-of-Knowledge Consistency*

The Depth-of-Knowledge (DOK) consistency results for grades 5 and 8 of the Science MAP are presented below. The tables present the results from the comparison between the depth-of-knowledge expected in the content GLEs and the depth-of-knowledge assessed by items. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the GLEs along with the corresponding standard deviations. GLEs with at least 50% of items at the same (or above) DOK level met the minimum indicator criterion.

**Table A-5. DOK Consistency for Science MAP, Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of GLEs Form 1**

Title of Strand	Mean Items per Strand	Depth-of-Knowledge Consistency						DOK Consistency Target Met
		% Items Below		% Items Same Level		% Items Above		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Properties and Principles of Matter and Energy (ME)	3.00	12.5	17.3	50.0	25.2	37.5	11.8	Yes
Properties and Principles of Force and Motion (FM)	2.38	16.7	23.6	64.6	24.3	18.8	20.8	Yes
Characteristics and Interactions of Living Organisms (LO)	3.38	20.8	30.5	70.0	28.1	9.2	17.1	Yes
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	7.38	32.1	16.2	54.4	19.5	13.5	14.7	Yes
Processes and Interactions of the Earth's Systems (ES)	3.38	0.0	0.0	81.3	27.4	18.8	27.4	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	3.63	6.7	12.8	50.6	25.8	42.7	26.9	Yes
Scientific Inquiry (IN)	9.50	43.8	25.1	52.2	23.1	4.0	5.6	Yes
Science, Technology, and Human Activity (ST)	2.00	56.3	32.0	31.3	37.2	12.5	23.1	No
<b>Percentage of strands with 50% of item DOK at or above objective DOK: 88%</b>								

**Table A-6. DOK Consistency for Science MAP, Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of GLEs Form 2**

Title of Strand	Mean Items per Strand	Depth-of-Knowledge Consistency						DOK Consistency Target Met
		% Items Below		% Items Same Level		% Items Above		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Properties and Principles of Matter and Energy (ME)	3.00	12.5	23.1	33.3	17.8	54.2	17.3	Yes
Properties and Principles of Force and Motion (FM)	3.00	8.3	25.9	54.2	35.4	37.5	37.5	Yes
Characteristics and Interactions of Living Organisms (LO)	4.00	19.8	15.4	67.9	16.2	12.3	17.4	Yes
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	7.13	22.9	41.7	67.3	13.9	9.8	14.0	Yes
Processes and Interactions of the Earth's Systems (ES)	4.88	43.8	7.6	37.5	51.8	18.8	25.9	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	3.00	8.3	22.7	72.9	25.1	18.8	20.8	Yes
Scientific Inquiry (IN)	8.00	40.3	15.4	55.0	25.3	4.7	6.5	Yes
Science, Technology, and Human Activity (ST)	2.00	62.5	17.3	18.8	25.9	18.8	25.9	No
<b>Percentage of strands with 50% of item DOK at or above objective DOK: 88%</b>								

**Table A-7. DOK Consistency for Science MAP, Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of GLEs Form 1**

Title of Strand	Mean Items per Strand	Depth-of-Knowledge Consistency						DOK Consistency Target Met
		% Items Below		% Items Same Level		% Items Above		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Properties and Principles of Matter and Energy (ME)	4.25	41.0	9.0	46.5	19.6	12.5	18.3	Yes
Properties and Principles of Force and Motion (FM)	1.75	12.5	23.1	31.3	25.9	56.3	32.0	Yes
Characteristics and Interactions of Living Organisms (LO)	4.00	29.2	28.2	51.0	29.0	19.8	12.5	Yes
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	4.25	28.5	9.4	15.0	22.7	56.5	23.9	Yes
Processes and Interactions of the Earth's Systems (ES)	6.00	19.6	10.6	58.8	14.8	21.7	11.5	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	3.75	19.8	12.5	70.8	16.1	9.4	12.9	Yes
Scientific Inquiry (IN)	10.25	12.5	9.1	70.6	8.1	16.9	11.9	Yes
Science, Technology, and Human Activity (ST)	1.75	50.0	37.8	50.0	37.8	0.0	0.0	Yes
<b>Percentage of strands with 50% of item DOK at or above objective DOK: 100%</b>								

**Table A-8. DOK Consistency for Science MAP, Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of GLEs Form 2**

Title of Strand	Mean Items per Strand	Depth-of-Knowledge Consistency						DOK Consistency Target Met
		% Items Below		% Items Same Level		% Items Above		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Properties and Principles of Matter and Energy (ME)	5.00	23.7	13.8	37.2	16.2	39.1	12.2	Yes
Properties and Principles of Force and Motion (FM)	1.75	12.5	35.4	43.8	49.6	43.8	49.6	Yes
Characteristics and Interactions of Living Organisms (LO)	4.50	9.4	12.9	82.5	11.0	8.1	11.3	Yes
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	4.50	53.8	10.3	15.4	13.4	30.8	10.5	No
Processes and Interactions of the Earth's Systems (ES)	4.13	32.7	21.9	58.5	26.0	8.8	12.2	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	4.25	6.3	11.6	58.8	24.0	35.0	24.2	Yes
Scientific Inquiry (IN)	9.88	8.6	7.7	77.1	14.6	14.4	12.3	Yes
Science, Technology, and Human Activity (ST)	2.00	27.1	29.5	41.7	41.8	31.3	37.2	Yes
<b>Percentage of strands with 50% of item DOK at or above objective DOK: 88%</b>								

### Range-of-Knowledge Correspondence

The results for Range-of-Knowledge correspondence for grades 5 and 8 for the Science MAP are presented below. The tables include the mean number, standard deviation, and percentage of GLEs by content strand. For acceptable range-of-knowledge correspondence, a minimum of 50% of content GLEs within each strand should be matched to at least one item.

**Table A-9. Range-of-Knowledge for Science MAP, Grade 5: Mean Percent of GLEs per Strand Linked with Items Form 1**

Title of Strand	Number of GLEs	Mean Items per Strand	Range of GLEs		% of Total GLEs per Strand	Range-of-Knowledge Target Met
			GLEs with At Least One Item			
			<i>M</i>	<i>SD</i>		
Properties and Principles of Matter and Energy (ME)	32	3.0	3.0	0.0	9%	No
Properties and Principles of Force and Motion (FM)	17	2.4	2.1	0.8	13%	No
Characteristics and Interactions of Living Organisms (LO)	12	3.3	2.6	0.5	22%	No
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	15	3.8	3.8	0.7	34%	No
Processes and Interactions of the Earth's Systems (ES)	22	1.9	1.6	0.5	13%	No
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	16	3.9	2.9	0.8	18%	No
Scientific Inquiry (IN)	40	9.5	7.0	0.9	26%	No
Science, Technology, and Human Activity (ST)	17	2.0	1.5	0.5	14%	No
Total	171					
<b>Percentage of strands with 50% of GLEs linked to at least one item: 0%</b>						

**Table A-10. Range-of-Knowledge for Science MAP, Grade 5: Mean Percent of GLEs per Strand Linked with Items Form 2**

Title of Strand	Number of GLEs	Mean Items per Strand	Range of GLEs		% of Total GLEs per Strand	Range-of-Knowledge Target Met
			GLEs with At Least One Item			
			<i>M</i>	<i>SD</i>		
Properties and Principles of Matter and Energy (ME)	32	3.0	3.0	0.0	9%	No
Properties and Principles of Force and Motion (FM)	17	3.0	2.8	0.5	16%	No
Characteristics and Interactions of Living Organisms (LO)	12	4.0	3.1	0.6	26%	No
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	15	5.0	4.1	0.6	38%	No
Processes and Interactions of the Earth's Systems (ES)	22	1.5	1.5	0.5	12%	No
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	16	3.1	3.0	0.0	19%	No
Scientific Inquiry (IN)	40	8.1	6.4	0.7	24%	No
Science, Technology, and Human Activity (ST)	17	2.1	1.8	0.5	16%	No
Total	171					
<b>Percentage of strands with 50% of GLEs linked to at least one item: 0%</b>						

**Table A-11. Range-of-Knowledge for Science MAP, Grade 8: Mean Percent of GLEs per Strand Linked with Items Form 1**

Title of Strand	Number of GLEs	Mean Items per Strand	Range of GLEs		% of Total GLEs per Strand	Range-of-Knowledge Target Met
			GLEs with At Least One Item			
			<i>M</i>	<i>SD</i>		
Properties and Principles of Matter and Energy (ME)	58	4.3	4.1	0.6	7%	No
Properties and Principles of Force and Motion (FM)	18	1.8	1.8	0.5	10%	No
Characteristics and Interactions of Living Organisms (LO)	34	4.0	3.9	0.4	11%	No
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	15	4.4	2.8	0.9	18%	No
Processes and Interactions of the Earth's Systems (ES)	39	5.8	5.6	0.7	14%	No
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	27	3.8	3.3	0.9	12%	No
Scientific Inquiry (IN)	55	11.0	6.1	1.0	11%	No
Science, Technology, and Human Activity (ST)	18	1.8	1.6	0.5	9%	No
Total	264					
<b>Percentage of strands with 50% of GLEs linked to at least one item: 0%</b>						

**Table A-12. Range-of-Knowledge for Science MAP, Grade 8: Mean Percent of GLEs per Strand Linked with Items Form 2**

Title of Strand	Number of GLEs	Mean Items per Strand	Range of GLEs		% of Total GLEs per Strand	Range-of-Knowledge Target Met
			GLEs with At Least One Item			
			<i>M</i>	<i>SD</i>		
Properties and Principles of Matter and Energy (ME)	58	5.5	4.5	0.5	8%	No
Properties and Principles of Force and Motion (FM)	18	1.8	1.8	0.5	10%	No
Characteristics and Interactions of Living Organisms (LO)	34	4.3	4.0	0.0	12%	No
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	15	4.5	4.1	0.6	28%	No
Processes and Interactions of the Earth's Systems (ES)	39	4.3	3.4	0.9	9%	No
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	27	4.3	3.6	0.7	13%	No
Scientific Inquiry (IN)	55	10.0	5.9	0.6	11%	No
Science, Technology, and Human Activity (ST)	18	2.0	1.8	0.5	10%	No
Total	264					
<b>Percentage of strands with 50% of GLEs linked to at least one item: 0%</b>						

### *Balance-of-Knowledge Representation*

The results for Balance-of-Knowledge representation for grades 5 and 8 of the Science MAP are presented below. The tables also include the percentage of items linked to each strand. The minimum acceptable balance index is 70 out of 100.

**Table A-13. Balance-of-Knowledge Representation for Science MAP, Grade 5: Mean Balance Index per Strand Form 1**

Title of Strand	GLEs per Strand	Balance-of-Knowledge Representation					Balance Index Target Met
		Mean GLEs Linked with Items	Mean Items per Strand	Mean % of Items (of total) Linked to Strand	Mean Balance Index		
		<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>SD</i>	
Properties and Principles of Matter and Energy (ME)	32	3.0	3.0	11%	100	0.0	Yes
Properties and Principles of Force and Motion (FM)	17	2.1	2.4	8%	100	0.0	Yes
Characteristics and Interactions of Living Organisms (LO)	12	2.6	3.3	11%	92	8.5	Yes
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	15	3.8	3.8	13%	100	0.0	Yes
Processes and Interactions of the Earth's Systems (ES)	22	1.6	1.9	7%	96	7.7	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	16	2.9	3.9	13%	93	8.1	Yes
Scientific Inquiry (IN)	40	7.0	9.5	33%	82	4.2	Yes
Science, Technology, and Human Activity (ST)	17	1.5	2.0	7%	100	0.0	Yes
<b>Total</b>	<b>171</b>						
<b>Percentage of standards with a balance of representation index of 70 or greater: 100%</b>							

**Table A-14. Balance-of-Knowledge Representation for Science MAP, Grade 5: Mean Balance Index per Strand Form 2**

Title of Strand	GLEs per Strand	Balance-of-Knowledge Representation					Balance Index Target Met
		Mean GLEs Linked with Items	Mean Items per Strand	Mean % of Items (of total) Linked to Strand	Mean Balance Index		
					M	SD	
Properties and Principles of Matter and Energy (ME)	32	3.0	3.0	10%	100	0.0	Yes
Properties and Principles of Force and Motion (FM)	17	2.8	3.0	10%	96	7.7	Yes
Characteristics and Interactions of Living Organisms (LO)	12	3.1	4.0	13%	86	5.8	Yes
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	15	4.1	5.0	17%	89	7.2	Yes
Processes and Interactions of the Earth's Systems (ES)	22	1.5	1.5	5%	100	0.0	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	16	3.0	3.1	10%	98	5.9	Yes
Scientific Inquiry (IN)	40	6.4	8.1	27%	86	4.4	Yes
Science, Technology, and Human Activity (ST)	17	1.8	2.1	7%	98	5.9	Yes
<b>Total</b>	<b>171</b>						
<b>Percentage of standards with a balance of representation index of 70 or greater: 100%</b>							

**Table A-15. Balance-of-Knowledge Representation for Science MAP, Grade 8: Mean Balance Index per Strand Form 1**

Title of Strand	GLEs per Strand	Balance-of-Knowledge Representation					Balance Index Target Met
		Mean GLEs Linked with Items	Mean Items per Strand	Mean % of Items (of total) Linked to Strand	Mean Balance Index		
		<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>SD</i>	
Properties and Principles of Matter and Energy (ME)	58	4.1	4.3	12%	98	5.3	Yes
Properties and Principles of Force and Motion (FM)	18	1.8	1.8	5%	100	0.0	Yes
Characteristics and Interactions of Living Organisms (LO)	34	3.9	4.0	11%	98	5.3	Yes
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	15	2.8	4.4	12%	80	9.0	Yes
Processes and Interactions of the Earth's Systems (ES)	39	5.6	5.8	16%	98	5.3	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	27	3.3	3.8	10%	92	8.9	Yes
Scientific Inquiry (IN)	55	6.1	11.0	30%	74	2.9	Yes
Science, Technology, and Human Activity (ST)	18	1.6	1.8	5%	100	0.0	Yes
<b>Total</b>	<b>264</b>						
<b>Percentage of standards with a balance of representation index of 70 or greater: 100%</b>							

**Table A-16. Balance-of-Knowledge Representation for Science MAP, Grade 8: Mean Balance Index per Strand Form 2**

Title of Strand	GLEs per Strand	Balance-of-Knowledge Representation					Balance Index Target Met
		Mean GLEs Linked with Items	Mean Items per Strand	Mean % of Items (of total) Linked to Strand	Mean Balance Index		
		<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>SD</i>	
Properties and Principles of Matter and Energy (ME)	58	4.5	5.5	15%	90	8.2	Yes
Properties and Principles of Force and Motion (FM)	18	1.8	1.8	5%	100	0.0	Yes
Characteristics and Interactions of Living Organisms (LO)	34	4.0	4.3	12%	96	6.9	Yes
Changes in Ecosystems and Interactions of Organisms with their Environments (EC)	15	4.1	4.5	12%	95	7.5	Yes
Processes and Interactions of the Earth's Systems (ES)	39	3.4	4.3	12%	86	5.8	Yes
Composition and Structure of the Universe and the Motion of Objects Within It (UN)	27	3.6	4.3	12%	90	8.5	Yes
Scientific Inquiry (IN)	55	5.9	10.0	27%	81	2.6	Yes
Science, Technology, and Human Activity (ST)	18	1.8	2.0	6%	100	0.0	Yes
<b>Total</b>	<b>264</b>						
<b>Percentage of standards with a balance of representation index of 70 or greater: 100%</b>							

### *GLEs Matched to Items by Panelists*

Tables A-17 and A-18 present the GLEs, along with the mean number of items, matched by panelists. Column 1 includes the Missouri Learning Standards code corresponding to the GLEs. One note of caution when reading these tables, the same item may not be represented by the mean number of items. For example, Missouri GLE 'ME.1.D.3.a' in the first row shows that 6 panelists matched a mean number of 1 items to this GLE on form 1. This does not mean/assume that the 1 item matched to the GLE by the panelists is the same item across panelists.

**Table A-17. Grade 5 MAP: Grade Span GLEs Matched to Items by Panelists**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
ME.1.D.3.a	311401	6	1.00	0.00	6	1.00	0.00
ME.1.D.3.b	311402						
ME.1.D.3.c	311403				7	1.00	0.00
ME.1.D.3.d	311404						
ME.1.D.3.e	311405						
ME.1.D.3.f	311406						
ME.1.D.3.g	311407				1	1.00	n/a
ME.2.A.3.a	312101						
ME.2.A.3.b	312102						
ME.2.A.3.c	312103						
ME.2.A.3.d	312104						
ME.2.C.3.a	312301						
LO.1.A.3.a	331101						
LO.1.B.3.a	331201						
LO.1.D.3.a	331401	8	1.00	0.00	8	1.00	0.00
LO.2.C.3.a	332301						
LO.3.D.3.a	333401						
EC.2.A.3.a	342101	5	1.00	0.00	1	1.00	n/a
EC.2.A.3.b	342102						
EC.2.A.3.c	342103	2	1.00	0.00	3	1.00	0.00
EC.2.A.3.d	342104	8	1.00	0.00	7	1.00	0.00
ES.1.C.3.a	351301						
ES.1.C.3.b	351302						
ES.1.C.3.c	351303						
ES.2.E.3.a	352501						
UN.1.A.3.a	361101						
UN.1.A.3.b	361102						
UN.2.A.3.a	362101						
UN.2.B.3.a	362201						
UN.2.B.3.b	362202	3	1.00	0.00	3	1.00	0.00
UN.2.C.3.a	362301						
UN.2.C.3.b	362302						
UN.2.C.3.c	362303						

(continued)

**Table A-17. Grade 5 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
IN.1.A.3.a	371101						
IN.1.A.3.b	371102						
IN.1.B.3.a	371201						
IN.1.B.3.b	371202						
IN.1.B.3.c	371203						
IN.1.B.3.d	371204						
IN.1.B.3.e	371205						
IN.1.C.3.a	371301						
IN.1.C.3.b	371302						
IN.1.C.3.c	371303						
IN.1.C.3.d	371304						
IN.1.D.3.a	371401						
ST.1.A.3.a	381101	7	1.43	0.53	8	1.25	0.46
ST.1.B.3.a	381201						
ST.2.A.3.a	382101						
ST.3.A.3.a	383101						
ST.3.A.3.b	383102	1	1.00	n/a			
ME.1.A.4.a	411101	8	1.00	0.00	8	1.00	0.00
ME.1.A.4.b	411102	1	1.00	n/a			
ME.1.A.4.c	411103	7	1.00	0.00			
ME.1.A.4.d	411104						
ME.1.B.4.a	411201						
ME.1.B.4.b	411202						
ME.1.B.4.c	411203						
ME.1.B.4.d	411204						
ME.2.A.4.a	412101						
ME.2.A.4.b	412102						
ME.2.A.4.c	412103						
ME.2.F.4.a	412601						
ME.1.I.4.a	412901						
FM.1.A.4.a	421101						
FM.1.A.4.b	421102	2	1.00	0.00			
FM.2.A.4.a	422101						
FM.2.A.4.b	422102						
FM.2.A.4.c	422103	2	1.00	0.00	2	1.00	0.00
FM.2.A.4.d	422104	3	1.00	0.00			
FM.2.B.4.a	422201						
FM.2.D.4.a	422401						
FM.2.D.4.b	422402						
FM.2.D.4.c	422403	3	1.00	0.00			

(continued)

**Table A-17. Grade 5 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
FM.2.D.4.d	422404						
EC.1.A.4.a	441101	1	1.00	n/a	4	1.25	0.50
EC.1.A.4.b	441102				1	1.00	n/a
EC.1.D.4.a	441401	5	1.00	0.00			
EC.2.A.4.a	442101	1	1.00	n/a	4	1.50	0.58
EC.2.A.4.b	442102	8	1.00	0.00	8	1.13	0.35
EC.2.A.4.c	442103				8	1.00	0.00
EC.3.A.4.a	443101						
EC.3.C.4.a	443301						
EC.3.C.4.b	443302						
EC.3.C.4.c	443303						
EC.3.C.4.d	443304						
ES.1.A.4.a	451101						
ES.1.A.4.b	451102						
ES.2.A.4.a	452101						
ES.2.A.4.b	452102						
ES.2.A.4.c	452103						
ES.2.A.4.d	452104						
ES.2.A.4.e	452105						
ES.3.A.4.a	453101						
ES.3.A.4.b	453102	5	1.40	0.55	4	1.00	0.00
IN.1.A.4.a	471101						
IN.1.A.4.b	471102						
IN.1.A.4.c	471103						
IN.1.B.4.a	471201						
IN.1.B.4.b	471202						
IN.1.B.4.c	471203						
IN.1.B.4.d	471204						
IN.1.B.4.e	471205						
IN.1.C.4.a	471301						
IN.1.C.4.b	471302						
IN.1.C.4.c	471303						
IN.1.C.4.d	471304						
IN.1.D.4.a	471401						
ST.1.A.4.a	481101						
ST.1.B.4.a	481201						
ST.1.C.4.a	481301	1	1.00	n/a	1	1.00	n/a
ST.2.A.4.a	482101						

(continued)

**Table A-17. Grade 5 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
ST.3.A.4.a	483101				1	1.00	n/a
ST.3.A.4.b	483102						
ME.1.C.5.a	511301						
ME.1.D.5.a	511401	2	1.00	0.00	2	1.00	0.00
ME.1.D.5.b	511402						
ME.1.I.5.a	511901						
ME.2.A.5.a	512101						
ME.2.A.5.b	512102						
ME.2.C.5.a	512301						
FM.2.A.5.a	522101						
FM.2.D.5.a	522401	6	1.33	0.52	6	1.00	0.00
FM.2.F.5.a	522601	1	1.00	n/a			
FM.2.F.5.b	522602				8	1.25	0.46
FM.2.F.5.c	522603						
FM.2.F.5.d	522604				6	1.00	0.00
LO.1.D.5.a	531401	2	1.00	0.00	5	1.00	0.00
LO.1.E.5.a	531501	2	1.50	0.71	2	1.00	0.00
LO.1.E.5.b	531502				5	1.00	0.00
LO.1.E.5.c	531503	5	1.00	0.00	6	1.00	0.00
LO.1.E.5.d	531504	5	1.20	0.45	4	1.50	0.58
LO.1.E.5.e	531505						
LO.2.C.5.a	532301	1	2.00	n/a			
ES.1.B.5.a	551201				8	1.00	0.00
ES.1.C.5.a	551301						
ES.2.E.5.a	551501	8	1.00	0.00			
ES.2.E.5.b	551502						
ES.2.F.5.a	552601						
ES.2.F.5.b	552602						
ES.3.A.5.a	553101						
ES.3.A.5.b	553102						
ES.3.A.5.c	553103						
UN.1.A.5.a	561101	1	2.00	n/a			
UN.1.A.5.b	561102						
UN.1.A.5.c	561103				1	1.00	n/a
UN.1.B.5.a	561201				8	1.00	0.00
UN.2.B.5.a	562201	8	1.00	0.00	5	1.00	0.00
UN.2.C.5.a	562301	7	1.14	0.38	1	2.00	n/a
UN.2.C.5.b	562302	1	2.00	n/a			

(continued)

**Table A-17. Grade 5 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
UN.2.C.5.c	562303	3	1.00	0.00	6	1.00	0.00
IN.1.A.5.a	571101	8	2.00	0.00	8	2.00	0.00
IN.1.A.5.b	571102	6	1.17	0.41	8	1.00	0.00
IN.1.A.5.c	571103	6	1.00	0.00	1	2.00	n/a
IN.1.A.5.d	571104	2	1.00	0.00			
IN.1.B.5.a	571201						
IN.1.B.5.b	571202						
IN.1.B.5.c	571203						
IN.1.B.5.d	571204	7	1.00	0.00	7	1.00	0.00
IN.1.B.5.e	571205				2	1.00	0.00
IN.1.B.5.f	571206						
IN.1.C.5.a	571301	4	1.00	0.00	6	1.33	0.82
IN.1.C.5.b	571302	8	1.63	0.74	7	1.43	0.53
IN.1.C.5.c	571303	7	1.00	0.00			
IN.1.C.5.d	571304				4	1.00	0.00
IN.1.D.5.a	571401	8	1.75	1.04	8	1.00	0.00
ST.1.A.5.a	581101						
ST.1.B.5.a	581201						
ST.1.C.5.a	581301	4	1.25	0.50	4	1.00	0.00
ST.2.A.5.a	582101						
ST.3.A.5.a	583101				1	1.00	n/a
ST.3.A.5.b	583102						

**Table A-18. Grade 8 MAP: Grade Span GLEs Matched to Items by Panelists**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
ME.1.A.6.a	611101						
ME.1.A.6.b	611102				4	1.00	0.00
ME.1.A.6.c	611103	3	1.33	0.58			
ME.1.A.6.d	611104						
ME.1.B.6.a	611201						
ME.1.B.6.b	611202						
ME.1.B.6.c	611203						
ME.1.C.6.a	611301	2	1.00	0.00			
ME.1.D.6.a	611401				2	1.00	0.00
ME.1.G.6.a	611701	2	1.00	0.00			
ME.1.G.6.b	611702	6	1.00	0.00	4	1.00	0.00
ME.1.G.6.c	611703						
ME.1.I.6.a	611901	5	1.00	0.00			
ME.2.A.6.a	612101						
ME.2.A.6.b	612102						
ME.2.A.6.c	612103				4	1.00	0.00
ME.2.A.6.d	612104						
ME.2.A.6.e	612106						
ME.2.A.6.f	612107						
ME.2.A.6.g	612108						
ME.2.A.6.h	612109						
ME.2.A.6.i	612110						
ME.2.A.6.j	612111						
ME.2.A.6.k	612301						
ME.2.C.6.a	612302						
ME.2.C.6.b	631101						
LO.1.A.6.a	631301						
LO.1.C.6.a	631501						
LO.1.E.6.a	631502				2	1.00	0.00
LO.1.E.6.b	631502						
LO.2.A.6.a	632101	5	1.00	0.00			
LO.2.A.6.b	632102						
LO.2.B.6.a	632201						
EC.1.A.6.a	641101				6	1.00	0.00
EC.1.B.6.a	641201						
EC.1.B.6.b	641202	8	2.50	0.76	7	1.00	0.00
EC.1.B.6.c	641203						
EC.1.D.6.a	641401	1	1.00	n/a	1	1.00	n/a
EC.1.D.6.b	641402	2	1.00	0.00			
EC.1.D.6.c	641403	1	1.00	n/a			

(continued)

**Table A-18. Grade 8 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
EC.2.A.6.a	642101	6	1.00	0.00			
EC.2.A.6.b	642102	2	1.00	0.00			
EC.3.A.6.a	643101						
EC.3.C.6.a	643301	2	1.00	0.00	7	1.00	0.00
EC.3.C.6.b	643302				7	1.29	0.49
ES.1.A.6.a	651101				2	1.00	0.00
ES.1.B.6.a	651201						
ES.2.A.6.a	652101	3	1.00	0.00			
ES.2.A.6.b	652102	1	1.00	n/a			
ES.2.A.6.c	652103	3	1.00	0.00			
ES.2.A.6.d	652104				1	1.00	n/a
ES.2.B.6.a	652201				4	1.50	0.58
ES.2.D.6.a	652401						
ES.2.D.6.b	652402				1	1.00	n/a
ES.3.A.6.a	653101						
ES.3.A.6.b	653102				7	1.00	0.00
ES.3.A.6.c	653103						
IN.1.A.6.a	671101						
IN.1.A.6.b	671102						
IN.1.A.6.c	671103						
IN.1.A.6.d	671104						
IN.1.A.6.e	671105						
IN.1.B.6.a	671201						
IN.1.B.6.b	671202						
IN.1.B.6.c	671203						
IN.1.B.6.d	671204						
IN.1.B.6.e	671205						
IN.1.B.6.f	671206						
IN.1.C.6.a	671301						
IN.1.C.6.b	671302						
IN.1.C.6.c	671303						
IN.1.C.6.d	671304						
IN.1.C.6.e	671305						
IN.1.D.6.a	671401						
ST.1.A.6.a	681101						
ST.1.B.6.a	681201						
ST.2.B.6.a	682201						
ST.2.B.6.a	682202						
ST.3.B.6.a	683201						
ST.3.B.6.b	683202						
ME.1.D.7.a	711401						
ME.1.I.7.a	711901						

(continued)

**Table A-18. Grade 8 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
ME.2.A.7.a	712101						
ME.2.A.7.b	712102						
ME.2.A.7.c	712103				6	1.00	0.00
ME.2.A.7.d	712104						
ME.2.A.7.e	712105				1	2.00	n/a
ME.2.A.7.f	712106						
ME.2.A.7.g	712107						
ME.2.A.7.h	712108						
ME.2.A.7.i	712109						
ME.2.A.7.j	712110						
ME.2.A.7.k	712111						
ME.2.A.7.l	712112						
ME.2.A.7.m	712113						
ME.2.A.7.n	712114						
ME.2.C.7.a	712310						
ME.2.F.7.a	712601	8	1.00	0.00	8	1.00	0.00
ME.2.F.7.b	712602				1	1.00	n/a
ME.2.F.7.c	712603						
FM.1.A.7.a	721101						
FM.1.A.7.b	721102						
FM.1.A.7.c	721103						
FM.1.A.7.d	721104	6	1.00	0.00			
FM.2.A.7.a	722101						
FM.2.A.7.b	722102						
FM.2.B.7.a	722201						
FM.2.B.7.b	722202				2	1.00	0.00
FM.2.B.7.c	722203						
FM.2.D.7.a	722401						
FM.2.D.7.b	722402						
FM.2.D.7.c	722403						
FM.2.D.7.d	722404	8	1.00	0.00			
FM.2.F.7.a	722601						
FM.2.F.7.b	722602						
FM.2.F.7.c	722603				7	1.00	0.00
FM.2.F.7.d	722604						
FM.2.F.7.e	722605				1	1.00	n/a
ES.1.C.7.a	751301						
ES.1.C.7.b	751302						
ES.2.E.7.a	752501	5	1.00	0.00			
ES.2.E.7.b	752502	2	1.00	0.00			
ES.2.E.7.c	752503						
ES.2.F.7.a	752601						

(continued)

**Table A-18. Grade 8 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
ES.2.F.7.b	752602						
ES.2.F.7.c	752603						
ES.2.F.7.d	752604						
ES.2.F.7.e	752605						
ES.2.F.7.f	752606						
ES.2.F.7.g	752607						
ES.2.F.7.h	752608						
ES.3.A.7.a	753101	8	1.00	0.00			
ES.3.A.7.b	753102						
UN.1.A.7.a	761101	6	1.00	0.00	1	1.00	n/a
UN.1.A.7.b	761102	1	1.00	n/a			
UN.1.A.7.c	761103	8	1.50	0.53	4	1.25	0.50
UN.1.B.7.a	761201						
UN.1.B.7.b	761202						
UN.1.C.7.a	761301	3	1.00	0.00	7	1.43	0.53
UN.1.C.7.b	761302						
UN.2.A.7.a	762101	7	1.00	0.00	7	1.00	0.00
UN.2.A.7.b	762102						
UN.2.A.7.c	762103						
UN.2.A.7.d	762104						
UN.2.A.7.e	762105						
UN.2.B.7.a	762201						
UN.2.B.7.b	762202						
UN.2.B.7.c	762203						
UN.2.B.7.d	762204						
UN.2.B.7.e	762205						
UN.2.B.7.f	762206						
UN.2.C.7.a	762301	1	1.00	n/a	6	1.17	0.41
UN.2.C.7.b	762302						
UN.2.C.7.c	762303						
UN.2.C.7.d	762304				2	1.00	0.00
UN.2.C.7.e	762305						
UN.2.C.7.f	762306						
UN.2.D.7.a	762401				1	1.00	n/a
UN.2.D.7.b	762402				1	1.00	n/a
UN.2.D.7.c	762403						
IN.1.A.7.a	771101						
IN.1.A.7.b	771102						
IN.1.A.7.c	771103	1	1.00	n/a			
IN.1.A.7.d	771104						
IN.1.A.7.e	771105						
IN.1.A.7.f	771106						

(continued)

**Table A-18. Grade 8 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
IN.1.B.7.a	771201						
IN.1.B.7.b	771202						
IN.1.B.7.c	771203	1	1.00	n/a			
IN.1.B.7.d	771204	1	1.00	n/a			
IN.1.B.7.e	771205						
IN.1.B.7.f	771206						
IN.1.B.7.g	771207						
IN.1.C.7.a	771301						
IN.1.C.7.b	771302						
IN.1.C.7.c	771303						
IN.1.C.7.d	771304						
IN.1.C.7.e	771305						
IN.1.D.7.a	771401						
ST.1.A.7.a	781101						
ST.1.B.7.a	781201	2	1.00	0.00			
ST.2.B.7.a	782201				1	1.00	n/a
ST.2.B.7.b	782202						
ST.3.B.7.a	783201				1	2.00	n/a
ST.3.B.7.b	783202	1	1.00	n/a			
ME.1.A.8.a	811101				1	2.00	n/a
ME.1.A.8.b	811102						
ME.1.C.8.a	811301						
ME.1.D.8.a	811401	6	1.00	0.00	1	1.00	n/a
ME.1.D.8.b	811402						
ME.1.D.8.c	811403				6	1.00	0.00
ME.1.F.8.a	811601				1	1.00	n/a
ME.1.I.8.a	811901	1	1.00	n/a	2	1.50	0.71
ME.1.I.8.b	811902						
ME.1.I.8.c	811903						
ME.2.A.8.a	812101						
ME.2.F.8.a	812601						
LO.1.A.8.a	831101						
LO.1.D.8.a	831401	2	1.00	0.00	1	1.00	n/a
LO.2.A.8.a	832101						
LO.2.A.8.b	832102	1	2.00	n/a			
LO.2.B.8.a	832201	7	1.00	0.00	7	1.29	0.49
LO.2.B.8.b	832202				6	1.00	0.00
LO.2.B.8.c	832203						
LO.2.C.8.a	832301						
LO.2.C.8.b	832302				1	1.00	n/a
LO.2.C.8.c	832303						
LO.2.C.8.d	832304						

(continued)

**Table A-18. Grade 8 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
LO.2.C.8.e	832305						
LO.2.C.8.f	832306						
LO.2.C.8.g	832307						
LO.2.F.8.a	832601						
LO.2.G.8.b	832702						
LO.2.G.8.c	832703						
LO.2.G.8.d	832704						
LO.3.A.8.a	833101	6	1.00	0.00			
LO.3.A.8.b	833102						
LO.3.A.8.c	833103				8	1.00	0.00
LO.3.A.8.d	833104	8	1.00	0.00	7	1.00	0.00
LO.3.C.8.a	833301	1	1.00	n/a			
LO.3.C.8.b	833302	1	1.00	n/a			
LO.3.C.8.c	833303						
LO.3.D.8.a	833401						
LO.3.D.8.b	833402						
EC.1.D.8.a	841401	1	1.00	n/a	6	1.00	0.00
EC.2.B.8.a	842201						
EC.2.B.8.b	842202						
ES.1.A.8.a	851101						
ES.1.A.8.b	851102	4	1.00	0.00			
ES.1.A.8.c	851103	4	1.00	0.00			
ES.1.A.8.d	851104						
ES.2.B.8.a	852201						
ES.2.B.8.b	852202						
ES.2.B.8.c	852203	6	1.00	0.00	5	1.80	0.45
ES.2.C.8.a	852301						
ES.2.C.8.b	852302						
ES.3.C.8.c	852303	1	1.00	n/a			
ES.2.D.8.a	852401	8	1.00	0.00	1	2.00	n/a
ES.2.D.8.b	852402				6	1.00	0.00
IN.1.A.8.a	871101	8	1.00	0.00	8	2.25	0.46
IN.1.A.8.b	871102	8	3.88	0.35	8	2.75	0.46
IN.1.A.8.c	871103	1	2.00	n/a			
IN.1.A.8.d	871104						
IN.1.A.8.e	871105						
IN.1.A.8.f	871106						
IN.1.B.8.a	871201						
IN.1.B.8.b	871202	8	1.00	0.00	8	1.13	0.35
IN.1.B.8.c	871203	1	2.00	n/a	5	1.00	0.00
IN.1.B.8.d	871204	1	2.00	n/a	2	1.50	0.71
IN.1.B.8.e	871205	3	1.00	0.00			

(continued)

**Table A-18. Grade 8 MAP: Grade Span GLEs Matched to Items by Panelists (continued)**

Missouri GLEs	HumRRO GLE Item Codes	Form 1			Form 2		
		Number of Panelists	Mean Number of Items per GLE	SD	Number of Panelists	Mean Number of Items per GLE	SD
IN.1.B.8.f	871206						
IN.1.B.8.g	871207	8	1.00	0.00	8	1.00	0.0.0
IN.1.C.8.a	871301						
IN.1.C.8.b	871302	2	1.50	0.71			
IN.1.C.8.c	871303						
IN.1.C.8.d	871304						
IN.1.C.8.e	871305						
IN.1.D.8.a	871401	8	2.25	0.46	8	1.88	0.35
ST.1.A.8.a	881101	2	1.50	0.71			
ST.1.B.8.a	881201	4	1.00	0.00			
ST.2.B.8.a	882201				6	1.00	0.00
ST.2.B.8.b	882202						
ST.3.B.8.a	883201				1	1.00	n/a
ST.3.B.8.b	883202	4	1.00	0.00	6	1.00	0.00



## Appendix B. Summary of Panelist Comments on Items

Tables B-1 and B-2 present a synopsis of panelists’ comments on the individual items of the Science MAP. To maintain test security, individual item identifiers are not presented, nor are any comments that would reveal the content of an item. Column 2 indicates the number of items receiving such comments, and Column 3 reports how many panelists included this type of comment.

**Table B-1. Grade 5 Science MAP: Summary of Panelists’ (N=8) Comments on Items by Topic**

Comment	Number of items with comment	Number of panelists with comment
• Item content does not align with the GLE.	19	5
• Item content aligns with multiple GLEs.	6	3
• Wording of the item or response options is confusing or misleading.	2	3
• Picture or graphic is not correctly labeled.	1	1

**Table B-2. Grade 8 Science MAP: Summary of Panelists’ (N=8) Comments on Items by Topic**

Comment	Number of items with comment	Number of panelists with comment
• Item addresses more than one GLE.	8	4
• Wording of the item or response options is confusing or misleading.	4	3
• Item requires knowledge beyond the GLE expectations.	3	2
• Item is poorly aligned with GLE and captures tangent knowledge.	2	5
• Item should be edited for clarity (e.g., add or remove words).	2	2
• Picture/graphic is confusing or is needed.	2	2



## Appendix C. Sample Alignment Review Materials

Panelists received the following instruction sheet as a reference guide corresponding with verbal instructions from HumRRO facilitators.

### MAP Science Grade Spans 3-5 and 6-8 Panelist Instructions

Rating Task	Documents Needed	File Format
1 MAP Science GLE (Consensus)	(1) Science GLEs_Grades x-x	Print copy
	(2) Panelist Instructions	Print copy
	(3) Science GLE DOK_Grades x-x	Excel
	(4) Science GLEs DOK Worksheet_Gradesx-x	Print copy
2 MAP Science Items (Individual)	(5) Science GLEs_Grades x-x	Print copy
	(6) Panelist Instructions	Print copy
	(7) Science Items – Grade x	Print copy
	(8) Science Gx-x_ItemRating	Excel
3 Whole Test (Consensus)	(1) Science Items – Grade x	Print copy
	(2) Science WholeTest Gx-x	Excel
	(3) Panelist Instructions	Print copy

#### Prior to alignment steps, train:

- (1) Review handouts, particularly the MAP Panelist Instructions
- (2) Access HumRRO item rating forms:
  - a. Locate form on desktop, double click to open.
  - b. “Save As” the file name and add **underscore and your 3 initials** (e.g., Science G6-8\_ItemRating\_eas).

#### 1 Review MAP GLE for grade span and provide Depth of Knowledge (DOK) rating (Consensus)

##### Train Task:

- (1) Receive the Science GLEs DOK Worksheet\_Gradesx-x paper copy.
  - a. You will handwrite your DOK rating on this form.
- (2) Make DOK ratings
  - a. The facilitator will ask for a volunteer to record the panel’s ratings and will discuss the 4 DOK levels. See the Support Materials section in this document for the information. Refer to this section as needed.

##### Conduct Task:

- (1) Provide individual ratings on the paper copy.
- (2) Determine if everyone provided the same rating. If not, share your reasons for your rating.
- (3) The group will come to a consensus on the rating and majority will rule if necessary.
- (4) The volunteer will enter the group’s consensus rating in the Science GLE DOK\_Grades x-x Excel form.

## 2 Rate MAP Science Items

### Train Task:

- (1) You will review MAP test items, assign a DOK level, select the GLE that the item is targeting, and provide ratings regarding the text, graphic (if appropriate), and overall item quality.
- (2) The facilitator will discuss the columns in the Excel form, including any other tabs toward the bottom of the screen for additional items.
  - a. Columns A and B: The sequence number and item ID
  - b. Column C: Assign the DOK level
  - c. Column D-G: Item Linkage and Overall Alignment
    - D: Select the grade level GLE that best covers the content measured by the item
    - E: Indicate how well the content measured by the item aligns (matches or links) with the selected GLE using the following rating scale.

Rating	Overall Alignment for Item and GLE Rating Descriptions
1	Not aligned to any GLE (No GLE was entered in column C)
2	Weakly aligned (item does not assess the content of the GLE well)
3	Highly aligned (item assesses GLE core content reasonably well)
4	Fully aligned (item assesses content that clearly matches with the GLE)

- F and G: If you rate the overall alignment as 1 or 2, describe exactly what content in the GLE is not covered by the item. Provide a secondary GLE if you feel the item equally assesses another GLE.
- d. Column H-I: Universal Design Characteristics for item text and graphics (when applicable)
    - Refer to the Universal Design Characteristics information on second page of the Support Materials Section.
    - Provide a rating of “Y” for yes if text and graphics meet universal design criteria or “N” for no if item needs revision.
  - e. Column J and K: Overall Item Quality and Explanation

Rating	Overall Item Quality Rating Descriptions
1	Item is of poor overall quality (Explanation required)
2	Item is of good quality, but has repairable flaws (Explanation required)
3	Item is of good quality (what you expect on this and similar tests)
4	Item is of exceptional quality

### Conduct the Task:

- (1) Save the Science Gx-x\_ItemRating on desktop with your 3 initials.
- (2) Rate 2 or so (facilitator will determine) items independently (DOK, alignment, and universal design/quality), then conduct consensus discussion.
- (3) Conduct individual ratings for each item in order. No consensus discussions.
- (4) Save the file regularly!!!

### 3 Rate 'Whole Test' only as consensus

#### Train and Conduct:

- (1) The facilitator will hand out the paper forms and get one volunteer to enter the consensus data.
- (2) Facilitator will provide any additional clarification on the question.
- (3) Conduct consensus discussion and capture results.

### Support Materials

#### DOK Definitions

- **Level 1 (recall)** Items or standards require student recall of information such as fact, definition, term or simple procedure as well as performance of a simple science process or procedure.

Keywords: Identify, define, determine, perform (simple procedure), list.

- **Level 2 (skill/concept)** Items or standards require student engagement of some mental processing beyond a habitual response. Students are required to make some decisions as to how to approach a problem or activity, such as selecting procedures, describing or giving examples of science concepts, deciding how to display or interpret data.

Keywords: Describe, observe, classify, confirm, organize, distinguish, compare.

- **Level 3 (strategic thinking)** Items or standards require student to use reasoning and evidence, plan, and make conjectures. Students should be able to explain phenomena in terms of scientific concepts, explain simple relationships, explain thought process and conclusions, solve non-routine problems, and develop research questions.

Keywords: Connect, explain, analyze, outline procedures, make conclusions, interpret.

- **Level 4 (extended thinking)** Items or standards require student to use complex and abstract reasoning and thinking, often over an extended period of time. Students must design and plan experimental studies, select and appropriate method among alternatives, or deduct the relationship among several variables.

Keywords: Design, plan, and develop experiments; make inferences from results; critique; predict; explain (complex) relationships or differences among variables.

## Universal Design Characteristics

### **Written Content - Does the item...**

- Have concise and readable text
- Commonly used words
- Vocabulary appropriate for grade level
- Minimum use of unnecessary words
- Idioms avoided unless idiomatic speech is being measured
- Technical terms and abbreviations avoided (or defined) if not related to the content being measured
- Sentence complexity is appropriate for grade level
- When time and setting are important to the sentence, place them at the beginning
- Question to be answered is clearly identifiable
- Instructions are presented in the exact order of occurrence

### **Figures/Graphics - Does the item...**

- Have clear pictures and graphics (when essential to item)
- Pictures are needed to respond to item
- Pictures with clearly defined features
- Dark lines (minimum use of gray scale and shading)
- Sufficient contrast between colors
- Color is not relied on to convey important information or distinctions
- Pictures and graphs are labeled
- Avoid graphics or decorative illustrations that are unrelated to the question

Panelists received the Missouri Learning Standards for Science coded for data entry into rating forms. The content of the standards was extracted exactly from the full Missouri Learning Standards document. Only a portion of the coded standards is replicated below for grade 5 as an example.

Grade	Strand	Big Idea	Concept followed by GLE(s)	HumRRO ID
3	1. Properties and Principles of Matter and Energy (ME)	1. Changes in properties and states of matter provide evidence of the atomic theory of matter	D. Physical changes in the state of matter that result from thermal changes can be explained by the Kinetic Theory of Matter	
			a. Compare the observable physical properties of solids, liquids, or gases (air) (i.e., visible vs. invisible, changes in shape, changes in the amount of space occupied)	311401
			b. Identify everyday objects/substances as solid, liquid, or gas (e.g., air, water)	311402
			c. Observe and identify that water evaporates (liquid water changes into a gas as it moves into the air)	311403
			d. Measure and compare the temperature of water when it exists as a solid to its temperature when it exists as a liquid	311404
			e. Investigate and observe that water can change from a liquid to a solid (freeze), and back again to a liquid (melt), as the result of temperature changes	311405
			f. Describe the changes in the physical properties of water (i.e., shape, volume) when frozen or melted	311406
			g. Predict and investigate the effect of heat (thermal energy) (i.e., change in temperature, melting, evaporation) on objects and materials	311407
		2. Energy has a source, can be stored, and can be transferred but is conserved within a system	A. Forms of energy have a source, a means of transfer (work and heat), and a receiver	
			a. Identify sources of thermal energy (e.g., Sun, stove, fire, body) that can cause solids to change to liquids, and liquids to change to gas	312101
			b. Identify sources of light energy (e.g., Sun, bulbs, flames)	312102

Panelists received the Missouri Learning Standards for Science in a rating form in which to make DOK ratings for each GLE. Panelists handwrote DOK ratings (1, 2, 3, or 4) in the last column of the table next to each GLE to facilitate the consensus discussion. The content of the standards was extracted exactly from the full Missouri Learning Standards document. Only a portion of the standards is replicated for grade 5 as an example.

Grade	Strand	Big Idea	Concept followed by GLE(s)	HumRRO ID	D.O.K. Rating
3	1. Properties and Principles of Matter and Energy (ME)	1. Changes in properties and states of matter provide evidence of the atomic theory of matter	D. Physical changes in the state of matter that result from thermal changes can be explained by the Kinetic Theory of Matter		
			a. Compare the observable physical properties of solids, liquids, or gases (air) (i.e., visible vs. invisible, changes in shape, changes in the amount of space occupied)	311401	
			b. Identify everyday objects/substances as solid, liquid, or gas (e.g., air, water)	311402	
			c. Observe and identify that water evaporates (liquid water changes into a gas as it moves into the air)	311403	
			d. Measure and compare the temperature of water when it exists as a solid to its temperature when it exists as a liquid	311404	
			e. Investigate and observe that water can change from a liquid to a solid (freeze), and back again to a liquid (melt), as the result of temperature changes	311405	
			f. Describe the changes in the physical properties of water (i.e., shape, volume) when frozen or melted	311406	
			g. Predict and investigate the effect of heat (thermal energy) (i.e., change in temperature, melting, evaporation) on objects and materials	311407	
		2. Energy has a source, can be stored, and can be transferred but is conserved within a system	A. Forms of energy have a source, a means of transfer (work and heat), and a receiver		
			a. Identify sources of thermal energy (e.g., Sun, stove, fire, body) that can cause solids to change to liquids, and liquids to change to gas	312101	
			b. Identify sources of light energy (e.g., Sun, bulbs, flames)	312102	

Panelists reviewed the individual Science MAP items using the following rating form in electronic format. The format of the rating form was identical for each grade span. The number of items listed per rating form did differ for each grade test.

MAP Science Performance Event Item Review for Grade Span 3-5 Form 3									
	Item DOK	Item Linkage and Overall Alignment				Universal Design Characteristics		Item Quality	
Item Number	Depth Of Knowledge	Linked GLE	Overall Alignment	Explanation	GLE 2	Written Content	Figures/Graphics	Overall Item Quality	Explanation
	1-Recall 2-Skills/concepts 3-Strategic thinking 4-Extended thinking	Enter GLE ID Code	Enter Scale of 1 to 4	If not highly or fully aligned, describe what the item measures that does not match with the GLE	Enter Secondary GLE ID Code	Y=Yes universal N=needs revision	Y=Yes universal N=needs revision	Enter Scale of 1 to 4	If you entered Overall Item Quality rating of '1' or '2', please describe your reasoning
1	2000416								
2	2000464								
3	2000319								
4	2000467								
5	2000560								
6	2000808								
7	2000335								
8	2000373								
9	2000773								
10	2000534								
11	2000580								
12	2000407								
13	2000759								
14	2000620								
15	2000621								