



JENNIFER M. GRANHOLM
GOVERNOR

STATE OF MICHIGAN
DEPARTMENT OF EDUCATION
LANSING



MICHAEL P. FLANAGAN
SUPERINTENDENT OF
PUBLIC INSTRUCTION

TO: State Board of Education

FROM: Mike Flanagan 

DATE: December 21, 2005

SUBJECT: **APPROVAL OF PROPOSED PERFORMANCE STANDARDS FOR
THE MEAP AND MI-ACCESS ASSESSMENTS**

This memo gives advance notice of the tentative performance standards that will be recommended to the State Board of Education on January 10, 2006. The content of the Board item will be as presented here, with an additional attachment that lists the recommended standards and their impact in terms of student performance. Therefore, the board item that follows is presented as if the standard setting activities had already occurred.

STATE BOARD ITEM AS IT WILL APPEAR AT THE JANUARY 10 MEETING

The purpose of this State Board of Education item is to present the tentative performance standards that were set for the new assessments in the Michigan Educational Assessment Program (MEAP) and MI-Access, which is Michigan's alternate assessment program for students with significant disabilities unable to participate in the MEAP assessments. The procedures used to set the performance standards are described in this memorandum. The results of the standard-setting activities will be presented to the State Board of Education at its January 2006 meeting, since standard setting is scheduled to occur January 4-7, 2006. The Michigan Technical Advisory Committee, comprised of nationally-recognized measurement and statistics experts, is scheduled to review the procedures used and standard setting outcomes on January 8-9, 2006.

Because of the No Child Left Behind (NCLB) law, both MEAP and MI-Access had to initiate grade level assessments in grades 3-8 this fall. As a result, each program has had to set performance standards (cut scores) for each content area and grade level assessed. The table below shows the grades and subject areas for which performance standards were set:

STATE BOARD OF EDUCATION

KATHLEEN N. STRAUS – PRESIDENT • JOHN C. AUSTIN – VICE PRESIDENT
CAROLYN L. CURTIN – SECRETARY • MARIANNE YARED MCGUIRE – TREASURER
NANCY DANHOF – NASBE DELEGATE • ELIZABETH W. BAUER
REGINALD M. TURNER • EILEEN LAPPIN WEISER

608 WEST ALLEGAN STREET • P.O. BOX 30008 • LANSING, MICHIGAN 48909
www.michigan.gov/mde • (517) 373-3324

Subject Area	Grade Levels for Which Standards Were Set	
	MEAP	Mi-Access
English language arts	3 - 8	3 - 8
Mathematics	3 - 8	3 - 8
Science	5, 8	--
Social Studies	6, 9	--

The performance standards will define the levels of performance for the statewide assessments used in Michigan. For MEAP, these are Level 1: Exceeds State Standards; Level 2: Met State Standards; Level 3: Basic; and, Level 4: Apprentice. For MI-Access, the three levels are labeled as Surpassed Standard; Attained Standard; and, Emerging Toward the Standard.

Standard setting activities were carried out for each grade assessed in MI-Access and MEAP, even those that were previously assessed, since the tests at these grades changed from 2004 to 2005. In these cases, the goal was to set standards that are similar to those set before, so as not to change dramatically the AYP determinations based on these performance standards. For newly assessed grades, the goal was to set performance standards that are consistent with the standards set for the grades previously assessed.

Standard setting was carried out by panels of educators working under the direction of the contractors for MEAP and MI-Access and staff of the Department. Each panel spent three or more days reviewing the assessment instrument(s) assigned to them, individually judging the level of performance that students would need to achieve for each of the four performance levels for each assessment, discussing these within their panel, and repeating this process up to three times, with additional performance information provided during each round.

Panelists made their final judgments on their own, and the resulting recommendations are a compilation of these individual judgments. Panelists were then asked to indicate their support for the standards that they set and the processes used to set them. This summary information will also be presented when the results of standard setting are presented to the State Board of Education.

The result of this effort is that each panel recommends performance standards for each program, grade level and content area. These performance standards are presented to the State Board of Education for your discussion and approval.

Standard Setting Process

Overview of MEAP and MI-Access Standards

MEAP Standards

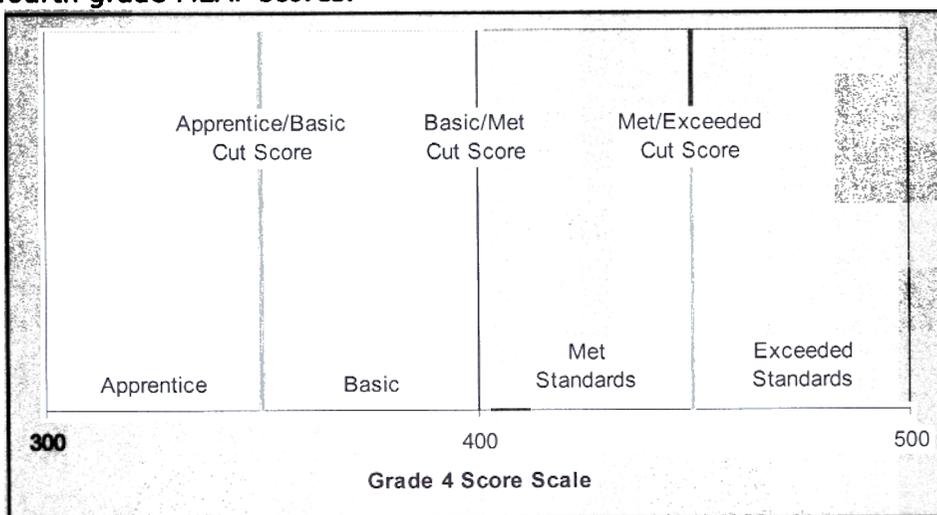
This attachment gives a broad description of the process used to set standards on the MEAP. A thorough description and plan is available upon request.

For the MEAP, standards have been set for the grades and subjects shown in Table 1. On the MEAP, four performance levels are defined by three cut scores as shown in Figure 1 for each of the grades and subjects listed in Table 1. In order to give MEAP scores content-based meaning, after standards were set, the cut score separating the *Basic* level of performance from the *Met Standards* level of performance was defined as $x \times 100$ for grade x . For statistical reasons, the other cut scores were not pre-defined.

Table 1. *Grades and Subjects for MEAP Standard Setting.*

Subject	Grade						
	3	4	5	6	7	8	9
				x			x

Figure 1. Cut Scores and Performance Levels for fourth grade MEAP scores.



Notice that in Table 1, Reading and Writing are listed separately for the MEAP. Because Michigan legislation requires students to meet standards both in Reading and Writing, standards are set separately on the two subjects. However, Michigan has chosen ELA as

the English/Reading indicator for No Child Left Behind (NCLB). Therefore, ELA standards are also set for Michigan students. This is done by combining the Reading and Writing cut scores at each grade level. The combination used to create the ELA scores weights the separate scores such that Reading scores account for two thirds of the overall ELA score and Writing scores account for one third of the overall ELA score.

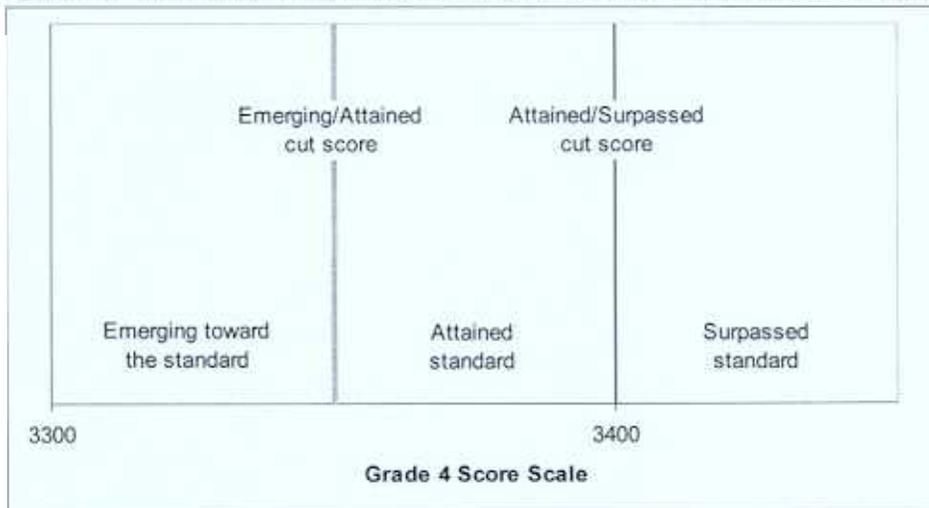
MI-Access Standards

Michigan’s assessment program for students with disabilities (SWD) is comprised of three types of assessments: (1) Functional Independence assessments for students with mild cognitive disabilities, (2) Supported Independence assessments for students with significant cognitive disabilities, and (3) Participation assessments for students with severe cognitive disabilities.

For Functional Independence, standards were set for English Language Arts (ELA) and Mathematics in grades 3-8, defining three performance levels as defined by two cut scores as shown in Figure 2. As with the MEAP, the cut score between *Emerging Toward the Standards* and *Attained Standard* for the Functional Independence assessment is defined as $3000+(x*100)$ for grade x . For statistical reasons, the other cut score is not pre-defined as a certain number.

For the Supported Independence and Participation assessments, the same type of cut scores are used, but the score scale is defined in a different manner to be consistent with the unique characteristics of those assessments. In addition, for Supported Independence and Participation, the cut scores were based on a single scale rather than on a Mathematics and ELA scale.

Figure 2. Cut Scores and Performance Levels for fourth grade MI-Access scores.



Overview of Standard Setting Panelists

For MEAP and each MI-Access assessment, two panels were seated for each subject. One panel examines student performance in grades 3-5, and the other in grades 6-8 for ELA and Mathematics, and one panel examines student performances in grades 5 or 6 and with the other examining student performance in grades 8 or 9 in Science and Social Studies, respectively.

The number of panelists was between 10 and 20 for each MI-Access panel. For MEAP, there were between 20 and 22 panelists for each panel. Because of the small numbers of students taking some of the MI-Access assessments, it was difficult to locate more panelists to participate in standard setting.

Panelists were carefully selected to represent the diversity of Michigan teachers and other stakeholders. Care was taken to balance the panels in terms of grade level of teachers, ethnicity, gender, income, geographical region, and education. Panelists experienced with students with disabilities and English language learners were also carefully selected. In addition, panelists with curriculum and assessment expertise were recruited as were community members and parents of school children. Significant efforts were made to recruit from each of these groups to represent the diversity of Michigan MEAP stakeholders.

The Processes

The process used to set cut scores depended upon the subject and assessment. Because the Writing assessment is very different from the other assessments, a different process was used. For all other subjects and assessments, a method known as Bookmarking or Item Mapping was used. For MEAP Writing, a method known as the Body of Work was used.

Both processes depended upon detailed descriptions of student performance, or Performance Level Descriptors (PLDs). These PLDs are required by NCLB, and indicate what it is that a student at each of the performance levels can do on the assessments.

PLDs were previously developed for MEAP through a process involving Michigan stakeholders. These PLDs define the progression of what children can do at each performance level from one level to the next, and from one grade to the next within each performance level. In other words, these PLDs provided a guide to the MEAP standard setting panelists to determine into which performance level students with certain scores fit. The MEAP standard setting panelists used the PLDs to guide their work, but had the prerogative to modify the PLDs slightly to be aligned more completely with the content of the MEAP.

For MI-Access, the standard setting panelists developed the PLDs based upon student performance on the assessments.

The reason for the difference between the MI-Access and MEAP approaches is logistical. The large amount of work the MEAP panelists were asked to do required that some of the process be performed before actual student results were available.

Bookmarking or Item Mapping

The first activity in Bookmarking was having the panelists actually take and score their own assessments. In this manner they become intimately acquainted with the assessments both in terms of content and difficulty. Following the assessment, panelists became familiar with the PLDs by studying the PLDs and the tasks on the assessment that elicit the knowledge and skills described by the PLDs. Panelists defined and became familiar with what it means to "just meet standards," "just exceed standards," and "just meet basic" requirements.

Following familiarization with the PLDs, the panelists engaged in a practice activity that mimicked the activities in which they engaged for standard setting. This familiarized them with the process they used before they actually made any judgments about student performance.

At this point, panelists were given a booklet of assessment items from a specific grade that had been ordered by the difficulty of the items, from easiest to hardest, based on actual student performance on those items. The first task of the panelists was to independently determine in the ordered assessment booklet where students who “just met standards” move from being more likely to respond correctly to an item to being more likely to respond incorrectly to an item. At this point in the booklet, the panelist placed a “bookmark” indicating that he or she judged that this is the point where that transition is made.

Following this first round of judgments, the panelists were brought back together as a grade-level group to review and discuss the entire set of recommendations made by the panelists, including their rationales for making the judgments they made. They were also presented with impact data that indicated what percentage of students fell into each performance category based upon the median of the group’s recommended cut scores. Panelists were also given impact data for significant student subgroups (e.g. male/female, black/white/Hispanic, SWD, and *etcetera*). Given the discussion and impact data, the panelists then independently revised their recommendations, and were again brought back together for discussion. During that discussion, both grade-level groups were brought together to discuss the alignment of the standards set on the two different grade-level assessments to discuss the premise from which the two groups are working. The panelists are again given the opportunity to revise their recommendations based upon the discussions and impact data as shown across grade levels. Panelists then repeat this process for the other cut scores they are to set for each assessment. This entire process is then repeated for each other grade level assigned to each panel. As a final task, the entire group of panelists for a content area was brought back together to discuss the alignment of the various cut scores that were set across the various grades.

Because of the high-stakes nature of the decisions arising from these panels, and a need to consider previous standards set for the MEAP, the booklet given first to each MEAP panel was from a grade level in which students were assessed in previous years. To assist the panels in setting standards in consideration of previously defined standards, a “reference bookmark” was placed in the ordered booklet to show where a standard equivalent to the previous standards would be (based on the same percentages of students falling into each performance category as in previous years). Panelists were informed what this bookmark represents, but that they were free to recommend different cut scores based upon their study of the PLDs and the content of the assessments. On subsequent standard setting tasks for grades in which assessments were not previously given, “reference bookmarks” were again placed in the book to show where the standards would be placed to maintain smooth trends in proficiency across grade levels. Again, panelists were informed what the standards represent, and that they were free to recommend different cut scores based upon their analysis of the PLDs and the content of the assessments.

In the unexpected case of significant anomalies, such as large changes from grade to grade in the impact of the standards are identified, panelists were presented with that information to inform their decisions.

Body of Work

The process followed for the Body of Work method was similar to that followed for the Bookmarking procedure. However, in the Body of Work method, the panelists were presented with many samples of actual student work. For the Writing MEAP, the work samples consisted of two student essays and five multiple choice editing items. Panelists were asked to identify the performance levels of each of the work samples. Cut scores were then recommended in several rounds, and the median of the final recommendations was taken as the cut score.

All other procedures and rounds were similar to the Bookmarking method used for the other assessments.

Results

Prior to this State Board of Education meeting, the results were presented to the Technical Advisory Committee (TAC) of the Office of Educational Assessment and Accountability (OEAA). The TAC, which is comprised of nationally-known assessment and statistics experts, reviewed and approved the full results.

The results presented to the State Board of Education are comprised of the recommended cut scores given by the panels of stakeholders, as well as the impact data (the percentages of students in each performance category in each subject, grade level, and student subgroup). Other information is available to answer any questions including the complete final standard setting plan and technical report. Psychometricians from OEAA, Pearson Educational Measurement, and the Standard Setting contractor are present as well to answer any questions.