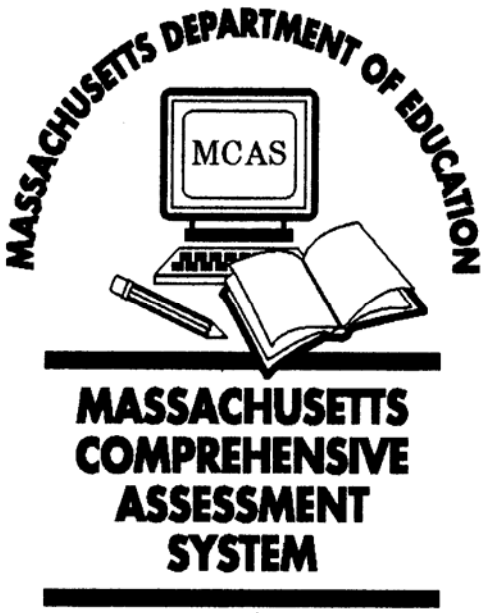
APPENDIX G

2006 MCAS STANDARD SETTING REPORT



2006 MCAS
Standard Setting Report

Reading Grade 3
ELA Grades 5, 6 and 8
Mathematics Grades 3, 5 and 7

Table of Contents

[Editor's Note: Page numbers given below are not the pages found in the original *2006 MCAS Standard Setting Report*; they apply to this Appendix G of the *2006 MCAS Technical Report*.]

Overview

History of MCAS Standard Setting

The MCAS tests have been administered to students in Massachusetts since 1998. At that time, Math, Science and Technology and ELA were the subjects administered. In subsequent years, additional grades and content areas have been added. After the initial administration of each of these new tests, performance standards were set. For grade 3 through 8, Table 1 displays the history of when standards were set for different content areas and the standard-setting method used; as can be seen in Table 1, standard setting has been done by several different contractors over the years: Harcourt Educational Measurement (HEM), Beck Evaluation and Testing Associates (BETA), and Advanced Systems in Measurement and Evaluation (ASME, now Measured Progress).

Table 1

Grade	Content	Year	Method	Contractor
3	Reading	2001	Bookmark	HEM, BETA
4	Math	1998	Body of Work	ASME
	ELA	1998/2001	Body of Work	ASME/HEM, BETA
5	Science	2003	Body of Work	HEM, BETA
6	Math	2001	Body of Work	HEM, BETA
7	ELA	2001	Body of Work	HEM, BETA
8	Math	1998	Body of Work	ASME
	Science	2003	Body of Work	HEM, BETA

Current Context

As part of NCLB requirements, in 2006, several new grades and content area tests were added to MCAS. As was the case previously, it was necessary to establish performance standards for these newly added tests. Additionally, for Grade 3 Reading, when initial standards were set in 2001, only three performance levels were established. To be consistent with the other grades and in compliance with NCLB, a fourth performance level needed to be added to Grade 3 Reading. This new performance level in grade 3, *Above Proficient*, is the highest level students can achieve; in the other grades the highest level is called *Advanced*. Table 2 displays the grades and content areas for which standards needed to be and were set in the summer of 2006.

Table 2

Grade	Content	Notes
3	Reading	Establish *Above Proficient*
	Math	
5	ELA	
	Math	
6	ELA	
7	Math	
8	ELA	

Appendix G

The new grades and content areas assessed in 2006 'fill in the gaps' of grades and contents so that all grades, 3 through 8, are tested in the areas of Mathematics and Reading or ELA.

Vertically Moderated Standards

A result of the addition of the new tests is that MCAS now assesses Math and Reading in all grades, 3 through 8. The standard-setting process used in 2006 for the new operational MCAS tests, was designed not only to establish cut scores between the performance levels for the new tests but also to create a system of logically consistent and coherent performance standards across the grades.

Establishing a system of vertically moderated standards requires that the content standards assessed by the tests are designed as a continuum across the grades, and the definitions of the performance levels are also developed to be consistent across the grades. The MCAS tests in grades 4 through 8 were designed specifically to meet these requirements.

Because students are now being tested in all adjacent grades 3-8, and because there is consistency in the performance level descriptors, it is important that the resulting distribution of students across the performance levels is coherent and makes sense across grades (Kane, 2001). For example, it would make little sense (and likely would be indefensible) to have 10% of the students *Proficient* or *Advanced* in one grade, 60% *Proficient* or *Advanced* in the next, and 30% *Proficient* or *Advanced* in the one after that. Within the current environment of testing at adjacent grades and having consequences (at least at the school level) associated with student performance on those tests, several authors (e.g., Mitzel, 2005, Ferrara, Johnson, & Chen, 2004) have argued that cross-grade consistency of results should be an explicit part of the standard-setting enterprise.

Mitzel (2005) describes three scenarios under which standard-setting activities occur within state departments of education. In the first of his scenarios, performance standards on existing tests are to remain intact, and performance standards are to be established for new tests added to the assessment system. This scenario describes the present MCAS state of affairs. To set standards in this situation, Mitzel and others (e.g., Lissitz and Huynh, 2003) describe a process for achieving consistency of results by starting with the percentages of students in performance levels for existing grades then fitting a line or a growth function to 'smooth' those percentages across grades. This fitted line or function would then be used to determine the percentage of students in each performance level at the new grades. The observed scores that produced those percentages would be the cut scores on the new tests. The outcome of this smoothing process is referred to as 'vertically moderated standards'.

This is similar to the process that was implemented for Massachusetts, with two major differences. First, grade 3 was excluded from the vertical moderation process (the reasons for this are discussed below), and second the 'vertically moderated standards' were used as initial cut scores that were validated by a panel in an abbreviated standard-setting meeting (more correctly termed a standards-validation meeting). The process achieved consistency of results across grades by examining the percentages of students in each performance level on the existing tests then estimating than expected percentages of students in each performance level on the new tests. It was determined that this underlying relationship does not exist for grade 3 as it does for the other grades making grade 3 inappropriate for inclusion in the vertical moderation process.

Grade 3

The premise on which vertically moderated standards are built is that there is a consistency in the underlying construct measured at each grade. This consistency is reflected in the performance level descriptors across grades as well as the content specifications for adjacent grade tests. Consequently, there is an expectation of equivalence of the meaning of the performance levels across grades.

Discussions with MDOE staff and inspections of the performance level definitions indicated that this underlying relationship does not exist for grade 3 as it does for the other grades, making grade 3 inappropriate for inclusion in the vertical moderation process. The content area experts assert that at grade 3, and below, the constructs of reading and math are fundamentally different than above grade 3. These differences stem from the developmental stages of reading and math acquisition through which students progress up through grade 3. These stages are typically completed by grade 4.

The recognition that grade 3 is different than the later grades was evidenced in the initial setting of standards of Grade 3 Reading where only three performance levels were defined. As stated earlier in this document, for Grade 3 Reading a fourth performance level needed to be added. An additional consequence of the differences associated with grade 3 is that Grade 3 Math did not go through a standards-validation process, but instead used a standard-setting process for which initial cut scores were not identified.

Due to the differences associated with grade 3, in the standard setting process initial cut-scores were not identified as they were for the other grades. Rather, panelists simply classified each body of student work into one of the performance levels. For grade 3 Math student work was categorized as *Warning, Needs Improvement, Proficient, or Above Proficient.* As noted earlier, for grade 3 Reading, a fourth performance level, *Above Proficient,* was added to the existing performance levels. Consequently, here, panelists only classified work as being in the *Above Proficient* category or being in the *Proficient* category. Student bodies of work in the *Warning* or *Needs Improvement* categories were not categorized by panelists because these cut-scores were already established. Also, an additional change to grade 3 Reading this year is the inclusion of open response (OR) scores in the students' total scores. In past years, OR scores were reported but not included in students' total scores, so they played no role in categorizing students into performance levels.

Standard-Setting Meetings

The standard-setting meeting to establish cut scores for the MCAS Grade 3 Reading and Grades 5, 6, and 8 English Language Arts (ELA) tests was held Tuesday and Wednesday, July 11th and 12th; the standard-setting meeting for MCAS Mathematics tests in grades 3, 5, and 7 was held Tuesday and Wednesday, August 22nd and 23rd. Standard setting followed the procedures specified in the proposal submitted to the Massachusetts Department of Education (MDOE) by Measured Progress in May 2006. A copy of that proposal is included as Appendix A.

As mentioned above, with the exception of the two grade 3 panels (for Reading and Mathematics), each panel performed a standards-validation process in which panelists were given starting cut points and either validated those starting cuts or recommended changes to them. The starting cut points were determined by interpolating from existing cut points in other grades (4 and 7 for ELA; 4, 6, and 8 for

Mathematics). For the two grade 3 panels, however, starting cut points were not provided, so the panels completed a full standard-setting process.

The standard-setting method implemented for all grade spans is a modified version of the Body of Work (BoW) method (Kingston, et al., 2001). In the original BoW method, panelists review different folders of student work in each round; for MCAS, the primary modification is presentation of the same student folders during multiple rounds of ratings. A second modification is the inclusion of an item-mapping process for multiple-choice items. Details of the BoW method and the item-mapping activity are provided in the Tasks Completed During the Meeting section below. To help ensure consistency of procedures between panels, each panel was led through the standard-setting or standards-validation process by trained facilitators from Measured Progress.

Tasks Completed Prior to the Meeting

Creation of Performance Level Descriptors

The Performance Level Descriptors provided panelists official descriptions of the knowledge, skills, and abilities students are expected to be able to demonstrate to be classified into each performance level. These Performance Level Descriptors are provided in Appendix C of this document.

Preparation of Materials for Panelists

The following materials were assembled into folders for presentation to panelists and are described in greater detail in the Tasks Completed During the Meeting subsection of this chapter:

- Meeting agenda
- Confidentiality agreement
- Student test booklet
- Answer key/scoring rubrics
- Item map
- Performance Level Descriptors
- Training set of student folders
- Student folders for standard setting
- Rating forms
- Evaluation form

The meeting agendas, Performance Level Descriptors, sample rating forms, and results of the evaluations are provided in the appendices to this report.

Preparation of Presentation Materials

The PowerPoint presentations used in the plenary sessions were prepared jointly by Measured Progress and MDOE staff prior to the meetings. Copies of the PowerPoint slides are included in Appendix D of this document.

Preparation of Instructions for Facilitators

Two versions of a document, "General Instructions for MCAS Standard Setting Group Facilitators," were created for group facilitators' reference during the meetings: one for Grade 3 Reading and Mathematics, and another for ELA and Mathematics in all other grades. Copies of these instructions are included in Appendix E of this document.

Preparation of Systems and Materials for Analysis During the Meeting

The programming of all analyses to be conducted during the standard-setting meetings was completed and systems were thoroughly tested ahead of time.

Determination of Starting Cut Points (Grades 5 through 8)

As described in the Overview section of this report, one of the purposes of the 2006 standard-setting meetings was to establish cut points on the tests for grades 4 through 8 that are coherent across grade levels, resulting in vertically moderated standards across those grades. Several authors (see, for example, Mitzel, 2005; Ferrara, Johnson, & Chen, 2004) suggest that establishing consistency of results across grades should be built into the standard-setting process. Both Mitzel and Lissitz and Huynh (2003) recommend achieving this goal by fitting a line to existing standards to smooth those standards and to find cut points for the intervening grade levels. For MCAS, this process was modified by using the cut points calculated in this way as the starting cuts for a validation process. The only exceptions were Grade 3 Reading and Mathematics, which were excluded from the vertical moderation process for the reasons discussed above.

More specifically, the MCAS starting cuts were determined using the five steps listed below. This process was presented to, and approved by, the Massachusetts Technical Advisory Committee (TAC) prior to the standard-setting meetings. Copies of the Powerpoint slides presented at the TAC meeting are included in Appendix F.

1. Find the percentage of students who fell below each raw score cut for grades 4 and 7 (ELA) or grades 4, 6, and 8 (Mathematics), based on the 2006 MCAS results.
2. Standardize the percent-below values using the *z*-transformation.
3. Calculate a line of best fit across grades.
4. Use the inverse-*z*-transformation to translate the *z*'s back into percent-below values.
5. For the intervening grades (5, 6, and 8 for ELA; 5 and 7 for Mathematics), find the raw score associated with the observed percent-below value closest to, but not lower than, the smoothed value.

These five steps were repeated for each cut point. The starting cuts and the associated percentages of students in each performance level are presented in tables 3 and 4 in the last section of this report.

Recruitment and Selection of Panelists

The composition of the panels will be composed of primarily grade-specific teachers, with a sprinkling of administrators, business and community leaders, and another education advocates including representation from special populations (LEP, ELL, SPED).

Appendix G

Panelists were selected prior to the standard-setting meeting by MDOE. A listing of these panelists by Grade and Content Area may be found in Appendix H. The total number of panelists who participated was 116, distributed as follows:

- Grade 3 Reading: 20
  - Teacher 17
  - Administrator 2
  - Business/Community Leader 1
- Grade 5 ELA: 19
  - Teachers 17
  - Administrator 1
  - Business/Community Leader 1
- Grade 6 ELA: 21
  - Teacher 18
  - Administrator 3
  - Business/Community Leader 0
- Grade 8 ELA: 20
  - Teacher 14
  - Administrator 5
  - Business/Community Leader 1
- Grade 3 Mathematics: 20
  - Teacher 14
  - Administrator 6
  - Business/Community Leader 0
- Grade 5 Mathematics: 21
  - Teacher 17
  - Administrator 3
  - Business/Community Leader 1
- Grade 7 Mathematics: 21
  - Teacher 16
  - Administrator 4
  - Business/Community Leader 1

The sample of panelists was chosen to be as geographically representative of Massachusetts's diverse educationally oriented concerned citizen population as possible. Copies of the application forms potential panelists were asked to submit are included in Appendix G, and a list of the panelists with their affiliations is included in Appendix H.

The sample of panelists was chosen to be as geographically representative of Massachusetts's population as possible. Copies of the application forms potential panelists were asked to submit are included in Appendix G, and a list of the panelists with their affiliations is included in Appendix H.

Tasks Completed During the Meeting

General Orientation

Each standard-setting meeting began with a plenary session on the first morning, attended by all panelists. This session, which was presented jointly by Measured Progress and MDOE staff, provided a general

orientation, including review of the meeting agenda (see Appendix B), background information, and an introduction to the issues of standard setting, and explained the activities that would occur during standard setting. The Plenary Session Power Point Presentation is found in Appendix D. At the conclusion of the plenary session the floor was opened to questions about the standard-setting process.
After the general orientation, the panelists assembled into their content/grade panels. Each panel met in a separate room with a trained room facilitator from Measured Progress. The remainder of the standard-setting tasks were done in the content/grade panel groups.

Orientation to Assessment

Once the panelists assembled into their content/grade panels, they took the same test the students took; this gave panelists the opportunity to become familiar with the assessment and with what students needed to do to score well. Panelists were asked to try to take on the perspective of the students as they took the test. Once panelists had completed the test, the test's answer key and scoring rubrics were distributed and panelists were allowed to self-score their tests. The panelists then discussed any issues or questions that arose regarding the test items or scoring rubrics.

Completing the Item Map for Multiple-Choice Items

Prior to starting the item-mapping activity, the room facilitator led a review of the multiple-choice item summary and the item map documents with the panelists.

In each student folder was a multiple-choice (MC) item summary that listed the test's multiple-choice items in order from the easiest to the most difficult, based on each item's *p*-value, or percentage of students who got the item correct. The summary showed the following information for each MC item: 1) the item's rank order, where item #1 was the easiest; 2) the item's position in the test booklet; 3) a brief summary of the item's text; 4) the student's answer to the item (either a plus sign, if the student got the item right, or the option he/she chose); and 5) the item's *p*-value. The summary also organized the items in terms of the passage (for Reading/ELA) or the content strand (for Mathematics) with which the item was associated.

The first three of the four columns of the item map were the same as those of the MC item summary, indicating each item's rank order and position in the test booklet, along with a summary of the item's text. The final column of the item map was left blank for panelist's notes.

Each panel reviewed the multiple-choice item summary for its grade/test, item by item, discussing the knowledge, skills, and abilities students needed to complete each item, and referencing the scoring rubrics and Performance Level Descriptors. Panelists also discussed why each item was more difficult than the previous item. Panelists wrote the knowledge, skills, and abilities the item measured onto their item maps. They were also advised to include any other information on the item map that might help them as they rated items.

Discussion of the MCAS Performance Level Descriptors

Next, the panelists reviewed the Performance Level Descriptors to ensure that they thoroughly understood the knowledge, skills, and abilities that students needed to demonstrate in order to be classified as *Needs*

Improvement, *Proficient*, and *Advanced* (or *Above Proficient* for grade 3). The panelists began by individually reviewing the general and content- and grade-specific descriptors. They then participated in a group discussion lead by the facilitator to further clarify what knowledge, skills, and abilities were specified by each Performance Level Descriptor. For Grade 3 Reading, the panelists reviewed all of the performance levels, but focused their discussion on *Proficient* and *Above Proficient*, since that was the cut point they were establishing. Bulleted lists of characteristics for each level were generated, based on the whole group discussion, and were posted in the room for panelists to refer to throughout the standard-setting process.

Training Round

Before beginning the individual rating process, each panel completed a training round that consisted of classifying a set of five training folders into the four performance levels. The purpose of the training round was to ensure that all panelists had a complete understanding of the rating task before they began their actual review of the full set of student folders.

To begin the training round, the facilitator briefly reviewed the Performance Level Descriptors and the set of five training folders, which were selected to represent performance across the range of possible raw scores. The facilitator first reviewed each open-response item, and then reviewed the multiple-choice item summaries. The facilitator emphasized that multiple-choice items should be considered carefully by the panelists in making their ratings since the majority of points on each test come from multiple-choice items.

The panelists then individually reviewed all five of the training sets—which were presented in random order—and placed them in order from lowest to highest. Once this was completed, the facilitator tallied the extent to which the panelists agreed about the order of the folders. The facilitator then led a group discussion of the characteristics of the folders, starting with the lowest scoring and pointing out why it was classified in the *Warning* performance level. For the discussion of each folder in turn, the facilitator pointed out the connections between the knowledge, skills, and abilities demonstrated and the Performance Level Descriptors for its performance level.

Individual Ratings

In the first step of the actual rating process, each panelist made an initial judgment of how each student folder should be categorized. Panelists used the Performance Level Descriptors, their completed item maps, and the student test booklet to rate the student folders for their grade and content area. Fifty student folders were assembled and presented to panelists for all tests except Grade 3 Reading. Because the panelists for Grade 3 Reading were setting only one cut point, only 30 folders were needed. Starting with the first folder (corresponding to the lowest overall raw score), the panelists individually considered the knowledge, skills, and abilities demonstrated by the student. Panelists for grade 3 then decided into which performance level each folder should be placed, while panelists for the higher grades considered whether the folder was accurately classified according to the given starting cut point. Panelists used this same process to rate all folders, and recorded their initial classification for each folder in the "Individual Rating" section of the Round 1 rating form. A sample of the Round 1 rating form is provided in Appendix I.

For Grade 3 Reading, three of the 30 student folders were pre-categorized into the *Warning* performance level and three were categorized as *Needs Improvement*; panelists did not have the option of changing these six folders' categorizations. Panelists were asked to categorize the remaining 24 folders into either the *Proficient* or *Above Proficient* performance level.

Revised Ratings after Group Discussion (Round 1)

Once the panelists completed their individual reviews and initial classifications of all of the student folders, they discussed each folder as a group, starting with the first folder. Panelists discussed the knowledge, skills, and abilities demonstrated in each folder and how they corresponded to the Performance Level Descriptors. The facilitator focused the discussion on any student folders for which there was disagreement among the panelists about categorization, making it clear that while panelists did not need to come to consensus about how to categorize the folders, they should express their own opinions while listening to the opinions of the other panelists. As the panelists completed their group discussion of each folder, each panelist entered a rating for that folder in the "Revised Rating After Discussion" section of the Round 1 rating form. Facilitators emphasized that each panelist, in each round of rating, should indicate her or his *individual* judgment on the rating form.

Tabulation of Round 1 Results

After the panelists had completed the group discussions and recorded their Round 1 ratings, the rating forms were returned to the Research and Analysis staff and the results were analyzed. Prior to beginning Round 2, panelists were given feedback on the Round 1 ratings. The information consisted of the group average cut scores based on the Round 1 ratings, which were determined using logistic regression. Specifically, for a given cut, each panelist's rating for each student folder was dichotomized (i.e., given a "score" of zero or one, where zero indicates that the panelist rated the folder as being below the cut). A logistic function was fit to the data for that cut, and the point of inflection on this curve was used to establish each panelist's cut point on the raw score scale. The cuts were then averaged across the panelists to come up with the overall group average cut score.

Final Ratings (Round 2)

During Round 2, the panelists in each panel examined the results from Round 1 and discussed their ratings. Focusing on student folders near the cut points, panelists discussed any folders for which there was either disagreement about classification, or (for grades 4–8) for which the Round 1 classification differed from the initial classification based on the given starting cut points. The panelists were encouraged to share their classification rationales, in terms of the knowledge, skills, and abilities students must be able to demonstrate. Again, panelists were told that they did not need to come to consensus, but that they should both participate in the discussion and listen to other panelists' points of view.

After all discussions had been completed, panelists recorded their final ratings on the Round 2 rating forms. For Grade 3 Reading, the first six student folders were pre-entered; panelists were only able to enter ratings of *Proficient* or *Above Proficient* for the remaining 24 folders. For the higher grades and Grade 3 Mathematics, panelists could classify any folder into any of the four performance levels. Appendix I includes both grade 3 rating forms, as well as the Grade 5 ELA rating form; rating forms for grades 5–8 follow the same format as the Grade 5 ELA rating form.

Evaluation

Upon completion of the rating process, panelists anonymously completed an evaluation form. The evaluation forms and a tabulation of the results for each panel are included in Appendix J. Analysis of the evaluations may be found in the following section.

Tasks Completed After the Meeting

Upon the conclusion of the standard-setting meeting, the meeting was reviewed and analyzed, and anomalies that may have occurred in the process or in the results were addressed.

Analysis and Review of Panelists' Feedback

After completing the standard-setting activities, panelists' evaluation feedback was reviewed. The review of the evaluation forms from the 2006 standard-setting meetings did not reveal any anomalies in the standard-setting process or indicate any reason that a particular panelist's data should not be incorporated in the final results. It appeared that all panelists understood the rating task and attended to it appropriately. Tables 3 and 4 below provide demographic information for the groups of panelists who completed the evaluation for each content area.

Table 3
Demographic Information on Panelists Completing Evaluations – Reading/ELA

	Grade 3	Grade 5	Grade 6	Grade 8
Panelist is a:				
Classroom teacher	9	10	14	11
K-12 administrator	2	0	0	4
University-level educator	1	0	0	0
Business/community representative	1	0	1	0
Other	3	3	4	2
Experience teaching students:				
with disabilities	9	9	14	14
with limited English proficiency	5	0	1	1
Neither of these	2	4	3	2
Total*	16	13	19	17

*Note: not all panelists completed these items, so numbers may not sum to the total

Table 4
Demographic Information on Panelists Completing Evaluations – Mathematics

	Grade 3	Grade 5	Grade 7
Panelist is a:			
Classroom teacher	8	10	10
K-12 administrator	3	2	1
University-level educator	0	1	2
Business/community representative	0	0	1
Other	4	3	2
Gender:			
Male	4	2	4
Female	12	14	12
Total*	16	16	16

*Note: not all panelists completed these items, so numbers may not sum to the total

Question 4 on both evaluation forms asked the panelists whether they relied primarily on the open-response or the multiple-choice items, or on both equally, in determining their ratings. The majority of the panelists indicated that they relied on both equally; the percentage ranged from 76% for Grade 8 ELA to 100% for Grade 3 Reading and Grades 3 and 5 Math. Overall, 89% relied on both equally for ELA/Reading and 96% for Math. Of the remaining panelists, all but one (in the Grade 8 ELA panel) indicated that they relied primarily on the open-response items.

Questions 5 through 12 asked panelists their opinions about the organization and process of the standard setting, as well as the final results. The response options were "Strongly Disagree," "Disagree," "Agree," and "Strongly Agree." Table 5 below shows the percent of panelists who indicated they either agreed or strongly agreed with each statement by content area.

Table 5
Summary of Responses to Evaluation by Content Area – Questions 5 through 12

	Percent Agree or Strongly Agree	
	Reading/ELA	Mathematics
5. Overall environment and accommodations were comfortable and appropriate.	97%	98%
6. Background information provided improved my ability to set standards.	95%	98%
7. Taking and discussing the exam helped me understand the purpose and process	100%	98%
8. By the end of the calibration training, I could distinguish among Performance Level Descriptors.	92%	92%
9. Overall, I was provided with clear instructions.	98%	91%
10. The group discussions after the first round improved my ability to set standards.	97%	98%
11. I am confident that the ratings I provided were consistent with the Performance Level Descriptors.	100%	98%
12. The standard-setting process provided for a reliable classification of student work.	100%	100%

As seen in Table 5, a large majority of panelists said they either agreed or strongly agreed with each statement. Looking at the results by grade level (see Appendix J), in all but three cases, the percentage who agreed or strongly agreed ranged from 89% to 100%. For question 8, 72% of the panelists in the Grade 6 ELA group and 75% of panelists in the Grade 3 Math group indicated that they agreed or strongly agreed that they could distinguish among the Performance Level Descriptors by the end of the training calibration. One panelist in each group indicated that he/she strongly disagreed with the statement, while the rest disagreed.

For question 9, 73% of the panelists in the Grade 7 Mathematics group agreed or strongly agreed that overall, they were provided with clear instructions. Four panelists disagreed with the statement, and none strongly disagreed.

Overall, there were only three instances in which a panelist indicated he/she strongly disagreed with a statement: the two for question 8 described above, and one for question 6 in the Grade 6 ELA group.

Questions 13 through 20 asked the panelists to indicate whether they felt the amount of time allotted to each of the various standard setting steps was appropriate. The response options were "Far too short," "Too short," "Approximately right," "Too long," and "Far too long." Table 6 below indicates the percentage of panelists who felt the time allowed for each step was "Approximately right," by content area.

Appendix G

Table 6 Summary of Responses to Evaluation by Content Area – Questions 13 through 20		
	Percent Responding "Approximately Right"	
Panelists' perception as to the time allotted for:	Reading/ELA	Mathematics
13. Initial background information	78%	77%
14. Taking and discussing the exam	81%	91%
15. Learning about and discussing Performance Level Descriptors	73%	77%
16. Ranking, discussing, and classifying students' work (calibration)	81%	68%
17. Initial individual classification of student work	81%	79%
18. Group discussion regarding initial ratings	81%	61%
19. Rating student work for the second time	87%	74%
20. Final rating of student work	86%	63%

Overall, as can be seen in Table 6, the percentage of panelists indicating that the time allowed was approximately right ranged from 61% to 91%. Among responses other than approximately right, responses of either too long or far too long outnumbered responses of too short or far too short by about two to one overall. However, there was quite a bit of variability across grades (see Appendix J).

For ELA/Reading, there were only 3 instances in which fewer than two-thirds of the panelists thought the amount of time was approximately right: question 13 for Grade 5 and questions 15 and 18 for Grade 6. For Grade 5, question 13, 58% said the time was approximately right, and the remaining 42% said it was too long. For Grade 6, question 15, 58% responded approximately right, 16% responded too short, and 26% responded too long or far too long. For Grade 6, question 18, 61% responded approximately right, and the remaining 39% responded too long or far too long.

For Mathematics, there were 6 instances in which fewer than two-thirds of the panelists answered approximately right: questions 16 and 20 for Grade 3, questions, 15, 17, and 18 for Grade 5, and question 20 for Grade 7. Of these, in only one case was the percentage who thought too little time was allowed greater than the percentage who thought too much time was allowed: Grade 5, question 17 (62.5% approximately right, 25% too short, and 12.5% too long). In the remaining instances, the percentages for too short vs. too long were either equal (Grade 3, question 16), or more panelists responded that the time was too long or far too long (Grade 3, question 20; Grade 5, question 15; Grade 5, question 18; Grade 7, question 20).

In general, responses of too short or far too short are of greater concern than responses of too long or far too long since having too little time is presumably more likely to have a negative effect on the panelists' ability to perform the task. As mentioned above, there were approximately half as many responses indicating not enough time than responses indicating too much time. There were only two responses of far too short: for Grade 3 Math, one panelist indicated far too short for both questions 18 and 20. In addition, in general, responses indicating too little time tended to be scattered around rather than concentrated for particular panels or particular questions. The main exception to this statement was for the Grade 8 ELA panel: for each of questions 14 through 18, there were between 1 and 5 responses of too short, and no responses of too long or far too long. Of greatest concern is question 15, for which 5 of the panelists (29%) indicated that the time allowed for learning about and discussing the Performance Level Descriptors was too short.

Summary

Overall, panelists' responses on questions 5 through 12 were generally positive. Some panelists indicated that they disagreed or strongly disagreed that they could distinguish among the Performance Level Descriptors by the end of the calibration training, and that overall they were provided with clear instructions. However, the responses to questions 11 and 12 were overwhelmingly positive, indicating that the panelists felt confident in their final classifications.

For questions 13 through 20, there were some indications that some panelists felt rushed at certain parts of the process. However, in most cases, those responses did not appear to be systematic. The one exception was Grade 8 ELA, as described above. In general, the responses suggest that some members of the group grasped the tasks more quickly while other members took a little longer. This is a very common occurrence, especially when some members of the group have participated in standard setting in the past. It is important to find a reasonable balance between moving too quickly vs. allowing the process to stall. The responses to questions 13 through 20, for the most part, provide evidence that this balance was successfully achieved.

Preparation of Recommended Cut Scores

After each round of ratings, the raw score cut points were calculated based on the average across all panelists' cuts, where the cuts for each panelist were calculated as described in the Tabulation of Round 1 Results section above. In addition, the percentage of students who would be classified into each performance level was determined. The results for all content areas and grades are presented in tables 7 and 8 below, along with the starting cut points and associated percentages. As mentioned previously, one of the goals of the standard setting was to establish coherent cut points across grades 4 through 8. The Round 2 results for these grades were found to meet this goal sufficiently, so all Round 2 results for grades 5 through 8 were approved by MDOE for use in scaling and reporting. This cross-grade coherence is illustrated in figures 1 and 2, which show the final approved cut points for grades 4 through 8.

The Round 2 results for Grade 3 Reading were also approved by MDOE and used to compute the scaled scores used for reporting. For Grade 3 Mathematics, however, some adjustments to the Round 2 results were determined to be necessary. The cut score determined by the standard setting panel was not consistent with the vertically moderated MCAS system for Grade 3 Math. Upon further consideration by the Department it was decided that Grade 3 Math scores should be vertically linked with other MCAS scores to avoid confusion in the interpretation of results by schools and districts. That is, it was determined that the percent of students categorized into the various proficiency levels should not be dramatically different compared to the other grade levels assessed in MCAS.

Therefore, it was necessary to adjust the cutscores for Grade 3 Math such that a higher raw score was needed to be classified into either *Needs Improvement*, *Proficient*, or *Above Proficient*. For the lower two cutscores the adjustment was based on the SEM of the panelists' recommendation. While the highest cutscore was limited by ceiling effect, this cutscore was moved up between the raw scores of 39 and 40. Ongoing test development efforts will provide more substantial performance information around the highest cutscore. The final approved cut points and associated percentages for Grade 3 Mathematics are presented in the final columns of table 8.

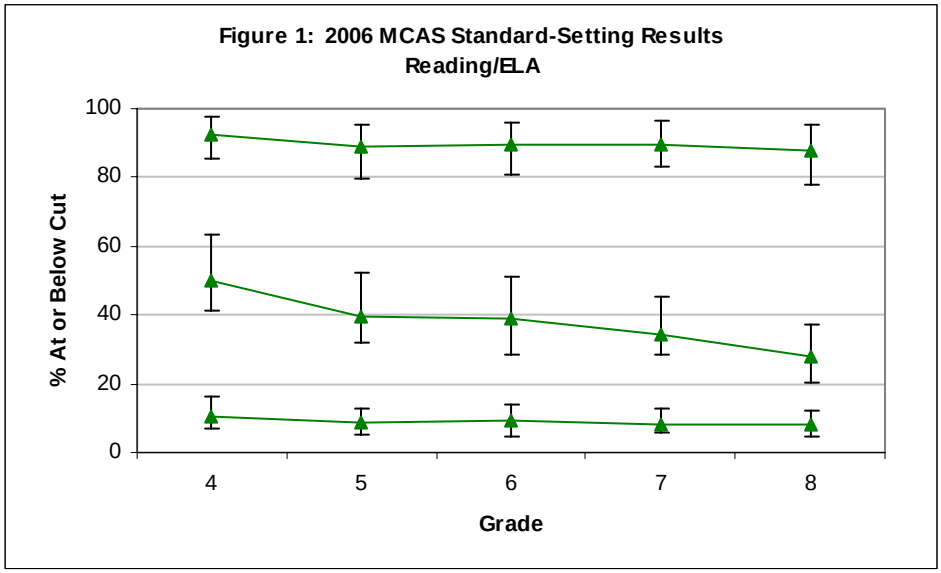
Appendix G

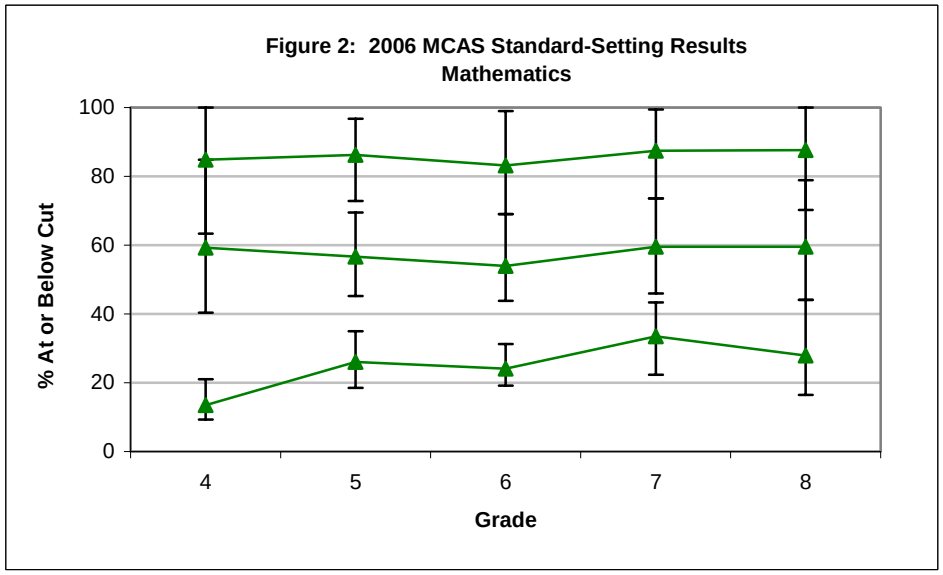
Table 7
2006 MCAS Standard-Setting Results: Reading/ELA

Grade	Performance Level	Starting Cuts		Panelist Cuts: Round 1		Panelist Cuts: Round 2		Final Adopted Cuts	
		Raw Score Range	% in Level	Raw Score Range	% in Level	Raw Score Range	% in Level	Raw Score Range	% in Level
3	*Warning*			0-20	7.1	0-20	7.1	Same as Round 2	
	Needs Improvement			21-36	33.9	21-36	33.9		
	Proficient			37-44	47.2	37-44	47.2		
	Above Proficient			45-48	11.8	45-48	11.8		
5	*Warning*	0-22	8.6	0-22	8.6	0-22	8.6	Same as Round 2	
	Needs Improvement	23-36	34.8	23-35	30.6	23-35	30.6		
	Proficient	37-45	45.5	36-44	45.1	36-45	49.6		
	Advanced	46-52	11.1	45-52	15.7	46-52	11.1		
6	*Warning*	0-21	8.1	0-22	9.3	0-22	9.3	Same as Round 2	
	Needs Improvement	22-34	30.7	23-33	25.8	23-34	29.5		
	Proficient	35-45	50.9	34-45	54.5	35-45	50.9		
	Advanced	46-52	10.3	46-52	10.3	46-52	10.3		
8	*Warning*	0-21	7.3	0-22	8.3	0-22	8.3	Same as Round 2	
	Needs Improvement	22-33	20.5	23-33	19.5	23-33	19.5		
	Proficient	34-46	60.1	34-46	60.1	34-46	60.1		
	Advanced	47-52	12.1	47-52	12.1	47-52	12.1		

Table 8
2006 MCAS Standard-Setting Results: Mathematics

Grade	Performance Level	Starting Cuts		Panelist Cuts: Round 1		Panelist Cuts: Round 2		Final Adopted Cuts	
		Raw Score Range	% in Level	Raw Score Range	% in Level	Raw Score Range	% in Level	Raw Score Range	% in Level
3	*Warning*			0-17	6.3	0-16	5.3	0-23	15.1
	Needs Improvement			18-25	13.3	17-25	14.3	24-32	32.3
	Proficient			26-38	69.8	26-37	62.4	33-39	48.4
	Above Proficient			39-40	10.5	38-40	18.0	40	4.2
5	*Warning*	0-26	26.1	0-26	26.1	0-26	26.1	Same as Round 2	
	Needs Improvement	27-41	40.0	27-39	33.7	27-38	30.5		
	Proficient	42-47	20.1	40-47	26.5	39-47	29.6		
	Advanced	48-54	13.8	48-54	13.8	48-54	13.8		
7	*Warning*	0-29	38.2	0-26	31.1	0-27	33.5	Same as Round 2	
	Needs Improvement	30-40	29.6	27-36	25.6	28-37	26.0		
	Proficient	41-48	22.3	37-47	30.7	38-47	28.0		
	Advanced	49-54	9.9	48-54	12.5	48-54	12.5		





Summary

This report summarizes the rationale, methods and results of the standard setting process that was used in the summer of 2006 to establish performance level cut scores for the new MCAS assessments in Grades 5, 6, and 8 ELA, and Grades 3, 5 and 7 Mathematics. In addition, a single cut score was determined for Grade 3 Reading to define a fourth performance level, *Above Proficient*. Implementation of the new assessments in 2006 completed the formation of a comprehensive assessment program across all grades, 3 through 8.

Throughout the standard setting process a primary goal was to establish a set of coherent performance standards across the entire assessment program. To this end, initial cut scores were calculated for the new assessments in grades 5 through 8 by interpolating from the cut points established in previous years on the already-existing MCAS assessments. Panelists participating

in the 2006 standard setting meetings were then asked to validate or recommend modifications to those starting cuts. Because Massachusetts had developed vertically-articulated achievement standards across grades, this process for establishing cut points results in a system that can be used to evaluate student growth. The process used for the 2006 standard setting, which built upon the strong foundation of standards set in previous years, ensures that Massachusetts' comprehensive assessment system can be used to fulfill NCLB accountability requirements.

References

Ferrara, Steve, Eugene Johnson, and Wen-Hung (Lee) Chen. 2004. Vertically moderated standards: logic, procedures, and likely classification accuracy of judgmentally articulated performance standards. Paper presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA.

Kane, Michael T. 2001. So much remains the same: conception and status of validation in setting standards. In G.J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives,* pp. 249-281. Mahwah, NJ: Lawrence Erlbaum.

Kingston, Neal M., Stuart R. Kahl, Kevin P. Sweeney, and Luz Bay. 2001. Setting performance standards using the body of work method. In G.J. Cizek (Ed.), *Setting performance standards: concepts, methods, and perspectives*, pp. 219-248. Mahwah, NJ: Lawrence Erlbaum.

Lissitz, Robert W., and Huynh Huynh. 2003. Vertical equating for state assessments: issues and solutions in determination of adequate yearly progress and school accountability. *Practical Assessment, Research & Evaluation,* 8(10). http://PAREonline.net/getvn.asp?v=8&n=10.

Mitzel, Howard C. 2005. Consistency for state achievement standards under NCLB. The Council of Chief State School Officers, http://www.ccsso.org/content/pdfs/CAS%20Series%202.Final.pdf.

Appendix A

Recommendation for Standard-Setting Procedures & Schedule of Tasks

MCAS STANDARD SETTING
Recommendation for
Standard-Setting Procedures &
Schedule of Tasks

Overview
History of MCAS Standard Setting

The MCAS tests have been administered to students in Massachusetts since 1998. At that time, Math and ELA were the subjects administered. In subsequent years, additional grades and content areas have been added. After the initial administration of each of these new tests, performance standards have been set. For grade 3 through 8, Table 1 displays the history of when standards were set for different content areas, the standard setting method used, and the responsible contractor.

Table 1

Grade	Content	Year	Method	Contractor
3	Reading	2001	Bookmark	HEM, BETA
4	Math	1998	Body of Work	ASME
4	ELA	1998/2001	Body of Work	ASME/HEM, BETA
5	Science	2003	Body of Work	HEM,BETA
6	Math	2001	Body of Work	HEM, BETA
7	ELA	2001	Body of Work	HEM, BETA
8	Math	1998	Body of Work	ASME
8	Science	2003	Body of Work	HEM, BETA
8	ELA	1998	Body of Work	ASME

Current Context

As part of NCLB requirements, in 2006, several new grades and content area tests were added to MCAS. As was the case previously, performance standards now need to be established for these newly added tests. Additionally, for grade 3 reading, when initial standards were set in 2001, only three performance levels were established. To be consistent with the other grades and in compliance with NCLB, a fourth performance level needs to be added to grade 3 reading. This new performance level in grade 3 will be the highest level students can achieve and will be called *Above Proficient*; in the other grades the highest level is called *Advanced*. Table 2 displays the grades and content areas for which standards need to be set in 2006, as well as the proposed standard setting method.

Table 2

Grade	Content	Method	Notes
3	Reading	Body of Work	Establish *Above Proficient*
3	Math	Body of Work	
5	Language & Literature	Body of Work	
5	Math	Body of Work	
6	Language & Literature	Body of Work	
7	Math	Body of Work	
8	Language & Literature	Body of Work	

The new grades and content areas assessed in 2006 'fill in the gaps' of grades and contents so that all grades, 3 through 8, are tested in the areas of Mathematics and ELA (or subset of the ELA content standards: Reading in grade 3 and Language & Literature in grades 5, 6, & 8).

Appendix G

Vertically Moderated Standards

Because students will be tested in all adjacent grades 3-8, and because there is consistency in the performance level descriptors, it is important that the resulting distribution of students across the performance levels are coherent and make sense across grades (Kane, 2001). For example, it would make little sense (and likely would be indefensible) to have 10%of the students *Proficient* or *Advanced* in one grade, 60% *Proficient* or *Advanced* in the next, and 30% *Proficient* or *Advanced* in the one after that. Within the current environment of testing at adjacent grades and having consequences (at least at the school level) associated with student performance on those tests, several authors (e.g., Mitzel, 2005, Ferrara, Johnson, & Chen, 2004) have argued that cross-grade consistency of results should be an explicit part of the standard setting enterprise.

Mitzel (2005) describes three scenarios under which standard setting activities occur within state departments of education. In the first of his scenarios, performance standards on existing tests are to remain intact, and performance standards are to be established for new tests added to the assessment system. This scenario describes the present MCAS state of affairs. To set standards in this situation, Mitzel and others (e.g., Lissitz and Huyunh, 2003) describe a process for achieving consistency of results by starting with the percentages of students in performance levels for existing grades then fitting a line or a growth function to 'smooth' those percentages across grades. This fitted line or function would then be used to determine the percentage of students in each performance level at the new grades. The observed scores that produced those percentages would be the cut scores on the new tests. The outcome of this smoothing process is referred to as 'vertically moderated standards'.

This is similar to the process we are recommending for Massachusetts, with two major differences. First, we are recommending excluding grade 3 from the vertical moderation process (the reasons for this are discussed below), and second we are recommending using the 'vertically moderated standards' as initial cut scores that will be validated by a panel in an abbreviated standard setting meeting (more correctly termed a standards validation meeting).

Grade 3

The premise on which vertically moderated standards are built is that there is a consistency in the underlying construct measured at each grade. This consistency is reflected in the performance level descriptors across grades as well as the content specifications for adjacent grade tests. Consequently, there is an expectation of equivalence of the meaning of the performance levels across grades.

Discussions with MDOE staff and inspections of the performance level definitions indicated that this underlying relationship does not exist for grade 3 as it does for the other grades, making grade 3 inappropriate for inclusion in the vertical moderation process. The content area experts assert that at grade 3, and below, the constructs of reading and math are fundamentally different than above grade 3. These differences stem from the developmental stages of reading and math acquisition through which students progress up through grade 3. These stages are typically completed by grade 4.

The recognition that grade 3 was different than the later grades was evidenced in the initial setting of standards of grade 3 reading where only three performance levels were initially defined. As stated earlier in this document, for grade 3 reading a fourth performance level needs to be added. It is worth noting that an additional change to grade 3 reading this year is the inclusion of open response (OR) scores in the total score. Previously, the OR scores were reported back to students, but did not contribute to the total score. Consequently, they played no role in categorizing students into performance levels.

An additional consequence of the differences associated with grade 3 is that grade 3 math will not go through a standards validation process, but will go through a standard setting process where initial cut scores will not be identified.

The addition of the 'Above Proficient' performance level for grade 3 reading and the setting of performance standards for grade 3 math will be accomplished through the implementation of the Body of Work standard setting procedure.

The remainder of this document details the recommended procedures for accomplishing the standard setting or validation tasks for 2006.

Recommendation for Standards Validation/Setting

Standards Validation Procedures – Math 5, 7 and L&L 5, 6, 8

Existing MCAS cut-points were first set in 1998 and continue to accurately reflect grade and content performance levels. The intent of the current process described here is to establish initial cut points and then convene panels of educators to validate these.

Because: 1) the existing MCAS standards are an integrated component of the schools vernacular; 2) the MCAS standards are stable and an integral part of current curriculum and 3) the MDOE finds the cut-points for the existing grades and contents reasonable, we recommend using the current existing cut-points to establish initial cut-points in Math and Language & Literature. Using these initial cut-points, we then recommend conducting a standards validation procedure using a modification of the Body of Work (BOW) method.

STANDARD VALIDATION MEETINGS

Implementation of the standard validation process for each test will be handled by Measured Progress in coordination with the MDOE. A process facilitator will be in charge of the general implementation of the process including assigning the tasks and following the agenda. Additionally, a content expert for each grade level and content area will be available to respond to panelists' concerns related to content and performance levels. Measured Progress staff members with extensive experience in facilitating standard setting meetings will serve as the process facilitator; the content experts will be provided by the MDOE. Additional Measured Progress staff will be present for the duration of the standard setting process, including the lead psychometrician for MCAS who will address the technical concerns of standard setting panelists, as well as the primary data analyst for MCAS and the MCAS project director.

MEETING LOGISTICS

All standard-setting meetings will take place in Danvers, Massachusetts. Measured Progress will assume responsibility for all tasks and costs associated with the planning and facilitation of each meeting, including:

a. procuring standard-setting meeting sites approved by the Department
b. notifying and pre-registering standard-setting panelists
c. preparing and producing standard-setting materials
d. registering panelists and distributing materials to panelists prior to and during the meetings
e. coordinating with site staff prior to and during the meetings
f. providing dinner for panelists, facilitators, and Department staff on the evening of the first meeting day
g. providing continental breakfast, lunch, and light afternoon refreshments for panelists, facilitators, and Department staff on the second and third days of the meeting
h. reimbursing panelists' travel expenses
i. paying lodging costs for participants for two nights
j. maintaining security of materials prior to, during, and following standard-setting meetings

Establishing Initial Cut-Scores

For the validation process, we propose using information from grades with existing standards to establish the initial cut-points for those grades on which standards need to be set. The initial cut-points will be determined by first calculating a line of best fit to the percent of students exceeding each cut-point based on the grades with existing standards. Then using this linear relationship across grades, determine the percent of students expected in each performance level for the grades on which standards need to be established. The raw score that then most closely produces the desired percentage is taken as the initial cut-score. Data from the 2006 MCAS test results will be

used for this process. Once initial cut-scores have been determined, standards validation will commence beginning with the BOW range finding phase (phase 2).

Performance Level Descriptors

Using existing performance level descriptors as 'anchors', general content performance level descriptors have been developed for the 2006 standard validation activities; these will be used during the standards validation process. The general level descriptors will be further refined to be grade level specific after the standards validation meetings are held.

Panel Membership

Each standard-validation panel will consist of 15 to 20 members representing educators and administrators. The exact composition will be determined jointly with the MDOE. Panelists will be familiar not only with the subject matter but also with the grade level for which they would be setting standards.

Daily Schedule

Each standard-setting meeting will last two days. The first day's activities will begin in the morning with an orientation to MCAS and the MCAS performance standards and standard-validation process, including an overview of the procedure that will be implemented. In the afternoon of the first day, panelists will receive training in the performance level descriptions and will learn to apply their understanding of those descriptions to specific bodies of student work. Panelists will then review all bodies of work, noting into which performance level each is initially placed. Panelists will individually either agree with, or re-assign, each BOW. Panelists will end the first day with their individual ratings. In the beginning of the second day, panelists will be given new classifications of students based on the entire rooms' ratings. They will discuss the appropriateness of the initial classifications and their individual ratings. On the basis of those discussions panelists may revise their categorizations of students. This will take until the end of the second day. Before leaving, panelists will complete an evaluation form.

Body of Work Method

For Massachusetts we are recommending a modification to the BOW methodology that has been used previously in Massachusetts. Previously, judges looked at collections of student work first in a range-finding round and then in a pin-pointing round. The range-finding round required several examples of student work at specific, fairly widely spaced points along the performance continuum; the pin-pointing round then honed-in on a narrow area of the performance continuum by providing student work within a restricted range, based on the results of range-finding. There have proven to be three potentially problematic unintended effects of this procedure. First, if one (or more panelists) is particularly high or low relative to others and they believe the cut score is out of the range of student work presented in the pin-pointing round (i.e., all student work is classified in one performance level) a cut score cannot be calculated for that panelist. Second, because it is unknown beforehand where the cut-scores will emerge from the range-finding round, materials need to be prepared for every possible contingency in the pin-pointing round, which means that a great deal of materials are prepared and never used. This increases the logistical complexity of the BOW procedure, increases the necessary timelines for materials preparation and increases the possibility of mistakes in materials production. Third, because the student work used in the range-finding and pin-pointing rounds are different, panelists have limited opportunity to discuss and revise their classifications of student work.

The recommended modification to the BOW process is to expand the range-finding round and eliminate the pin-pointing round. The expansion of the range-finding round calls for an increase in the number of papers included in the range-finding round and that these papers be equally spaced in small increments throughout the performance continuum. So for example, suppose we had a 50 point test. Previously, a range-finding set of papers may have had 3 examples of student work at every 5th score point for a total of 30 sets of student work. The pin-pointing sets of student work to be prepared in advance may have been 2 sets of papers at every score point, for a total of 100 examples of student work. Of these 100, one third to one half would likely be used in the pin-pointing round (the others would be outside the possible cut-score area determined in range-finding). Thus, a total of 130 sets of student work would need to be prepared for every panelist. Assuming 15 panelists, this would be 1950 sets of

student work for each grade content area. Of the 130 sets of student work, approximately 80 to 97 would be reviewed by the panelists.

In contrast, the current recommendation would provide 60 sets of student work, all of which get reviewed and discussed by the panelists. Analyses conducted previously in a study of one states' standard setting data using only the range-finding ratings to calculate the cut-scores, indicated little difference (typically less than 1 point) between the cut scores based only on range-finding and the final cut-scores. Additionally, the current recommendation provides panelists ample opportunity to review and discuss their ratings with other panelists and be able to adjust their ratings on the basis of that review and discussion.

Because the thresholds are based on the panelists' classification of student responses, the selection of those responses is a crucial part of the preparation. Additionally, materials to be prepared for each standard-setting meeting include rating and tally forms as well as sets of student responses that will be classified by the panelists.

The student responses to be classified by panelists will be selected in advance based on the following criteria:

- Relationship between OR scores and MC scores. Students should not be selected that have low OR scores and high MC scores, or vice versa.
- Consistency of OR scores. Papers that have identical scores by two raters are preferred over papers that have do not have identical scores (which may indicate some difficulty in scoring those papers).

The body of work of each student includes responses to open-response items and a "multiple-choice display" providing the following information for each multiple-choice item on the test:

- the stem of the item and the correct response;
- the difficulty of the item; and
- an indicator whether the student's response was correct or incorrect.

Templates used for entering panelists data will be prepared prior to the meeting. When data have been entered and proofread, they will be analyzed on-site to obtain the initial set of cut-scores as determined by the panelists. After discussion panelists final cut-scores will be calculated. Analysis programs will be written using SAS statistical software and be tested prior to the meeting.

Each two-day standard-validation meeting begins with a large group orientation on the first day. The orientation provides panelists with an understanding of the purpose of the meeting and the procedures to be followed in validating the performance levels.

After completion of orientation, panelist will disperse into their specific grade/content groups. They will then take a form of the test for which they will be validating standards; they will review their responses relative to the scoring guides.

Panelists will then be introduced to, and become familiar with, the performance level definitions.

Calibration Exercise

Panelists will be presented with six BOWs with previously assigned performance levels. They will review these and discuss the characteristics of this work as it relates to each performance level descriptor. The facilitator will point out characteristics of each BOW that were instrumental in its classification.

Rating Student Work

As described earlier, panelists will review the bodies of student work and either agree with or re-assign students to different performance levels.

Analysis

Appendix G

Analysis of standard-setting data involves analysis of rating data and of evaluation data. Rating data will be analyzed by computing threshold scores using logistic regression. SAS statistical software programs will be written to compute overall and individual cutscores. Overall threshold scores and individual threshold scores are computed. The average and standard deviation of the individual threshold scores will be computed. These computations are performed for each round of ratings. Results of these computations show the changes in the panelists' overall cutpoints from round to round. The results also show the change in inter-rater consistency through the different classification rounds. These results are presented in table and graph form to show the convergence of the group during the process.

Analysis of evaluation data results in the summary of the panelists' responses to evaluation questions. For each question, the frequency of each possible response will be computed. The average response of panelists is also computed for each question as appropriate.

Facilitator Training

Approximately two weeks prior to the standard-validation meeting, Kevin Sweeney and other Measured Progress staff will conduct a one-day facilitator-training meeting. The purpose of the meeting is to have all Measured Progress and Department personnel and identified facilitators involved in the standard-setting meeting walk through the standard-setting procedures that will be implemented. The participants in the meeting will also review all the materials that will be used for standards validation. All details of standards validation are finalized in this meeting.

Security of Materials

Throughout the standard-setting process, security of the materials will be of the utmost concern. Participants will be required to sign a confidentiality and non-disclosure statement before they are allowed to see any MCAS materials. All test-related materials will be collected and accounted for at the end of each day, and Measured Progress will have designated staff responsible for material security during meals and break times throughout the standard-setting meeting.

Participant Evaluation

To provide evidence of the participants' views of the standard setting process, judges will be given an opportunity to participate in an anonymous evaluation of the standard setting process. This evaluation will consist of judges completing a questionnaire after they have completed all other standard-setting activities. Results will be evaluated and provided to the Department.

Standards Setting Procedures – Grade 3 Math and Reading

The Procedures for setting standards for grade 3 in reading and math are essentially the same as for the standards validation, except that initial classifications of student work will not be provided to the panelists. For grade 3 reading, only those BOWs that score high enough to be included in the *Proficient* category will be included. Panelists will need to establish only one cut score here. Consequently, it is likely that grade 3 reading will complete the standard setting tasks well before the other groups and perhaps in a single day.

References

Ferrara, Steve, Eugene Johnson & Wen-Hung (Lee) Chen (2004). Vertically moderated standards: logic, procedures, and likely classification accuracy of judgmentally articulated performance standards. Paper presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA.

Kane, Michael T. (2001). So much remains the same: Conception and status of validation in setting standards. In Cizek, Gregory J., (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 249-281). Mahwah, NJ: Lawrence Erlbaum.

Lissitz, Robert W. & Huynh Huynh (2003). Vertical equating for state assessments: issues and in determination of adequate yearly progress and school accountability. *Practical Assessment, Research & Evaluation,* 8(10). Retrieved from http://PAREonline.net/getvn.asp?v=8&n=10.

Mitzel, Howard C. (2005). Consistency for state achievement standards under NCLB. The Council of Chief State Schools Officers. Retrieved from http://www.ccsso.org/content/pdfs/CAS%20Series%202.Final.pdf.

Appendix G

Schedule of Tasks - ELA

TIMEFRAME	TASKS	DATES
Prior to Meeting	1. Make meeting arrangements (i.e., hotel, travel, meeting room logistics)	June 7
	2. Conduct Performance Level definition meeting	June 30
	3. Prepare materials for Standard Setting practice day for group leaders and Department staff	June 5 – June 12
	4. Prepare Standard Setting folders/materials for panelists a. Meeting agenda b. Confidentiality agreement c. Performance level descriptions d. Set of items on which standards will be set e. Rating forms f. Student profiles g. Evaluation forms	June 1 – July 10
	5. Prepare Standard Setting presentation materials (i.e., PowerPoint)	
	6. Prepare systems, written procedures, and Standard Setting materials for analysis during the meeting	May 15 – July 10
	7. Select panelists – Standard Setting (DOE)	June 21
	8. Conduct practice day for group leaders and Department staff for Standard Setting	June 12
	9. Establish initial cut points and review with Department.	June 14
During the Standard Setting Meeting	1. Orientation	July 11
	2. Reviewing assessment materials	July 11
	3. Discussion of performance level descriptions	July 11
	4. Round 2: Table group discussion of panelist judgments and opportunity for revised judgments	July 11
	5. Tabulation of Round 2 results	July 12
	6. Round 3: Whole group discussion of panelist results and impact data and final opportunity to revise judgments	July 12
	7. Evaluation	July 12
After the Standard Setting Meeting	1. Analyze and review panelists' feedback	July 17 – 21
	2. Review panel's recommended cut scores for cross-grade-span consistency	July 17 – 21
	3. Prepare recommended cut scores with DOE	July 26
	4. Preparation of standard-setting report	October 2006

Schedule of Tasks - Math

TIMEFRAME	TASKS	DATES
Prior to Meeting	1. Make meeting arrangements (i.e., hotel, travel, meeting room logistics)	Aug 7
	2. Conduct Performance Level definition meeting	June 30
	3.Prepare materials for Standard Setting practice day for group leaders and Department staff	Aug 7 - 11
	4. Prepare Standard Setting folders/materials for panelists a. Meeting agenda b. Confidentiality agreement c. Performance level descriptions d. Benchmarks and outcomes e. Set of items on which standards will be set f. Rating forms g. Student profiles h. Evaluation forms	July 17 – August 18
	5. Prepare Standard Setting presentation materials (i.e., PowerPoint)	July 24 – Aug 20
	6. Prepare systems, written procedures, and Standard Setting materials for analysis during the meeting	June 9 – 30
	7. Select panelists – Standard Setting	June 21
	8. Conduct practice day for group leaders and Department staff for Standard Setting	July 11
	9. Establish initial cut points and review with Department.	Aug 8
During the Standard Setting Meeting	10. Orientation	August 21
	11. Reviewing assessment materials	August 21
	12. Discussion of performance level descriptions	August 21
	13. Round 2: Table group discussion of panelist judgments and opportunity for revised judgments	August 21
	14. Tabulation of Round 2 results	August 22
	15. Round 3: Whole group discussion of panelist results and impact data and final opportunity to revise judgments	August 22
	16. Evaluation	August 22
After the Standard Setting Meeting	17. Analyze and review panelists' feedback	Aug 28 – Sept 1
	18. Review panel's recommended cut scores for cross-grade-span consistency	Aug 28 – Sept 1
	19. Prepare recommended cut scores	Sept 6
	20. Preparation of standard-setting report	October 2006

Appendix B

Standard Setting Agendas

Massachusetts Comprehensive Assessment System

ELA Standard Setting

Sheraton Ferncroft, Danvers, MA

Agenda

Tuesday, July 11, 2006

- 8:00 – 9:00 Registration
  Continental Breakfast available in the Living Room outside of North Shore "B"

- 9:00 – Noon Welcome and Introductions (Department of Education and Measured Progress)
  - Overview of Standard Setting Process (Measured Progress)
  - Break
  - Move to Grade-level/content area work rooms
  - Materials Orientation
  - Take the MCAS Test

- 12:00 – 1:00 Lunch (*Living Room outside of North Shore "B")*

- 1:00 – 4:00 Work Session

Wednesday, July 12, 2006

- 7:30 – 8:30 Continental Breakfast

- 8:30 – 12:00 Work Session

- 12:00 – 12:45 Lunch (*North Shore "B"*)

- 12:45 – Conclusion
  - Work Session
  - Standard Setting Evaluation

Appendix G

Massachusetts Comprehensive Assessment System

Mathematics Standard Setting

August 22–23, 2006
Sheraton Ferncroft, Danvers, MA

AGENDA

Tuesday, August 22

8:00 a.m. – 9:00 a.m.	Registration *Continental Breakfast available in Living Room outside of North Shore "B"*
9:00 a.m. – 10:30 a.m.	Plenary Session • Welcome and introductions • Overview of standard-setting process
10:30 a.m. – 10:45 a.m.	Break
10:45 a.m. – 12:00 p.m.	Grade-Level Work Session • Materials orientation • Take the MCAS test
12:00 p.m. – 1:00 p.m.	Lunch *(Living Room outside of North Shore "B")*
1:00 p.m. – 4:00 p.m.	Work Session

Wednesday, August 23

8:00 a.m. – 9:00 a.m.	Continental Breakfast *(Meeting rooms)*
9:00 a.m. – 12:00 p.m.	Work Session
12:00 p.m. – 12:45 p.m.	Lunch *(North Shore "B")*
12:45 p.m. – Conclusion	Work Session and Standard-Setting Evaluation

Appendix C

Standard Setting Performance Level Descriptors

Appendix G

Grade 3 - Reading			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**[Above Proficient]** **On MCAS, a student at this level:**
Language/Vocabulary	*Vocabulary/Words in Context* Demonstrates **partial** knowledge of common prefixes (e.g., "un," "dis"), dictionary skills, and context clues to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **partial** knowledge of parts of speech (e.g., verb, noun), punctuation, capitalization, and other grammatical conventions (e.g., quotation marks to set off spoken words, contractions) to understand text	*Vocabulary/Words in Context* Demonstrates **solid** knowledge of common prefixes (e.g., "un," "dis"), dictionary skills, and context clues to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **solid** knowledge of parts of speech (e.g., verb, noun), punctuation, capitalization, and other grammatical conventions (e.g., quotation marks to set off spoken words, contractions) to understand text	*Vocabulary/Words in Context* Demonstrates **substantial** knowledge of common prefixes (e.g., "un," "dis"), dictionary skills, and context clues to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **substantial** knowledge of parts of speech (e.g., verb, noun), punctuation, capitalization, and other grammatical conventions (e.g., quotation marks to set off spoken words, contractions) to understand text

Grade 3 – Reading			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Above Proficient** **On MCAS, a student at this level:**
Comprehension	*Grasp of facts and main idea* Demonstrates **partial** understanding of basic facts and main idea(s) in literary and non-literary texts *Comprehension of connections and skill at making comparisons within and across texts* Makes simple comparisons and shows **partial** understanding of similarities and differences between two stories or story elements *Understanding Of Actions And Motivations Of Characters* Demonstrates **partial** understanding of actions and motivations of characters in literary texts	*Grasp of facts and main idea* Demonstrates **solid** understanding of basic facts and main idea(s) in literary and non-literary texts *Comprehension of connections and skill at making comparisons within and across texts* Makes comparisons of **varied** complexity between and within texts, and shows **solid** understanding of similarities and differences between two stories or story elements *Understanding Of Actions And Motivations Of Characters* Demonstrates **solid** understanding of actions and motivations of characters in literary texts	*Grasp of facts and main idea* Demonstrates **substantial** understanding of basic facts and main idea(s) in literary and non-literary texts *Comprehension of connections and skill at making comparisons within and across texts* Makes some **complex** comparisons between and within texts, and shows **substantial** understanding of similarities and differences between two stories or story elements *Understanding Of Actions And Motivations Of Characters* Demonstrates **substantial** understanding of actions and motivations of characters in literary texts

Appendix G

Grade 3 - Reading			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Above Proficient** **On MCAS, a student at this level:**
Text Elements and Techniques	*Knowledge and Understanding of Text Elements* Demonstrates **partial** awareness of graphic and textual features (e.g., font style) *Knowledge of characteristics associated with literary and non-literary texts* Demonstrates **partial** knowledge and understanding of the distinctions among forms of literary and non-literary texts such as poetry, prose, fiction, nonfiction, biography and drama, and uses this knowledge to support reading comprehension	*Knowledge and Understanding of Text Elements* Demonstrates **solid** awareness of graphic and textual features (e.g., font style) *Knowledge of characteristics associated with literary and non-literary texts* Demonstrates **solid** knowledge and understanding of the distinctions among forms of literary and non-literary texts such as poetry, prose, fiction, nonfiction, biography and drama, and uses this knowledge to support reading comprehension	*Knowledge and Understanding of Text Elements* Demonstrates **substantial** awareness of graphic and textual features (e.g., font style) *Knowledge of characteristics associated with literary and non-literary texts* Demonstrates **substantial** knowledge and understanding of the distinctions among forms of literary and non-literary texts such as poetry, prose, fiction, nonfiction, biography and drama, and uses this knowledge to support reading comprehension

Grade 5 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Language/Vocabulary	*Vocabulary/Words in Context* Demonstrates **partial** knowledge of common Greek and Latin roots, suffixes, prefixes, context clues (e.g., definitions, examples, explanations in the text), and dictionary/thesaurus skills to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **partial** knowledge of parts of speech (e.g., noun, pronoun, verb, adverb, adjective, conjunction, preposition), grammatical conventions (e.g., verb tenses, simple and compound sentences), mechanics (e.g., apostrophes, quotation marks, comma use in compound sentences, paragraph indentations), and sentence structure	*Vocabulary/Words in Context* Demonstrates **solid** knowledge of common Greek and Latin roots, suffixes, prefixes, context clues (e.g., definitions, examples, explanations in the text), and dictionary/thesaurus skills to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **solid** knowledge of parts of speech (e.g., noun, pronoun, verb, adverb, adjective, conjunction, preposition), grammatical conventions (e.g., verb tenses, simple and compound sentences), mechanics (e.g., apostrophes, quotation marks, comma use in compound sentences, paragraph indentations), and sentence structure	*Vocabulary/Words in Context* Demonstrates **comprehensive** knowledge of common Greek and Latin roots, suffixes, prefixes, context clues (e.g., definitions, examples, explanations in the text), and dictionary/thesaurus skills to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **comprehensive** knowledge of parts of speech (e.g., noun, pronoun, verb, adverb, adjective, conjunction, preposition), grammatical conventions (e.g., verb tenses, simple and compound sentences), mechanics (e.g., apostrophes, quotation marks, comma use in compound sentences, paragraph indentations), and sentence structure

Appendix G

Grade 5 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Comprehension	*Grasp of Facts and Main Idea* Demonstrates **partial** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections between and within Texts* Makes **simple** comparisons of **limited** complexity within texts and shows **partial** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and drawing conclusions about the main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **simple** comparisons of **limited** complexity between texts and shows **partial** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **partial** understanding of actions, motivations, and traits of characters in literary texts	*Grasp of Facts and Main Idea* Demonstrates **solid** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections between and within Texts* Makes **solid** comparisons of **moderate** complexity within texts and shows **solid** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and drawing conclusions about the main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **solid** comparisons of **moderate** complexity between texts and shows **solid** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **solid** understanding of actions, motivations, and traits of characters in literary texts	*Grasp of Facts and Main Idea* Demonstrates **comprehensive** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections between and within Texts* Makes **complex** comparisons within texts and shows **comprehensive and in-depth** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and drawing conclusions about the main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **complex** comparisons between texts and shows **comprehensive and in-depth** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **in-depth** understanding of actions, motivations, and traits of characters in literary texts

Grade 5 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Text Elements and Techniques	*Knowledge of Purpose/Function of Graphic and Textual Features* Demonstrates **partial** awareness of graphic and textual features (e.g., title, headings, key words, captions, table of contents, index; charts, graphs, diagrams, illustrations) and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **partial** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essay, myth, and traditional narrative) *Understanding of Style and Language* Demonstrates **partial** understanding of how an author uses sensory details and figurative language	*Knowledge of Purpose/Function of Graphic and Textual Features* Demonstrates **solid** awareness of graphic and textual features (e.g., title, headings, key words, captions, table of contents, index; charts, graphs, diagrams, illustrations) and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **solid** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essay, myth, and traditional narrative) *Understanding of Style and Language* Demonstrates **solid** understanding of how an author uses sensory details and figurative language	*Knowledge of Purpose/Function of Graphic and Textual Features* Demonstrates **comprehensive** awareness of graphic and textual features (e.g., title, headings, key words, captions, table of contents, index; charts, graphs, diagrams, illustrations) and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **comprehensive** and **in-depth** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essay, myth, and traditional narrative) *Understanding of Style and Language* Demonstrates **comprehensive** and **in-depth** understanding of how an author uses sensory details and figurative language

Appendix G

Grade 6 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Language/Vocabulary	*Vocabulary/Words in Context* Demonstrates **partial** knowledge of common Greek and Latin roots, suffixes, prefixes, context clues (e.g., definitions, examples, explanations in the text), and dictionary/thesaurus skills to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **partial** knowledge of parts of speech (e.g., noun, pronoun, verb, adverb, adjective, conjunction, preposition, interjection), grammatical conventions (e.g., verb tenses, simple and compound sentences), mechanics (e.g., apostrophes, quotation marks, comma use in compound sentences, paragraph indentations), and sentence structure	*Vocabulary/Words in Context* Demonstrates **solid** knowledge of common Greek and Latin roots, suffixes, prefixes, context clues (e.g., definitions, examples, explanations in the text), and dictionary/thesaurus skills to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **solid** knowledge of parts of speech (e.g., noun, pronoun, verb, adverb, adjective, conjunction, preposition, interjection), grammatical conventions (e.g., verb tenses, simple and compound sentences), mechanics (e.g., apostrophes, quotation marks, comma use in compound sentences, paragraph indentations), and sentence structure	*Vocabulary/Words in Context* Demonstrates **comprehensive** knowledge of common Greek and Latin roots, suffixes, prefixes, context clues (e.g., definitions, examples, explanations in the text), and dictionary/thesaurus skills to derive meaning of unfamiliar words in text *Grammatical Conventions* Demonstrates **comprehensive** knowledge of parts of speech (e.g., noun, pronoun, verb, adverb, adjective, conjunction, preposition, interjection), grammatical conventions (e.g., verb tenses, simple and compound sentences), mechanics (e.g., apostrophes, quotation marks, comma use in compound sentences, paragraph indentations), and sentence structure

Grade 6 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Comprehension	*Grasp of Facts and Main Idea* Demonstrates **partial** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections between and within Texts* Makes **simple** comparisons of **limited** complexity within texts and shows **partial** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and analyzing the main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **simple** comparisons of **limited** complexity between texts and shows **partial** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **partial** understanding of actions, motivations, and traits of characters in literary texts	*Grasp of Facts and Main Idea* Demonstrates **solid** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections between and within Texts* Makes **solid** comparisons of **moderate** complexity within texts and shows **solid** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and analyzing the main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **solid** comparisons of **moderate** complexity between texts and shows **solid** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **solid** understanding of actions, motivations, and traits of characters in literary texts	*Grasp of Facts and Main Idea* Demonstrates **comprehensive** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections between and within Texts* Makes **complex** comparisons within texts and shows **comprehensive** and **in-depth** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and analyzing the main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **complex** comparisons between texts and shows **comprehensive** and **in-depth** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **in-depth** understanding of actions, motivations, and traits of characters in literary texts

Appendix G

Grade 6 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Text Elements and Techniques	*Knowledge of Purpose/Function of Graphic and Textual Features* Demonstrates **partial** awareness of graphic and textual features (e.g., title, headings, key words, captions, table of contents, index; charts, graphs, diagrams, illustrations) and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **partial** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essay, myth, and traditional narrative) *Understanding of Style and Language* Demonstrates **partial** understanding of how an author uses sensory details and figurative language	*Knowledge of Purpose/Function of Graphic and Textual Features* Demonstrates **solid** awareness of graphic and textual features (e.g., title, headings, key words, captions, table of contents, index; charts, graphs, diagrams, illustrations) and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **solid** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essay, myth, and traditional narrative) *Understanding of Style and Language* Demonstrates **solid** understanding of how an author uses sensory details and figurative language	*Knowledge of Purpose/Function of Graphic and Textual Features* Demonstrates **comprehensive** awareness of graphic and textual features (e.g., title, headings, key words, captions, table of contents, index; charts, graphs, diagrams, illustrations) and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **comprehensive** and **in-depth** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essay, myth, and traditional narrative) *Understanding of Style and Language* Demonstrates **comprehensive** and **in-depth** understanding of how an author uses sensory details and figurative language

Grade 8 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Language/Vocabulary	*Language and Vocabulary* Demonstrates a **partial** knowledge of prefixes, suffixes, root words, dictionary skills, and a variety of context clues (e.g., contrast, cause and effect) to derive the meaning of unfamiliar words in text *Conventions of Grammar and Mechanics* Demonstrates **partial** knowledge of parts of speech (e.g., verbs, nouns, adjectives, adverbs, interjections, conjunctions, pronouns), grammatical conventions (e.g., prepositional phrases, simple, compound, and complex sentences, pronoun references), and mechanics (e.g., apostrophes, quotation marks, commas) to understand text	*Language and Vocabulary* Demonstrates a **solid** knowledge of prefixes, suffixes, root words, dictionary skills, and a variety of context clues (e.g., contrast, cause and effect) to derive the meaning of unfamiliar words in text *Conventions of Grammar and Mechanics* Demonstrates **solid** knowledge of parts of speech (e.g., verbs, nouns, adjectives, adverbs, interjections, conjunctions, pronouns), grammatical conventions (e.g., prepositional phrases, simple, compound, and complex sentences, pronoun references), and mechanics (e.g., apostrophes, quotation marks, commas) to understand text	*Language and Vocabulary* Demonstrates a **comprehensive** knowledge of prefixes, suffixes, root words, dictionary skills, and a variety of context clues (e.g., contrast, cause and effect) to derive the meaning of unfamiliar words in text *Conventions of Grammar and Mechanics* Demonstrates **comprehensive** knowledge of parts of speech (e.g., verbs, nouns, adjectives, adverbs, interjections, conjunctions, pronouns), grammatical conventions (e.g., prepositional phrases, simple, compound, and complex sentences, pronoun references), and mechanics (e.g., apostrophes, quotation marks, commas) to understand text

Appendix G

Grade 8 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Comprehension	*Grasp of Facts and Main Idea(s)* Demonstrates **partial** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections within and between Texts* Makes **simple** comparisons of **limited** complexity within texts and shows **partial** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and analyzing the author's purpose, main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **simple** comparisons of **limited** complexity between texts and shows **partial** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **partial** understanding of actions, motivations, and traits of characters in literary texts	*Grasp of Facts and Main Idea(s)* Demonstrates **solid** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections within and between Texts* Makes **solid** comparisons of **moderate** complexity within texts and shows **solid** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and analyzing the author's purpose, main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **solid** comparisons of **moderate** complexity between texts and shows **solid** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **solid** understanding of actions, motivations, and traits of characters in literary texts	*Grasp of Facts and Main Idea(s)* Demonstrates **comprehensive** understanding of basic facts and main idea(s) in literary and non-literary grade-level texts *Comprehension of Connections within and between Texts* Makes **insightful** comparisons of **rich** complexity within texts and shows **sophisticated** understanding of the similarities and differences among elements within a text (e.g., comparing characters within a story, or the same character at different points in a story; identifying and analyzing the author's purpose, main idea or theme from evidence within a literary or nonfiction text; analyzing the interrelations between ideas in a nonfiction text); makes **insightful** comparisons of **rich** complexity between texts and shows **sophisticated** understanding of the similarities and differences between two texts *Understanding of Actions and Motivations of Characters* Demonstrates **insightful** understanding of actions, motivations, and traits of characters in literary texts

Grade 8 - ELA			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Advanced** **On MCAS, a student at this level:**
Text Elements and Techniques	*Function of Graphic and Textual Features* Demonstrates **partial** awareness of graphic and textual features and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **partial** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essay, myth, traditional narrative, and editorials) *Understanding of Style and Language* Demonstrates **partial** understanding of how an author uses sensory details and figurative language, suggests mood, and sets tone	*Function of Graphic and Textual Features* Demonstrates **solid** awareness of graphic and textual features and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **solid** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essays, myth, traditional narrative, and editorials) *Understanding of Style and Language* Demonstrates **solid** understanding of how an author uses sensory details and figurative language, suggests mood, and sets tone	*Function of Graphic and Textual Features* Demonstrates **comprehensive** awareness of graphic and textual features and organizational structures to comprehend text *Knowledge of Elements and Characteristics of Literary and Non-Literary Texts* Demonstrates **in-depth** knowledge of characteristics associated with forms of literary and non-literary grade-level texts (e.g., poetry, autobiography, biography, informational text, drama, personal essay, myth, traditional narrative, and editorials) *Understanding of Style and Language* Demonstrates **insightful** understanding of how an author uses sensory details and figurative language, suggests mood, and sets tone

Appendix G

Grade 3 Performance Descriptors -Math			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Proficient with Distinction** **On MCAS, a student at this level:**
Number Sense	• Reads numbers and identifies the magnitude of numbers through thousands • Adds and subtracts without regrouping • Identify fractional parts of a whole (with denominators through 10) • Rounds two-digit numbers to the nearest 10	• Reads, writes, and interprets different place value representations through thousands • Adds and subtracts with regrouping • Knows multiplication and division basic facts • Identifies fractional parts of a group and models mixed numbers • Rounds three-digit numbers to the nearest 10 and 100	• Uses understanding of numeration system to explain solution to problems (intentionally the same as Grade 4) • Understands the concept of multiplication and multiplies two-digit numbers by one-digit numbers • Selects and uses appropriate operations to solve problems • Compares fractions and locates fractions on a number line • Uses rounding to estimate whole number computations (addition, subtraction, multiplication)
Patterns, Relations and Algebras	• Extends addition and subtraction patterns and geometric patterns	• Describes addition and subtraction patterns and geometric patterns • Uses equations to represent mathematical situations • Finds missing numbers in equations • Compares numbers using >, <, =	• Applies patterns to the solution of problems • Compares expressions using >, <, =
Geometry	• Identifies two-dimensional	• Identifies three-dimensional	• Describes three-dimensional

	shapes	shapes • Describes two-dimensional shapes • Recognizes right angles, perpendicular lines, and parallel lines • Uses ordered pairs to locate and identify points on a grid	shapes
Measurement	• Tells time at five-minute intervals on analog and digital clocks • Understands and accurately measures length in metric and US Customary units	• Carries out simple unit conversions • Finds perimeter and area of rectangles using grids • Tells time to the nearest minute on analog and digital clocks • Understands and accurately measures weight and temperature	• Applies the concepts of length, weight, temperature, elapsed time, perimeter, and area to the solution of problems
Data Analysis, Statistics and Probability	• Constructs and reads tallies, bar graphs, tables, and pictographs (where key represents one)	• reads pictographs (where key represents number other than one) and reads and constructs line plots. • Determines number of possible combination of objects from two sets	• Constructs pictographs (where key represents number other than one) • Draws conclusions from representations of data sets

Appendix G

Grade 5 Performance Descriptors - Math			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Proficient with Distinction** **On MCAS, a student at this level:**
Number Sense	• Demonstrate an understanding of place value from thousandths to millions • Demonstrates an understanding of fractions as parts of whole units (intentionally same as grade 6) • Identifies and determines common equivalent fractions (with denominators 2, 4, 5, 10) • Adds, subtracts, multiplies, and divides whole numbers; adds and subtracts positive decimals and fractions; multiplies positive decimals and fractions with whole numbers with partial accuracy • Estimate sums, differences , and products of whole numbers	• Represents and compares positive numbers in various forms such as expanded notation without exponents • Demonstrates and understanding of fractions as a ratio of whole numbers (intentionally same as grade 6) • Identifies and determines common equivalent fractions and mixed numbers (with denominators 2, 4, 5, 10) and decimals • Finds common factors and common multiples and uses divisibility rules (for 2, 3, 5, and 10) • Adds, subtracts, multiples and divides whole numbers; adds and subtracts positive decimals and fractions; multiplies positive decimals and fractions with whole numbers with few errors • Estimate sums and differences of positive fractions and positive decimals	• Applies understanding of numeration system and number theory concepts to the solution of problems • Compare and order whole numbers, positive fractions, positive mixed numbers, and positive decimals • Adds, subtracts, multiples and divides whole numbers; adds and subtracts positive decimals and fractions; multiplies positive decimals and fractions with whole numbers with a high level of accuracy • Selects and uses appropriate operations to solve non-routine and multi-step problems • Estimate products of positive decimals with whole numbers

Patterns, Relations and Algebras	• Extends simple symbolic, arithmetic, and geometric patterns and progressions (intentionally same as grade 6) • Evaluates simple expressions when values are given for the variables (intentionally same as grade 6) • Represents simple mathematical relationships with concrete models and tables	• Explains the rules for extending a variety of patterns and progressions (intentionally same as grade 6) • Evaluates expressions when values are given for the variables (intentionally same as grade 6) • Represents simple mathematical relationships with graphs and rules • Solve simple problems involving proportional relationships	• Determines the rules for extending patterns and progressions and applies the rules to the solution of problems (intentionally same as grade 6) • Applies properties of equality to the solution of problems (intentionally same as grade 6) • Represents mathematical relationships with tables, graphs, and rules • Applies proportional reasoning to the solution of problems • Interpret graphs that represent the relationship between two variables
Geometry	• Identifies special types of triangles and quadrilaterals based on properties of their sides • Identifies relationships among lines (e.g., parallel, perpendicular, and intersecting) • Graphs points and find coordinates of points using whole numbers (intentionally the same as Grade 6) • Describes simple transformations on two	• Compares special types of triangles and quadrilaterals • Identifies and describes cubes and prisms based on properties of faces • Describe paths on the Cartesian coordinate plane • Performs simple transformations on two-dimensional shapes • Identify and describe multiple lines of symmetry in two dimensional shapes	• Recognizes and describes relationships among special types of triangles and quadrilaterals • identifies and describes cubes and prisms based on number of edges; identifies and describes pyramids based on properties of faces and number of edges • Uses manipulatives to predict and explain the results of taking apart and combining shapes

	dimensional shapes • Identify and describe line symmetry in two-dimensional shapes		
Measurement	• Calculates the perimeter of triangles and the perimeter and area of rectangles • Calculates volume of rectangular prisms • Classifies and draws various angles	• Calculates the area of triangles • Calculates surface area of rectangular prisms • Solve problems involving simple unit conversions within a system of measurement • Find the sum of the measures of interior angles of triangles	• Applies concepts of perimeter, area, volume, and surface area to the solution of problems
Data Analysis, Statistics and Probability	• Identifies mode, maximum and minimum for a given set of data • Constructs and interprets bar graphs	• Identifies mean, median, and range for a given set of data • Constructs and interprets line graphs and line plots; interpret circle graphs • Computes the probability of outcomes of simple experiments with one event (intentionally the same as grade 6 needs improvement)	• Applies concepts of central tendency and spread to the solution of problems • Applies concepts of probability to the solution of problems

Grade 7 Performance Descriptors - Math			
	Needs Improvement **On MCAS, a student at this level:**	**Proficient** **On MCAS, a student at this level:**	**Proficient with Distinction** **On MCAS, a student at this level:**
Number Sense and Operations	• Represents and compares integers, fractions, mixed numbers, decimals and percents • Computes and estimates with whole numbers, fractions, and decimals	• Demonstrates an understanding of scientific notation (positive powers of ten only), absolute value and positive integer exponents • Sets up ratios and proportions • Computes and estimates with integers and percents	• Applies numeration concepts to the solution of problems • Applies concepts of ratio, proportions and percent to the solution of problems (intentionally same as grade 8) • Applies operations on rational numbers to the solution of problems
Patterns, Relations and Algebras	• Identifies numeric and/or geometric patterns (intentionally same as Grade 8) • Determines the value of an algebraic expression by substituting variables with given values (intentionally same as Grade 8)	• Generate rules or general terms to describe numeric and geometric patterns (intentionally same as Grade 8) • Solves linear equations with one variable (intentionally same as Grade 8)	• Represents theoretical and practical situations using equations and solves the problems graphically and algebraically (intentionally same as Grade 8)
Geometry	• Performs transformations on two dimensional shapes • Graphs and identifies points on the coordinate plane	• Identifies congruent and similarity relationships, performs transformation on the coordinate plane • Demonstrates an understanding of the relationships of angles formed by two parallel lines cut by a transversal	• Applies concepts of congruence, similarity, transformations, and geometric relationships to the solution of problems • Demonstrates an understanding of the relationships of angles in more complex diagrams and in problem solving situations

		• Identifies congruent and similarity relationships, performs transformation on the coordinate plane • Demonstrates an understanding of the relationships of angles formed by two parallel lines cut by a transversal	• Applies concepts of congruence, similarity, transformations, and geometric relationships to the solution of problems • Demonstrates an understanding of the relationships of angles in more complex diagrams and in problem solving situations
Measurement	• Computes perimeter, circumference, and area of geometric figures • Performs simple conversions within a system	• Computes surface area and volume of rectangular prisms, cylinders • Performs simple conversions from one system to another	• Applies computations of surface are, and volume to the solution of problems • Performs conversions of square and/or cubic units
Data Analysis, Statistics and Probability	• Computes mean, median, and mode for a given data set • Displays data using tables and charts • Computes the probability of outcomes of simple experiments with one event (intentionally the same as Grade 6)	• describes data sets using multiple measures of central tendency and spread • displays data using circle graphs, Venn diagrams, and stem-and-leaf plots •	• Uses mean, median and mode to analyze and compare sets of data (intentionally same as grade 8) • Constructs circle graphs using calculated angle measures • Computes compound probabilities for independent events

Appendix D

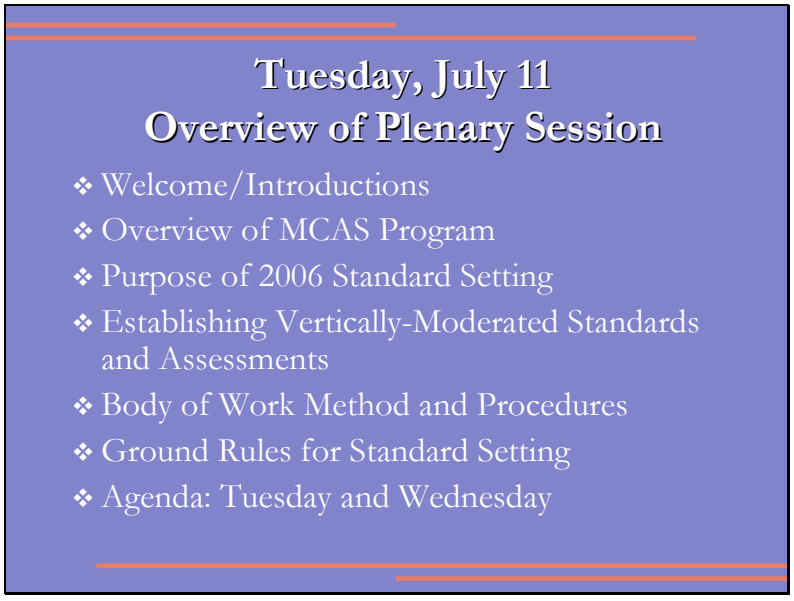
Standard Setting Plenary Session
Power Point Presentations
English Language Arts
Mathematics

Appendix G

Plenary Session PowerPoint Presentations

Slide 1

Slide 2



Slide 3

Slide 4

Slide 5

Standard Setting Facilitators

- Grade 3 – Jan Katien
- Grade 5 – Lisa Ehrlich and Eric Wigode
- Grade 6 – Brenda Thomas
- Grade 8 – Sharman Price

Slide 6

Welcome Grade 3 Reading Panelists

Karen Alexopoulos, Grade 3 Teacher, Somerville Public Schools
Gayla Berry, ELL Teacher, Holyoke Public Schools
Marie Champion, Title 1 Teacher – ELA, Medford Public Schools
Meghan Coleman, Grade 3 Teacher, Cohasset Public Schools
Madelyn Farrell, Grade 3 Teacher, Woburn Public Schools
Nancy Fogg, Grade 4 Teacher, Brockton Public Schools
Robin Gazelian, Paraprofessional, Methuen Public Schools
Andrea Hallion, Adjunct Professor, Framingham State College
Kathleen Jankins, Reading Specialist, Bridgewater Raynham Regional
Sheila Kukstis, Principal, Taunton Public Schools
Kerri Laurenzo, Reading Specialist, East Longmeadow Public Schools
Margaret Martinez, Director of Curriculum & Instruction, Berkley Public Schools
Elaine McNamara, Title I Director/Title I Reading Teacher, Dracut Public Schools
Linda Mros, Grade 3 Teacher, Taunton Public Schools
Kevinetta O'Brien, Retired Elementary Teacher, Charlton District
Gina Patti, Reading Specialist, Oak Bluffs Public Schools
Linda Volpicelli, Language Arts Coordinator K-5, Reading Specialist 3-5, Bedford Public Schools
Joyce Welch, Special Education Teacher, Springfield Public Schools
Pamela Westmoreland, Representative, National Association for the Education of Young Children

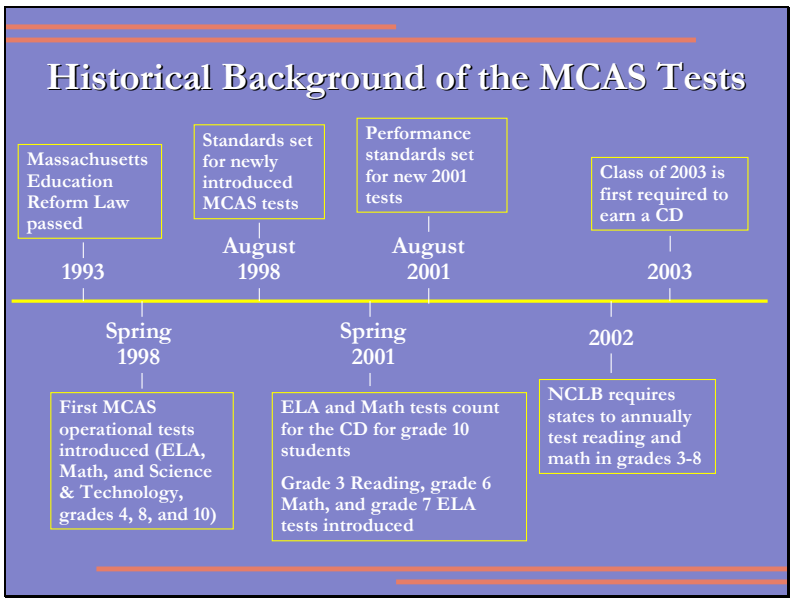
Slide 7

Slide 8

Appendix G

Slide 9

Slide 10



Slide 11

Purpose of MCAS Program

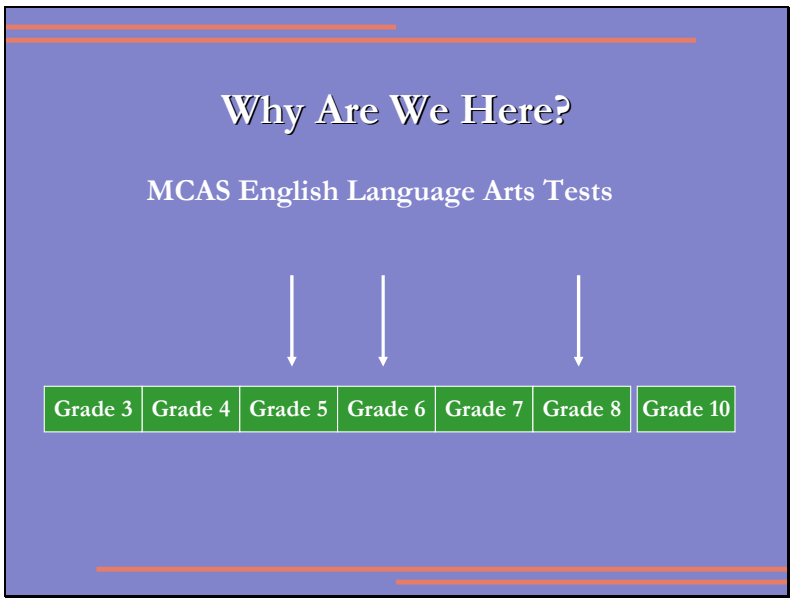
- Inform/improve curriculum and instruction
- Evaluate student, school, and district performance according to *Curriculum Framework* content standards and MCAS performance standards
- Determine eligibility for high school Competency Determination

Slide 12

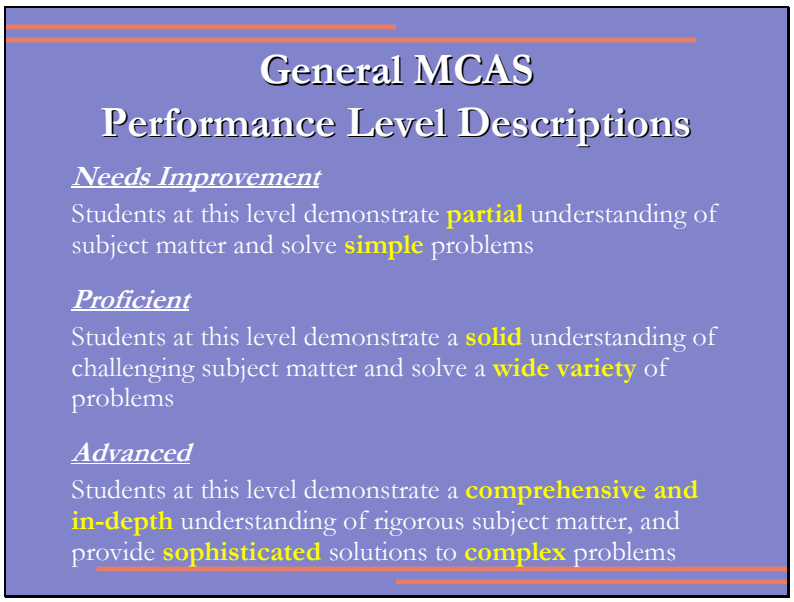
Selected Features of MCAS

- Custom developed based on *Massachusetts Curriculum Framework* content standards and MCAS performance standards
- 100% of questions used to determine student scores released annually
- Measures performance of ALL students educated with public funds
- Results reported according to scaled scores and performance levels

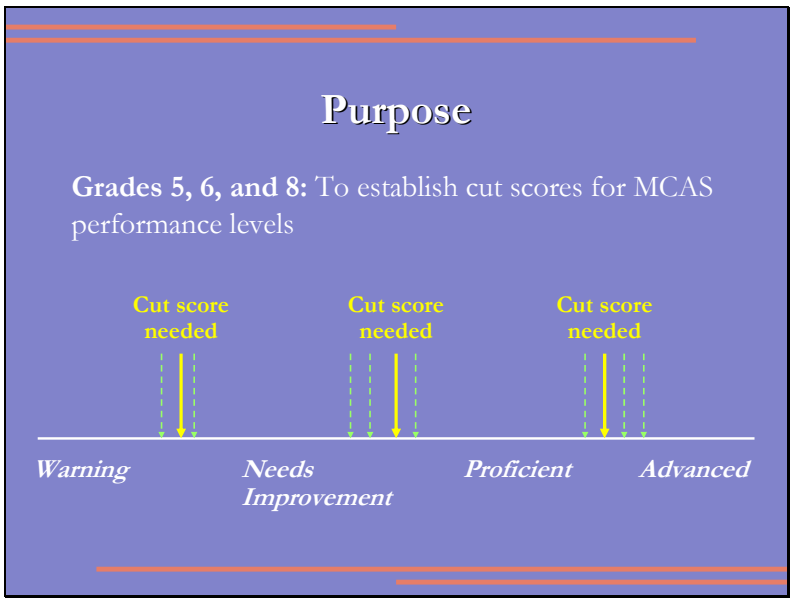
Slide 13



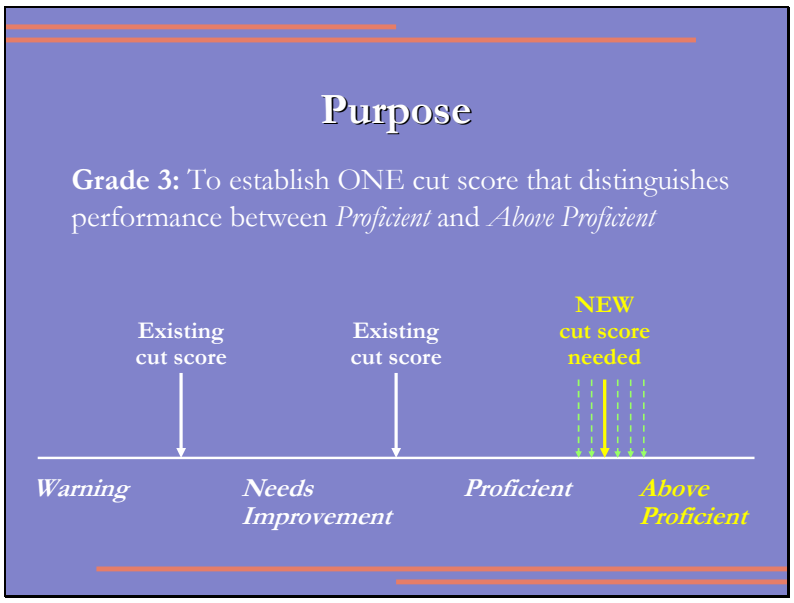
Slide 14



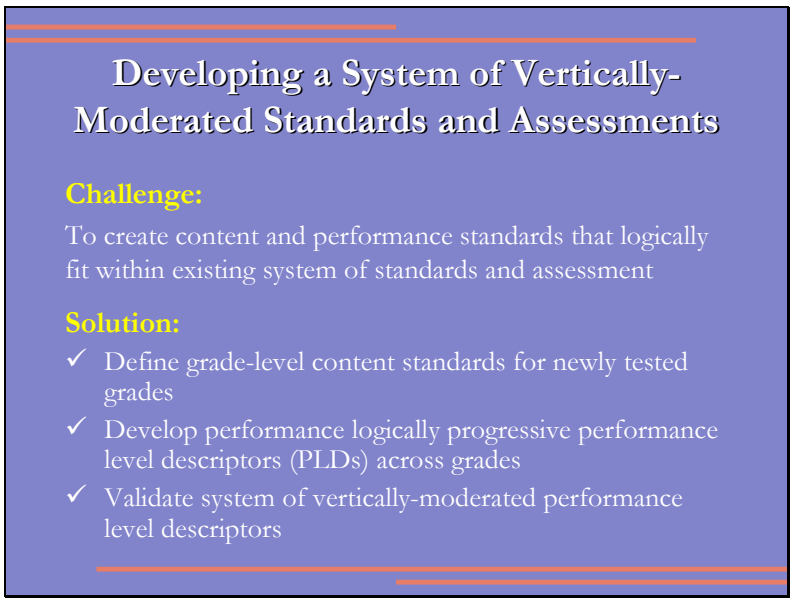
Slide 15



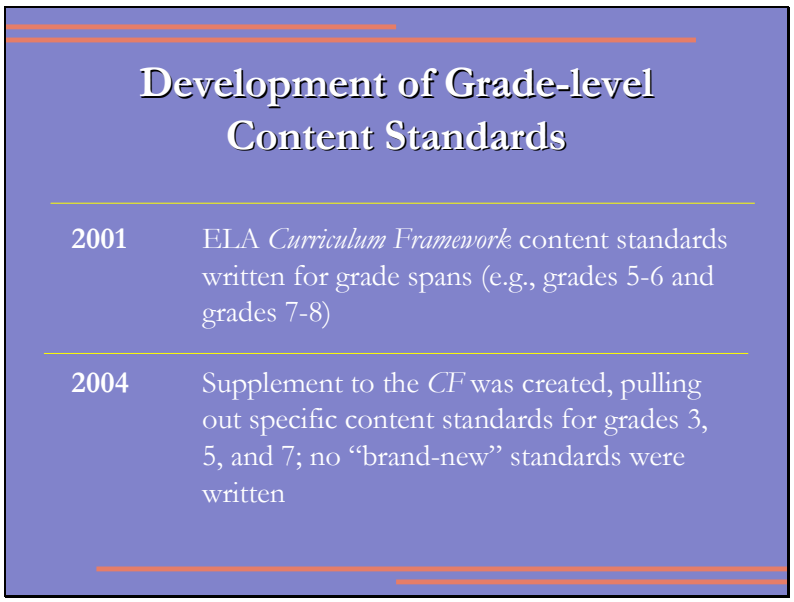
Slide 16



Slide 17



Slide 18



Slide 19

Development of Grade-Specific Performance Level Descriptions for New Assessments

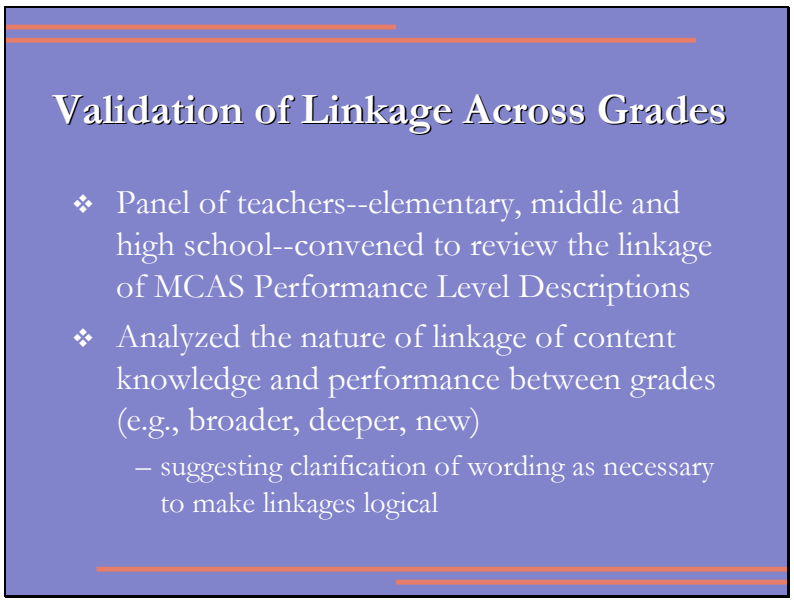
- Created for new tests at grades tested 5, 6, and 8
- New performance level—*Above Proficient*—created for grade 3
- Linked to general MCAS descriptions of student work at *Needs Improvement*, *Proficient*, and *Advanced* levels (except grade 3)
- Describe performance across levels within grade *and* link performance across grades

Slide 20

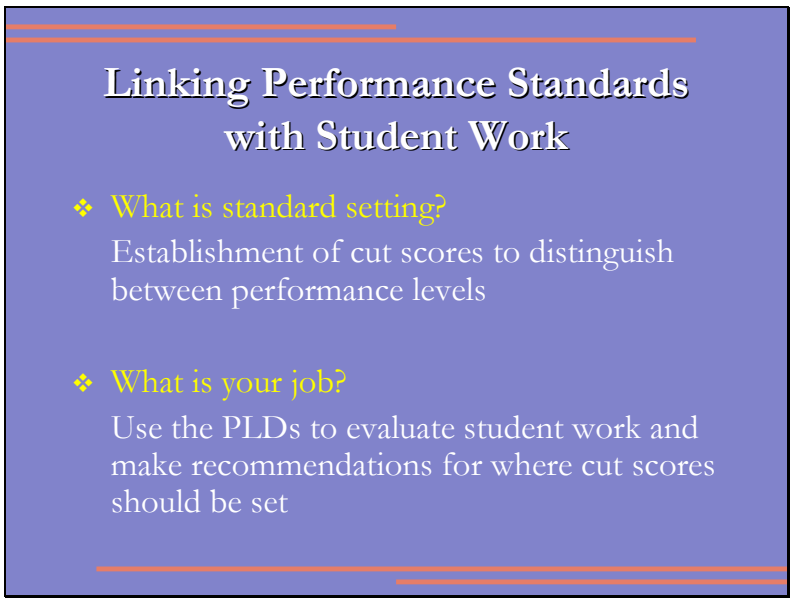
Content Standards *vs.* Performance Standards

- Content standards = "What"
  Describe the knowledge and skills students should acquire in a particular content and grade
- Performance Standards = "How well"
  Describe student work on MCAS tests at the *Needs Improvement*, *Proficient*, and *Advanced* and *Above Proficient* levels

Slide 21



Slide 22



Slide 23

Purpose of Standard Setting

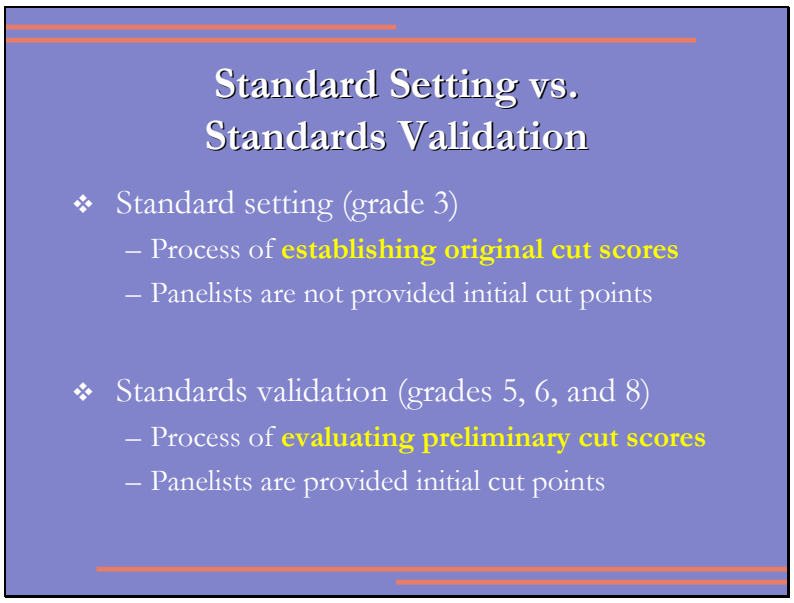
- Determine cut scores for reporting assessment results
- Answer the question:
  - How much is enough?

Slide 24

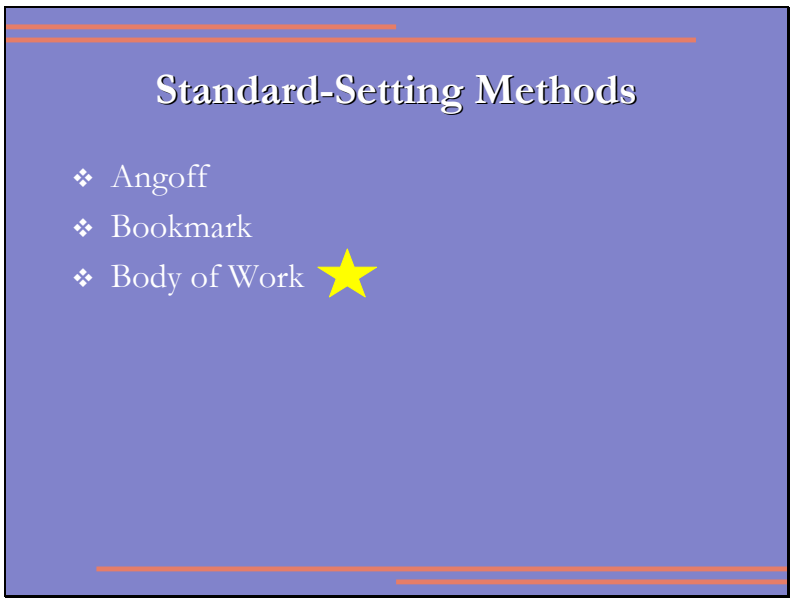
General Phases of Standard Setting/ Standards Validation

- Data-collection phase ★
- Policy-making/decision-making phase

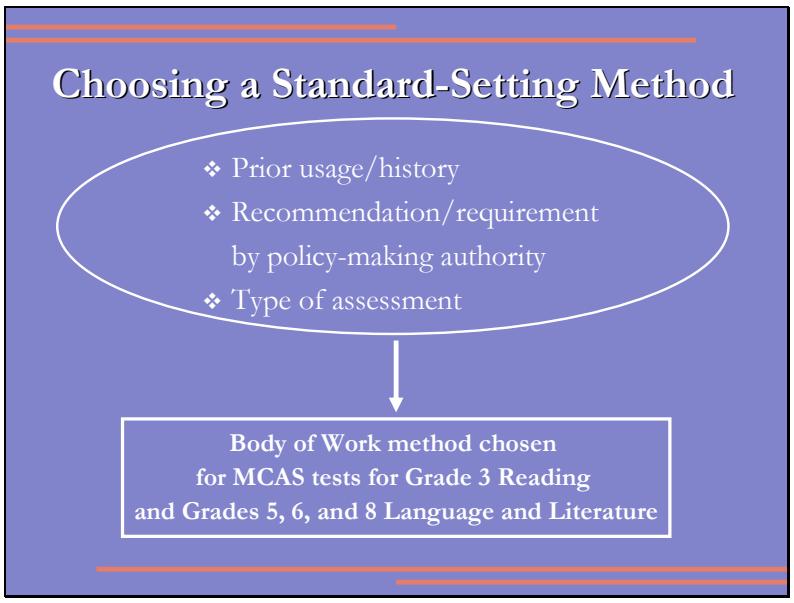
Slide 25



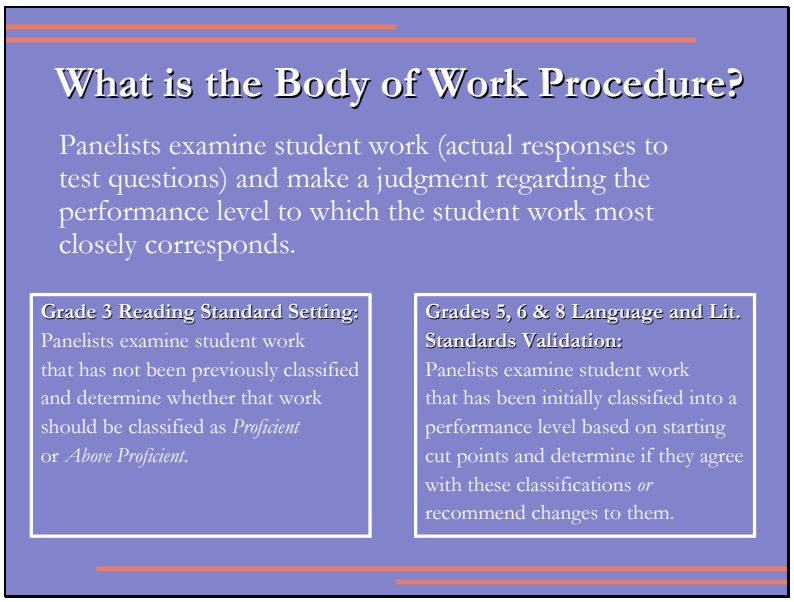
Slide 26



Slide 27



Slide 28



Slide 29

Initial Classification of Student Work

Initial classification of student work in grades 5, 6, and 8 based on most recent grade 4 and 7 performance.

Step 1: Use data from grades 4 and 7 to **estimate** percent of students at each performance level in grades 5, 6, and 8.

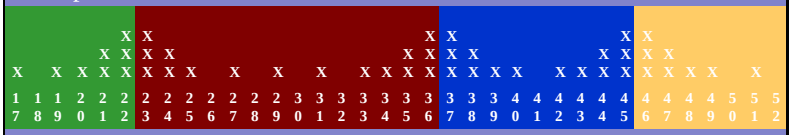
Step 2: Determine scores on the grades 5, 6, and 8 tests that produced the percentages from Step 1.

Step 3: Select student work with scores ranging from very low to very high; classify them into performance levels based on **preliminary** cut points found in Step 2.

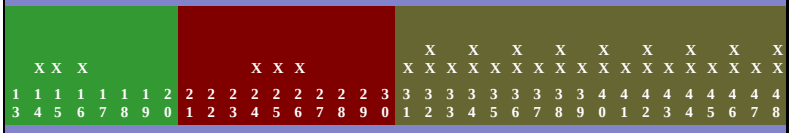
Slide 30

Selected Student Work

Example Distribution of Selected Student Work: Grades 5, 6, & 8



Example Distribution of Selected Student Work: Grade 3



Slide 31

Initial Classification of Student Work

- Determining the initial classifications of student work was a purely statistical process
- Need to verify that the statistical results are reasonable and make sense (that's your job!)

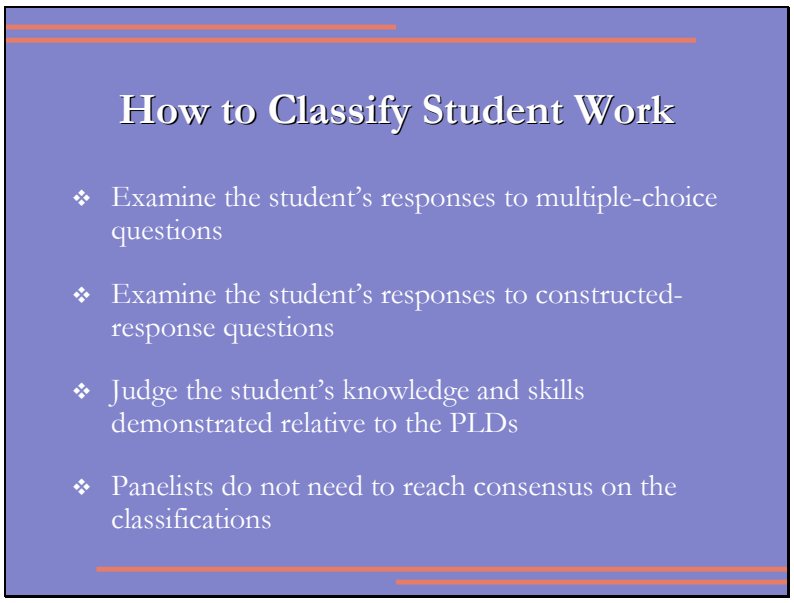
Slide 32

How to Classify Student Work

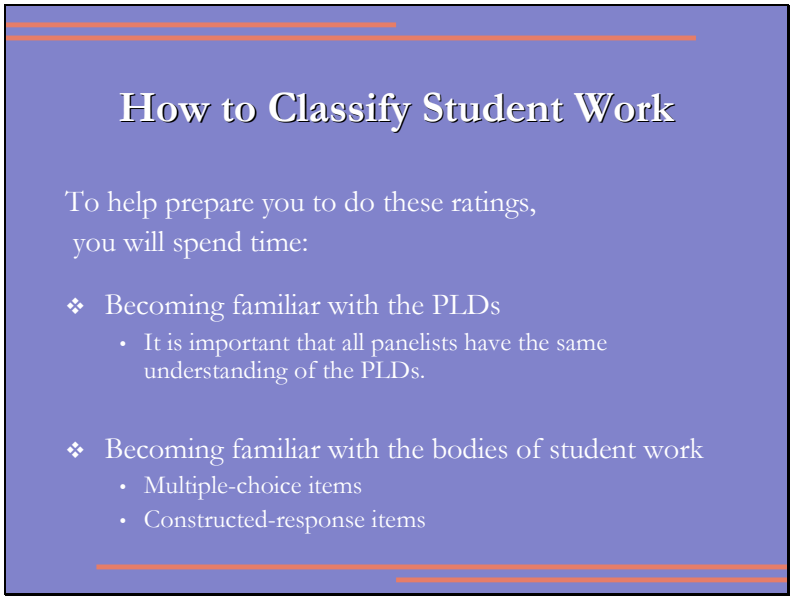
Materials you will need:

- Performance Level Definitions
  - General
  - Grade and content specific
- Bodies of Student Work
  - Responses to constructed-response questions
  - Multiple-choice summary sheet
- Rating Form

Slide 33



Slide 34



Slide 35

How to Classify Student Work

- You will have the opportunity to discuss your classifications and change them if desired.
- Don't worry! We have procedures, materials, and staff to assist you in this process.

Slide 36

What Next?

Break into grade-level groups:

- Take the assessment
- Discuss the Performance Level Definitions
- Complete the Item Map
- Complete training round
- Complete individual ratings
- Discuss ratings and revise
- Receive feedback from revised ratings
- Discuss feedback and provide final ratings
- Complete an evaluation form

Slide 37

Top 8 Most Misunderstood Things about Standard Setting

8. Standard setting is a great opportunity to rewrite *Curriculum Framework* standards.
7. The process is rigged.
6. This is a good time to vent about all the things you hate about MCAS.
5. We should use this time to rework ELA performance level definitions.

Slide 38

Top 8 Most Misunderstood Things about Standard Setting

4. Standard setting is scoring.
3. Only literary scholars should be doing this work.
2. Only teachers should be doing this work.
1. Disagreement is bad.

Slide 39

Ground Rules

- Role of facilitator is to "facilitate" and keep process on track
- Process *solely* focused on recommending performance standards (cut scores) for MCAS
- MCAS performance level definitions are integral to process but are not up for debate
- Panelists' recommendations are vital; however, final cut scores determined by the MDOE
- Each panelist must be in attendance for the duration of the process for his/her judgments to be considered
- Cell phones off, please!

Slide 40

Agenda

Tuesday, July 11

Plenary	9:00 – 10:00 am
Break	10:00 – 10:15 am
Grade-level work	10:15 am –12:00 pm
Lunch	12:00 – 12:45 pm
Grade-level work	12:45 – 4:30 pm

Wednesday, July 12

Breakfast	7:30 – 8:30 am
Grade-level work	8:30 am – 12:00 pm
Lunch	12:00 – 12:45 pm
Grade-level work	12:45 – 4:30 pm

Slide 41

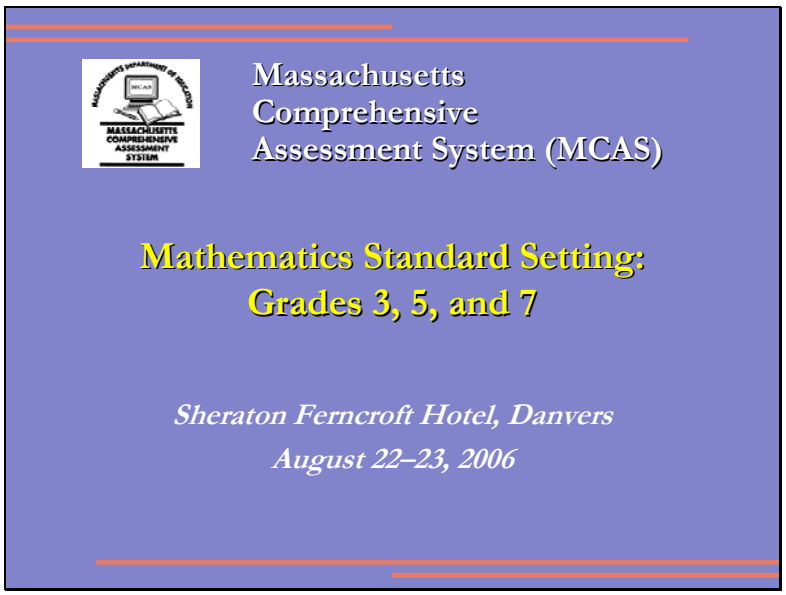
Room Assignments

Grade 3 - Marblehead Room
Grade 5 - Newburyport Room
Grade 6 - Gloucester Room
Grade 8 - Ipswich Room

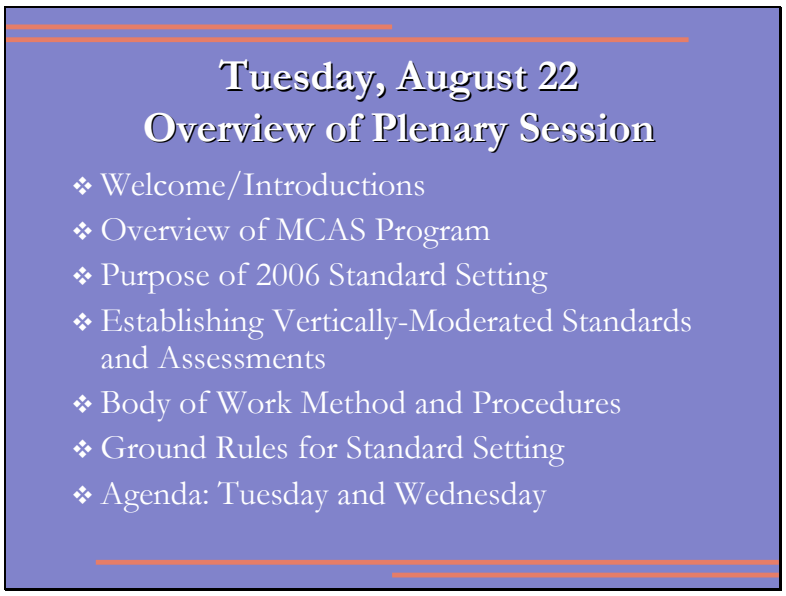
Slide 42

Questions?

Slide 1



Slide 2



Appendix G

Slide 3

Slide 4

Slide 5

Standard Setting Facilitators

- Grade 3 – Sally Blake
- Grade 5 – Margaret Hill
- Grade 7 – Donna Carlins

Slide 6

Welcome Grade 3 Panelists

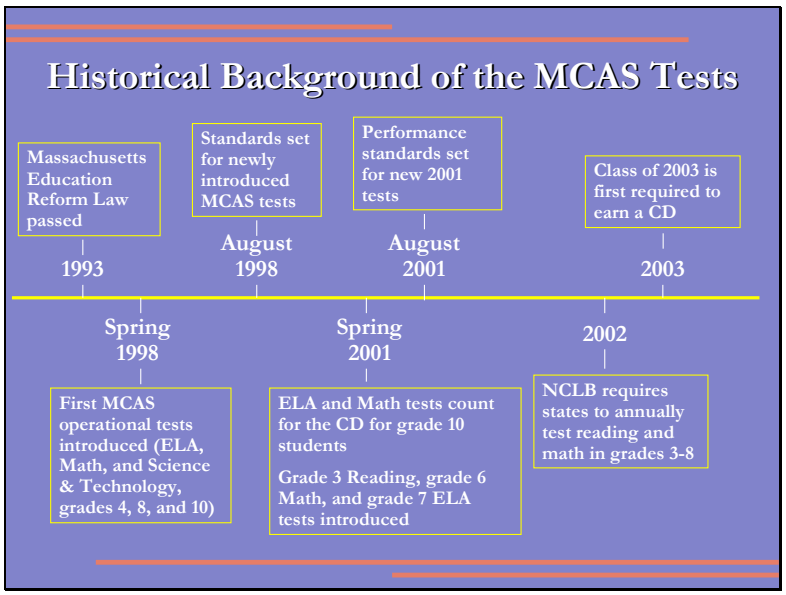
John Cardoza, Middleborough Public Schools
Adriana Gallo-Grimaldi, Agawam Public Schools
Linda Gauthier, Saugus Public Schools
Cheryl Goguen, Framingham Public Schools
Rev. Ray Hammond, Boston Ten Point Coalition
Kristine Klumpp, Duxbury Public Schools
Carol LaPolice, Springfield Public Schools
Susan Mello, Fall River Public Schools
Kathleen Millett, West Springfield Public Schools
Lyudmila Moiseyeva, Brookline Public Schools
Judy Moore, Harvard Public Schools
Stephanie Morris, Central Berkshire Regional School District
Stephanie Murchison-Brown, Holyoke Public Schools
Arthur Norman, Fitchburg Public Schools
Misael Ramos, Springfield Public Schools
Jen Rubera, Haverhill Public Schools
Victoria Sapko, Framingham State College
Michael Stanton, Walpole Public Schools
Elizabeth Sweeney, Boston Public Schools
Denise Young, Natick Public Schools

Appendix G

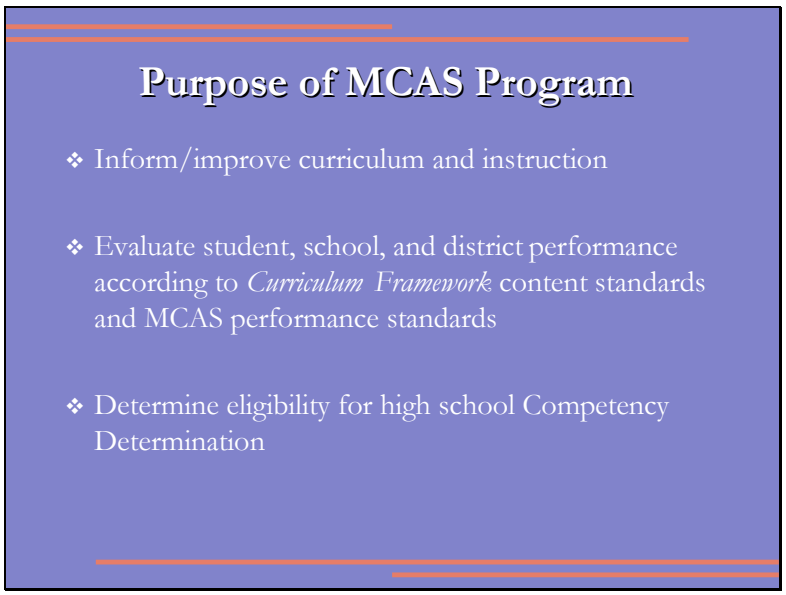
Slide 7

Slide 8

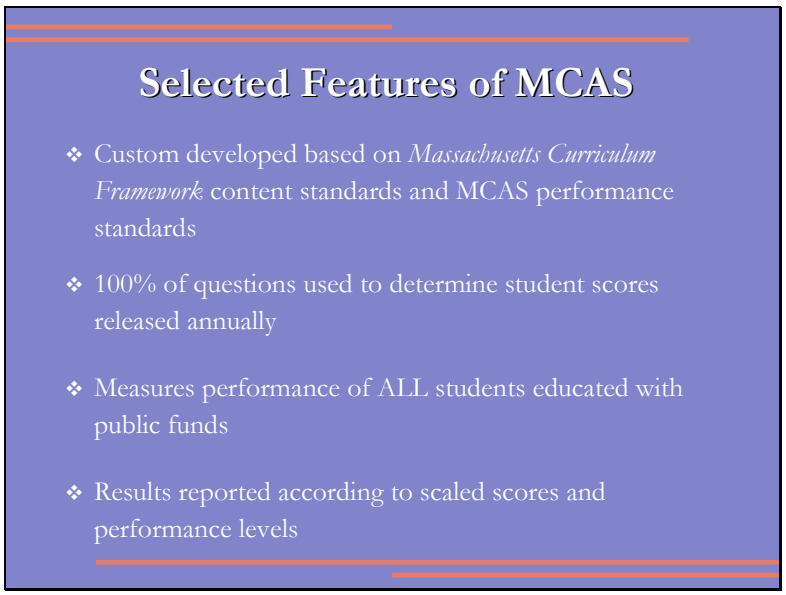
Slide 9



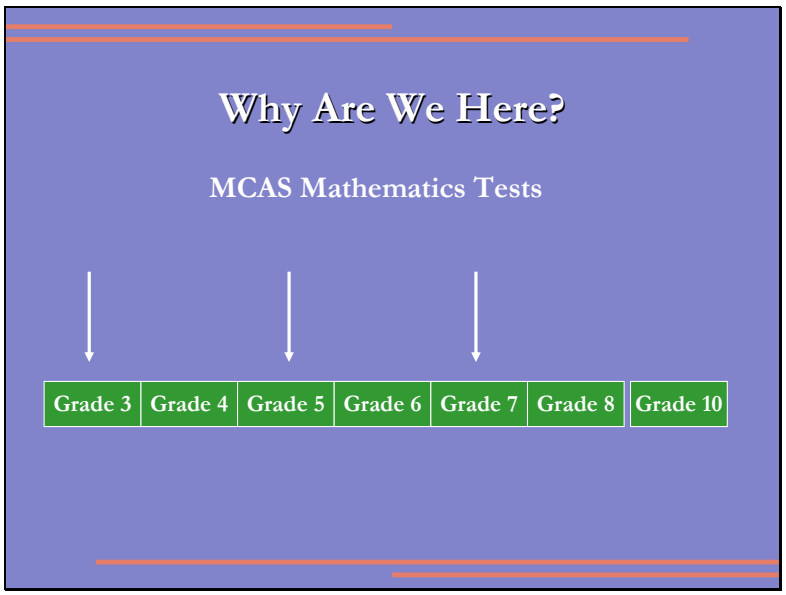
Slide 10



Slide 11



Slide 12



Slide 13

General MCAS Performance Level Descriptions

Needs Improvement

Students at this level demonstrate **partial** understanding of subject matter and solve **simple** problems

Proficient

Students at this level demonstrate a **solid** understanding of challenging subject matter and solve a **wide variety** of problems

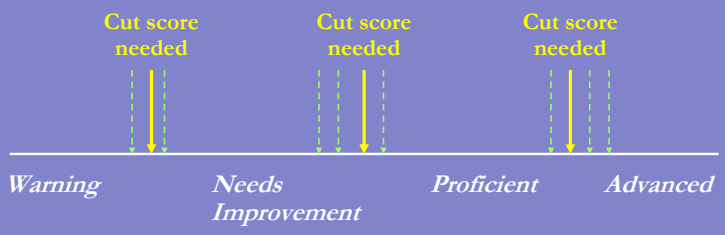
Advanced

Students at this level demonstrate a **comprehensive and in-depth** understanding of rigorous subject matter, and provide **sophisticated** solutions to **complex** problems

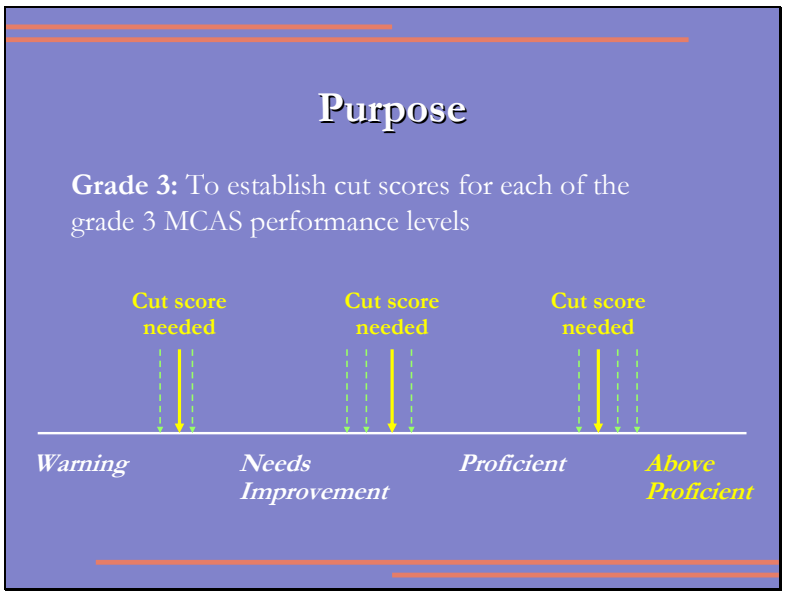
Slide 14

Purpose

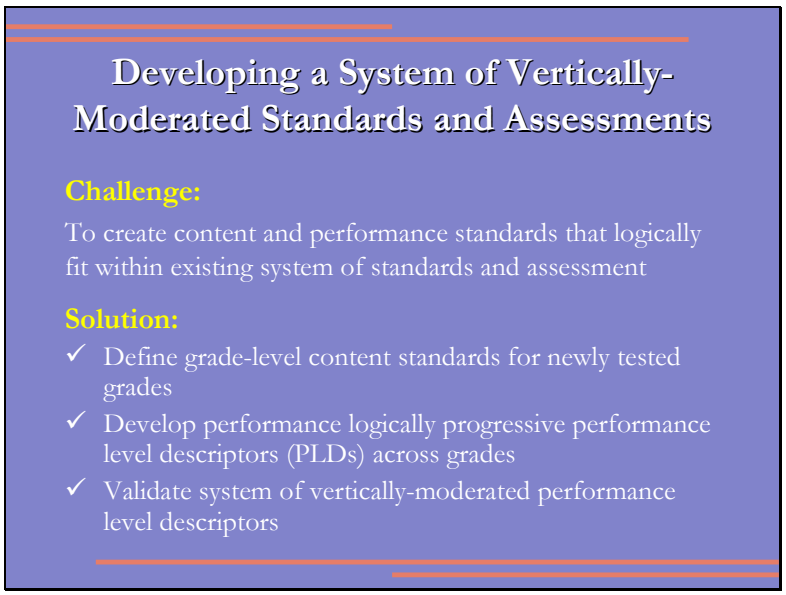
Grades 5 and 7: To establish cut scores for MCAS performance levels



Slide 15



Slide 16



Slide 17

Development of Grade-level Content Standards

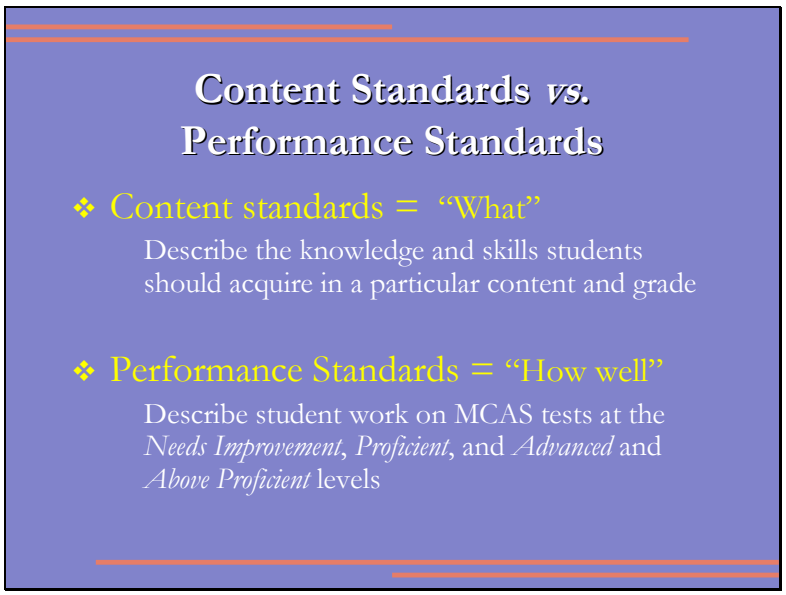
2000	Mathematics *Curriculum Framework* content standards written for grade spans (e.g., grades 5-6 and grades 7-8)
2004	Supplement to the *CF* was created, pulling out specific content standards for grades 3, 5, and 7; no "brand-new" standards were written

Slide 18

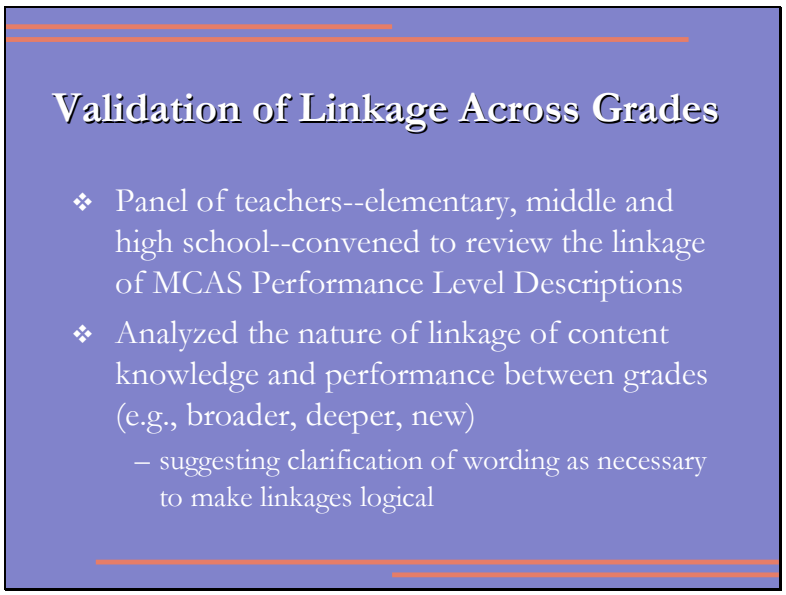
Development of Grade-Specific Performance Level Descriptions for New Assessments

- Created for new tests at grades tested (3, 5, and 7)
- New performance level—*Above Proficient*—created for grade 3
- Linked to general MCAS descriptions of student work at *Needs Improvement*, *Proficient*, and *Advanced* levels (except grade 3)
- Describe performance across levels within grade *and* link performance across grades

Slide 19



Slide 20



Slide 21

Linking Performance Standards with Student Work

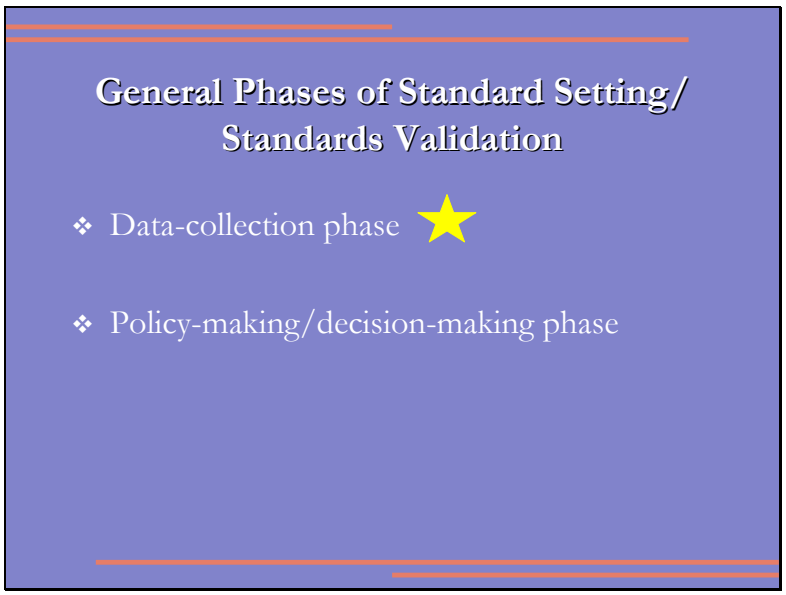
- What is standard setting?
  Establishment of cut scores to distinguish between performance levels
- What is your job?
  Use the PLDs to evaluate student work and make recommendations for where cut scores should be set

Slide 22

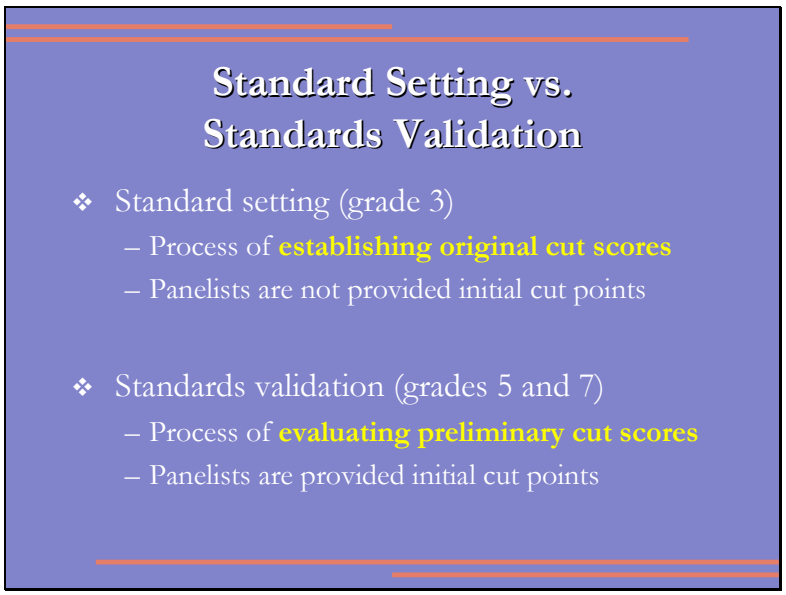
Purpose of Standard Setting

- Determine cut scores for reporting assessment results
- Answer the question:
  - How much is enough?

Slide 23

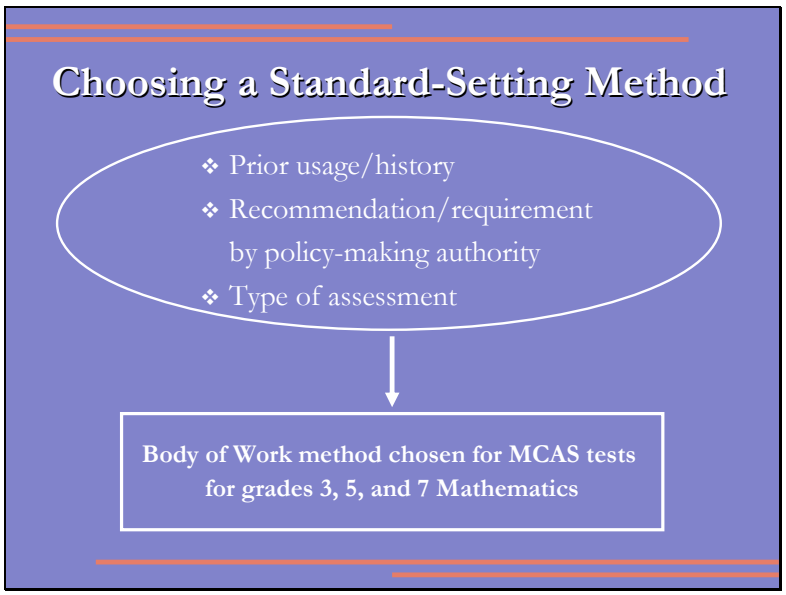


Slide 24

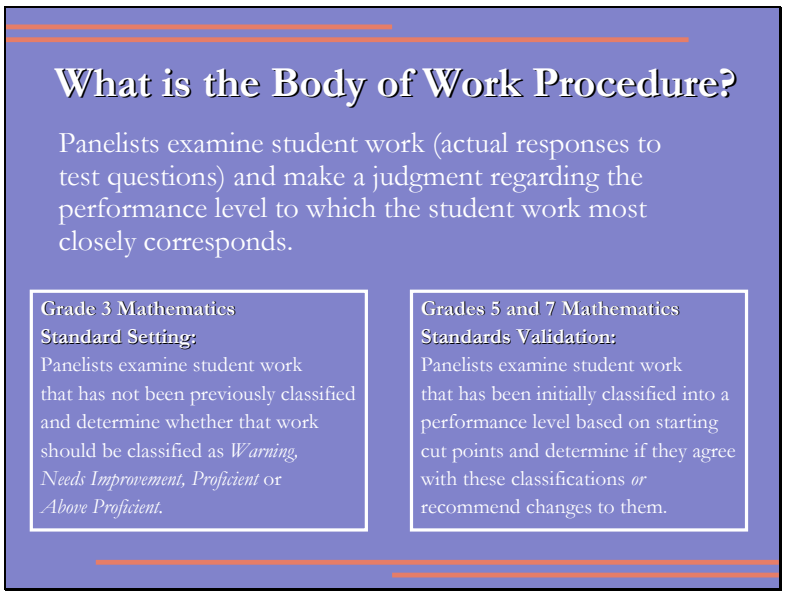


Slide 25

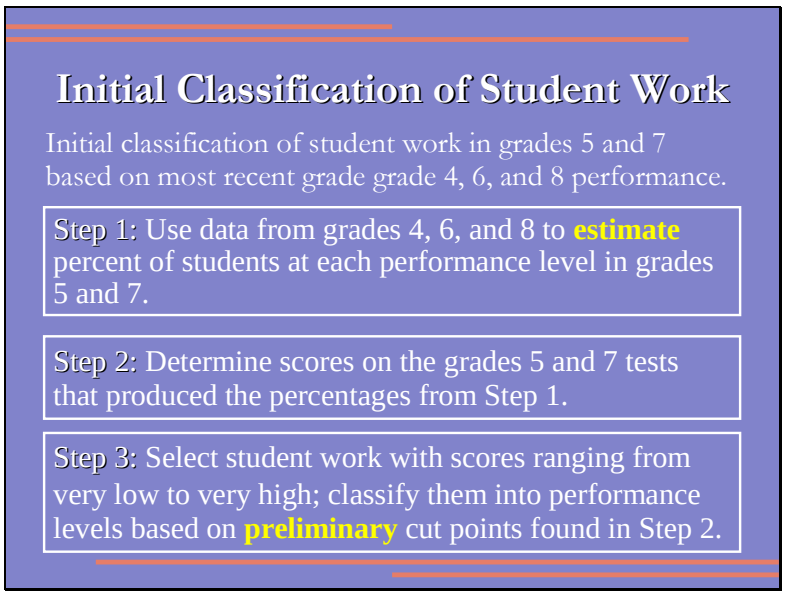
Slide 26



Slide 27



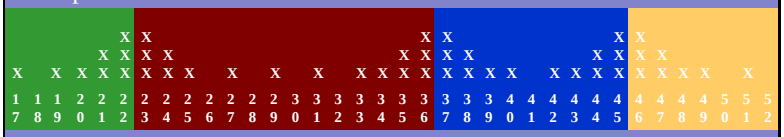
Slide 28



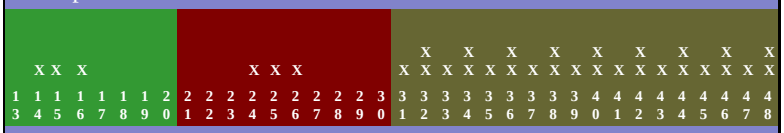
Slide 29

Selected Student Work

Example Distribution of Selected Student Work: Grades 5 and 7



Example Distribution of Selected Student Work: Grade 3



Slide 30

Initial Classification of Student Work

- Determining the initial classifications of student work was a purely statistical process
- Need to verify that the statistical results are reasonable and make sense (that's your job!)

Slide 31

How to Classify Student Work

Materials you will need:

- Performance Level Definitions
  - General
  - Grade and content specific
- Bodies of Student Work
  - Responses to constructed-response questions
  - Multiple-choice summary sheet
- Rating Form

Slide 32

How to Classify Student Work

- Examine the student's responses to multiple-choice questions
- Examine the student's responses to constructed-response questions
- Judge the student's knowledge and skills demonstrated relative to the PLDs
- Panelists do not need to reach consensus on the classifications

Slide 33

How to Classify Student Work

To help prepare you to do these ratings, you will spend time:

- Becoming familiar with the PLDs
  - It is important that all panelists have the same understanding of the PLDs.
- Becoming familiar with the bodies of student work
  - Multiple-choice items
  - Constructed-response items

Slide 34

How to Classify Student Work

- You will have the opportunity to discuss your classifications and change them if desired.
- Don't worry! We have procedures, materials, and staff to assist you in this process.

Slide 35

What Next?

Break into grade-level groups:

- Take the assessment
- Discuss the Performance Level Definitions
- Complete the Item Map
- Complete training round
- Complete individual ratings
- Discuss ratings and revise
- Receive feedback from revised ratings
- Discuss feedback and provide final ratings
- Complete an evaluation form

Slide 36

Top 8 Most Misunderstood Things about Standard Setting

8. Standard setting is a great opportunity to rewrite *Curriculum Framework* standards.
7. The process is rigged.
6. This is a good time to vent about all the things you hate about MCAS.
5. We should use this time to rework Mathematics performance level definitions.

Slide 37

Top 8 Most Misunderstood Things about Standard Setting

4. Standard setting is scoring.
3. Only Mathematics scholars should be doing this work.
2. Only teachers should be doing this work.
1. Disagreement is bad.

Slide 38

Ground Rules

- Role of facilitator is to "facilitate" and keep process on track
- Process *solely* focused on recommending performance standards (cut scores) for MCAS
- MCAS performance level definitions are integral to process but are not up for debate
- Panelists' recommendations are vital; however, final cut scores determined by the MDOE
- Each panelist must be in attendance for the duration of the process for his/her judgments to be considered
- Cell phones off, please!

Slide 39

Agenda

Tuesday, August 22

Plenary	9:00 – 10:00 am
Break	10:00 – 10:15 am
Grade-level work	10:15 am –12:00 pm
Lunch	12:00 – 12:45 pm
Grade-level work	12:45 – 4:30 pm

Wednesday, August 23

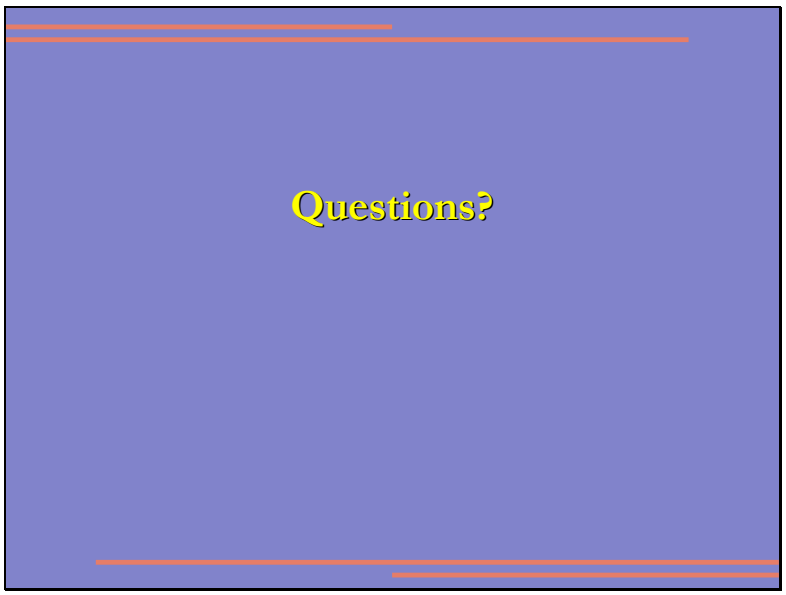
Breakfast	7:30 – 8:30 am
Grade-level work	8:30 am – 12:00 pm
Lunch	12:00 – 12:45 pm
Grade-level work	12:45 – 4:30 pm

Slide 40

Room Assignments

Grade 3 - Marblehead Room
Grade 5 - Gloucester Room
Grade 7 - Newburyport Room

Slide 41



Appendix E

Standard Setting Group Facilitator Instructions
Grade 3 – Reading & Math
Grades 5, 6, 8 – Language & Literature
Grades 5 & 7 - Math

Group Facilitator Instructions

GENERAL INSTRUCTIONS FOR MCAS STANDARD SETTING GROUP FACILITATORS

READING GRADE 3
MATH GRADE 3

Prior to Round 1 Ratings

Introductions:

1. Welcome group, introduce yourself (name, affiliation, a little selected background information).
2. Have each participant introduce him/herself.

Take the Test

Overview: In order to establish an understanding of the MCAS test items and for panelists to gain an understanding of the experience of the students who take the test, each participant will take the test for their content area. Panelists may wish to discuss or take issue with the items in the test. Tell them we will gladly take their feedback to the DOE. However, this is the actual assessment that students took and it is the set of items on which we must set standards.

Activities:

1) Introduce the MCAS test and convey/do each of the following:
    a. Tell panelists that they are about to take the actual MCAS assessment;
    b. The purpose of the exercise is to help panelists establish a good understanding of the test items and to gain an understanding of the experience of the students who take the assessment;
2) Give each panelist a test booklet;
3) Tell panelists to try to take on the perspective of a student as they complete the test.
4) When the majority of the panelists have finished, pass out the answer key and scoring rubrics for the OR items
5) Allow panelists to self-score the test
6) Once they are done scoring the test, give panelists a few minutes to discuss any questions or issues that arose as they were taking the test.

Appendix G

Fill Out Item Map

Overview: The primary purpose of filling out the item map is for panelists to think about and document the knowledge, skills, and abilities students need to answer each multiple-choice question. Panelists should have an understanding of what makes one test item harder or easier than another. The notes panelists take here will be useful in helping them in understanding performance on the MC items are related to the performance level descriptors and in discussions during the rounds of ratings.

Activities:

1. Make sure panelists have the following materials:
   a. Item map
   b. MC item summary sheet
   c. Test Booklet
   d. Reference sheet/toolkit
2. Review MC item summary sheet and item map with the panelists. Explain what each is, and point out the correspondence of the ordered items between the two. Explain that the items are organized from easiest to hardest, with the content strand for Mathematics or passage for Language & Literature indicated on the summary sheet.
3. Provide an overview of the task paraphrasing the following:
   a. The primary purpose of this activity is for panelists to think about what makes one MC question harder or easier than another. There are several factors that contribute to the how easy or how difficult a particular question is, including the concept tested and the wording of the question. For example, it may be that the concept tested is a difficult concept, or that the concept isn't difficult but that the particular wording of the question makes it a difficult question. Similarly, the concept may be a difficult one, but the wording of the question makes it easier. For Language and Literature, the difficulty of the passage also plays a part in making an item easy or difficult.
   b. Panelists should take notes about their thoughts regarding each question. These will be useful in the rating activities and later discussions.
4. Panelists should work on this as a group, discussing each item with their colleagues. They should take notes as they work through each item.
5. Panelists will begin the item mapping process with the first ordered item on the MC Summary sheet.
6. Each panelist will begin with the starting ordered item and compare it to the next ordered item. What makes the second item harder than the first? Panelists should not agonize over these decisions. It may be that the second item is only slightly harder than the first.

Discuss Performance Level Descriptors

Overview: In order to establish an understanding of the relationship between the work exhibited by a student and the performance level to which it most closely corresponds, panelists must have a clear understanding of the definition of the four performance levels. The main purpose of this activity is for the panelists to obtain a common understanding of the attributes associated with each Performance Level Descriptor.

For Reading, the panelists will review each of the performance levels, but will focus on *Proficient* and *Above Proficient* since that is the cut they will be establishing.

This activity is critical since the ratings panelists will be making in Rounds 1 and 2 will be based on these understandings.

Activities:

1) Introduce task. In this activity they will:
   a. Individually review the general and the content & grade specific Performance Level Descriptors;
   b. discuss Descriptors as a group;
   c. generate a bulleted list of the kinds of things students in each performance level can do.
2) Individually Review Performance Level Descriptors. Have panelists individually review the content & grade specific Performance Level Descriptors for all four performance levels, again for grade 3 Reading, panelists will focus on *Proficient* and *Above Proficient* since that is the cut they will be establishing. They can make notes if they like. The goal here is for the panelists to come to a common understanding of the knowledge, skills, and abilities of students in each performance level. It is not unusual for panelists to disagree with the descriptions they will see; almost certainly there will be some panelists who will want to change them. However, the task at hand is for panelists to have a common understanding of what knowledge, skills, and abilities are described by each Performance Level Descriptor.
3) Room Level Discussion of Performance Level Descriptors. After individually reviewing the Descriptors, have panelists in the entire room discuss each one, starting with *Needs Improvement*, then *Proficient* and finally *Above Proficient*. The facilitator should act as a note taker for the room to capture the points made as well as to make sure that any questions that may arise are resolved. The purpose of this is to have a collegial discussion in which to bring up/clarify any issues or questions that any individual may have regarding the descriptors and to reach a consensus on an understanding of the descriptors.
4) After each discussion, post the notes taken by the facilitator on the wall of the room.

Appendix G

Training Round

Overview of Training Round: The primary purpose of the Training Round is for panelists to become familiar with the task of classifying student work into one of the MCAS performance levels. The facilitator will briefly review the performance level descriptors and then review the 5 training BOWs with the entire room. The facilitator will point out characteristics of each BOW and lead the group through a discussion of classifying these. Panelists should note the increasing sophistication demonstrated in the student work and its correspondence to the performance level descriptors.

Throughout the remainder of this document, the body of work for a given student will be referred to as a student folder.

Activities:

1. Make sure panelists have the following materials:
    a. Training Set of student folders
        i. The training set of student folders are not initially ordered from highest to lowest, but are in random order.
    b. Performance level descriptors
2. Orient panelists to the set of student folders.
    a. Review the OR questions that students are responding to.
    b. Review the multiple-choice summary sheet at the end of each student folder, and explain that it should be carefully evaluated when judgments are being made about the student work because the majority of points come from the multiple-choice items. Point out that the items are organized by difficulty from easiest to hardest with the content strand for mathematics or the associated passage for language and literature indicated on the summary sheet. Call attention to the data associated with the multiple-choice display:
        i. The average score across the state for each strand (in math)or each passage (in language and literature)
        ii. The student's score for each strand (in math)or each passage (in language and literature)
        iii. The p-value for each item strand
        iv. The student's answer for each item
3. Give the panelists a few minutes to read through each student folder. Once they have finished their review, have panelists sort the folders from lowest to highest.
4. The facilitator leads the discussion of reviewing each student folder.
    a. Doing a tally on chart paper, indicate the agreement of the panelists in their sort order. There should be little disagreement here.
    b. Point out characteristics of the lowest scoring student folder that indicate why it is classified as *Warning*. Draw a connection between the performance level descriptor and the student work demonstrated in the folder.
    c. Work through the remaining folders with the panelists. Ask the panelists where each should be classified. Make a tally on chart paper of their initial classifications. During the discussion, draw connections between the works demonstrated in the folders and the performance level descriptors. There should be little disagreement about where the student folders get classified. For grade 3 Reading, the folders classified below Proficient are intended to show panelists the full range of student responses, although they will not be

making judgments about those. This will help panelists set realistic expectations for *Proficient* and *Above Proficient* work.

Appendix G

Round 1

Overview of Round 1: The primary purpose of Round 1 is to ask the panelists to review the student folders and place each into a performance level. The student folders will be pre-sorted from lowest score to highest score

The first step in the process will be for the panelists to individually review the entire set of student folders and make their initial judgments as to the performance level each folder belongs in. These initial judgments will be bubbled in the "Individual Rating" column of the rating sheet. Once all panelists have finished rating each student folder, the group will go back to the first folder and discuss their initial ratings. The panelists will proceed through the entire set of student folders, discussing whether the knowledge, skills and abilities demonstrated in each are consistent with the description of the performance level into which it has been initially categorized by each panelist. Once the discussion of each student folder is finished, each panelist will individually complete the "Revised Rating after Discussion" column on the rating form, indicating the performance level category into which they feel each student folder should be categorized. These ratings, in the "Revised Rating after Discussion " column, are the panelists' official Round 1 ratings and will be used for the Round 1 data analyses. Note that, though they will have discussed each student folder as a group, panelists do not need to reach consensus. We are looking for each panelist's best professional judgment.

For Reading, three student folders will be pre-classified into the *Warning* category and three will be pre-classified into the *Needs Improvement* category. These are to be used for illustrative purposes to demonstrate the full range of student responses. Panelists should not classify these student folders, but should use them as anchors for the lower categories.

Activities:

1. Make sure panelists have the following materials:
    a. Round 1 rating form
    b. Set of student folders
    c. Performance level descriptors
2. Orient panelists to the set of student folders.
    a. Tell panelists these student folders are structured the same way as the training sets.
    b. Review the OR questions that students are responding to.
    c. Review the multiple-choice display at the end of each student folder, and explain that it should be carefully evaluated when judgments are being made about the student work since the majority of points come from the multiple-choice items.
3. Orient panelists to the rating forms
    a. Have panelists write their ID number on the rating form. The ID number is on their name tags.
    b. Lead panelists through a step-by-step demonstration of how to fill in the rating form.
    c. Answer questions the panelists may have.
4. Have panelists individually review the entire set of student folders and indicate into which performance level they believe each folder should be categorized by bubbling in a rating in the "Individual Rating" section on the rating sheet. As they are reviewing the student folders, the panelists should keep in mind the Performance level descriptors. They should consider the

knowledge, skills and abilities demonstrated in each student folder and how they relate to the definitions of the performance levels.

   a. The primary purpose of this activity is for panelists to make initial categorizations of the student folders.
   b. Each panelist's judgments need to be based on his/her experience with the content of the test and understanding of the students' work. If panelists are struggling with categorizing a particular student folder, they should use their best judgment and move on. They will have an opportunity to revise their categorizations.
   c. Panelists should feel free to take notes if there are particular points about a certain student folder, and how they think it should be categorized, that they would like to make during discussions of the ratings.

5. Once the panelists have finished their individual initial categorizations, panelists will discuss these classifications as a whole group. Beginning with the first student folder, the panelists will begin discussing their initial categorization of each.
   a. Using a show of hands ask the panelists into which category they classified each paper (i.e., "How many of you classified this students' work as *Warning*?", "How many classified it as *Needs Improvement*?", and so on. Where there was disagreement, ask the panelists for the reasons they thought the student folders should be classified into the different categories.
   b. Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
   c. If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
   d. On the basis of the discussions and the feedback presented, panelists should make any necessary adjustments to their initial categorizations.
   e. Make sure panelists know that they should not feel compelled to change their initial ratings.
6. Tell panelists that they will be discussing each student folder with the other panelists in the room, but that they will be categorizing the folders individually. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating they disagree with.
7. Once discussions have been completed for a student folder, each panelist will fill out the "Revised Rating After Discussion" section of the rating form, making any changes to the categorizations they feel are appropriate. **Important: Panelists must fill out a rating for each folder for both the *Individual Rating* and *Revised Rating after Discussion.* These ratings are the only documentation of the panelists intentions; having an explicit rating removes any ambiguity from the panelists classifications.**
8. As panelists complete the task, ask them to carefully inspect their rating forms to ensure they are filled out properly.
   a. The ID number must be filled in.
   b. Each student folder must be assigned to one and only one performance level.
   c. Although the student folders are presented in order from lowest- to highest-scoring, the panelists' category assignments do not need to be in strictly increasing order.

Appendix G

Tabulation of Round 1 Results

Tabulation of Round 1 results will be completed as quickly as possible after receipt of the rating forms.

Round 2

Overview of Round 2: The primary purpose of Round 2 is to ask the panelists to discuss their Round 1 placements as a whole group and to revise their ratings on the basis of that discussion. They will discuss their ratings in the context of the ratings made by other members of the group. A graphic showing the number of panelists who assigned each student folder to each performance level category will be provided. Also, which student folders will be assigned to each level according to the group average cut points from Round 1 will be provided. Focusing on the student folders for which there appears to be the most disagreement, the panelists will discuss why they categorized each folder as they did, making sure that all different points of view are included in the discussion.

Once panelists have reviewed and discussed the Round 1 categorizations, they will be given the opportunity to change or revise their Round 1 ratings.

Activities:

1. Make sure panelists have the following materials:
    a. The Rounds 1 & 2 rating form
    b. Set of student folders
    c. Performance level descriptors
    d. Data based on the room average ratings from Round 1
2. Have panelists write their ID number and table on the rating form.
3. Provide an overview of Round 2. Paraphrase the following:
    a. As in Round 1, the primary purpose is to categorize each student folder into the performance level category where you feel it belongs.
    b. Each panelist needs to base his/her judgments on his/her experience with the content area , understanding of students' work and discussions with other panelists.
4. Review the feedback information with the panelists.
    a. Show the panelists how the student folders will be categorized based on the room average Round 1 cut point placements. Based on their Round 1 rating form, panelists will know for which student folders their categorization disagrees with that based on the full group's ratings.
5. Give panelists an opportunity to ask questions about the feedback information or about the task for Round 2.
6. Starting with the first student folder, panelists should discuss those folders for which there appears to be the most disagreement in categorization based on the Round 1 ratings.
    f. Panelists only need to discuss those student folders for which there was disagreement as to how they should be categorized.
    g. Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
    h. If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
    i. On the basis of the discussions and the feedback presented, panelists should make a second round of ratings.
    j. When making their second round of categorizations, panelists should not feel compelled to change their ratings.

k. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating they disagree with.

Encourage the panelists to use the discussion and Round 1 results to assess how stringent or lenient a judge they are. If a panelist is categorizing student folders consistently higher or lower than the group, they may have a different understanding of the performance level descriptors than the rest of the group. **It is O.K. for panelists to disagree, but that disagreement should be based on a common understanding of the Performance level descriptors.**

7. When the group has completed their second ratings, collect the rating forms. When you collect the rating forms carefully inspect them to ensure they are filled out properly.
   a. The ID and table number must be filled in.
   b. Each student folder for Round 2 must have one (and only one) rating.

Tabulation of Round 2 Results

Round 2 results will be tabulated as soon as possible upon receipt of the rating forms.

Complete Evaluation Form

Upon completion of Round 2, have panelists fill out the evaluation form. Emphasize that their honest feedback is important. Review each evaluation form for completeness, making sure that no part of the evaluation form is inadvertently skipped or missed.

GENERAL INSTRUCTIONS FOR MCAS STANDARD SETTING GROUP FACILITATORS

LANGUAGE & LITERATURE GRADES 5, 6, AND 8 MATH GRADES 5 AND 7

Prior to Round 1 Ratings

Introductions:

3. Welcome group, introduce yourself (name, affiliation, a little selected background information).
4. Have each participant introduce him/herself.

Take the Test

Overview: In order to establish an understanding of the MCAS test items and for panelists to gain an understanding of the experience of the students who take the test, each participant will take the test for their content area and grade level. Panelists may wish to discuss or take issue with the items in the test. Tell them we will gladly take their feedback to the DOE. However, this is the actual assessment that students took and it is the set of items on which we must set standards.

Activities:

7) Introduce the MCAS test and convey/do each of the following:
    a. Tell panelists that they are about to take the actual MCAS assessment;
    b. The purpose of the exercise is to help them establish a good understanding of the test items and to gain an understanding of the experience of the students who take the assessment;
8) Give each panelist a test booklet;
9) Tell panelists to try to take on the perspective of a student as they complete the test.
10) When the majority of the panelists have finished, pass out the answer key and scoring rubrics for the OR items
11) Allow panelists to self-score the test
12) Once they are done scoring the test, give panelists a few minutes to discuss any questions or issues that arose as they were taking the test.

Appendix G

Fill Out Item Map

Overview: The primary purpose of filling out the item map is for panelists to think about and document the knowledge, skills, and abilities students need to answer each multiple-choice question. Panelists should have an understanding of what makes one test item harder or easier than another. The notes panelists take here will be useful in helping them in understanding performance on the MC items are related to the performance level descriptors and in discussions during the rounds of ratings.

Activities:

7. Make sure panelists have the following materials:
   a. Item map
   b. MC item summary sheet
   c. Test booklet
   d. Reference sheet/toolkit
   e. Calculator (Grade 7 only)
8. Review the MC item summary sheet and item map with the panelists. Explain what each is, and point out the correspondence of the ordered items between the two. Explain that the items are organized from easiest to hardest, with the content strand for Mathematics or passage for Language & Literature indicated on the summary sheet.
9. Provide an overview of the task paraphrasing the following:
   a. The primary purpose of this activity is for panelists to think about what makes one MC question harder or easier than another. There are several factors that contribute to the how easy or how difficult a particular question is, including the concept tested and the wording of the question. For example, it may be that the concept tested is a difficult concept, or that the concept isn't difficult but that the particular wording of the question makes it a difficult question. Similarly, the concept may be a difficult one, but the wording of the question makes it easier. For Language and Literature, the difficulty of the passage also plays a part in making an item easy or difficult.
   b. Panelists should take notes about their thoughts regarding each question. These will be useful in the rating activities and later discussions.
10. Panelists should work on this as a group, discussing each item with their colleagues. They should take notes as they work through each item.
11. Panelists will begin the item mapping process with the first ordered item on the MC Summary sheet.
12. Each panelist will begin with the starting ordered item and compare it to the next ordered item. What makes the second item harder than the first? Panelists should not agonize over these decisions. It may be that the second item is only slightly harder than the first.

Discuss Performance Level Descriptors

Overview: In order to establish an understanding of the relationship between the work exhibited by a student and the performance level to which it most closely corresponds, panelists must have a clear understanding of the definition of the four performance levels. The main purpose of this activity is for the panelists to obtain a common understanding of the attributes associated with each Performance Level Descriptor.

This activity is critical since the ratings panelists will be making in Rounds 1 and 2 will be based on these understandings.

Activities:

5) Introduce task. In this activity they will:
   a. Individually review the general and the content & grade specific Performance Level Descriptors;
   b. discuss Descriptors as a group;
   c. generate a bulleted list of the kinds of things students in each performance level can do.
6) Individually Review Performance Level Descriptors. Have panelists individually review the content & grade specific Performance Level Descriptors all four performance levels. They can make notes if they like. The goal here is for the panelists to come to a common understanding of the knowledge, skills, and abilities of students in each performance level. It is not unusual for panelists to disagree with the descriptions they will see; almost certainly there will be some panelists who will want to change them. However, the task at hand is for panelists to have a common understanding of what knowledge, skills, and abilities are described by each Performance Level Descriptor.
7) Room Level Discussion of Performance Level Descriptors. After individually reviewing the Descriptors, have panelists in the entire room discuss each one, starting with *Needs Improvement*, then *Proficient* and finally *Advanced* (or *Above Proficient* for grade 3). The facilitator should act as a note taker for the room to capture the points made as well as to make sure that any questions that may arise are resolved. The purpose of this is to have a collegial discussion in which to bring up/clarify any issues or questions that any individual may have regarding the descriptors and to reach a consensus on an understanding of the descriptors.
8) After each discussion, post the notes taken by the facilitator on the wall of the room.

Appendix G

Training Round

Overview of Training Round: The primary purpose of the Training Round is for panelists to become familiar with the task of classifying student work into one of the MCAS performance levels. The facilitator will briefly review the performance level descriptors and then review the 5 training folders with the entire room. The facilitator will point out characteristics of the folder and lead the group through a discussion of classifying these. Panelists should note the increasing sophistication demonstrated in the student work and its correspondence to the performance level descriptors.

Throughout the remainder of this document, the body of work for a given student will be referred to as a student folder.

Activities:

5. Make sure panelists have the following materials:
    a. Training Set of student folders
        i. The training set of student folders are not initially ordered from highest to lowest, but are in random order.
    b. Performance level descriptors
6. Orient panelists to the set of student folders.
    a. Review the OR questions that students are responding to.
    b. Review the multiple-choice summary sheet at the end of each student folder, and explain that it should be carefully evaluated when judgments are being made about the student work because the majority of points come from the multiple-choice items. Point out that the items are organized by difficulty from easiest to hardest with the content strand for mathematics or the associated passage for language and literature indicated on the summary sheet. Call attention to the data associated with the multiple-choice display:
        i. The average score across the state for each strand (in math)or each passage (in language and literature)
        ii. The student's score for each strand (in math)or each passage (in language and literature)
        iii. The p-value for each item strand
        iv. The student's answer for each item
7. Give the panelists a few minutes to read through each student folder. Once they have finished their review, have panelists sort the folders from lowest to highest.
8. The facilitator leads the discussion of reviewing each student folder.
    a. Doing a tally on chart paper, indicate the agreement of the panelists in their sort order. There should be little disagreement here.
    b. Point out characteristics of the lowest scoring student folder that indicate why it is classified as *Warning*. Draw a connection between the performance level descriptor and the student work demonstrated in the folder.
    c. Work through the remaining student folders with the panelists. Ask the panelists where each should be classified. Make a tally on chart paper of their initial classifications. During the discussion, draw connections between the works demonstrated in the folders and the performance level descriptors. There should be little disagreement about where the BOWs get classified

Round 1

Overview of Round 1: The primary purpose of Round 1 is to ask the panelists to review the student folders and determine agreement with the initial categorizations and whether any student folders should be re-assigned to a different performance level category. The student folders will be pre-sorted from lowest score to highest score, and pre-classified in performance levels based on the starting cut-points.

The first step in the process will be for the panelists to individually review the entire set of folders and make their initial judgments as to whether they are placed correctly according to the starting cut-points. The panelists will indicate their initial judgments by bubbling in the performance level they believe is most appropriate in the "Individual Rating" column of the rating sheet. Once all panelists have finished reviewing the student folders, the group will go back to the first folder and discuss whether they feel its initial classification is accurate and, if not, why not. The panelists will proceed through the entire set of student folders, discussing whether the knowledge, skills and abilities demonstrated in each are consistent with the description of the performance level into which it has been initially categorized.

Once the discussion of each student folder is finished, each panelist will individually complete the "Revised Rating After Discussion" column on the rating form, indicating the performance level category they feel each folder should be categorized into. These ratings are the panelists' official Round 1 ratings and will be used for the Round 1 data analyses. Note that, though they will have discussed each student folder as a group, they do not need to reach consensus. We are looking for each panelist's best professional judgment.

Activities:

9. Make sure panelists have the following materials:
    a. Round 1 rating form, with the initial categorization of each student folder
    b. Set of student folders (60 folders of student work)
    c. Performance level descriptors
10. Orient panelists to the set of student folders.
    a. Tell panelists these student folders are structured the same way as the training sets.
    b. Review the OR questions that students are responding to.
    c. Review the multiple-choice summary sheet at the end of each student folder, and explain that it should be carefully evaluated when judgments are being made about the student work since the majority of points come from the multiple-choice items.
11. Orient panelists to the rating forms
    a. Have panelists write their ID number on the rating form. The ID number is on their name tags.
    b. Lead panelists through a step-by-step demonstration of how to fill in the rating form.
    c. Answer questions the panelists may have.
12. Have panelists individually review the entire set of student folders and indicate whether they agree or disagree with the classification by bubbling in a rating in the "Individual Rating" section on the rating sheet. As they are reviewing the student folders, the panelists should keep in mind the Performance level descriptors. They should consider the knowledge, skills and abilities demonstrated in each student folder and how they relate to the definitions of the performance levels. Panelists should focus their attention on those folders near the initial cuts.

a. The primary purpose of this activity is for panelists to identify whether they think the initial categorizations of the student folders are accurate, or if they believe some of the folders should be re-categorized. Panelists should focus their attention on folders around the initial cuts, since this is where there is likely to be the most disagreement.
b. Each panelist's judgments needs to be based on his/her experience with the content of the test and understanding of the students' work. If panelists are struggling with categorizing a particular student folder, they should use their best judgment and move on. They will have an opportunity to revise their categorizations.
c. Panelists should feel free to take notes if there are particular points about a certain folder, and how they think it should be categorized, that they would like to make during discussions of the ratings.

13. Once the panelists have finished their individual review and initial categorizations, panelists will discuss initial and individual classifications as a whole group. Beginning with the first student folder the panelists will begin discussing whether the initial categorization of each is accurate and, if not, why not.
    l. Using a show of hands ask the panelists whether they agreed with the classification of each paper. Where there was disagreement, ask the panelists for the reasons they thought the student folders should be classified into a different category. Next ask the panelists who agreed with the initial categorization why they thought that initial classification was correct.
    m. Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
    n. If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
    o. On the basis of the discussions and the feedback presented, panelists should make any necessary adjustments to their initial categorizations.
    p. Make sure panelists know that they should not feel compelled to change their initial ratings.
14. Tell panelists that they will be discussing each student folder with the other panelists in the room, but that they will be categorizing the folders individually. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating they disagree with.
15. Once discussion has been completed for a student folder, each panelist will fill out the "Revised Rating After Discussion" section of the rating form, making any changes to the categorizations they feel are appropriate. **Important: Panelists must fill out a rating for each folder for both the *Individual Rating* and *Revised Rating after Discussion.* These ratings are the only documentation of the panelists intentions; having an explicit rating removes any ambiguity from the panelists classifications.**
16. As panelists complete the task, ask them to carefully inspect their rating forms to ensure they are filled out properly.
    a. The ID number must be filled in.
    b. Each student folder must be assigned to one and only one performance level.
    c. Although the folders are presented in order from lowest- to highest-scoring, the panelists' category assignments do not need to be in strictly increasing order.

Tabulation of Round 1 Results

Tabulation of Round 1 results will be completed as quickly as possible after receipt of the rating forms.

Appendix G

Round 2

Overview of Round 2: The primary purpose of Round 2 is to ask the panelists to discuss their Round 1 placements as a whole group and to revise their ratings on the basis of that discussion. They will discuss their ratings in the context of the ratings made by other members of the group. A graphic showing the number of panelists who assigned each student folder to each performance level category will be provided. Also, which student folders will be assigned to each level according to the group average cut points from Round 1 will be provided. Focusing on the student folders that are near the cut points, the panelists will discuss why they categorized each BOW as they did, making sure that all different points of view are included in the discussion.

Once panelists have reviewed and discussed the Round 1 categorizations, they will be given the opportunity to change or revise their Round 1 ratings by bubbling in their final ratings on the final rating sheet.

Activities:

8. Make sure panelists have the following materials:
    a. The Rounds 1 & 2 rating forms
    b. Set of student folders
    c. Performance level descriptors
    d. Data based on the room average ratings from Round 1
9. Have panelists write their ID number and table on the rating form.
10. Provide an overview of Round 2. Paraphrase the following:
    a. As in Round 1, the primary purpose is to categorize each student folder into the performance level category where you feel it belongs.
    b. Each panelist needs to base his/her judgments on his/her experience with the content area , understanding of students' work and discussions with other panelists.
11. Review the feedback information with the panelists.
    a. Show the panelists how the student folders will be categorized based on the room average Round 1 cut point placements. Based on their Round 1 rating form, panelists will know for which BOWs their categorization disagrees with that based on the full group's ratings. The round 2 Rating form will also display the initial classifications of the student folders. Consequently, panelists can see whether, as a group, they differed from the initial classifications of the student folders and whether, as an individual, they differed from the either group or the initial classifications.
12. Give panelists an opportunity to ask questions about the feedback information or about the task for Round 2.
13. Beginning with the *Warning/Needs Improvement* cut, have panelists review and discuss the student folders for which there is discrepancy between the categorizations based on the room average and the initial categorizations. Panelists only need to discuss those folders for which there was disagreement as to how they should be categorized.
    a. Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
    b. If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
    c. On the basis of the discussions and the feedback presented, panelists should make their final ratings.

d. When making their final round of categorizations, panelists should not feel compelled to change their ratings.
e. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating they disagree with.

Encourage the panelists to use the discussion and feedback to assess how stringent or lenient a judge they are. If a panelist is categorizing student folders consistently higher or lower than the group, they may have a different understanding of the performance level descriptors than the rest of the group. **It is O.K. for panelists to disagree, but that disagreement should be based on a common understanding of the Performance level descriptors.**

14. Once panelists have finished discussing the discrepant student folders for the *Warning/ Needs Improvement* cut, the panelists discussion should then move onto the *Needs Improvement/Proficient* cut and finally, the *Proficient/Advanced* cut.

15. When the group has completed their final ratings, collect the rating forms. When you collect the rating forms carefully inspect them to ensure they are filled out properly.
    a. The ID and table number must be filled in.
    b. Each student folder for the final rating must have one (and only one) rating.

Tabulation of Round 2 Results

Round 2 results will be tabulated as soon as possible upon receipt of the rating forms.

Complete Evaluation Form

Upon completion of their final ratings, have panelists fill out the evaluation form. Emphasize that their honest feedback is important. Review each evaluation form for completeness, making sure that no part of the evaluation form is inadvertently skipped or missed.

Appendix F
Standard Setting Panel Applications

2006 Standard-Setting Panel Application
for Higher Education Personnel, Members of the Business Community, Members of Community Organizations, and the General Public

1. Background Information (Please Print)

Name: (First) _____**(Last)** _____

Home Address: (Street) _____

(City/Town) _____**(State)**_____(**Zip)**_____

Home Phone: _____ **Business Phone:** _____

Cell Phone: _____

Occupation: _____**Employer (*if applicable)*:** _____

Business Address: (Street) _____

(**City/Town)** _____ **(State)** _____(**Zip)** _____

Race (optional): _____

Please indicate the constituent group you will be representing (check one):

Higher Education:_____

Business: _____

Community Organization: _____

General Public: _____

2. Subject/Grade Selection

I am applying to serve on the following standard-setting panel (*please indicate 1st and 2nd choice*):

Grade 3 Reading_____ Grade 3 Mathematics _____
Grade 5 Lang. and Literature_____ Grade 5 Mathematics _____
Grade 6 Lang. and Literature_____ Grade 7 Mathematics _____
Grade 8 Lang. and Literature_____

3. Required Attachments

Please include the following as part of your application:

Your current résumé

A brief statement of interest

Your completed application must be received by 5:00 p.m., **Friday, May 19,** at the following address:

MCAS Standard-Setting Applications
Student Assessment
Massachusetts Department of Education
350 Main Street
Malden, MA 02148
Fax: (781) 338-3630

2006 Standard-Setting Panel Application for Educators

1. **Background Information (Please Print)**

Name: (First) _____ (Last) _____

Current Position: _____

Home Address: (Street) _____

City/Town:**_____ State:_____ Zip:_____**

Home Phone: _____ E-mail Address: _____

Cell Phone: _____

School Address: (Street) _____

City/Town: _____

School Phone: _____ School Fax: _____

Race (optional): _____

If currently a teacher, please indicate years of teaching experience: _____

Do you have experience teaching special education students? (*check one*)

Yes _____ No _____

If yes, please explain: _____

_____

Do you have experience teaching limited English proficient students? (*check one*)

Yes _____ No _____

If yes, please explain: _____

2. Subject/Grade Selection

I am applying to serve on the following standard-setting panel
(*please indicate 1st and 2nd choice*):

Grade 3 Reading _____	Grade 3 Mathematics_____
Grade 5 Lang. and Literature_____	Grade 5 Mathematics_____
Grade 6 Lang. and Literature _____	Grade 7 Mathematics _____
Grade 8 Lang. and Literature _____	

Principal or Superintendent Recommendation

I support the application of _____ to serve on the

_____(subject *and grade*) standard-setting panel.

Name: _____

Title: _____

Signature: _____

3. Required Attachments

Please include the following as part of your application:

Your current résumé

A brief statement of interest

Your completed application must be received by 5:00 p.m., **Friday, May 19,** at the following address:

MCAS Standard-Setting Panel Applications
Student Assessment
Massachusetts Department of Education
350 Main Street
Malden, MA 02148
Fax: 781/338-3630

Appendix G

Standard Setting Panel of Participants

Appendix G

2006 Standard-Setting Panel of Participants

Grade 3 Reading

First Name	Last Name	Subject/Grade Selection	Position	Affiliation
Karen	Alexopoulos	Gr 3 Reading	Teacher	Somerville Public Schools
Gayla	Berry	Gr 3 Reading	ELL Teacher	Holyoke Public Schools
Marie	Champion	Gr 3 Reading	Title 1 Teacher - ELA	Medford Public Schools
Meghan	Coleman	Gr 3 Reading	Grade 3 Teacher	Cohasset Public Schools
Meghan	Connolly	Gr 3 Reading		
Madelyn	Farrell	Gr 3 Reading	Grade 3 Teacher	Woburn Public Schools
Nancy	Fogg	Gr 3 Reading	Grade 4 Teacher	Brockton Public Schools
Robin	Gazelian	Gr 3 Reading	Paraprofessional	Methuen Public Schools
Andrea	Hallion	Gr 3 Reading	Adjunct Professor	Framingham State College
Kathleen	Jankins	Gr 3 Reading	Reading Specialist	Bridgewater Raynham Regional
Sheila	Kukstis	Gr 3 Reading	Principal	Taunton Public Schools
Kerri	Laurenzo	Gr 3 Reading	Reading Specialist	East Longmeadow Public Schools
Margaret	Martinez	Gr 3 Reading	Director of Curriculum & Inspection	Berkley Public Schools
Elaine	McNamara	Gr 3 Reading	Title I Director/Title 1 Reading Teacher	Dracut Public Schools
Linda	Mros	Gr 3 Reading	Grade 3 Teacher	Taunton Public Schools
Kevinetta	O'Brien	Gr 3 Reading	Retired Elementary Teacher	Charlton District
Gina	Patti	Gr 3 Reading	Reading Specialist	Oak Bluffs Public Schools
Linda	Volpicelli	Gr 3 Reading	Language Arts Coordinator K-5, Reading Specialist 3-5	Bedford Public Schools

Joyce	Welch	Gr 3 Reading	Special Education Teacher	Springfield Public Schools
Pamela	Westmoreland	Gr 3 Reading	Representative	National Association for the Education of Young Children

Appendix G

Grade 5 ELA

First Name	Last Name	Subject/Grade Selection	Position	Affiliation
Alyssa	Kaplan	Gr 5 Language & Literature	Literacy Specialist	Sudbury Public Schools
Paul	McKenna	Gr 5 Language & Literature	Grade 5 Teacher	Granville Public Schools
Anne	Pramas	Gr 5 Language & Literature	Teaching Principal	Dracut Public Schools
Karen	Swan	Gr 5 Language & Literature	Consulting Teacher of Reading	Plymouth Public Schools
Matthew	D'Andrea	Gr 5 Language & Literature	Grade 5 Teacher	Wareham Public Schools
Rev. Gregory	Groover	Gr 5 Language & Literature	Chairman, Education Committee	Black Ministerial Alliance
Germaine	McManus	Gr 5 Language & Literature	DID ATTEND	
Toni	Kanes	Gr 5 Language & Literature	Grade 5 Teacher	Braintree Public Schools
Joyce	Koss-McGregor	Gr 5 Language & Literature	Grade 5 Teacher	Auburn Public Schools
Mary-Anne	Ryan	Gr 5 Language & Literature	Grade 5 Teacher	Brewster Public Schools
Deborah	Trapp	Gr 5 Language & Literature	Literacy Specialist	Framingham Public Schools
Ricki	Goldberg	Gr 5 Language & Literature	Reading Teacher/Title I	Methuen Public Schools
Ann	Dudley	Gr 5 Language & Literature	Grade 5 Teacher, ELA and Science	Greenfield Public Schools
Carol	Jacobson	Gr 5 Language & Literature	Reading Specialist	Burlington Public Schools
Richard	Bettano	Gr 5 Language & Literature	Grade 5 Teacher	Townsend Public Schools
Mary Beth	Witkavitch	Gr 5 Language & Literature	DID ATTEND	
Heather	Wassall	Gr 5 Language & Literature	Grade 5 Teacher	Mystic Valley Regional Charter School
Wendy	Darcy	Gr 5 Language & Literature	Literacy Coach & Reading Specialist	Wareham Public Schools
Janice	Gauthier	Gr 5 Language & Literature	Director of Curriculum	Everett Public Schools

Grade 6 ELA

First Name	Last Name	Subject/Grade Selection	Position	Affiliation
Julie	Gorman-Porter	Gr 6 Language & Literature	Grade 6 Language Arts, Grade 6 Literacy Teacher and L.A. Curriculum Resource Teacher	Framingham Public Schools
Jeanne	Goranson	Gr 6 Language & Literature	Reading Specialist	Lincoln Public Schools
Katie	Bourne	Gr 6 Language & Literature	Grade 6 Language Arts Teacher	Cambridge Public Schools
Jacqueline	Haley	Gr 6 Language & Literature	Grade 6 Language Arts Teacher	Palmer Public Schools
Karen	Havener Ruiz	Gr 6 Language & Literature	Grade 6,7,8 Language Arts Teacher	Wachusett Regional School District
Lyudmila	Moiseyeva	Gr 6 Language & Literature	ELL Teacher	Brookline Public Schools
Brenda	Steeves	Gr 6 Language & Literature	Grade 6 English Language Arts Teacher	Whitinsville Public Schools
Cindy	Olson	Gr 6 Language & Literature	Reading Specialist (Grade 4-6)	Easton Public Schools
Rae-Ann	Trifilo	Gr 6 Language & Literature	Title I Director/Teacher	Narragansett Regional School District
Alicia	O'Brien	Gr 6 Language & Literature	Grade 6 Teacher - Social Studies and Reading	Taunton Public Schools
Holly	Goodrich	Gr 6 Language & Literature	Grade 6 English Teacher	Pittsfield Public Schools
Jean	Silva	Gr 6 Language & Literature	Grade 6 Reading Teacher	Canton Public Schools
Amy	Fitzgerald	Gr 6 Language & Literature	ELA Teacher	Holyoke Public Schools

Nancy	Meagher	Gr 6 Language & Literature	Grade 6 Reading Specialist	Barnstable Horace Mann Charter School
Patricia	Desmond	Gr 6 Language & Literature	Director, English/Lang Arts, K-12	Medford Public Schools
Natalie	Toporowski	Gr 6 Language & Literature	Grade 6 English Language Arts Teacher	West Springfield Public Schools
Suzanne	Dunn	Gr 6 Language & Literature	Grade 6 Teacher-all subjects	Hopedale Public Schools
Charlene	Bobek	Gr 6 Language & Literature	Guidance Counselor	Lawrence Public Schools
Jessica	Greenfield	Gr 6 Language & Literature	Middle School ELA Teacher (grades 6-8)	Framingham Public Schools
G. Sanford	Bogage	Gr 6 Language & Literature	Grade 6 Educator	Wellesley Public Schools
Ann	Galvani	Gr 6 Language & Literature	Grade 7 ELA Teacher	Natick Public Schools

Grade 8 ELA

First Name	Last Name	Subject/Grade Selection	Position	Affiliation
Meghan	Harrison	Gr 8 Language & Literature	Grade 8 ELL Support	Holyoke Public Schools
Kerry	Winer	Gr 8 Language & Literature	Grade 8 Teacher & ELA Curriculum Specialist	Framingham Public Schools
Karen	Dorgan	Gr 8 Language & Literature	Assistant Principal	Fairhaven Public Schools
Kathleen	Moore	Gr 8 Language & Literature	Grade 8 ELA Teacher/Curriculum Leader	Carver Public Schools
Lori	DiGisi	Gr 8 Language & Literature	Grade 8 Literacy Specialist	Framingham Public Schools
Alan	Keller	Gr 8 Language & Literature	Grade 7&8 Teacher	Foxborough Public Schools
Sarah	Redman	Gr 8 Language & Literature	Grade 7 ELA Teacher/Curriculum Coordinator	Tewksbury Public Schools
George	Biggs	Gr 8 Language & Literature	Grade 7 Language Arts/Team Leader	Hampshire Public Schools
Catherine	Symonds	Gr 8 Language & Literature	English Curriculum Team Leader	Wilmington Public Schools
Janet	Norris	Gr 8 Language & Literature	Special Education Teacher	Hamilton-Wenham Regional School District
Evelyn	Ford-Connors	Gr 8 Language & Literature	Instructor	Boston University School of Education
Deborah	MacDonald	Gr 8 Language & Literature	Associate Principal	Haverhill Public Schools

Cynthia	Winfield	Gr 8 Language & Literature	Grade 8 Teacher	Arlington Public Schools
Jeanne	Noyes	Gr 8 Language & Literature	English Teacher	Canton Public Schools
Charles	Baker	Gr 8 Language & Literature	Education Consultant	Calliope Magazine
Monique	Greilich	Gr 8 Language & Literature	Special Education Teacher	Salem Public Schools
Pamela	Cangemi	Gr 8 Language & Literature	High School English Teacher	Longmeadow Public Schools
James	Keefe	Gr 8 Language & Literature	Department Head	Lynn Public Schools
Sherri	Travers	Gr 8 Language & Literature	Grade 8 ELA Teacher, ELA Department Head, and Co-Director of the Central MA Writing Proj.	Whitinsville Public Schools
Laura	Miceli	Gr 8 Language & Literature	Director of Curriculum, Instruction and Personal	Hanover Public Schools

2006 Standard-Setting Panel of Participants

Grade 3 Math

First Name	Last Name	Subject/Grade Selection .1	Position	Affiliation
John	Cardoza	Gr 3 Mathematics	Grade 4 Inclusion Teacher	Middleborough Public Schools
Adriana	Gallo-Grimaldi	Gr 3 Mathematics	'Math Educator	Agawam Public Schools
Linda	Gauthier	Gr 3 Mathematics	'Curriculum Coordinator K - 5	Saugus Public Schools
Cheryl	Goguen	Gr 3 Mathematics	4th Grade General Educator	Framingham Public Schools
Rev. Ray	Hammond	Gr. 3 Mathematics	Chairman	Boston Ten Point Coalition
Kristine	Klumpp	Gr 3 Mathematics	Grade 3 Teacher	Duxbury Public Schools
Carol	LaPolice	Gr 3 Mathematics	Mathematics Collaborative Professional Development Teacher	Springfield Public Schools
Susan	Mello	Gr 3 Mathematics	Elementary Mathematics Coach	Fall River Public Schools
Kathleen	Millett	Gr 3 Mathematics	Elementary Math Facilitator	West Springfield Public Schools
Lyudmila	Moiseyeva	Gr 3 Mathematics	ELL Teacher	Brookline Public Schools
Judy	Moore	Gr 3 Mathematics	Grade 3 Teacher	Harvard Public Schools
Stephanie	Morris	Gr 3 Mathematics	Grade 4 Teacher	Central Berkshire Regional School District
Stephanie	Murchison-Brown	Gr 3 Mathematics	Mathematics Teacher	Holyoke Public Schools
Arthur	Norman	Gr 3 Mathematics	Assistant Principal	Fitchburg Public Schools
Misael	Ramos	Gr 3 Mathematics	Collaborative Professional Development Teacher	Springfield Public Schools
Jen	Rubera	Gr 3 Mathematics	Grade 4 Teacher	Haverhill Public Schools
Victoria	Sapko	Gr 3 Mathematics	Assistant Professor	Framingham Public Schools
Michael	Stanton	Gr 3 Mathematics	Principal	Walpole Public Schools

Appendix G

Elizabeth	Sweeney	Gr 3 Mathematics	Assistant Program Director	Boston Public Schools
Denise	Young	Gr 3 Mathematics	Grade 3 Teacher	Natick Public Schools

Grade 5 Math

First Name	Last Name	Subject/Grade Selection .1	Position	Affiliation
Deborah	Allard	Gr 5 Mathematics	Grade 5 Math/Soc. Studies Teacher	Greenfield Public Schools
Geraldine	Carter	Gr 5 Mathematics	Grade 5 Teacher	Pentucket Public Schools
Dianne	Connolly	Gr 5 Mathematics	Math Specialist	Haverhill Public Schools
Maureen	Coughlin	Gr 5 Mathematics	Grade 5 Math/Science/Reading Teacher	Taunton Public Schools
Valerie	Daniel	Gr 5 Mathematics	Math Coach/Learning Site Coordinator	Boston Public Schools
Charlene	D'Onofrio	Gr 5 Mathematics	Grade 5 Teacher	Franklin Public Schools
Jacqueline	Figueiredo	Gr 5 Mathematics	Elementary School Principal	Dartmouth Public Schools
Suzanne	Hickey	Gr 5 Mathematics	Math Coach Grades 5-8	Lawrence Public Schools
Andrea	Hume	Gr 5 Mathematics	Special Education Teacher	Dedham Public Schools
Ann Marie	Laduzenski	Gr 5 Mathematics	Grade 6 Math Teacher & Dept. Chair	Springfield Public Schools.
Brian	Ledbetter	Gr 5 Mathematics	Grade 5 Math/Science Teacher	Sturbridge Public Schools
Julie	Levandosky	Gr 5 Mathematics	Assistant Math Professor	Framingham State College
Rosemary	Macek	Gr 5 Mathematics	Mathematics Coach Grades K-6	Fall River Public Schools
Bruce	Michitson	Gr 5 Mathematics	Grade 5 Teacher	Haverhill Public Schools
Lisa	Mikus	Gr 5 Mathematics	Grade 4 Teacher	Newton Public Schools
Christine	Panarese	Gr 5 Mathematics	Grade 6 Math Instructor	Wareham Public Schools
Glen	Panciocco	Gr 5 Mathematics	Grade 5 Spanish Immersion Teacher	Millis Public Schools

Appendix G

Nancy	Pasquaretta	Gr 5 Mathematics	Grade 5 Teacher	Ayer Public Schools
Laura	Raposa	Gr 5 Mathematics	Grade 5 Teacher	Littleton Public Schools
Virginia	Willims	Gr 5 Mathematics	Chief Executive Officer	Urban League of Eastern Massachusetts
Virginia	Young	Gr 5 Mathematics	Math/Science Curriculum Coord. PreK-8	Taunton Public Schools

Grade 7 Math

First Name	Last Name	Subject/Grade Selection .1	Position	Affiliation
Linda	Coffey	Gr 7 Mathematics	Middle School Math Curriculum Coordinator/Grade 7 Math Teacher	Westwood Public Schools
Lois	Cole	Gr 7 Mathematics	Grade 6 & 8 Math Teacher	Lynn Public Schools
Dianne	Costello	Gr 7 Mathematics	Special Education Coordinator Grade 7 Math Teacher	Lexington Public Schools
Harold	Dickert	Gr 7 Mathematics	Grade 7 Math Teacher/Secondary Math Specialist	Hopkinton Public Schools
Nancy	Farrell	Gr 7 Mathematics	Grade 7 Math Teacher	Agawam Public Schools
Kathy	Favazza	Gr 7 Mathematics		Reading Public Schools
Paula	Fay	Gr 7 Mathematics	Middle School Math Teacher	Barnstable Public Schools
Noreen	Flanagan	Gr 7 Mathematics	Mathematics Teacher	Haverhill Public Schools
Joseph	Gillis, Jr.	Gr 7 Mathematics	Director of Information Technology	Plymouth Sheriff Department
Erin	Houghton	Gr 7 Mathematics	Grade 8 SPED Teacher	Woburn Public Schools
Sylvia	Leonard	Gr 7 Mathematics	Math Teacher	Wareham Public Schools
James	Liptak	Gr 7 Mathematics	Grade 7 Math Teacher, Co-Curriculum Coordinator for Math Department	Hampshire Regional District
Alan	MacDonald	Gr. 7 Mathematics	Executive Director	MA Defense Technology Initiative
Katherine	Madden	Gr 7 Mathematics	Math Support Specialist	Springfield Public Schools
Maura	Mast	Gr 7 Mathematics	Professor of Mathematics	Umass Boston

Appendix G

Andrew	Perry	Gr 7 Mathematics	Professor of Mathematics	Springfield College
Sherry	Sajdak	Gr 7 Mathematics	Program Director, Elementary Mathematics	Boston Public Schools
Susan	Skaff	Gr 7 Mathematics	Math/Science Coach	Lawrence Public Schools
Barbara	Swidler	Gr 7 Mathematics	Curriculum Coordinator (5-8 Math)	Shrewsbury Public Schools
Ryan	Toal	Gr 7 Mathematics	Math Department Head	Boston Public Schools
Stephanie	Wooley	Gr 7 Mathematics	Grade 7 Math Teacher	Cohasset Public Schools

Appendix H

Standard Setting Sample Rating Forms

Appendix G

Sample Rating Forms

ELA Grade 3 **Rater ID: _____**

BoW No.	ID	Individual Rating				Revised Rating After Discussion			
		Warning	Needs Improvement	Proficient	Above Proficient	Warning	Needs Improvement	Proficient	Above Proficient
1	1303012853	●	(N)			●	(N)		
2	1304012015	●	(N)			●	(N)		
3	1316011205	●	(N)			●	(N)		
4	1323012622	(W)	●			(W)	●		
5	1301010855	(W)	●			(W)	●		
6	1302010362	(W)	●			(W)	●		
7	1301010579			(P)	(A)			(P)	(A)
8	1322011994			(P)	(A)			(P)	(A)
9	1304010972			(P)	(A)			(P)	(A)
10	1315010699			(P)	(A)			(P)	(A)
11	1301010639			(P)	(A)			(P)	(A)
12	1319012358			(P)	(A)			(P)	(A)
13	1301012332			(P)	(A)			(P)	(A)
14	1311014211			(P)	(A)			(P)	(A)
15	1301011040			(P)	(A)			(P)	(A)
16	1317011954			(P)	(A)			(P)	(A)
17	1301011307			(P)	(A)			(P)	(A)
18	1316011820			(P)	(A)			(P)	(A)
19	1310014145			(P)	(A)			(P)	(A)
20	1321011343			(P)	(A)			(P)	(A)
21	1301011136			(P)	(A)			(P)	(A)
22	1310012566			(P)	(A)			(P)	(A)
23	1306010694			(P)	(A)			(P)	(A)
24	1321011659			(P)	(A)			(P)	(A)
25	1301010604			(P)	(A)			(P)	(A)
26	1301012443			(P)	(A)			(P)	(A)
27	1301011487			(P)	(A)			(P)	(A)
28	1322011091			(P)	(A)			(P)	(A)
29	1314011323			(P)	(A)			(P)	(A)
30	1318013346			(P)	(A)			(P)	(A)

ELA Grade 3 **Rater ID: _____**

BoW No.	ID	Final Rating			
		Warning	Needs Improvement	Proficient	Above Proficient
1	1303012853	●	Ⓝ		
2	1304012015	●	Ⓝ		
3	1316011205	●	Ⓝ		
4	1323012622	Ⓦ	●		
5	1301010855	Ⓦ	●		
6	1302010362	Ⓦ	●		
7	1301010579			Ⓟ	Ⓐ
8	1322011994			Ⓟ	Ⓐ
9	1304010972			Ⓟ	Ⓐ
10	1315010699			Ⓟ	Ⓐ
11	1301010639			Ⓟ	Ⓐ
12	1319012358			Ⓟ	Ⓐ
13	1301012332			Ⓟ	Ⓐ
14	1311014211			Ⓟ	Ⓐ
15	1301011040			Ⓟ	Ⓐ
16	1317011954			Ⓟ	Ⓐ
17	1301011307			Ⓟ	Ⓐ
18	1316011820			Ⓟ	Ⓐ
19	1310014145			Ⓟ	Ⓐ
20	1321011343			Ⓟ	Ⓐ
21	1301011136			Ⓟ	Ⓐ
22	1310012566			Ⓟ	Ⓐ
23	1306010694			Ⓟ	Ⓐ
24	1321011659			Ⓟ	Ⓐ
25	1301010604			Ⓟ	Ⓐ
26	1301012443			Ⓟ	Ⓐ
27	1301011487			Ⓟ	Ⓐ

Appendix G

ELA Grade 3 **Rater ID: _____**

BoW No.	ID	Final Rating			
		Warning	Needs Improvement	Proficient	Above Proficient
28	1322011091			Ⓟ	Ⓐ
29	1314011323			Ⓟ	Ⓐ
30	1318013346			Ⓟ	Ⓐ

MATH Grade 3 **Rater ID: _____**

BoW No.	ID	Individual Rating				Revised Rating After Discussion			
		Warning	Needs Improvement	Proficient	Above Proficient	Warning	Needs Improvement	Proficient	Above Proficient
1	1309332714	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
2	1313335447	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
3	1301330661	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
4	1302331681	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
5	1302333503	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
6	1307335287	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
7	1308331086	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
8	1315333350	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
9	1305334178	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
10	1306330318	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
11	1312330918	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
12	1306332747	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
13	1307334992	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
14	1309330071	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
15	1306331886	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
16	1311331848	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
17	1312330754	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
18	1305331494	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
19	1310334315	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
20	1311334231	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
21	1311331225	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
22	1311331237	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
23	1312335527	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
24	1303331920	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
25	1310334577	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
26	1311335122	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
27	1301332797	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ

Appendix G

MATH Grade 3 **Rater ID: _____**

BoW No.	ID	Individual Rating				Revised Rating After Discussion			
		Warning	Needs Improvement	Proficient	Above Proficient	Warning	Needs Improvement	Proficient	Above Proficient
28	1304335094	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
29	1309334109	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
30	1301331124	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
31	1303333446	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
32	1314331371	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
33	1304335789	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
34	1308333572	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
35	1310333834	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
36	1310333577	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
37	1313332442	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
38	1313333109	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
39	1309333650	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
40	1312336301	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
41	1314334256	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
42	1304335747	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
43	1305334198	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
44	1306333660	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
45	1303335686	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
46	1306336956	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
47	1311334968	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
48	1306337364	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
49	1312336583	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
50	1315333813	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ

MATH Grade 3 **Rater ID: _____**

BoW No.	ID	Final Rating			
		Warning	Needs Improvement	Proficient	Above Proficient
1	1309332714	Ⓦ	Ⓝ	Ⓟ	Ⓐ
2	1313335447	Ⓦ	Ⓝ	Ⓟ	Ⓐ
3	1301330661	Ⓦ	Ⓝ	Ⓟ	Ⓐ
4	1302331681	Ⓦ	Ⓝ	Ⓟ	Ⓐ
5	1302333503	Ⓦ	Ⓝ	Ⓟ	Ⓐ
6	1307335287	Ⓦ	Ⓝ	Ⓟ	Ⓐ
7	1308331086	Ⓦ	Ⓝ	Ⓟ	Ⓐ
8	1315333350	Ⓦ	Ⓝ	Ⓟ	Ⓐ
9	1305334178	Ⓦ	Ⓝ	Ⓟ	Ⓐ
10	1306330318	Ⓦ	Ⓝ	Ⓟ	Ⓐ
11	1312330918	Ⓦ	Ⓝ	Ⓟ	Ⓐ
12	1306332747	Ⓦ	Ⓝ	Ⓟ	Ⓐ
13	1307334992	Ⓦ	Ⓝ	Ⓟ	Ⓐ
14	1309330071	Ⓦ	Ⓝ	Ⓟ	Ⓐ
15	1306331886	Ⓦ	Ⓝ	Ⓟ	Ⓐ
16	1311331848	Ⓦ	Ⓝ	Ⓟ	Ⓐ
17	1312330754	Ⓦ	Ⓝ	Ⓟ	Ⓐ
18	1305331494	Ⓦ	Ⓝ	Ⓟ	Ⓐ
19	1310334315	Ⓦ	Ⓝ	Ⓟ	Ⓐ
20	1311334231	Ⓦ	Ⓝ	Ⓟ	Ⓐ
21	1311331225	Ⓦ	Ⓝ	Ⓟ	Ⓐ
22	1311331237	Ⓦ	Ⓝ	Ⓟ	Ⓐ
23	1312335527	Ⓦ	Ⓝ	Ⓟ	Ⓐ
24	1303331920	Ⓦ	Ⓝ	Ⓟ	Ⓐ
25	1310334577	Ⓦ	Ⓝ	Ⓟ	Ⓐ
26	1311335122	Ⓦ	Ⓝ	Ⓟ	Ⓐ
27	1301332797	Ⓦ	Ⓝ	Ⓟ	Ⓐ

Appendix G

MATH Grade 3 **Rater ID: _____**

BoW No.	ID	Final Rating			
		Warning	Needs Improvement	Proficient	Above Proficient
28	1304335094	Ⓦ	Ⓝ	Ⓟ	Ⓐ
29	1309334109	Ⓦ	Ⓝ	Ⓟ	Ⓐ
30	1301331124	Ⓦ	Ⓝ	Ⓟ	Ⓐ
31	1303333446	Ⓦ	Ⓝ	Ⓟ	Ⓐ
32	1314331371	Ⓦ	Ⓝ	Ⓟ	Ⓐ
33	1304335789	Ⓦ	Ⓝ	Ⓟ	Ⓐ
34	1308333572	Ⓦ	Ⓝ	Ⓟ	Ⓐ
35	1310333834	Ⓦ	Ⓝ	Ⓟ	Ⓐ
36	1310333577	Ⓦ	Ⓝ	Ⓟ	Ⓐ
37	1313332442	Ⓦ	Ⓝ	Ⓟ	Ⓐ
38	1313333109	Ⓦ	Ⓝ	Ⓟ	Ⓐ
39	1309333650	Ⓦ	Ⓝ	Ⓟ	Ⓐ
40	1312336301	Ⓦ	Ⓝ	Ⓟ	Ⓐ
41	1314334256	Ⓦ	Ⓝ	Ⓟ	Ⓐ
42	1304335747	Ⓦ	Ⓝ	Ⓟ	Ⓐ
43	1305334198	Ⓦ	Ⓝ	Ⓟ	Ⓐ
44	1306333660	Ⓦ	Ⓝ	Ⓟ	Ⓐ
45	1303335686	Ⓦ	Ⓝ	Ⓟ	Ⓐ
46	1306336956	Ⓦ	Ⓝ	Ⓟ	Ⓐ
47	1311334968	Ⓦ	Ⓝ	Ⓟ	Ⓐ
48	1306337364	Ⓦ	Ⓝ	Ⓟ	Ⓐ
49	1312336583	Ⓦ	Ⓝ	Ⓟ	Ⓐ
50	1315333813	Ⓦ	Ⓝ	Ⓟ	Ⓐ

ELA Grade 5 **Rater ID: _____**

BoW No.	ID	Individual Rating				Revised Rating After Discussion			
		Warning	Needs Improvement	Proficient	Advanced	Warning	Needs Improvement	Proficient	Advanced
1	1500067689	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
2	1500052817	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
3	1500039886	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
4	1500078265	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
5	1500066844	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
6	1500069724	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
7	1500053059	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
8	1500060422	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
9	1500032000	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
10	1500051505	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
11	1500070324	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
12	1500092288	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
13	1500008142	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
14	1500015769	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
15	1500037307	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
16	1500066110	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
17	1500081005	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
18	1500092488	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
19	1500006118	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
20	1500073254	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
21	1500012559	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
22	1500043737	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
23	1500002637	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
24	1500051824	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
25	1500088368	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
26	1500013179	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ

Appendix G

ELA Grade 5 **Rater ID: _____**

BoW No.	ID	Individual Rating				Revised Rating After Discussion			
		Warning	Needs Improvement	Proficient	Advanced	Warning	Needs Improvement	Proficient	Advanced
27	1500029985	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
28	1500048729	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
29	1500074822	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
30	1500028020	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
31	1500060928	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
32	1500067614	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
33	1500078731	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
34	1500047920	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
35	1500064760	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
36	1500065063	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
37	1500064021	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
38	1500075700	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
39	1500017931	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
40	1500038447	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
41	1500042650	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
42	1500082704	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
43	1500039156	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
44	1500045213	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
45	1500058486	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
46	1500069503	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
47	1500002119	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
48	1500079212	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
49	1500020510	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ
50	1500072480	Ⓦ	Ⓝ	Ⓟ	Ⓐ	Ⓦ	Ⓝ	Ⓟ	Ⓐ

ELA Grade 5 **Rater ID: _____**

BoW No.	ID	Final Rating			
		Warning	Needs Improvement	Proficient	Advanced
1	1500067689	Ⓦ	Ⓝ	Ⓟ	Ⓐ
2	1500052817	Ⓦ	Ⓝ	Ⓟ	Ⓐ
3	1500039886	Ⓦ	Ⓝ	Ⓟ	Ⓐ
4	1500078265	Ⓦ	Ⓝ	Ⓟ	Ⓐ
5	1500066844	Ⓦ	Ⓝ	Ⓟ	Ⓐ
6	1500069724	Ⓦ	Ⓝ	Ⓟ	Ⓐ
7	1500053059	Ⓦ	Ⓝ	Ⓟ	Ⓐ
8	1500060422	Ⓦ	Ⓝ	Ⓟ	Ⓐ
9	1500032000	Ⓦ	Ⓝ	Ⓟ	Ⓐ
10	1500051505	Ⓦ	Ⓝ	Ⓟ	Ⓐ
11	1500070324	Ⓦ	Ⓝ	Ⓟ	Ⓐ
12	1500092288	Ⓦ	Ⓝ	Ⓟ	Ⓐ
13	1500008142	Ⓦ	Ⓝ	Ⓟ	Ⓐ
14	1500015769	Ⓦ	Ⓝ	Ⓟ	Ⓐ
15	1500037307	Ⓦ	Ⓝ	Ⓟ	Ⓐ
16	1500066110	Ⓦ	Ⓝ	Ⓟ	Ⓐ
17	1500081005	Ⓦ	Ⓝ	Ⓟ	Ⓐ
18	1500092488	Ⓦ	Ⓝ	Ⓟ	Ⓐ
19	1500006118	Ⓦ	Ⓝ	Ⓟ	Ⓐ
20	1500073254	Ⓦ	Ⓝ	Ⓟ	Ⓐ
21	1500012559	Ⓦ	Ⓝ	Ⓟ	Ⓐ
22	1500043737	Ⓦ	Ⓝ	Ⓟ	Ⓐ
23	1500002637	Ⓦ	Ⓝ	Ⓟ	Ⓐ
24	1500051824	Ⓦ	Ⓝ	Ⓟ	Ⓐ
25	1500088368	Ⓦ	Ⓝ	Ⓟ	Ⓐ
26	1500013179	Ⓦ	Ⓝ	Ⓟ	Ⓐ
27	1500029985	Ⓦ	Ⓝ	Ⓟ	Ⓐ

Appendix G

ELA Grade 5 **Rater ID: _____**

BoW No.	ID	Final Rating			
		Warning	Needs Improvement	Proficient	Advanced
28	1500048729	Ⓦ	Ⓝ	Ⓟ	Ⓐ
29	1500074822	Ⓦ	Ⓝ	Ⓟ	Ⓐ
30	1500028020	Ⓦ	Ⓝ	Ⓟ	Ⓐ
31	1500060928	Ⓦ	Ⓝ	Ⓟ	Ⓐ
32	1500067614	Ⓦ	Ⓝ	Ⓟ	Ⓐ
33	1500078731	Ⓦ	Ⓝ	Ⓟ	Ⓐ
34	1500047920	Ⓦ	Ⓝ	Ⓟ	Ⓐ
35	1500064760	Ⓦ	Ⓝ	Ⓟ	Ⓐ
36	1500065063	Ⓦ	Ⓝ	Ⓟ	Ⓐ
37	1500064021	Ⓦ	Ⓝ	Ⓟ	Ⓐ
38	1500075700	Ⓦ	Ⓝ	Ⓟ	Ⓐ
39	1500017931	Ⓦ	Ⓝ	Ⓟ	Ⓐ
40	1500038447	Ⓦ	Ⓝ	Ⓟ	Ⓐ
41	1500042650	Ⓦ	Ⓝ	Ⓟ	Ⓐ
42	1500082704	Ⓦ	Ⓝ	Ⓟ	Ⓐ
43	1500039156	Ⓦ	Ⓝ	Ⓟ	Ⓐ
44	1500045213	Ⓦ	Ⓝ	Ⓟ	Ⓐ
45	1500058486	Ⓦ	Ⓝ	Ⓟ	Ⓐ
46	1500069503	Ⓦ	Ⓝ	Ⓟ	Ⓐ
47	1500002119	Ⓦ	Ⓝ	Ⓟ	Ⓐ
48	1500079212	Ⓦ	Ⓝ	Ⓟ	Ⓐ
49	1500020510	Ⓦ	Ⓝ	Ⓟ	Ⓐ
50	1500072480	Ⓦ	Ⓝ	Ⓟ	Ⓐ

Appendix I

Standard Setting Evaluation Forms

Appendix G

Grade 3

MASSACHUSETTS DEPARTMENT OF EDUCATION
MCAS 2006

MCAS English Language Arts Standard Setting –
(July 11 and 12)

Standard Setting Evaluation Form

Please check the most appropriate category or fill in the blank for each of the following statements:

1. I participated in the following group:

16 Grade 3	_____ Grade 6
_____ Grade 5	_____ Grade 8

2. I am a
- 9 Classroom Teacher
- 2 K-12 Education Administrator
- 1 University-level Educator
- 1 Business and/or Community Representative
- 3 Other

3. I have substantial familiarity with the achievement of:
- 9 Students with disabilities
- 5 Students with limited English proficiency
- 2 Neither of these two groups of students.

4. Please circle the letter that best describes the extent to which your ratings were based on student responses to multiple-choice and open-response questions. (CIRCLE ONLY ONE)

A. 0 - Overall I relied primarily on open-response questions to determine my ratings.

B. 16 - Overall, I relied equally on open-response and multiple-choice questions to determine my ratings.

C. 0 - Overall, I relied primarily on multiple-choice questions to determine my ratings.

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating the degree to which you agree or disagree with each of the following statements:

1-Strongly Disagree 2-Disagree 3-Agree 4-Strongly Agree

Statements	Strongly Disagree	Disagree	Agree	Strongly Agree
5. The overall environment and accommodations were comfortable and appropriate for standard-setting activities.	**0**	**1**	**3**	**12**
6. The background information provided on Tuesday regarding the *Curriculum Framework*, MCAS ELA exams, and the purpose of standard setting improved my ability to set standards.	**0**	**0**	**3**	**13**
7. Taking and discussing the MCAS ELA exam during my orientation helped me understand the purpose and process of MCAS ELA standard setting.	**0**	**0**	**3**	**13**
8. By the end of the calibration training (ranking, discussing, and classifying sets of student work), I could distinguish among MCAS ELA Performance Level Descriptors.	**0**	**0**	**3**	**13**
9. Overall, I was provided with clear instructions for my standard-setting assignments.	**0**	**0**	**4**	**12**
10. The group discussions that took place after the first round of ratings improved my ability to set standards.	**0**	**0**	**8**	**8**
11. I am confident that the ratings I provided were consistent with the MCAS ELA Performance Level descriptors.	**0**	**0**	**1**	**15**
12. The MCAS ELA standard-setting process provided for a reliable classification of student work.	**0**	**0**	**3**	**13**

Appendix G

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating your perceptions as to the time allotted for each of the segments of standard setting:

1-Far too short 2-Too short 3-Approximately right 4-Too long 5-Far too long

Segments of Standard Setting	Far too short	Too short	Approximately right	Too long	Far too long
13. Initial background information provided on Tuesday morning	**0**	**0**	**15**	**1**	**0**
14. Taking and discussing the MCAS ELA exam	**0**	**0**	**14**	**2**	**0**
15. Learning about and discussing Performance Level descriptors	**0**	**2**	**13**	**1**	**0**
16. Ranking, discussing, and classifying students work (calibration)	**0**	**1**	**14**	**0**	**1**
17. Initial individual classification of student work	**0**	**0**	**15**	**1**	**0**
18. Group discussion regarding initial ratings	**0**	**0**	**14**	**1**	**1**
19. Rating student work for the second time	**0**	**1**	**14**	**1**	**0**
20. Final rating of student work	**0**	**1**	**12**	**3**	**0**

Please provide any additional comments on the back of this page.

Comments for grade 3 – ELA

- Thank you so much for letting me be a part of this! I will look forward to the results. Let me know if I can help again! ☺

- I would like additional training on scoring open response questions and how to improve the unity of open response questions with students! Everything was great! A true learning experience!

- This was my first experience with a DOE Panel. I enjoyed it as well as the women I worked with. I noticed there were no other Paraprofessionals involved and wished there were more because we are very involved with children, and it is a wonderful experience.

- More information needed on how the test is developed so participants would understand the questions and level of difficulty.

- Even though scoring was not supposed to be a part of this, it was still difficult for some participants to get a cut score number out of their discussion.

- Overall, these 2 days were well organized, orderly, and it was evident that prior preparation was thorough. The facilitator, Jan Kahen, make things more easily, kept discussions on track, and was a terrific facilitator. She allowed each to speak, was nonjudgmental, and very gracious. Thank you for this excellent opportunity to participate in such a worthwhile endeavor. Now everything makes a little more sense.

- Enjoyed experience

- This was a fantastic op. as an educator. If only all these discussions could happen daily.

-
    1. Event should be more centrally located in state,
    2. Far too many teachers that have a narrow perspective – made it difficult to set a state-wide standard.
    3. Classroom or U-shape set up would have been better – we felt like 3 small groups at round tables rather than 1 large group.
    4. Hotel rooms were horrible.

- This was such a great opportunity! I really enjoyed being on this panel and feel that I am leaving with a lot of insight regarding the standards.

Thank you for being a part the MCAS ELA 2006 standard-setting team.

Appendix G

Grade 5

MASSACHUSETTS DEPARTMENT OF EDUCATION
MCAS 2006

MCAS English Language Arts Standard Setting –
(July 11 and 12)

Standard Setting Evaluation Form

Please check the most appropriate category or fill in the blank for each of the following statements:

1. I participated in the following group:

_____ Grade 3 _____ Grade 6
_13___ Grade 5 _____ Grade 8

2. I am a
_10____ Classroom Teacher
_0____ K-12 Education Administrator
_0____ University-level Educator
_0____ Business and/or Community Representative
_3____ Other

3. I have substantial familiarity with the achievement of:
_9___ Students with disabilities
_0____ Students with limited English proficiency
_4____ Neither of these two groups of students.

4. Please circle the letter that best describes the extent to which your ratings were based on student responses to multiple-choice and open-response questions. (CIRCLE ONLY ONE)

A. 1 - Overall I relied primarily on open-response questions to determine my ratings.

B. 12 - Overall, I relied equally on open-response and multiple-choice questions to determine my ratings.

C. 0 - Overall, I relied primarily on multiple-choice questions to determine my ratings.

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating the degree to which you agree or disagree with each of the following statements:

1-Strongly Disagree 2-Disagree 3-Agree 4-Strongly Agree

Statements	Strongly Disagree	Disagree	Agree	Strongly Agree
5. The overall environment and accommodations were comfortable and appropriate for standard-setting activities.	**0**	**0**	**6**	**7**
6. The background information provided on Tuesday regarding the *Curriculum Framework*, MCAS ELA exams, and the purpose of standard setting improved my ability to set standards.	**0**	**0**	**7**	**6**
7. Taking and discussing the MCAS ELA exam during my orientation helped me understand the purpose and process of MCAS ELA standard setting.	**0**	**0**	**6**	**7**
8. By the end of the calibration training (ranking, discussing, and classifying sets of student work), I could distinguish among MCAS ELA Performance Level Descriptors.	**0**	**0**	**1**	**12**
9. Overall, I was provided with clear instructions for my standard-setting assignments.	**0**	**0**	**7**	**6**
10. The group discussions that took place after the first round of ratings improved my ability to set standards.	**0**	**0**	**8**	**5**
11. I am confident that the ratings I provided were consistent with the MCAS ELA Performance Level descriptors.	**0**	**0**	**4**	**9**
12. The MCAS ELA standard-setting process provided for a reliable classification of student work.	**0**	**0**	**5**	**8**

Appendix G

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating your perceptions as to the time allotted for each of the segments of standard setting:

1-Far too short 2-Too short 3-Approximately right 4-Too long 5-Far too long

Segments of Standard Setting	Far too short	Too short	Approximately right	Too long	Far too long
13. Initial background information provided on Tuesday morning	**0**	**0**	**7**	**5**	**0**
14. Taking and discussing the MCAS ELA exam	**0**	**0**	**10**	**2**	**0**
15. Learning about and discussing Performance Level descriptors	**0**	**0**	**11**	**1**	**0**
16. Ranking, discussing, and classifying students work (calibration)	**0**	**1**	**9**	**2**	**0**
17. Initial individual classification of student work	**0**	**0**	**8**	**3**	**1**
18. Group discussion regarding initial ratings	**0**	**0**	**11**	**1**	**1**
19. Rating student work for the second time	**0**	**1**	**11**	**0**	**0**
20. Final rating of student work	**0**	**0**	**11**	**1**	**0**

Please provide any additional comments on the back of this page.

Comments for grade 5 – ELA

- This was my first experience with a DOE – sponsored workshop on MCAS. I found the quality of the proceedings to be exceptional. The staff personnel were always professional and cordial. I'm glad that I made the trip!

- Our facilitator was SUPERB! She kept us focused and on track and moving along at an appropriate pace – Thank you for a GREAT experience!

Thank you for being a part the MCAS ELA 2006 standard-setting team.

Appendix G

Grade 6

MASSACHUSETTS DEPARTMENT OF EDUCATION
MCAS 2006

MCAS English Language Arts Standard Setting –
(July 11 and 12)

Standard Setting Evaluation Form

Please check the most appropriate category or fill in the blank for each of the following statements:

1. I participated in the following group:

_____ Grade 3	_19___ Grade 6
_____ Grade 5	_____ Grade 8

2. I am a

_14___	Classroom Teacher
_0____	K-12 Education Administrator
_0____	University-level Educator
_1____	Business and/or Community Representative
_4____	Other

3. I have substantial familiarity with the achievement of:

_14___	Students with disabilities
_1____	Students with limited English proficiency
_3____	Neither of these two groups of students.

4. Please circle the letter that best describes the extent to which your ratings were based on student responses to multiple-choice and open-response questions. (CIRCLE ONLY ONE)

A. 2 - Overall I relied primarily on open-response questions to determine my ratings.

B. 17 - Overall, I relied equally on open-response and multiple-choice questions to determine my ratings.

C. 0 - Overall, I relied primarily on multiple-choice questions to determine my ratings.

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating the degree to which you agree or disagree with each of the following statements:

1-Strongly Disagree 2-Disagree 3-Agree 4-Strongly Agree

Statements	Strongly Disagree	Disagree	Agree	Strongly Agree
5. The overall environment and accommodations were comfortable and appropriate for standard-setting activities.	**0**	**1**	**8**	**10**
6. The background information provided on Tuesday regarding the *Curriculum Framework*, MCAS ELA exams, and the purpose of standard setting improved my ability to set standards.	**1**	**1**	**9**	**8**
7. Taking and discussing the MCAS ELA exam during my orientation helped me understand the purpose and process of MCAS ELA standard setting.	**0**	**0**	**8**	**11**
8. By the end of the calibration training (ranking, discussing, and classifying sets of student work), I could distinguish among MCAS ELA Performance Level Descriptors.	**1**	**4**	**10**	**3**
9. Overall, I was provided with clear instructions for my standard-setting assignments.	**0**	**1**	**8**	**10**
10. The group discussions that took place after the first round of ratings improved my ability to set standards.	**0**	**1**	**14**	**4**
11. I am confident that the ratings I provided were consistent with the MCAS ELA Performance Level descriptors.	**0**	**0**	**8**	**11**
12. The MCAS ELA standard-setting process provided for a reliable classification of student work.	**0**	**0**	**11**	**8**

Appendix G

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating your perceptions as to the time allotted for each of the segments of standard setting:

1-Far too short 2-Too short 3-Approximately right 4-Too long 5-Far too long

Segments of Standard Setting	Far too short	Too short	Approximately right	Too long	Far too long
13. Initial background information provided on Tuesday morning	**0**	**0**	**14**	**4**	**0**
14. Taking and discussing the MCAS ELA exam	**0**	**3**	**14**	**2**	**0**
15. Learning about and discussing Performance Level descriptors	**0**	**3**	**11**	**3**	**2**
16. Ranking, discussing, and classifying students work (calibration)	**0**	**2**	**14**	**3**	**0**
17. Initial individual classification of student work	**0**	**1**	**13**	**4**	**1**
18. Group discussion regarding initial ratings	**0**	**0**	**11**	**5**	**2**
19. Rating student work for the second time	**0**	**1**	**17**	**1**	**0**
20. Final rating of student work	**0**	**1**	**18**	**3**	**0**

Please provide any additional comments on the back of this page.

Comments for grade 6 – ELA

- I am interested in doing this again on other subjects to gain a greater awareness. Thanks for the opportunity!

- I thoroughly enjoyed this process. It was helpful to me.

- It would have been more beneficial if the students OR scores were given to us…it would've saved so much time and "debate," Many of us were trying to score these (+ that's a whole other process). Too many people were debating kids' "or" scores & getting into their own "agendas."

- Clearer instructions should have been given. More use of small group discussions rather than open forum discussions would have made for better discussions.

- The item mapping process was very time consuming and not particularly helpful in the whole overall process. It was not clear how it fit into the standard setting process and it also focused much of the discussion on writing which got us off track when looking at the student samples.

-
    - Directions need to be clearer when given
    - Guidelines for NI, P, and A categories when developing them were too general - focus on reading not writing.
    - Everyone was pleasant and friendly

- Providing the 6th grade performance level definitions was confusing. This was both content/ performance based standards. As a result, members were correcting for conventions / structure along with content. The facilitator needed to move discussions along.

- Related to Statement 8: I was previously trained to score the Grade 7 ELA composition, so I felt that I could classify the open responses. However, among my panel, we did not make consistent scoring judgments for open response questions. I

  I felt that my facilitator did not effectively lead the group. She did not clearly explain the important considerations (i.e. that grammar and sentence structure is not considered in open responses). She did not moderate group discussions well.

Thank you for being a part the MCAS ELA 2006 standard-setting team.

Appendix G

Grade 8

MASSACHUSETTS DEPARTMENT OF EDUCATION
MCAS 2006

MCAS English Language Arts Standard Setting –
(July 11 and 12)

Standard Setting Evaluation Form

Please check the most appropriate category or fill in the blank for each of the following statements:

1. I participated in the following group:

_____ Grade 3	_____ Grade 6
_____ Grade 5	___17_ Grade 8

2. I am a
- _11____ Classroom Teacher
- _4____ K-12 Education Administrator
- _0____ University-level Educator
- _0____ Business and/or Community Representative
- _2____ Other

3. I have substantial familiarity with the achievement of:
- _14___ Students with disabilities
- _1____ Students with limited English proficiency
- _2____ Neither of these two groups of students.

4. Please circle the letter that best describes the extent to which your ratings were based on student responses to multiple-choice and open-response questions. (CIRCLE ONLY ONE)

A. 3 - Overall I relied primarily on open-response questions to determine my ratings.

B. 13 - Overall, I relied equally on open-response and multiple-choice questions to determine my ratings.

C. 1 - Overall, I relied primarily on multiple-choice questions to determine my ratings.

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating the degree to which you agree or disagree with each of the following statements:

1-Strongly Disagree 2-Disagree 3-Agree 4-Strongly Agree

Statements	Strongly Disagree	Disagree	Agree	Strongly Agree
5. The overall environment and accommodations were comfortable and appropriate for standard-setting activities.	**0**	**0**	**4**	**13**
6. The background information provided on Tuesday regarding the *Curriculum Framework*, MCAS ELA exams, and the purpose of standard setting improved my ability to set standards.	**0**	**1**	**4**	**12**
7. Taking and discussing the MCAS ELA exam during my orientation helped me understand the purpose and process of MCAS ELA standard setting.	**0**	**0**	**4**	**13**
8. By the end of the calibration training (ranking, discussing, and classifying sets of student work), I could distinguish among MCAS ELA Performance Level Descriptors.	**0**	**0**	**5**	**12**
9. Overall, I was provided with clear instructions for my standard-setting assignments.	**0**	**0**	**8**	**9**
10. The group discussions that took place after the first round of ratings improved my ability to set standards.	**0**	**1**	**9**	**7**
11. I am confident that the ratings I provided were consistent with the MCAS ELA Performance Level descriptors.	**0**	**0**	**8**	**9**
12. The MCAS ELA standard-setting process provided for a reliable classification of student work.	**0**	**0**	**8**	**9**

Appendix G

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating your perceptions as to the time allotted for each of the segments of standard setting:

1-Far too short 2-Too short 3-Approximately right 4-Too long 5-Far too long

Segments of Standard Setting	Far too short	Too short	Approximately right	Too long	Far too long
13. Initial background information provided on Tuesday morning	**0**	**0**	**13**	**4**	**0**
14. Taking and discussing the MCAS ELA exam	**0**	**3**	**14**	**0**	**0**
15. Learning about and discussing Performance Level descriptors	**0**	**5**	**12**	**0**	**0**
16. Ranking, discussing, and classifying students work (calibration)	**0**	**2**	**15**	**0**	**0**
17. Initial individual classification of student work	**0**	**1**	**15**	**0**	**0**
18. Group discussion regarding initial ratings	**0**	**2**	**14**	**0**	**0**
19. Rating student work for the second time	**0**	**0**	**13**	**3**	**0**
20. Final rating of student work	**0**	**0**	**13**	**2**	**1**

Please provide any additional comments on the back of this page.

Comments for grade 8 – ELA

-
  - ✓ Good experience over-all
  - ✓ Perhaps participant should be encouraged to read for flavor and the whole – reading of the large packet took too long – we are not correcting individual assignments as classroom teachers – Armed with the rubrics and class pastings, teachers should have been able to read faster and make decisions without agonizing over the work.

- I found the experience to be very valuable. I ill be able to utilize what I learned in my teaching.

- Teachers identify with "our" school systems. Our name badges list our name (fine) and our town of residence (CONFUSING!). I'd rather wear my name and be identified with my school system.

- This was extremely insightful and refreshing workshop / seminar. From accommodations to presenters, things were done very professionally and made me and my input feel valued.

- This process has been highly instructional and informative. I learned a lot and enjoyed the experience. Thank you!

-
  1. I wish that we had been given fewer student examples to read so that I could concentrate more intensely on each sample.
  2. I would have preferred to receive more guidance when I reviewed the student examples. I used my own methods which I found to be (quite) faulty.

Thank you for being a part the MCAS ELA 2006 standard-setting team.

Appendix G

Grade 3

MASSACHUSETTS DEPARTMENT OF EDUCATION
MCAS 2006

MCAS Mathematics Standard Setting
(August 22 and 23)

Standard Setting Evaluation Form

Please check the most appropriate category for each of the following statements:

1. I participated in the following group:

___16__ Grade 3 _____ Grade 5 _____ Grade 7

2. I am a
_8___ Classroom Teacher
_3____ K-12 Education Administrator
_0____ University-level Educator
_0____ Business and/or Community Representative
_4____ Other

3. I am
_4____ Male
_12____ Female

4. Please circle the letter that best describes the extent to which your ratings were based on student responses to multiple-choice and constructed response (short-answer and open-response) questions. (CIRCLE ONLY ONE)

A. 0 - Overall I relied primarily on constructed response (short-answer and open-response) questions to determine my ratings.

B. 16 - Overall, I relied equally on constructed response (short-answer and open-response) and multiple-choice questions to determine my ratings.

C. 0 - Overall, I relied primarily on multiple-choice questions to determine my ratings.

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating the degree to which you agree or disagree with each of the following statements:

1-Strongly Disagree 2-Disagree 3-Agree 4-Strongly Agree

Statements	Strongly Disagree	Disagree	Agree	Strongly Agree
5. The overall environment and accommodations were comfortable and appropriate for Standard-Setting activities.	**0**	**0**	**4**	**12**
6. The background information provided on Tuesday regarding the *Curriculum Framework*, MCAS Mathematics exams, and the purpose of Standard Setting improved my ability to set standards.	**0**	**0**	**10**	**6**
7. Taking and discussing the MCAS Mathematics exam during my orientation helped me understand the purpose and process of MCAS Mathematics Standard Setting.	**0**	**0**	**4**	**12**
8. By the end of the calibration training (ranking, discussing, and classifying sets of student work), I could distinguish among MCAS Mathematics Performance Level Descriptors.	**1**	**3**	**7**	**5**
9. Overall, I was provided with clear instructions for my Standard-Setting assignments.	**0**	**0**	**8**	**8**
10. The group discussions that took place after the first round of ratings improved my ability to set standards.	**0**	**0**	**8**	**8**
11. I am confident that the ratings I provided were consistent with the MCAS Mathematics Performance Level Descriptors.	**0**	**1**	**6**	**8**
12. The MCAS Mathematics Standard-Setting process provided for a reliable classification of student work.	**0**	**0**	**13**	**3**

Appendix G

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating your perceptions as to the time allotted for each of the segments of Standard Setting:

1-Far too short 2-Too short 3-Approximately right 4-Too long 5-Far too long

Segments of Standard Setting	Far too short	Too short	Approximately right	Too long	Far too long
13. Initial background information provided on Tuesday morning	**0**	**0**	**14**	**2**	**0**
14. Taking and discussing the MCAS Mathematics exam	**0**	**0**	**16**	**0**	**0**
15. Learning about and discussing Performance Levels Descriptors	**0**	**2**	**13**	**1**	**0**
16. Ranking, discussing, and classifying students work (calibration)	**0**	**3**	**10**	**3**	**0**
17. Initial individual classification of student work	**0**	**2**	**14**	**0**	**0**
18. Group discussion regarding initial ratings	**1**	**1**	**10**	**3**	**0**
19. Rating student work for the second time	**0**	**0**	**13**	**2**	**1**
20. Final rating of student work	**1**	**2**	**8**	**0**	**5**

Please provide any additional comments on the back of this page.

Comments for grade 3 – MATH

- We had an outstanding facilitator. A very valuable experience and the discussions / debate allowed for all viewpoints to be shared and considered. Thank you for the opportunity! Great accommodations!

- To prevent too long unnecessary discussions state the goals of the workshop more clear.

- Thank you for the opportunity to experience the process of standard setting.
  The descriptions of performance levels – are too vague and or to specific – I'd love to see a reunite of these.
  The group was super led to value the full range of agree / disagree respectfully – all participants worked hard, all day, both days.
  I am disappointed in the lack of direction in the (entire) group, not just in our section.

- See # 12
  The AP was a difficult item to deal with given the design of the test. The explanations / comments given by DOE personnel at end of day two was helpful.
- I was concerned about the Above Proficiency Rating given the original intent of the test.

- I am uncomfortable assigning an above proficient score to students who took a test written with the purpose of testing for proficiency only. I would feel much more comfortable keeping the 3 categories of warning, needs improvement & proficient.

  Overall – I really enjoyed the process – and view it as an excellent professional development opportunity.

- I feel that the Standard – Setting process would have worked better if the test had been written after the inclusion of the Above Proficient category.

Thank you for being a part the MCAS Mathematics 2006 Standard-Setting team.

Appendix G

Grade 5

MASSACHUSETTS DEPARTMENT OF EDUCATION
MCAS 2006

MCAS Mathematics Standard Setting
(August 22 and 23)

Standard Setting Evaluation Form

Please check the most appropriate category for each of the following statements:

1. I participated in the following group:

_____ Grade 3 __16___ Grade 5 _____ Grade 7

2. I am a
_10___ Classroom Teacher
_2____ K-12 Education Administrator
_1____ University-level Educator
_0____ Business and/or Community Representative
_3____ Other

3. I am
_2____ Male
_14____ Female

4. Please circle the letter that best describes the extent to which your ratings were based on student responses to multiple-choice and constructed response (short-answer and open-response) questions. (CIRCLE ONLY ONE)

A. 0 - Overall I relied primarily on constructed response (short-answer and open-response) questions to determine my ratings.

B. 16 - Overall, I relied equally on constructed response (short-answer and open-response) and multiple-choice questions to determine my ratings.

C. 0 - Overall, I relied primarily on multiple-choice questions to determine my ratings.

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating the degree to which you agree or disagree with each of the following statements:

1-Strongly Disagree 2-Disagree 3-Agree 4-Strongly Agree

Statements	Strongly Disagree	Disagree	Agree	Strongly Agree
5. The overall environment and accommodations were comfortable and appropriate for Standard-Setting activities.	**0**	**0**	**1**	**15**
6. The background information provided on Tuesday regarding the *Curriculum Framework*, MCAS Mathematics exams, and the purpose of Standard Setting improved my ability to set standards.	**0**	**0**	**6**	**10**
7. Taking and discussing the MCAS Mathematics exam during my orientation helped me understand the purpose and process of MCAS Mathematics Standard Setting.	**0**	**0**	**2**	**14**
8. By the end of the calibration training (ranking, discussing, and classifying sets of student work), I could distinguish among MCAS Mathematics Performance Level Descriptors.	**0**	**0**	**6**	**10**
9. Overall, I was provided with clear instructions for my Standard-Setting assignments.	**0**	**0**	**5**	**11**
10. The group discussions that took place after the first round of ratings improved my ability to set standards.	**0**	**1**	**5**	**10**
11. I am confident that the ratings I provided were consistent with the MCAS Mathematics Performance Level Descriptors.	**0**	**0**	**3**	**13**
12. The MCAS Mathematics Standard-Setting process provided for a reliable classification of student work.	**0**	**0**	**5**	**11**

Appendix G

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating your perceptions as to the time allotted for each of the segments of Standard Setting:

1-Far too short 2-Too short 3-Approximately right 4-Too long 5-Far too long

Segments of Standard Setting	Far too short	Too short	Approximately right	Too long	Far too long
13. Initial background information provided on Tuesday morning	**0**	**0**	**12**	**3**	**1**
14. Taking and discussing the MCAS Mathematics exam	**0**	**0**	**16**	**0**	**0**
15. Learning about and discussing Performance Levels Descriptors	**0**	**0**	**10**	**6**	**0**
16. Ranking, discussing, and classifying students work (calibration)	**0**	**1**	**12**	**2**	**1**
17. Initial individual classification of student work	**0**	**4**	**10**	**2**	**0**
18. Group discussion regarding initial ratings	**0**	**0**	**5**	**8**	**3**
19. Rating student work for the second time	**0**	**1**	**11**	**4**	**0**
20. Final rating of student work	**0**	**1**	**14**	**0**	**0**

Please provide any additional comments on the back of this page.

Comments for grade 5 – MATH

- I think that giving the pre-determined rating skewed some of my ratings. Maybe giving less student samples (lots less) and allowing participating members to pre-determine, and then compare to pre-determined rating, may be more beneficial to help determine cut-off points.

- A copy of the 5th grade math frameworks would have been helpful during these activities.

-
    1. Is 50 a good sample to use?
    2. How would we know if there were students who had accommodations during this test?
    3. How would we know if this sampling represents the diversity of students across our state?
    4. Are these students samples sorted by demographics asked in the first section of the MCAS Exam?

- We should have had 3 days or started earlier each day.

- I found this to be a highly valuable process. Having worked on student scoring, as well as on Assessment Development Committee work, I feel I have a lot more info on the whole MCAS process. This also showed me things to bring back to the classroom to aid students in constructing quality answers to open response questions.

- The grey bands hindered my ability to make decisions because they categorized student responses before I read them. During the 1st reading, I fell the grey bands should be left off.

Thank you for being a part the MCAS Mathematics 2006 Standard-Setting team.

Appendix G

Grade 7

MASSACHUSETTS DEPARTMENT OF EDUCATION
MCAS 2006

MCAS Mathematics Standard Setting
(August 22 and 23)

Standard Setting Evaluation Form

Please check the most appropriate category for each of the following statements:

1. I participated in the following group:

_____ Grade 3 _____ Grade 5 __16__ Grade 7

2. I am a
_10___ Classroom Teacher
_1____ K-12 Education Administrator
_2____ University-level Educator
_1____ Business and/or Community Representative
_2____ Other

3. I am
_4____ Male
_12____ Female

4. Please circle the letter that best describes the extent to which your ratings were based on student responses to multiple-choice and constructed response (short-answer and open-response) questions. (CIRCLE ONLY ONE)

A. 2- Overall I relied primarily on constructed response (short-answer and open-response) questions to determine my ratings.

B. 14 Overall, I relied equally on constructed response (short-answer and open-response) and multiple-choice questions to determine my ratings.

C. 0 - Overall, I relied primarily on multiple-choice questions to determine my ratings.

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating the degree to which you agree or disagree with each of the following statements:

1-Strongly Disagree 2-Disagree 3-Agree 4-Strongly Agree

Statements	Strongly Disagree	Disagree	Agree	Strongly Agree
5. The overall environment and accommodations were comfortable and appropriate for Standard-Setting activities.	**0**	**1**	**5**	**10**
6. The background information provided on Tuesday regarding the *Curriculum Framework*, MCAS Mathematics exams, and the purpose of Standard Setting improved my ability to set standards.	**0**	**1**	**10**	**5**
7. Taking and discussing the MCAS Mathematics exam during my orientation helped me understand the purpose and process of MCAS Mathematics Standard Setting.	**0**	**1**	**5**	**10**
8. By the end of the calibration training (ranking, discussing, and classifying sets of student work), I could distinguish among MCAS Mathematics Performance Level Descriptors.	**0**	**0**	**11**	**5**
9. Overall, I was provided with clear instructions for my Standard-Setting assignments.	**0**	**4**	**9**	**2**
10. The group discussions that took place after the first round of ratings improved my ability to set standards.	**0**	**0**	**8**	**8**
11. I am confident that the ratings I provided were consistent with the MCAS Mathematics Performance Level Descriptors.	**0**	**0**	**6**	**10**
12. The MCAS Mathematics Standard-Setting process provided for a reliable classification of student work.	**0**	**0**	**8**	**8**

Appendix G

Standard Setting Evaluation Form

DIRECTIONS: Please circle the number indicating your perceptions as to the time allotted for each of the segments of Standard Setting:

1-Far too short 2-Too short 3-Approximately right 4-Too long 5-Far too long

Segments of Standard Setting	Far too short	Too short	Approximately right	Too long	Far too long
13. Initial background information provided on Tuesday morning	**0**	**1**	**10**	**3**	**1**
14. Taking and discussing the MCAS Mathematics exam	**0**	**3**	**11**	**1**	**0**
15. Learning about and discussing Performance Levels Descriptors	**0**	**1**	**13**	**1**	**0**
16. Ranking, discussing, and classifying students work (calibration)	**0**	**1**	**10**	**4**	**0**
17. Initial individual classification of student work	**0**	**2**	**13**	**0**	**0**
18. Group discussion regarding initial ratings	**0**	**0**	**13**	**2**	**0**
19. Rating student work for the second time	**0**	**2**	**11**	**2**	**0**
20. Final rating of student work	**0**	**3**	**7**	**4**	**1**

Please provide any additional comments on the back of this page.

Thank you for being a part the MCAS Mathematics 2006 Standard-Setting team.