

# Proposed Design: The Missouri Assessment Program Science Standard Setting for July 2008

Karla Egan  
CTB/McGraw-Hill  
December 11, 2007

In July 2008, a modified Bookmark Standard Setting Procedure (MBSSP) will be implemented to establish cut scores for the Missouri Assessment Program (MAP) Science tests in Grades 5, 8, and 11. The MBSSP will establish four achievement levels on the Science MAPs: *Below Basic*, *Basic*, *Proficient*, and *Advanced*.

This document provides a brief overview of the Science MAP and highlights the MBSSP used to establish cut scores for the Science MAP.

## Overview of the Science MAP

Missouri currently administers the Science MAP in Grades 3, 7, and 10 as voluntary assessments. Districts choose whether or not to administer the test. In 2007, 372 school districts participated in either the Grade 3, 7, or 10 Science MAP. This represents nearly 2/3 of Missouri school districts that participated in the grade-level Communication Arts or Mathematics MAP.

For the voluntary Science MAP, cut scores were established in 1998. Table 1 summarizes the percentage of students in each achievement level from the 2007 administration for those districts that participated.

**Table 1. Percent of Students in each Achievement Level for Science MAP by Grade, 2007**

	Grade		
	3	7	10
<i>Step 1</i>	4.19	24.66	20.51
<i>Progressing</i>	12.43	37.53	35.79
<i>Nearing Proficient</i>	37.45	22.53	37.03
<i>Proficient</i>	34.51	12.24	5.29
<i>Advanced</i>	11.42	3.04	1.37
<i>Proficient or Advanced</i>	45.93	15.29	6.66

Field testing began in May 2006 for the new generation of the Science MAPs. In 2008, the new generation of the Science MAP will be administered operationally for Grades 5, 8, and 11, and the assessment will be required of all Missouri students in these grades. The Science MAP will no longer be administered in Grades 3, 7, and 10. The July 2008 standard setting will be based on data from the April 2008 administration.

Beginning with the first operational assessment in 2008, new grade levels will be assessed, and the Science MAP will be required by DESE as opposed to being an optional assessment. Note that different achievement levels (*Below Basic*, *Basic*, *Proficient*, and *Advanced*) will be used for the

new MAP Science. For these reasons, continuity would not be expected between the impact data from 2007 and the impact data for the new MAPs.

### ***Overview of the Science MAP Standard Setting***

As with the 2005 Standard Setting for Communication Arts and Mathematics, a modification of the BSSP is necessary because of Missouri Senate Bill 1080, which “amends certain policies with regard to MAP testing by stating that, no later than June 30, 2006, the State Board of Education shall:

- 1) Align the achievement standards of the MAP so that such indicators meet, but do not exceed, the achievement standards of the National Assessment of Educational Progress (NAEP) exam; and
- 2) Administer any other adjustments necessary in order to aid the state in conforming to federal standards, including the institution of yearly examination of students in the required subject areas where compelled by federal regulations.” (Missouri National Education Association, 2004).

Missouri Senate Bill 1080 was passed so that “Missouri schools are judged according to an appropriate national standard when determining Adequate Yearly Progress” (Missouri National Education Association, 2004) under the federal NCLB Act.

This law has been interpreted such that the *Proficient* achievement level will meet, but not exceed, the NAEP achievement standards. In other words, the percentage of students who attain *Proficient* on MAP should be similar to or slightly higher than the percentage attaining *Proficient* on NAEP. The percentage of students in the other three achievement levels may be allowed to vary between NAEP and MAP.

### ***Standard Setting Design***

The design of the MBSSP for the Science MAPs will be based on the design of the 2005 MBSSP for Communication Arts and Mathematics. This design involves three distinct steps:

1. Determining preliminary *Proficient* cut scores
2. Validating achievement standards on the MAP using the modified Bookmark Standard Setting Procedure
3. Smoothing validated achievement standards across grades

#### **Step 1. Determining Preliminary Proficient Cut Scores**

To comply with Missouri law, the MAP achievement standards across grades of a content area must reflect the corresponding achievement standards of the NAEP, specifically *Proficient*. In addition, the *Proficient* achievement standards must be well articulated across the grades of a content area. For the Communication Arts and Mathematics MAPs, DESE established a range of initial *Proficient* cut scores based on impact data. This same range will be used for the Science

MAPs. Participants in the Science MBSSP will be allowed to place their *Proficient* bookmark anywhere within this range of cut points.

The initial range of cut points for *Proficient* will be the same as those presented to the Communication Arts and Mathematics standard setters in December 2005. This range allows participants to use some discretion in placing the cut scores, yet restricts the cut scores so that they are in compliance with Missouri law. Initial cut points will not be set for *Basic* or *Advanced*.

The range of cut points will be based on percentages, which will be translated to cut points by CTB. The same range of percentages will be used for the three grades tested in Science.

- The high end of the range (in terms of scale score points) will be set for each grade/content area such that approximately 26% of students are at or above the cut score.
- The low end of the range (in terms of scale score points) will be set for each grade/content area such that approximately 43% of students are at or above the cut score.

The high end of the range was based on the percentage of students classified as *Proficient* or *Advanced* on the NAEP test. The low end of the range was based on the percentage of students classified as *Proficient* or *Advanced* on the Communication Arts and Mathematics MAP.

A different range could have been created using MAP and NAEP Science data. For example, the Grade 3 Science MAP data would have allowed the cut score associated with the low end of the range to be even lower than that allowed by the constraints for the 2005 standard setting. And the Grade 8 Science NAEP National Public data would have slightly lowered the high end of the range for which 27% of students would be allowed to be classified as *Proficient* or *Advanced* instead of the 26% allowed by last year's range. DESE has, however, elected to use the same range as was used for the Communication Arts and Mathematics standard setting in order to maintain consistency between the standard settings for the different content areas.

Different grades are being assessed by the new the generation of Science MAPs. This, of course, limits the comparability between the old and new tests. The same level of performance (46% *Proficient* and *Advanced*) achieved on the Grade 3 Science test would not be expected on the Grade 5 Science test, which has new content standards and has never been assessed by a large-scale testing instrument.

## **Step 2. Validating Achievement Standards on the MAP Using the Modified Bookmark Standard Setting Procedure**

The Science MAP Standard Setting is designed to be held over the course of four days, from July 8 to July 11, 2008. A proposed schedule for the standard setting workshop is presented in Table 2.

**Table 2. Proposed Schedule for the Science MAP Modified Bookmark Standard Setting**

<b>Day</b>	<b>Time</b>	<b>Activity</b>
July 8, 2008	PM	Table Leader Training
July 9, 2008	AM	Opening Session
	PM	Bookmark Activities
July 10, 2008	AM	Bookmark Activities – Round 1 vote
	PM	Bookmark Activities – Rounds 2 & 3 votes
July 11, 2008	AM	Description Writing/Smoothing
	PM	Description Writing

The standard setting will occur largely as specified in the *CTB Standard Setting Handbook* (2005); moreover, to the degree possible, this standard setting will follow the design used in the 2005 MAP standard setting. Differences between the 2005 and proposed 2008 standard setting will be noted.

The proposed modifications to the Bookmark Standard Setting Procedure (MBSSP) described in the *Handbook* include: (1) configuration of the standard setting committee; (2) presentation of the cut score ranges and policy model; and (3) presentation of the pre-set bookmarks on the item maps. This MBSSP method promotes the establishment of well-articulated achievement levels by informing participants of the need for a policy model and the rationale used to identify the selected policy model. A discussion of each of the modifications follows.

### **Configuration of the Standard Setting Committee**

The MAP serves widely diverse populations throughout Missouri. As described in the *CTB Standard Setting Handbook*, the standard setting committee should be composed of participants who mirror the major demographic groups in Missouri; a homogenous sample of participants could introduce bias into the resultant cut scores.

Standard setting participants will be recruited for Science in Grades 5, 8, and 11. Fifteen panelists will be recruited for each grade. Fifty percent of the panelists will be educators, 30% will be administrators, and 20% will be non-education professionals. CTB recommends smaller groups of participants whenever constraints are placed on the standard setting.

In addition, three Table Leaders will be selected by DESE for each grade. The Table Leader has additional roles and responsibilities that are detailed in the *CTB Standard Setting Handbook*. At the MAP Standard Setting, the Table Leaders will *not* be voting members of the standard setting committee. The selected Table Leaders will be educators with whom DESE has worked extensively in the past. Because of their close ties with the state, DESE has elected to have these Table Leaders serve as non-voting members.

Based on the inclusion of 15 *voting* participants per grade, 45 participants will be needed for the standard setting. DESE will also select the nine Table Leaders.

The standard setting workshop will be held in three breakout rooms, one for each grade. The participants within each grade level will be split into three tables of five participants each.

The number of voting participants is a key difference between the 2005 and 2008 standard settings. In 2005, there were approximately eight participants per grade/content area while in 2008 there will be 15 participants per grade/content area. Because of the increased number of participants, CTB will recommend some changes to the 2005 standard setting design.

### Presentation of the Cut Score Ranges and Policy Model

During orientation, participants will be shown the *Proficient* cut point ranges selected by DESE, and DESE will explain the rationale for this range as well as Senate Bill 1080. In addition, DESE should explain the need for a system of cut scores and impact data that makes sense across the grades. At the beginning of the workshop, DESE will explain that the participants should strive to set well-articulated data. If DESE has a particular policy model for articulating the data that they would like participants to consider, then participants should be shown this information during the opening session. Participants should understand that their input is vital to the standard setting, that the standard setting will add evidence of validity to the cut scores, and that the overall system of achievement standards must be considered when examining the cut scores for individual grades.

Throughout the standard setting, participants will be reminded that their recommendations will be reviewed by DESE and the technical advisory committee (TAC) before presentation to the body which will ultimately adopt the cut scores, the State Board of Education. DESE should explain to participants that the State Board of Education may alter the participant-recommended impact data and cut scores as they deem necessary.

### Presentation of the Pre-set Bookmarks on the Item Maps

During the workshop, the preliminary *Proficient* cut point range will be shown on the item maps, as illustrated in Figure 1. Participants will study the items in the ordered item booklets. They will then be given an opportunity to recommend a *Proficient* bookmark anywhere within the given range. CTB will help participants stay within this range of cut scores by printing the preliminary *Proficient* cut point ranges directly on the item maps.

**Figure 1. Sample item map with sample pre-set bookmark ranges.**

Order of difficulty (easy to hard)	Location	Form	Item No.	Item Type	Score Key	Content Strand	What does this item measure? That is, what do you know about a student who can respond successfully to this item/score point?	Why is this item more difficult than the preceding items?
1	548	12	1	MC	2	1		N/A
2	552	9	4	MC	3	4		
3	560	9	3	MC	2	5		
<b>Proficient Bookmark range begins on Page 4</b>								
4	563	12	2	MC	4	1		
5	571	12	4	MC	2	4		
6	580	9	6	MC	1	2		
7	584	12	5	MC	4	3		
<b>Proficient Bookmark range ends on Page 7</b>								
8	594	12	7	MC	3	2		

**Target Students** . Before the first round of voting, CTB typically has participants develop descriptions of the *Basic*, *Proficient*, and *Advanced* Target Students based on the state content standards. Because of Senate Bill 1080, the Missouri law that requires Missouri achievement levels to be similar to NAEP, participants will be given NAEP descriptors of *Basic*, *Proficient*, and *Advanced*. The NAEP descriptors, along with Missouri Grade-Level Expectations, will serve as the basis for developing the Target Student descriptors.

**Round 1.** The range for the *Proficient* bookmark will be pre-set on the item map. The *Basic* and *Advanced* cut scores will not be constrained. As participants study the Ordered Item Booklets in Round 1, they will describe the content measured by each item. Participants will use this information to judge where they want to place their *Proficient* bookmarks. It will also provide a sufficient foundation for participants to place bookmarks for *Basic* and *Advanced*.

**Round 2.** During this round, participants will be shown the median bookmarks for their grade-level as well as the associated impact data. Staff from DESE and CTB will facilitate the discussion of the grade-level impact data. Following the discussion of the impact data, the table leaders will lead a content-based discussion of the bookmarks at their table. After discussion, participants will place their Round 2 bookmarks.

This represents a departure from the 2005 standard setting design. During the 2005 MAP standard setting, cross-grade impact data was presented to the entire content group at the beginning of Round 2. The entire content group (all grades within a content area) participated in a large-group discussion of the cross-grade impact data. For the 2005 standard setting, this presentation worked well because of the small size of the grade-level groups (eight participants per grade-level).

This is not recommended for the 2008 standard setting due to the increased number of participants being invited. In 2008, the grade-level groups are too large for a meaningful cross-grade conversation about impact data. Moreover, the requirement that cross-grade impact data be shown at Round 2 necessitates that all three groups work at the same pace. This rarely occurs in a standard setting. Typically, the elementary groups work at a faster pace than high school groups, sometimes finishing an hour or two ahead of the high school groups. If cross-grade impact data is presented, then one or more groups will either be rushed to finish Round 1 activities or one or more groups will encounter a good deal of down time while they wait for a different group to complete Round 1 activities.

**Round 3.** At the beginning of Round 3, participants will be shown the cross-grade impact data and cut scores. The cross-grade impact data will be shown in each grade-level group instead of to the entire content area in one large room. Again, this is a departure from the 2005 standard setting design where the cross-grade impact data was presented to all grade-level groups at one time. In 2008, however, the size of the grade-level groups is not conducive to a discussion across the three groups.

During this round, CTB and DESE will show participants where the impact data and/or cut scores are poorly articulated. To enable participants to establish cut scores that are well articulated, CTB can provide each table with an item map that includes impact data, if desired by DESE.

Participants will be able to use this item map to place bookmarks in order to maintain consistent data with the other grades within their content area.

The cross-grade impact data will be based on the participants' Round 2 votes if they are available. If the Round 2 votes are not available, then the cross-grade impact data will be based on the Round 1 votes.

Once participants finish discussion of the impact data, they will be invited to discuss their bookmarks. This will be the participants' final opportunity to place bookmarks.

***Descriptor Writing.*** Following the Round 3 presentations, participants will write draft achievement-level descriptors. These descriptors will reflect the aggregate knowledge, skills, and abilities of the items needed to attain each achievement level. This process will be led by CTB Development staff.

Throughout the MBSSP, participants will discuss and define the Target Students. The Target Student definitions that are drafted, reviewed, and discussed by all participants in a grade group will be based only on participant expertise and on the NAEP performance standards and Missouri GLEs. Throughout the BSSP, participants will gain new knowledge and insights that will help them refine their Target Student definitions. For example, when participants study the OIB, they will actively discuss the skills of the Target Students. During this time, they may discover that a particular skill is more difficult for students than they originally thought or they may find the opposite to be true. They will further refine their Target Student definitions when they place and discuss their bookmarks and when they see impact data. The descriptor writing portion of the standard setting workshop represents the culmination of the discussions about Target Students.

In descriptor writing, participants will create draft descriptors based on their final recommended cut scores. The knowledge, skills, and abilities of the Target Students will be described from the content of the items. Unlike the original Target Student definitions that participants discuss at the beginning of the workshop, these descriptors reflect student performance and participant input on what it means to be classified in each achievement level.

### **Step 3: Smoothing Validated Achievement Standards, Across and Within Grades**

Once standard setting participants have made final recommendations, DESE may adopt the recommended achievement standards or may choose to smooth these standards.

***Smoothing Participant Recommendations.*** Following the standard setting workshop, it is possible that the cut scores participants recommend may need to be smoothed. If this is necessary, the Table Leaders will meet while the rest of the participants are drafting the achievement-level descriptors. The Table Leaders will examine the entire system of impact data and cut scores and make recommendations for smoothing the impact data and cut scores. During the opening session, the Table Leaders and participants should be informed that this smoothing may occur if necessitated by the results of the standard setting.

## References

CTB Standard Setting Handbook. (2005). CTB/McGraw-Hill LLC.

Missouri National Education Association. (2004, April). *Legislative Update 2004*. Retrieved July 18, 2005, from <http://www.mnea.org/publications/legislative/lu04-Apr23.htm>

Missouri National Education Association. (2004). *2004 MNEA Bills of Interest*. Retrieved July 18, 2005, from <http://www.mnea.org/capitol/BillSummSenate.htm>