

# Alignment Analysis of Michigan Grades 3-8 Mathematics 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. Thirteen reviewers, including mathematics content experts, district mathematics supervisors, mathematics teachers, and a mathematics education professor, met to analyze the agreement between the state's mathematics standards and Michigan Educational Assessment Program assessments for six grades. Ten reviewers were from Michigan, and 3 were experts brought in from other states. Twelve to 13 reviewers analyzed grades 3, 4, and 5 assessments, while 6 or 7 reviewers analyzed grades 6, 7, and 8 assessments. Because of time constraints the reviewers were divided into two groups to analyze the assessments for the higher grades. All of the reviewers participated in analyzing the depth-of-knowledge (DOK) levels of the standards.

## Summary

Overall, the alignment between the mathematics assessments and standards at five of the six grades is reasonable. The grade 6 assessment was fully aligned. Full alignment between the assessments at grades 4, 5, 7, and 8 and the previous grade standards could be achieved by replacing one item (grades 4, 5, and 8) or three items (grade 7) on each assessment. Full alignment for the grade 3 assessment and grade 2 standards would require replacing six items with items that measure content related to data and probability. Reviewers did have some problems coding items to specific standards because limits imposed on number size and type of number in the grade-level expectations did not fully coincide with the numbers used in the items. As a consequence, reviewers coded a relatively large number of items to the goal or standard rather than to

specific grade-level expectations. The lack of exact fit between an assessment item and a grade-level expectation could be due to the grade-level expectations being overly restrictive, or to test blueprint specifications that did not attend to the stated limits. Although, reviewers did code a relative high number of items to the goal or standard, this was not such a serious issue as to consider the assessments and standards not aligned.

## **Alignment Criteria Used for This Analysis**

For the purposes of this analysis, we have employed the convention of standards, goals, and objectives to describe three levels of expectations for what students are to know and do. Standard is used here as the most general (for instance, *Data and Probability*). There are four or five such standards for each grade level. Grades 2 through 5 do not have an algebra standard. The grade 6 standards include a standard for algebra and one for measurement. In grade 7, the measurement standard is dropped. Each standard is comprised of one to four goals, each of which is comprised of up to 15 objectives. These objectives are intended to span the content of the goals and standards under which they fall.

The number of items on the mathematics assessments varied some by grade. The grade 3 assessment had 63 items, the grade 8 assessment had 67 items, and the assessments for grades 4 through 7 had from 69 to 71 items. All of the items were multiple choice, each counted as one point. Thus, the number of points for each assessment was the same as the number of items.

Reviewers were trained to identify the DOK of objectives and assessment items. This training included reviewing the definitions of the four DOK levels and then reviewing examples of each. Then the reviewers participated in 1) a consensus process to determine the DOK levels of the Michigan objectives and 2) individual analyses of the assessment items of each of the assessments. Following individual analyses of the items, reviewers participated in a debriefing discussion in which they gave their overall impressions of the alignment between the assessment and the Michigan curriculum standards for grades 2-7.

To derive the results on the degree of agreement between the Michigan mathematics standards and each assessment, the reviewers' responses are averaged. Any variance among reviewers is considered legitimate, with the true DOK level for the item falling somewhere between two or more assigned values. Such variation could signify a lack of clarity in how the objectives were written, the robustness of an item that can legitimately correspond to more than one objective, and/or a depth of knowledge that falls between two of the four defined levels. Reviewers were allowed to identify one assessment item as corresponding to up to three objectives—one primary hit (objective) and up to two secondary hits. However, reviewers could only code one DOK level to each assessment item, even if the item corresponded to more than one objective. Finally, in addition to learning the process, reviewers were asked to provide suggestions for improving the process.

Reviewers were instructed to focus primarily on the alignment between the state standards and the various assessments. However, they were encouraged to offer their opinions on the quality of the standards, or of the assessment activities/items, by writing a note about the item. Reviewers could also indicate whether there was a source-of-challenge issue with the item—i.e., a problem with the item that might cause a student who knows the material to give a wrong answer, or enable someone who does not have the knowledge being tested to answer the item correctly. For example, a mathematics item that involves an excessive amount of reading may represent a source-of-challenge issue because the skill required to answer is more a reading skill than a mathematics skill.

The results produced from the institute pertain only to the issue of agreement between the Michigan state standards and the assessment instruments. Note that this alignment analysis does not serve as external verification of the general quality of the state’s standards or assessments. Rather, only the degree of alignment is discussed in these results. The averages of the reviewers’ coding were used to determine whether the alignment criteria were met. When reviewers did vary in their judgments, the averages lessened the error that might result from any one reviewer’s finding. Standard deviations are reported, which give one indication of the variance among reviewers.

To report on the results of an alignment study of Michigan’s curriculum standards and six assessments, the study addressed specific criteria related to the content agreement between the state standards and grade-level assessments. Four alignment criteria received major attention: categorical concurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance of representation. These are defined briefly in Table 1.

Table 1  
Criteria for Alignment

<b>Criterion</b>	<b>Definition</b>
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

Alignment between the assessments for grades 3–8 and the Michigan Mathematics Academic Content Standards for grades 2–7 was found to be reasonable and acceptable, except at grade 3 (Table 2). The match between the assessment and the standards on the four alignment criteria had an acceptable level in nearly all comparisons for all six grades. Reviewers did not find any items on the grade 3 assessment that measured content related to any of the three objectives under the grade 2 Data and Probability Standard. Two of the three objectives could be assessed on an on-demand assessment (D.RE.2, Read and interpret pictographs with scales, using scale factors of 2 and 3 and D.RE.3, Solve problems using information in pictographs; include scales such as “each case represents 2 apples”; avoid half cases).

The grade 6 assessment was fully aligned to the grade 5 academic content standards. The grade 5 and the grade 8 assessments were fully aligned with the previous grade’s standards except for a weak level for one standard and the Range-of-Knowledge Correspondence criterion. On the grade 5 assessment, reviewers only coded items as corresponding to 17 of the 37 objectives under the number and operations standard, 46% just below the acceptable level of 50%. The large number of objectives under this standard made it more difficult to achieve the acceptable level on the Range-of-Knowledge Correspondence criterion. Measuring 17 of the objectives indicates that the assessment is addressing some breadth in number and operations and is not considered a large concern for alignment.

On the grade 8 assessment, the items were found to have a slightly lower DOK level than the corresponding objectives under the data and probability standard. One of the objectives under data and probability was judged by the reviewers to have a DOK level of 3 (D.RE.1, Represent and interpret data using circle graphs, stem and leaf plots, histograms, and box-and-whisker plots, and select appropriate representation to address specific questions). Reviewers found items that corresponded to this objective, but rated these items with a DOK level of 2 or 1. The items required students to read or interpret information from these graphical representations, but did not go as far as to have students select appropriate representations, or to actually represent data. The minimum level on the Depth-of-Knowledge Consistency criterion could be fully met by just replacing one item with an item at a higher DOK level. As for grade 5, this is a very minor alignment issue.

The grade 4 assessment did not reach an acceptable level on the Range-of-Knowledge Correspondence criterion with the geometry standard. Reviewers on the average coded items corresponding to 2.77 of the 7 objectives, less than the 50% required to have an acceptable level on the Range criterion. The assessment measured students’ knowledge in identifying two- and three-dimensional shapes, but did not measure students’ knowledge of geometric properties associated with parallel lines and faces, perpendicular lines, and lines and distances. Full alignment could be achieved by just replacing one item that measured one of the objectives not measured on the existing assessment.

Table 2

*Summary of Acceptable Levels on the Four Alignment Criteria for Mathematics Grades 3–8 Mathematics Assessments for Michigan Alignment Analysis*

Standards	Alignment Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range of Knowledge	Balance of Representation
Grade 3				
N - Number & Operations	YES	YES	YES	YES
M - Measurement	YES	YES	YES	YES
G - Geometry	YES	YES	YES	YES
D - Data & Probability	NO	NO	NO	NO
Grade 4				
N - Number & Operations	YES	YES	YES	YES
M - Measurement	YES	YES	YES	YES
G - Geometry	YES	YES	NO	YES
D - Data & Probability	YES	YES	YES	YES
Grade 5				
N - Number & Operations	YES	YES	WEAK	YES
M - Measurement	YES	YES	YES	YES
G - Geometry	YES	YES	YES	YES
D - Data & Probability	YES	YES	YES	YES
Grade 6				
N - Number & Operations	YES	YES	YES	YES
M - Measurement	YES	YES	YES	YES
G - Geometry	YES	YES	YES	YES
D - Data & Probability	YES	YES	YES	YES

Table 2 (continued)

*Summary of Acceptable Levels on the Four Alignment Criteria for Mathematics Grades 3–8 Mathematics Assessments for Michigan Alignment Analysis*

Standards	Alignment Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range of Knowledge	Balance of Representation
Grade 7				
N - Number & Operations	YES	YES	YES	YES
A - Algebra	YES	YES	YES	YES
M - Measurement	NO	YES	WEAK	YES
G - Geometry	YES	YES	YES	YES
D - Data & Probability	NO	YES	YES	YES
Grade 8				
N - Number & Operations	YES	YES	YES	YES
A - Algebra	YES	YES	YES	YES
G - Geometry	YES	YES	YES	YES
D - Data & Probability	YES	WEAK	YES	YES

## Action Needed for Assessments and Standards to be Fully Aligned

In summary, because the alignment between the assessments for grades 3–8 and the standards for grades 2–7 is reasonable, very few changes are needed to achieve full alignment. To achieve full alignment would require these changes or modifications:

Grade 3. Six items need to be replaced or added that measure content related to data and probability. Six items currently measuring content related to number and operations or measurement could easily be replaced with data and probability, without changing the alignment on those two standards.

Grade 4. One item from number and operations needs to be replaced by an item that measures a geometry objective that currently does not have a corresponding item.

Grade 5. Two items currently measuring content knowledge related to number and operations should be replaced by items that measure one of the objectives under number and operations currently not measured. Items that could be replaced are those that relate to objectives N.ME.7 and N.MR.2.

Grade 6. No action required.

Grade 7. Three items need to be replaced. All three items to be replaced could be items related to number and operations and algebra that measure objectives with other related items. These three items should be replaced by one item that relates to measurement and two items that relate to data and probability.

Grade 8. One item that currently assesses content knowledge related to data and probability should be replaced by another item measuring similar content, but with a higher DOK level. For example, either item 21 or 39 could be replaced by an item at DOK Level 3.

Some attention should be given to the high number of items that were coded to generic objectives. This suggests more of a problem with the statement of the standards and the limits imposed on the size and type of numbers. The statement of the standards could be changed very slightly by replacing, for example, two-digit numbers with three-digit numbers. Such subtle changes would result in an even tighter alignment between the standards and assessments.