

Alignment Analysis of Michigan Elementary and Middle School Science Standards and Assessments

The alignment of expectations for student learning with assessments for measuring students' attainment of these expectations is an essential attribute for an effective standards-based education system. Alignment is defined as the degree to which expectations and assessments are in agreement and serve in conjunction with one another to guide an education system toward students learning what they are expected to know and do. As such, alignment is a quality of the relationship between expectations and assessments and not an attribute of any one of these two system components. Alignment describes the match between expectations and assessment that can be legitimately improved by changing either student expectations or the assessments. As a relationship between two or more system components, alignment is determined by using the multiple criteria described in detail in a National Institute for Science Education (NISE) research monograph, *Criteria for Alignment of Expectations and Assessments in Language Arts and Science Education* (Webb, 1997).

Dr. Norman Webb from the University of Wisconsin led a three-day alignment institute in Lansing, Michigan, September 21, 22, and 23, 2005. Nine reviewers, including science content experts, the state science coordinator, district science supervisors, and science teachers met to analyze the agreement between the state's science standards and Michigan Educational Assessment Program assessments for grades 5 and 8. Five reviewers were from Michigan, and four were experts brought in from other states. Nine reviewers analyzed four of the six assessments, while four reviewers analyzed the 2004 grade 8 assessment and six reviewers analyzed the 2005 grade 8 assessment. Because of time constraints, the reviewers were divided into two groups to analyze these two assessments. All of the reviewers participated in analyzing the depth-of-knowledge (DOK) levels of the standards.

Summary

The Michigan science standards and assessments for grades 5 and 8 lack full alignment because one standard is not assessed. Reviewers at most only coded three items to Standard II (*Reflecting on Scientific Knowledge*) on any of the six forms analyzed. On most forms, reviewers found no items that they judged to correspond to objectives under this standard. Many of the objectives under this standard seek to have students develop an awareness of the nature of science or an application of science, which are more difficult to measure on an on-demand assessment. Considering the assessments and the other four standards for both grade levels, the alignment is reasonable, with only a few changes needed to achieve full alignment. If the three forms at each grade level are considered in aggregate, then the combined test is fully aligned with the four standards.

If each assessment form is thought of as a separate assessment, then only a few changes to each form are needed to achieve acceptable alignment between the assessment and the science standards. Each grade 5 form would need to have only one or two items replaced or added to meet the minimal acceptable levels on all four alignment criteria. The grade 8 forms would require from three to five additional items, or replaced items, to achieve an acceptable alignment on the four alignment criteria; in each case, for each of the six forms, it would be possible to retain the total number of items and have full alignment if existing items were replaced by new items.

Alignment Criteria Used for This Analysis

This analysis, which judged the alignment between standards and assessments on the basis of four criteria, also reported on the quality of items by identifying items with sources of challenge and other issues. For each alignment criterion, an acceptable level was defined by what would be required to assure that a student had met the standards. These are defined briefly in Table 1.

Table 1
Criteria for Alignment

Criterion	Definition
Categorical Concurrence	<i>At least six items measuring content from a standard</i>
Depth-of-Knowledge Consistency	At least 50% of the items corresponding to a standard had to be at or above the level of knowledge of the standard
Range-of-Knowledge Correspondence	Fifty percent of the benchmarks for a standard had to have at least one related assessment item
Balance of Representation	Items/activities are distributed among all of the benchmarks at least to some degree

Alignment of Curriculum Standards and Assessments

The Michigan science assessments are designed to assess the full range of content in the grade 5 and grade 8 standards over multiple assessment forms. In this study, reviewers analyzed three assessment forms given successively in winter, fall, and then winter for each grade. By design, the science tests forms are not intended to assess all of the objectives under a standard, but should assess a sufficient number of the objectives over three forms.

At grade 5, the alignment was found to be acceptable for four of the five assessment standards across the three forms (Table 2a). Reviewers only coded items on one form (winter 2004) as corresponding to content under Standard II (*Reflecting on Scientific Knowledge*). On the other two forms, nearly all reviewers failed to find any items that measured content related to Standard II. Thus, the grade 5 assessments are not considered to be aligned to Standard II, considering the individual assessment forms and the set of three forms.

Table 2a
Summary of Acceptable Levels on the Four Alignment Criteria for Elementary Science Assessments—Grade 5 Forms, Winter 2004, Winter 2005, and Fall 2005—for Michigan Alignment Analysis

Standards	Alignment Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range of Knowledge	Balance of Representation
Grade 5 Winter 2004				
I – Constructing New Scientific Knowledge	YES	YES	YES	YES
II – Reflecting on Scientific Knowledge	NO	WEAK	NO	YES
III – Using Life Science Knowledge	YES	YES	YES	YES
IV – Using Physical Science Knowledge	YES	YES	WEAK	YES
V – Using Earth Science Knowledge	YES	YES	YES	YES
Grade 5 Fall 2005				
I – Constructing New Scientific Knowledge	NO	YES	WEAK	YES
II – Reflecting on Scientific Knowledge	NO	YES	NO	NO
III – Using Life Science Knowledge	YES	YES	YES	YES
IV – Using Physical Science Knowledge	YES	YES	WEAK	YES
V – Using Earth Science Knowledge	YES	YES	YES	YES
Grade 5 Winter 2005				
I – Constructing New Scientific Knowledge	NO	YES	YES	YES
II – Reflecting on Scientific Knowledge	NO	NA	NA	NA
III – Using Life Science Knowledge	YES	YES	YES	YES
IV – Using Physical Science Knowledge	YES	YES	YES	YES
Table 2a. (continued)				
V - Using Earth Science Knowledge	YES	YES	WEAK	YES

On four of the five science standards, the alignment is generally good, with only a few shortfalls. The average number of items coded by reviewers to Standard I was just short of the required six items to meet the Categorical Concurrence criterion for grade 5 forms, Fall 2005 (5.56) and Winter 2005 (5.44). This is not considered a serious alignment issue. All three grade 5 forms had an acceptable proportion of items with DOK levels that compared to those of the corresponding objectives under the standards. Each form only weakly met the Range-of-Knowledge Correspondence criterion for one or two standards. For all three forms, only one or two objectives under any one standard needed a corresponding item in order to meet an acceptable level on the Range criterion. Considering the set of three forms, a proportion of objectives under each standard, except Standard II, was assessed to have sufficient breadth of content. Across the three forms, reviewers coded items to 67% of the objectives under Standard I, 84% of the objectives under Standard III, 81% of the objectives under Standard IV, and 78% of the objectives under Standard V. All three forms met an acceptable level on the Balance-of-Representation criterion.

As was the case for grade 5, the three assessment forms and grade 8 science are generally aligned, except for Standard II (*Reflecting on Scientific Knowledge*) (Table 3b). Reviewers only coded, on the average, one item on each form as measuring content related to Standard II. This number of items is insufficient to make any judgment of students' knowledge related to Standard II, even by aggregating the three test forms. For the other four grade 8 standards, there is alignment if all three forms are considered. An acceptable level on the Range-of-Knowledge Correspondence criterion was obtained for each of the four standards by aggregating all three test forms. Across the three forms, the proportions of objectives with corresponding items are 83% of the objectives under Standard I, 75% of the objectives under Standard III, 76% of the objectives under Standard IV, and 88% of the objectives under Standard V.

Each of the three grade 8 science assessment forms presented too many issues to be considered fully aligned, in addition to having few items measuring content related to Standard II. The grade 8 Winter 2004 form only weakly met an acceptable level on the Depth-of-Knowledge Consistency criterion for Standard IV because 52% of the corresponding items were under the DOK level of the corresponding objective. Also, the Winter 2004 form only had items that corresponded to about one-third of the objectives under Standard IV, to low a proportion to meet an acceptable level for the Range-of-Knowledge Correspondence criterion.

The Fall 2005 form for grade 8 science and the grade 8 standards presented the greatest number of alignment problems of the three forms. Reviewers' analyses indicated that there were only about four items that measured content related to Standard I, below the minimal number of six required to have an acceptable level on the Categorical Concurrence criterion. These items only weakly met an acceptable level on the Depth-of-Knowledge Consistency criterion, with more than half of the items on the average being judged to have a lower DOK level than the corresponding objective. Also, the four items only measured content related to about one-third of the objectives under Standard I.

Table 2b

Summary of Acceptable Levels on the Four Alignment Criteria for Middle School Science Assessments—Grade 8 Forms Winter 2004, Winter 2005, and Fall 2005—for Michigan Alignment Analysis

Standards	Alignment Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range of Knowledge	Balance of Representation
Grade 8 Winter 2004				
I - Constructing New Scientific Knowledge	YES	YES	YES	YES
II - Reflecting on Scientific Knowledge	NO	NO	NO	YES
III - Using Life Science Knowledge	YES	YES	YES	YES
IV - Using Physical Science Knowledge	YES	WEAK	NO	YES
V - Using Earth Science Knowledge	YES	YES	YES	YES
Grade 8 Fall 2005				
I - Constructing New Scientific Knowledge	NO	WEAK	NO	YES
II - Reflecting on Scientific Knowledge	NO	NO	NO	NO
III - Using Life Science Knowledge	YES	YES	WEAK	YES
IV - Using Physical Science Knowledge	YES	YES	WEAK	YES
V - Using Earth Science Knowledge	YES	YES	YES	YES
Grade 8 Winter 2005				
I - Constructing New Scientific Knowledge	YES	YES	YES	WEAK
II - Reflecting on Scientific Knowledge	NO	NO	NO	WEAK
III - Using Life Science Knowledge	YES	YES	YES	YES
IV - Using Physical Science Knowledge	YES	YES	YES	YES
V - Using Earth Science Knowledge	YES	YES	WEAK	YES

Standards III and IV only weakly met an acceptable level on the Range-of-Knowledge Correspondence criterion because the items corresponded to less than 50% of the objectives, about 7 of 16 objectives under Standard III and 9 of 22 objectives under Standard IV.

Alignment between the Winter 2005 form for grade 8 science and the grade 8 standards was nearly acceptable for all of the standards, except for Standard II. The assessment only had items that measured about 7 of the 16 objectives under Standard V. Therefore, an acceptable level on the Range-of-Knowledge Correspondence criterion was only weakly met. The assessment only weakly met the Balance-of-Representation criterion for Standard I because of the 8 items that reviewers on average coded to objectives under this standard; about five items were coded to one objectives (CS1.5), while each of the other items corresponded to one objective.

Action Needed for Assessments and Standards to be Fully Aligned

In summary, from three to six items would need to be added to each science form for grade 5 and grade 8 to meet the acceptable level on the Categorical Concurrence criterion for Standard II. After that, very few changes are needed. These items would need to be selected with adequate DOK levels and corresponding to at least three of the objectives under the standard:

The Winter 2004 grade 5 form needs to have two items replaced by items that measure objectives under Standard IV, which is not currently assessed.

The Fall 2005 grade 5 form would need to have two items replaced. One of these items should be replaced by an item that measures an objective under Standard I. The other item should be replaced by an item that measures an objective under Standard IV, not currently assessed.

The Winter 2005 grade 5 form only needs one item replaced with one that measures content related to an objective under Standard V, not currently assessed.

The Winter 2004 grade 8 form would require three items to be replaced to achieve full alignment with four of the standards. One item measuring content related to Standard IV needs to be replaced by an item with a higher DOK level. Two additional items need to be replaced by items that measure objectives under Standard IV, not currently being assessed.

The Fall 2005 grade 8 form would need a total of five items to be replaced. Two items need to be added that measure content related to Standard I. If these items have a sufficiently high DOK level and measure content related to objectives not currently assessed, then this would solve the issues on the three criteria not currently met. In addition, one item needs to be replaced with an item measuring an objective under Standard IV that is not currently assessed and two items need to be replaced with items measuring objectives under Standard V, not currently assessed.

The Winter 2005 grade 8 form only needs two items to be replaced with items that measure objectives under Standard V, which are not currently assessed. One of these items could be one that currently measures Objective CS1.5. This then would remove the Balance issue.