



Massachusetts Department of  
**ELEMENTARY & SECONDARY  
EDUCATION**

## Ensuring Technical Quality: Policies and Procedures Guiding the Development of the MCAS Tests

---

September 2008

**Massachusetts Department of Elementary and Secondary Education**  
350 Main Street, Malden, MA 02148  
Phone 781-338-3000 TTY: N.E.T. Relay 800-439-2370  
[www.doe.mass.edu](http://www.doe.mass.edu)



This document was prepared by the  
Massachusetts Department of Elementary and Secondary Education

Mitchell D. Chester, Ed.D.  
Commissioner of Elementary and Secondary Education

The Massachusetts Department of Elementary and Secondary Education, an affirmative action employer, is committed to ensuring that all of its programs and facilities are accessible to all members of the public.  
We do not discriminate on the basis of age, color, disability, national origin, race, religion, sex or sexual orientation.  
Inquiries regarding the Department's compliance with Title IX and other civil rights laws may be directed to the Human Resources Director, 350 Main St., Malden, MA 02148 781-338-6105.

© 2008 Massachusetts Department of Elementary and Secondary Education  
*Permission is hereby granted to copy any or all parts of this document for non-commercial educational purposes. Please credit the "Massachusetts Department of Elementary and Secondary Education."*

*This document printed on recycled paper*

Massachusetts Department of Elementary and Secondary Education  
350 Main Street, Malden, MA 02148-5023  
Phone 781-338-3000 TTY: N.E.T. Relay 800-439-2370  
[www.doe.mass.edu](http://www.doe.mass.edu)



# Contents

Acknowledgement.....	4
Foreword.....	5
Introduction.....	6
Background of the MCAS Program .....	7
The MCAS Tests .....	9
MCAS Test Administration .....	13
Performance Standards.....	14
Reporting MCAS Results .....	17
Student Participation in the MCAS Tests.....	20
MCAS and High School Graduation.....	21
Additional Resource .....	23

## Acknowledgement

The Department of Elementary and Secondary Education wishes to thank Charles DePascale, Senior Associate at the National Center for the Improvement of Educational Assessment, Inc., for his contribution to this document.

## Foreword

In the decade since the first MCAS test administration, much evidence has been gathered and presented documenting the technical quality of the MCAS program, including the annual *MCAS Technical Report* produced by the Department and its assessment contractor; independent, external reviews of the MCAS tests and *Curriculum Frameworks*; and comparisons of MCAS results with the performance of Massachusetts students on nationally recognized tests such as the National Assessment of Educational Progress (NAEP). These documents and analyses have provided ongoing evidence that the MCAS tests are technically sound. Less attention, however, has been devoted to describing the guiding policies and operating procedures that lead to and support the technical quality of the MCAS program. The information presented here is intended to describe those policies and procedures.

This document provides educators and interested members of the community with an overview of the processes and procedures designed to make the MCAS program and individual MCAS tests fair, challenging, and technically sound. Key technical activities such as developing test items and forms, scoring constructed-response items, setting performance standards, and maintaining performance standards across years are presented and explained. The document also discusses the role of MCAS reports of results and reporting tools in promoting the appropriate interpretation and use of MCAS test results. In addition, the document includes a description of components of the MCAS program—such as the MCAS Alternate Assessment, retest opportunities, and the performance appeals process—designed to enhance the accuracy of the measurement of the performance of all students.

Since the initial administration of the MCAS tests in 1998, the Department has attempted to make the MCAS program as transparent as possible to increase the usefulness of MCAS results and to support their appropriate use. The annual release of all MCAS test items used to determine student scores, along with sample student responses to constructed-response items and the return of individual student compositions, are examples of this transparency. This document reflects the Department's continuing commitment to the openness and accessibility of the program.

## Introduction

Validity is the cornerstone upon which all measurement systems are built. Whether the measurement tool is a standardized MCAS test, a weekly test constructed and administered by a student’s classroom teacher, or a thermometer used to measure the outside air temperature, the goal is to use a tool that produces results that support valid inferences and, ultimately, actions. In everyday use, a *valid* statement is one that is well grounded on principles, facts, and evidence. A valid statement is sound and able to withstand criticism. Similarly, in educational measurement, the validity of inferences drawn from test results must be sound and well grounded on principles and empirical evidence.

The primary inferences drawn from the MCAS test results are conclusions about the level of students’ achievement of the standards contained in the Massachusetts *Curriculum Frameworks*. Therefore, the MCAS tests are custom-designed to support those conclusions. All items included on the MCAS tests are written to measure performance based on standards contained in the *Curriculum Frameworks*. Equally important, virtually all standards contained in the *Curriculum Frameworks* are measured by items on the MCAS tests.<sup>1</sup> This document will describe technical aspects of the MCAS test designed to enhance the validity of inferences drawn about student achievement of the standards.

### Uses of MCAS Results

MCAS results are used for a variety of purposes. Official uses of MCAS results include the following:

- determining school and district Adequate Yearly Progress (AYP) toward meeting federal No Child Left Behind (NCLB) requirements
- determining whether high school students have demonstrated the knowledge and skills required to earn a Competency Determination (CD)—one requirement for earning a high school diploma in Massachusetts
- providing information to support program evaluation at the school and district levels
- making decisions about scholarships, including the John and Abigail Adams Scholarship

Using MCAS results for any purpose requires an interpretation of scores. In all cases, it is important to be cautious in the interpretation and use of MCAS scores or any single piece of information about student performance. Decisions regarding classroom instruction, curriculum planning, and student placement should be based not only on MCAS results, but also on other relevant information about the school, district, or student.

---

<sup>1</sup> A small number of standards in the *Curriculum Frameworks* have been classified as not appropriate for paper-and-pencil large-scale assessment such as the MCAS tests. Examples of those standards from the English Language Arts framework include Language Standard 3, which requires students to make oral presentations, and Composition Standard 24, which requires students to conduct a research project. Standards such as those are assessed at the local level.

## Background of the MCAS Program

This section provides background information on the purposes of the MCAS program and a description of how it has evolved over time.

### Education Reform Mandate

The Massachusetts Comprehensive Assessment System (MCAS) is designed to meet requirements of the Education Reform Law of 1993. This law specifies that the testing program must

- test *all* students who are educated with Massachusetts public funds, including students with disabilities and limited English proficient students;
- measure performance based on the Massachusetts *Curriculum Framework* learning standards;
- report on the performance of individual students, schools, and districts.

As required by the Education Reform Law, students must pass the grade 10 tests in English Language Arts (ELA) and Mathematics as one condition of eligibility for a high school diploma (in addition to fulfilling local requirements). Students in the class of 2010 are the first students required to earn a CD in one of four high school Science and Technology/Engineering subjects (Biology, Chemistry, Introductory Physics, and Technology/Engineering). The MCAS program is also used to hold schools and districts accountable, on a yearly basis, for the progress they have made toward the objective of the No Child Left Behind (NCLB) Law that all students be proficient in reading and mathematics by 2014.

### A Custom-Designed Assessment Program

MCAS is a custom-designed program owned by the Commonwealth and built specifically to measure student performance on the Massachusetts *Curriculum Frameworks*. The ongoing development and implementation of MCAS is managed by staff of the Department of Elementary and Secondary Education with assistance and support from an assessment contractor. As described in this report, Massachusetts educators play a key role in MCAS through service on a variety of committees related to the development of MCAS test items and the establishment of MCAS performance standards. For technical issues, the Department is supported by a five-member national Technical Advisory Committee that meets three times each year. In addition, the Department consults with specialists in educational measurement and assessment at Massachusetts colleges and universities.

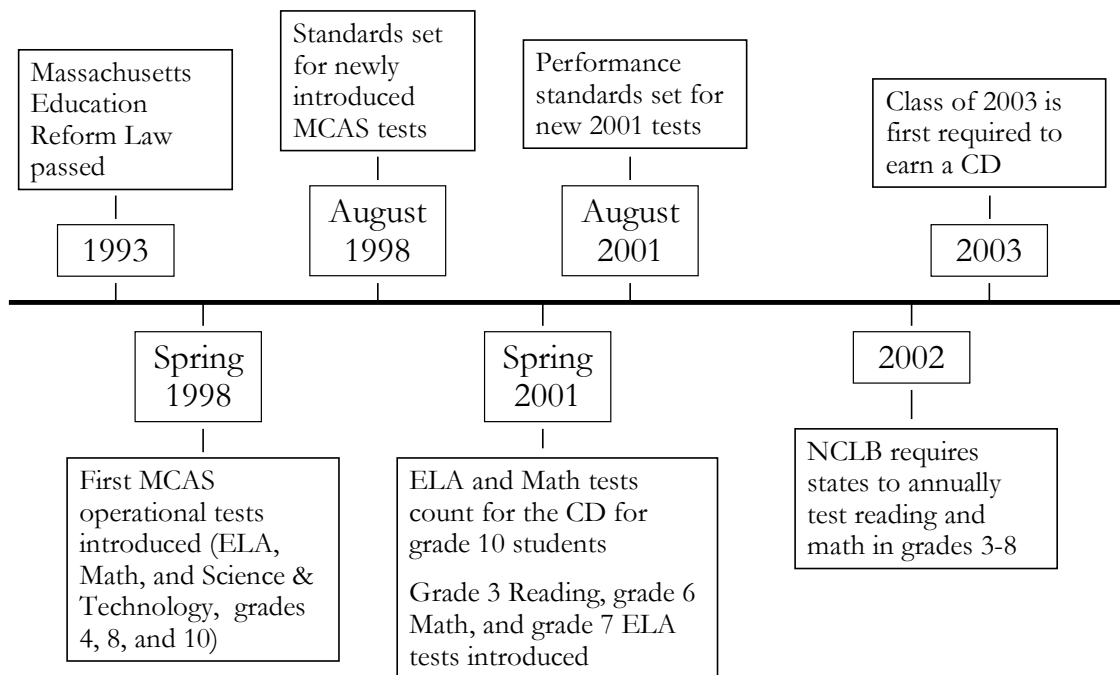
### History of the MCAS Program

MCAS tests have been administered to students in Massachusetts since 1998. In 1998, students were assessed in ELA and Mathematics at grades 4, 8, and 10. In subsequent years, additional grades and content areas were added to the testing program. Following the initial administration of each new test, performance standards were set.

Public school students in the graduating class of 2003 were the first students required to earn a CD in ELA and Mathematics as a condition for receiving a high school diploma. To fulfill the requirements of NCLB, in 2006 tests for several new grades and content areas were added to

MCAS. As a result, all students in grades 3 through 8 are assessed in both ELA and Mathematics. The time line below shows major milestones in the MCAS program from 1993 to 2003.

Figure 1. Historical Background of the MCAS Tests, 1993-2003



### Role of *Curriculum Frameworks*

All MCAS test items are custom developed based on the Massachusetts *Curriculum Framework* learning standards. Schools and districts have been provided with copies of the original frameworks and subsequent revisions and supplements. In the past ten years the Department has also offered training and technical assistance to schools in aligning their curriculum in each subject area to the grade-level learning standards specified in the corresponding framework. MCAS tests measure the degree to which students have mastered these learning standards, and can provide information about how well schools and districts have aligned their curriculum and instruction to the state standards.

### Spring 2007 MCAS Testing

In spring 2007, Massachusetts public school students participated in the tenth annual administration of the MCAS tests. More than one half million students in grades 3 through 12 participated in one or more MCAS tests. Virtually all public school students at grades 3 through 8 and 10 were tested in ELA and mathematics either through the standard MCAS tests or through the MCAS Alternate Assessment. Ninth and tenth grade students enrolled in introductory physics, biology, chemistry, or technology/engineering courses completed the first administration of the MCAS high school end-of-course tests in Science and Technology/Engineering (STE). Eleventh and twelfth grade students

who had not yet earned their CD in ELA and Mathematics had the opportunity to participate in one or more retest opportunities offered during the year.

## The MCAS Tests

During the 2006–2007 school year, 24 different MCAS tests (including retests) were administered in ELA, mathematics, and STE. Tests in History and Social Science are under development, and pilot tests were administered at grades 5, 7, and high school in spring 2008. History and Social Science tests will be fully operational beginning in spring 2009. Each individual MCAS test is custom designed in a multi-year process that involves content and assessment specialists from the Department of Elementary and Secondary Education (ESE); the Department’s assessment contractor; content, assessment, and bias and sensitivity specialists from Massachusetts and across the country; and, in particular, committees of Massachusetts public school teachers, curriculum directors, and educational specialists. The process is designed to produce MCAS tests that are aligned with the *Curriculum Frameworks* and that support valid inferences about student performance on the learning standards contained in the frameworks.

## Item Development Cycle

The MCAS test development process—from the selection of the learning standards that are included in each test to the development of test items (questions) to the production of test booklets—is designed to ensure that test results are valid and reliable. Validity and reliability are enhanced when test results meet two important conditions:

1. The results reflect a student’s knowledge and understanding of the content area being measured and not just the student’s knowledge and understanding of the particular items on one year’s test.
2. The results are representative of the student’s true performance and are not the result of chance or other irrelevant factors.

Keys to meeting those conditions and attaining valid and reliable test results include ensuring that a) there are enough test items to adequately sample from the entire content area and to obtain a solid example of student work, b) test items measure the full range of skills from recall to problem solving, and c) all test items are free of error and bias.

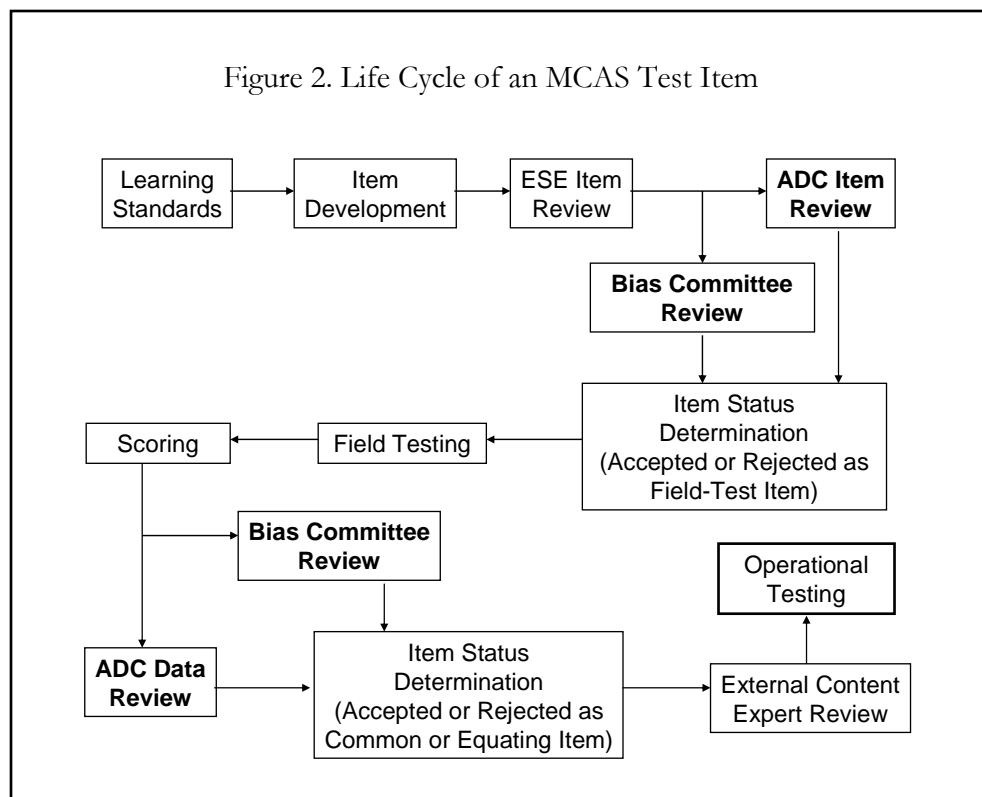
Items undergo extensive review and field testing before they become *common items* on which student results are based. Items are reviewed by an Assessment Development Committee (ADC), a Bias Committee, and content expert reviewers prior to being field tested. Field testing provides data for each item, and these data and the items are reviewed again by the committees prior to items being approved for use as common items.

Separate ADCs exist for each test (e.g., grade 5 Mathematics). They are composed of Massachusetts public school teachers and curriculum specialists who currently teach or are responsible for instruction in the grade and subject area for which items are being developed. ADCs for ELA also assist in the selection of reading passages upon which ELA items are based. The Bias Committee, composed of Massachusetts educators, ensures that no questions appearing on the tests may potentially disadvantage particular groups of students taking the test and no questions will likely

favor one group of students taking the test over another group of students for non-educationally relevant reasons.

As described above, Massachusetts educators play a critical role in the development, selection, and approval of each item included on the MCAS tests. Currently, 257 Massachusetts educators are serving as ADC members for the January to December 2008 term.

After review by the ADCs and Bias Committee, specific items are selected for the next MCAS administration. The item development cycle, from the beginning of item development to the inclusion of an item as a common item on an MCAS test, generally takes 18 to 24 months (see figure 2 below).



## Field Testing

The purpose of field testing questions before they are used to generate student, school, and district results is to ensure that the questions perform well and will contribute to reliability and validity. Field testing yields statistics that help test developers, ADCs, and the Bias Committee determine whether a particular item should be used on an MCAS test.

Statistics gathered from field testing include the following:

- *Item Difficulty*: indicates the proportion of students who answered a multiple-choice or short-answer question correctly and the average score of all students on an open-response

question. If very few students answer an item correctly, it is carefully reviewed for accuracy and appropriateness.

- *Item Discrimination*: indicates how well a test question distinguishes between students who performed well on the test and those who performed poorly on the test. Positive discrimination indicates that students who correctly answered a question received a higher average score on the overall test than students who answered the question incorrectly. Test questions with low discrimination are further reviewed for difficulty, accuracy, and appropriateness.
- *Differential Item Functioning*: indicates whether different groups of students with equivalent total test scores performed differently on a particular test item. For example, if a higher proportion of males than females, who both score at the same level on a test, answer a particular question correctly, the item statistics will capture this information. Differential item functioning statistics are carefully scrutinized by Bias Committee members.

Field testing for MCAS occurs as part of the annual spring MCAS test administration. Items that have been reviewed by the ADC and Bias Committee and approved for field testing are eligible to be embedded within each MCAS test. This type of field testing is designed to ensure that new items are tested under the same conditions in which they will be administered when they are included in MCAS scores for students, schools, and districts.

To enable field testing of new items, MCAS test forms administered to each student consist of both *common* and *matrix-sampled* items.

- **Common** test items are identical in all forms of the test administered at each grade level. Approximately 80% of the questions on any given test form are common questions. All MCAS performance level results for individual students, schools, and districts are based exclusively on common items. Consequently, the performance of each student at each grade is based on responses to the same questions.
- Approximately 20% of items in a test booklet are **matrix-sampled** items. These items differ across the test forms in each subject and at each grade level tested. They serve two purposes: to link the MCAS tests across years (**equating** items) and to field test new questions for use on future tests (**field-test** items).

After field testing, all items are, once again, analyzed and extensively reviewed by the Department, its assessment contractor, and committees of Massachusetts educators prior to being selected as common items used to evaluate student, school, or district performance.

## Test Blueprints

Test blueprints show the number and percentage of each type of test item (e.g., multiple-choice, open-response, and short-answer; common and matrix) that will be included on each MCAS test. They also show the percentage of items (or points) on each test that should be selected from each *Curriculum Framework* content strand. A content strand (or reporting category) is a set of related learning standards such as geometry in mathematics or physical science in science and technology/engineering.

## Scoring MCAS Test Items

A key step in maximizing reliability and validity is to ensure that all MCAS items are scored accurately. The accuracy of the measurement of student performance in a content area is determined, in part, by the accuracy and consistency of scoring.

MCAS tests include a variety of item types: multiple-choice items, short-answer items, open-response items, and writing prompts. Multiple-choice items are machine-scored by comparing student response selections (A, B, C, or D) to an answer key for each test. All other MCAS item types are hand-scored by trained scorers. Numerous steps are taken prior to and during scoring to ensure the accuracy of scoring.

Prior to scoring, benchmarking is conducted. Benchmarking is the process in which open-response and composition items are reviewed using the scoring guidelines and a range of student responses. Anchor sets of student responses are assembled to clearly define the full extent of each score point, including the upper and lower limits. After benchmarking, detailed scoring materials are prepared for each test item, including a scoring guide, samples of student papers representing each score point, training materials for scorers, and qualifying tests for scorers.

MCAS test items are scored at scoring sites across the country by trained scorers hired by the state's assessment contractor. To be eligible to score MCAS test items at grades 3 through 8, a scorer must have completed two or more years of college work in the content area being scored and demonstrate expertise in the content area. To be eligible to score high school MCAS test items, a scorer must have a Bachelor of Arts or Bachelor of Science degree with a concentration in the content area being scored or in a related area.

All scorers must pass a qualifying test to score particular MCAS items. After receiving training on a single item, eligible scorers complete a qualifying test to demonstrate that they are able to accurately score that item. Teams of scorers who have passed the qualifying test independently score student responses to the test item for which they qualified. After scoring of a particular item is completed, scorers are then trained and qualify to score subsequent items.

Throughout the scoring process, the accuracy of scoring is monitored both by trained scoring leaders and through the use of computer programs that generate daily accuracy reports for each scorer. Scorers who do not attain the established threshold for score accuracy are removed from the scoring process for retraining and/or replacement, and any student responses that were scored by such a scorer are rescored. At least 10% of all student responses to each MCAS test item at grades 3 through 8 are scored independently by two scorers. For the high-stakes high school MCAS tests, all student responses are scored independently by two scorers.

Following each test administration, the Department posts on its Web site ([www.doe.mass.edu/mcas/student](http://www.doe.mass.edu/mcas/student)) the scoring guides for each common open-response item and writing prompt, along with a sample of student work at each score point. Also available on this page is a link to the NCS Mentor for MCAS Professional Development Tool, which allows educators to learn how to score student responses to open-response questions and writing prompts.

## **MCAS Test Administration**

The validity of conclusions drawn about student and school performance are also influenced by the conditions under which the MCAS tests are administered. To maintain the integrity of results for all students, it is critical that local schools and districts are adequately prepared to administer the MCAS tests, and that they follow established test administration and security procedures.

### **Test Administration Schedule**

The Department posts test administration schedules to its Web site approximately two years in advance. These schedules indicate the order in which tests must be administered during each test administration window. In addition, the Department mandates specific dates for administering tests that students must pass to earn a CD. In determining the dates for testing windows, the Department considers a broad range of potential scheduling conflicts, including holidays and major educational professional events, by surveying a sample of schools and districts and by consulting with the state's major educational and professional organizations.

Schools may design and implement their particular administration schedule to meet local needs, provided that they administer the test(s) within the specified testing window and comply with all other MCAS requirements.

### **Test Security Requirements**

Strict test security measures are implemented during all phases of the development, production, and administration of MCAS tests in order to maintain the fairness and integrity of the MCAS program. During test development and construction, access to secure materials is limited, and materials are kept in locked storage when not in use. Test booklet drafts, item review sets, and other secure materials are either shipped or faxed to a secure, dedicated machine with limited access. Secure materials and information regarding secure materials are not transmitted via e-mail.

Since the quality and validity of the assessment data generated by MCAS depends, in large part, on uniformity of test administration and security of test materials, each person directly involved in MCAS test administration in schools is responsible for immediately reporting any violation or suspected violation of test security to the school principal, who must report all incidents to the Department.

### **Administration Requirements and Training for School Officials**

MCAS tests are administered in untimed sessions. The guidelines for scheduling test sessions are based on an estimate that each test session requires approximately 45 to 60 minutes, depending on the subject. The Department recommends that schools schedule a two-hour block of time for each test session to allow sufficient opportunity for all students to complete their work. However, schools are allowed to provide additional time, within the school day, to students who require more time to complete a session. The Department recommends that principals schedule a maximum of two test sessions on a single day.

Workshop presentations are offered each fall and spring. The spring workshops, which are offered via teleconference and in face-to-face meetings at regional sites across Massachusetts, focus on test administration procedures for upcoming MCAS testing and address participation guidelines, accommodations policies, test scheduling, test security, special populations, and other topics and

updates related to MCAS testing. The fall workshops, which are offered via teleconference, address issues related to MCAS results.

### **Toll-Free Phone Support for Schools**

Through the MCAS Service Center, the Department’s testing contractor provides schools with year-round, toll-free telephone support from Monday through Friday, 7:00 a.m. to 5:00 p.m. The Service Center staff receives training in all operational aspects of MCAS test administration, including completing online forms, ordering materials, and returning materials. Department content experts and test administration staff are available to answer policy questions Monday through Friday during business hours.

### **Performance Standards**

It is the goal of education reform in Massachusetts as well as the goal of NCLB that all students demonstrate proficiency in selected content areas. A key purpose of the MCAS tests is to determine and report on whether students have attained proficiency on the content standards contained in the *Curriculum Frameworks*. It is not enough, however, to simply report whether student performance is proficient or not proficient. As student performance moves toward proficiency, at any given time, there is likely to be a wide range of student performance within a classroom, in a school, and across the state. To reflect the full range of student performance, results on the MCAS tests are reported in terms of four performance levels: *Warning*<sup>2</sup>, *Needs Improvement*, *Proficient*, and *Advanced*.<sup>3</sup> The four reporting levels capture the varying levels of student performance on the grade-level content standards.

The content knowledge and skills associated with each of these four levels (across subjects and grade levels) is shown below in Table 1.

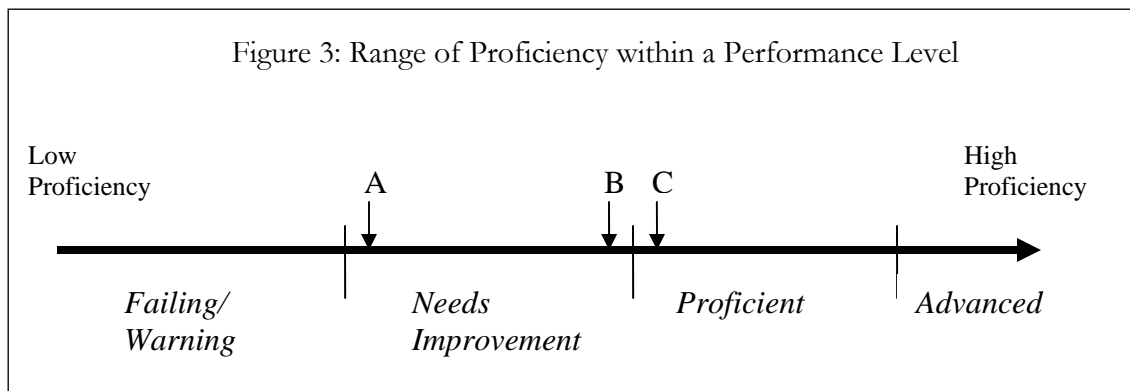
Table 1. General Performance Level Descriptions

<b>Performance Level</b>	<b>General Description of Associated Student Performance</b>
<i>Advanced</i>	Students at this level demonstrate a comprehensive and in-depth understanding of rigorous subject matter, and provide sophisticated solutions to complex problems.
<i>Proficient</i>	Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.
<i>Needs Improvement</i>	Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.
<i>Failing/Warning</i>	Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.

<sup>2</sup> For high school tests, the lowest performance level is labeled *Failing* rather than *Warning*. This label indicates that the student has failed to demonstrate the minimum level of knowledge and skills required for high school graduation.

<sup>3</sup> At grade 3, the highest level of performance is labeled *Above Proficient* rather than *Advanced*. The grade 3 MCAS tests do not contain items that require students to demonstrate the complex knowledge and skills required at the *Advanced* level. At grade 3, students performing at the *Above Proficient* level demonstrate mastery of challenging subject matter and the ability to construct solutions to challenging problems.

At each grade level and content area, the four performance levels describe ranges of performance along the scale. Figure 3 on the next page shows the performance of three fictitious students, Angel (A), Bailey (B), and Cam (C), along the scale. As shown by the figure, each performance level encompasses a wide range of student performance. Angel (A) and Bailey (B) are performing at the *Needs Improvement* level, and Cam (C) is performing at the *Proficient* level. However, Bailey’s level of performance is much closer to Cam’s than it is to Angel’s. In terms of the types of things they know and are able to do in the classroom, we would expect Bailey to be more similar to Cam than to Angel.



## Standard Setting

For a particular content area, such as fourth grade mathematics, the process of establishing the range of test performance that corresponds to each performance level is known as *standard setting*. This process requires a judgment to determine when performance along the scale changes from one performance level to the next. Standard setting is a multi-step process that begins with the establishment of subject- and grade-specific performance level definitions and concludes with the identification of *cutscores* that separate each performance level from the next.

As in the test development process, Massachusetts educators play a central role in setting performance standards for the MCAS tests. First, Department content specialists and Massachusetts educators use the *Curriculum Frameworks* and knowledge of the range of student performance measured by each test to describe specific content knowledge and skills associated with each general performance level definition. The grade-specific descriptions, known as content area performance level definitions, provide lists of the content knowledge and skills that a typical student at each performance level is expected to exhibit.

Second, panels composed primarily of classroom teachers<sup>4</sup> and other Massachusetts educators are assembled to recommend a set of performance level cutscores for each MCAS test. The panels compare the content area performance level definitions to student performance on an MCAS test using a process called the *Body of Work* method of standard setting<sup>5</sup>. In this method, panelists review

<sup>4</sup> Teacher panelists are required to have experience teaching the grade and subject for which standards are being set, as well as familiarity with the learning standards and the Massachusetts *Curriculum Frameworks*.

<sup>5</sup> Because scores on the original grade 3 Reading test were based only on multiple-choice questions, a modified Bookmark method was considered more appropriate for standard setting on that test.

actual (anonymous) student responses from a sample of students with scores ranging from very low to very high on the test. The panelists examine each student’s responses in order to compare the level of performance demonstrated by the student to the content- and grade-specific performance level definitions. Panelists then classify the student’s performance on the test into the performance level category that it most closely resembles. Because not all students will demonstrate all of the content knowledge and skills listed in a particular performance level definition on a single MCAS test, panelists must determine which performance level definition the student’s performance most closely matches.

At the end of the process, the panelists’ ratings are combined to produce a set of recommended performance level cutscores. The Department reviews the panelists’ recommendations along with other relevant information about student performance within and across grade levels. Final performance level cutscores for each MCAS test are approved by the state commissioner of elementary and secondary education.

### Maintaining Performance Standards across Years

Performance standards are set through standard setting the first time an MCAS test is administered and scores are reported. The table below shows, for each current MCAS test, the year in which standard setting took place.

Table 2. Year of Standard Setting for the Current MCAS Tests

Grade	MCAS Test	Year
3	ELA (Reading Comprehension)	2001*
	Mathematics	2006*
4	ELA (Reading Comprehension and Composition)	2001
	Mathematics	1998
5	Science and Technology/Engineering	2003
	ELA (Reading Comprehension)	2006
	Mathematics	2006
6	Mathematics	2001
	ELA (Reading Comprehension)	2006
7	ELA (Reading Comprehension and Composition)	2001
	Mathematics	2006
8	ELA (Reading Comprehension)	2006
	Mathematics	1998
	Science and Technology/Engineering	2003
10	ELA (Reading Comprehension and Composition)	1998
	Mathematics	1998
9/10	Biology	2007
9/10	Chemistry	2007
9/10	Introductory Physics	2007
9/10	Technology/Engineering	2007
* The <i>Above Proficient</i> cutscores for the grade 3 tests were established in 2006 for Reading Comprehension and 2007 for Mathematics		

Whenever performance standards are established, it is important that they remain the same across years, since it would be unfair to hold students and schools to different standards each year. Shifting performance standards would also make it impossible to measure improvement in performance over time.

If the same MCAS test were administered each year, maintaining performance standards would not be a challenge. However, because all MCAS tests are released to the public after being administered, each MCAS test can only be used one time. Even if the test were not released, administering the same test year after year is not an ideal approach. As teachers and students became familiar with individual items on the test, instruction might focus on solutions to particular test items rather than on the general content standards that those items are intended to represent.

Because MCAS tests are different each year, it is necessary to link the tests through a process called *equating*. Equating is used to determine the cutscores on each subsequent MCAS test that will maintain the performance standards established on the original test. If the original and subsequent tests were identical in difficulty, the cutscores would be the same on each test. However, it is very challenging to develop two tests that are exactly equal in difficulty. Each subsequent test is likely to be slightly easier or slightly more difficult than the original test. If a subsequent test is easier, the cutscores needed to move from one performance level to the next will be slightly higher. If a subsequent test is more difficult than the original, the cutscores needed to move from one performance level to the next will be slightly lower. In general, the cutscores on MCAS tests change by no more than one or two points in a year.

For MCAS, tests are equated each year using an *anchor test* that is embedded within the actual MCAS test. The items on the anchor test do not count toward a student's score on the test. The items are only used to link performance standards on the original and subsequent MCAS tests. Items on the anchor test are not released to the public after each test administration so that they can be used again to link the tests. The anchor test measures the same content and skills measured by the actual MCAS test.

Because of the critical importance of maintaining the same performance standards across years, the Department takes extra precautions to ensure the accuracy of the equating process. Each year, psychometricians from the Department's assessment contractor and the Center for Educational Assessment at UMASS Amherst independently and simultaneously equate the MCAS tests. The Department, with assistance as needed from its national Technical Advisory Committee, analyzes the results of the two independent equating analyses prior to reporting any MCAS test results.

## Reporting MCAS Results

MCAS results are reported in a variety of formats and modes using several different types of scores. Various MCAS reports are produced on paper, posted electronically (either to school officials through a secure portal or on the Department's website for public reports), or delivered as data files to be used for further analysis by local school personnel or educational researchers. The primary student and school results reported on the MCAS tests are performance levels and scaled scores. Additional item statistics, such as the percent correct or percentage of students selecting each option on a multiple-choice item, are reported for each MCAS test item. The reason for producing multiple

reports with multiple scores is to provide results that are useful for a variety of audiences and purposes.

## MCAS Scaled Scores

As described previously, performance levels describe how well a student performs in relation to the *Curriculum Framework* learning standards. Results on the grades 4–10 MCAS tests are reported according to four performance levels: *Advanced*, *Proficient*, *Needs Improvement*, and *Warning/Failing*. Because performance levels encompass a range of student performance, results for grades 4–10 are also reported as scaled scores, which range from 200–280 and are linked to particular performance levels. At grade 3, results are reported according to four performance levels—*Above Proficient*, *Proficient*, *Needs Improvement*, and *Warning*. Because the grade 3 performance levels are different from other MCAS tests, only raw score points rather than scaled scores are reported at grade 3.

The use of scaled scores assists in the interpretation of MCAS results across tests and years. In particular, scaled scores provide information about the location of a student’s performance within a performance level, as shown in Table 3 below.

Table 3. Scaled Score Ranges and Corresponding Performance Levels

Scaled Score Range	Performance Level
200–218	<i>Warning/Failing</i>
220–238	<i>Needs Improvement</i>
240–258	<i>Proficient</i>
260–280	<i>Advanced</i>

## Reports and Interpretive Material

Annual reports include individual student results produced for parents/guardians, student rosters for schools and districts, summary reports of school and district results, reports of state results across MCAS tests, and reports of the performance of subgroups of students based on factors such as race/ethnicity and gender. The data files and spreadsheets can be imported into software packages for additional analysis or imported into reporting software packages for customized displays of the results.

In addition to the reports themselves, the Department produces interpretive material to support the proper interpretation and use of MCAS results ([www.doe.mass.edu/mcas/results.html](http://www.doe.mass.edu/mcas/results.html)). School- and district-level results, including results of individual MCAS test items, are available to the public through the School and District Profiles (<http://profiles.doe.mass.edu/>).

The most widely produced printed MCAS report is the *Parent/Guardian Report* that is distributed by schools and districts to the parents/guardians of all students participating in MCAS testing. The *Parent/Guardian Report* contains MCAS test results for the individual student and displays graphically the range of scores the student might receive if he or she took the test many times. The report also compares the individual student’s results to school, district, and state results. To accompany the *Parent/Guardian Report*, the Department produces an interpretive guide for parents/guardians that provides background information on the MCAS tests as well as descriptions and explanations of all results contained in the report. The interpretive guide is translated into the ten languages most commonly spoken at home by students whose first language is not English, as identified through the

Department's data collection system (Student Information Management System). In 2007, those languages were Arabic, Cape Verdean, Chinese (Traditional), Haitian Creole, Khmer, Korean, Portuguese, Russian, Spanish, and Vietnamese.

### **Annual Release of MCAS Test Items and Sample Student Work**

Since the first MCAS administration in 1998, a major part of the reporting of MCAS results has been the release of all common test items included on each MCAS test. Also released with the test items are scored samples of actual student responses to the constructed-response test items. Released items and sample student work are posted to [www.doe.mass.edu/mcas/testitems.html](http://www.doe.mass.edu/mcas/testitems.html) and are available to download.

The annual release of all common test items serves several purposes intended to enhance validity. Most importantly, the release of all common items is the centerpiece of the Department's policy to make the assessment and performance expectations as transparent as possible. The items also enable school teachers and administrators to more effectively interpret MCAS performance within a year and across years. A review of actual test items over time assists in the evaluation of the type of content knowledge and skills on which the school is consistently performing well, and also those with which the school has difficulties. Also, the release of all MCAS test items is intended to support the broadening of instruction by demonstrating the depth and breadth of knowledge and skills assessed on the MCAS tests.

A major use of released student work is in professional development activities. Practice in using scoring guidelines to evaluate student compositions and written responses helps teachers improve instruction in their classrooms. They can also use released questions and responses to help students prepare for the tests.

Since the beginning of the program in 1998, more than 5,000 MCAS test items have been released. Print versions of released tests are available for all years since 1998. In addition, released items are maintained for five years in the Department's electronic search tool, which currently allows users to browse and search through a database of over 3,000 MCAS test items.

### **Test Reporting and Analysis Software**

Since the early years of the MCAS program, local school districts have used a variety of test reporting and analysis software packages to conduct their own analyses of MCAS results. When it was discovered that the majority of districts in the state had selected the same software package on their own, the Department purchased a statewide license to make the tool available to all school districts free-of-charge. Ongoing development of the state's online data warehouse system includes additional tools that enhance the capability of local school districts to evaluate and analyze the performance of their students.

The purpose of supporting districts' use of data analysis tools is to enable local school officials to track student performance across years, and most importantly, include other indicators of student or school performance in their analyses.

## Student Participation in the MCAS Tests

Ensuring that all students have access to quality instruction and high standards is a fundamental principle of education reform in Massachusetts and across the country. Monitoring student achievement and holding schools and districts accountable for all students is one key part of the social compact established by the Education Reform Law and reinforced by the No Child Left Behind law. These processes help confirm that **all** students are tested, regardless of disability, English language proficiency, or other factors.

It is equally important, however, that each individual student is able to participate in MCAS in a way that is most appropriate for him or her. To ensure that as many students as possible are able to participate in the general MCAS tests, numerous test accommodations are allowable. A student with a disability who is unable to participate in the standard MCAS tests, even with accommodations, is assessed via the MCAS Alternate Assessment (MCAS-Alt).

For each student with a disability, the student's IEP or 504 team determines during its annual meeting how the student will participate in MCAS for each subject scheduled for assessment. The team decides which test accommodations, if any, are appropriate for the student—generally these are accommodations that the student normally receives during classroom instruction and testing. As stated above, the team may also decide that it is most appropriate for the student to participate in MCAS through the MCAS-Alt.

### Standard and Nonstandard Accommodations

Based on the recommendation of the IEP or 504 team, a student with a disability is eligible to participate in the standard MCAS tests using appropriate test accommodations. Standard test accommodations are defined as changes in the routine conditions under which students take MCAS tests. Standard accommodations are grouped into the following four categories:

- changes in timing or scheduling of the test; for example, administering the test in short intervals or at a specific time of day
- changes in test setting; for example, administering the test in a small group or a separate setting
- changes in test presentation; for example, using a large-print or Braille edition of the test
- changes in how the student responds to test questions; for example, dictating responses to a scribe

Nonstandard test accommodations are defined as those which change the way an MCAS test is presented or change the way a student responds to test questions, and may alter a portion of what the test is intended to measure.

The Department publishes a list of over two dozen commonly used standard test accommodations and a half dozen commonly used nonstandard test accommodations, and IEP teams may, under certain conditions, suggest additional accommodations not listed.

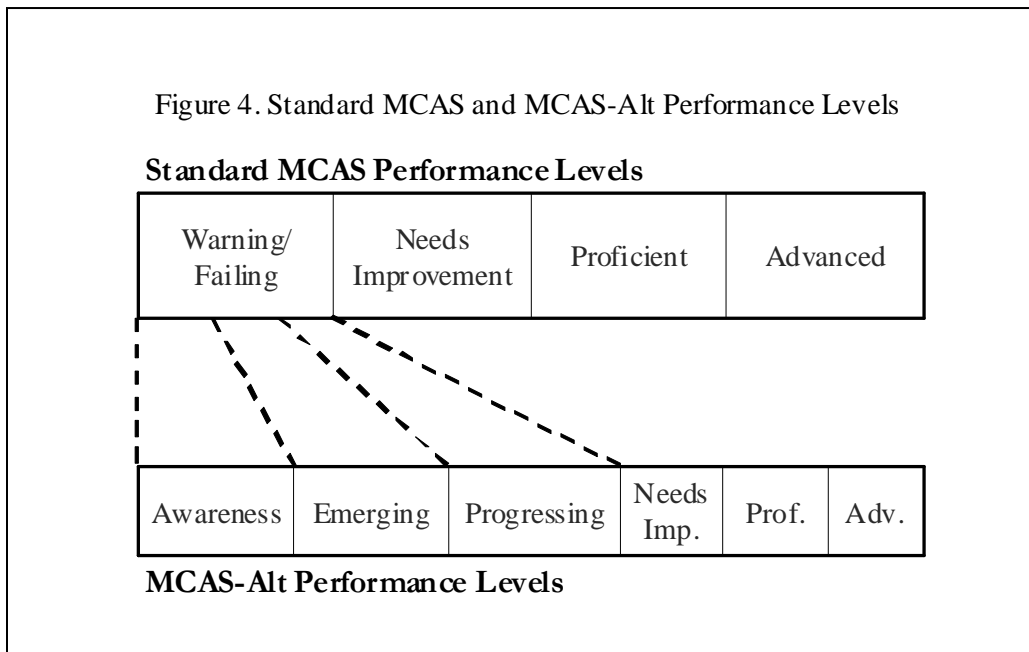
In most large-scale assessment programs like MCAS, the most commonly used test accommodation is to allow extra time for testing. However, because there are no limits on the time allowed for the administration of MCAS tests for any student, an accommodation to allow extra time is not needed.

## MCAS Alternate Assessment (MCAS-Alt)

Even with the allowable accommodations, a relatively small number of students are unable to participate in the standard MCAS tests. In general, these students account for approximately one to two percent of all of the students at each grade level. Many of these students have significant cognitive disabilities that prevent them from achieving the same grade level performance standards as other students. Some students, however, may have physical or other disabilities that prevent their participation in a paper-and-pencil test. Instead, these students participate in MCAS through the MCAS-Alt.

The MCAS-Alt is a portfolio assessment program. Consistent with the standard MCAS tests, the MCAS-Alt measures the student's performance and achievement of the learning standards in the Massachusetts *Curriculum Frameworks*. However, in the case of the MCAS-Alt, the content of the assessment is customized by the student's teacher to meet the individual student's instructional program and needs. For each student who is scheduled to take the MCAS-Alt in one or more subjects, the student, the student's teacher, and other adults who work with the student develop a portfolio over the course of the school year. The student's portfolio must include data and other evidence of the student's performance in the subject being assessed.

The performance levels for students who take the MCAS-Alt include *Awareness*, *Emerging*, and *Progressing*; these levels are contained within the *Warning/Failing* level, as shown below:



## MCAS and High School Graduation

Although this document focuses on the MCAS tests rather than on the uses of MCAS test results, it would be incomplete without a discussion of the use of MCAS test results as part of the high school graduation qualifying process. This is arguably the most highly publicized use of MCAS test results and represents the area of the program where emphasis is on **student** accountability rather than on

school and district accountability. Because of the high stakes use of results for individual students, additional steps are taken to ensure reliability and validity. As described in previous sections, there are institutionalized enhancements in the scoring and administration of the high school tests designed to increase reliability and validity. In addition, there are multiple opportunities and methods provided for students to demonstrate their level of performance.

### **High School Competency Determination Requirement**

The Massachusetts Education Reform Law of 1993 mandates that, as one condition for earning a high school diploma, students must demonstrate “competency” in the areas of mathematics, science and technology/engineering, history and social science, foreign languages, and English. The Competency Determination (CD) is based on the standards for high school students contained in the *Curriculum Frameworks* and must reflect a determination that a particular student has demonstrated mastery of a common core of content knowledge, skills, and competencies, as measured by the state assessment. The current passing standard required to earn a CD is a scaled score of 220, or the very bottom of the *Needs Improvement* level. Beginning with the graduating class of 2003, students were required to demonstrate competency in ELA and mathematics in order to be eligible to earn a high school diploma. Beginning with the graduating class of 2010, students must also meet the CD standard in science and technology/engineering in order to be eligible for a high school diploma. Current plans call for history and social science to be added to the CD requirement beginning with the graduating class of 2012.

In 2006, the Board voted to amend the regulation governing the CD. The amendment changes the criteria for earning a CD by requiring that students, beginning with the class of 2010, either meet or exceed a scaled score of 240 on the grade 10 ELA and Mathematics tests, or meet or exceed a scaled score of 220 on both tests and fulfill the requirements of an Educational Proficiency Plan.

The Education Reform Law identifies the CD as one of several requirements for earning a high school diploma. In addition to demonstrating competency on the tenth grade standards, students are required to successfully complete eleventh and twelfth grade and meet all local school district requirements for graduation. By design, the tenth grade CD should be the first and lowest hurdle that students must clear on their way toward earning a high school diploma. The tenth grade standards represent the Commonwealth’s minimum level of performance that all students graduating from high school are expected to attain across multiple content areas. It is fully expected that virtually all students completing the twelfth grade and earning a high school diploma will progress beyond the tenth grade standards, especially in the content area(s) in which they have demonstrated an interest for future study and employment.

### **Retesting Opportunities**

Students who fail to earn the passing score to meet the CD requirement in a particular content area are provided multiple opportunities to retake the test in that content area. Students must only retake the MCAS test in those content areas in which they have not met the passing score for a CD. For example, a student who has met the CD requirement in ELA and STE, but not in Mathematics, is only required to take the Mathematics test again.

For ELA and Mathematics, there are two regularly scheduled retest administrations each year in the fall and spring. Prior to their scheduled graduation in twelfth grade, students have at least five

opportunities to earn a CD through the MCAS tests. However, more than 90 percent of students earn their CD by the second retest.

There is no limit on the number of times that a student may participate in MCAS retests. Individuals may continue to participate in MCAS retests to earn a CD and high school diploma after they have completed high school. On the whole, as the percentage of students meeting the CD requirements on the tenth grade test has increased steadily since 2001, the overall number of students who require multiple retesting opportunities has declined accordingly.

In STE, students are required to meet the passing score on one of the four content area tests offered beginning in grade 9. Students who do not meet the passing score have two options for retesting. In subsequent years, students may take a test in the same content area or take a test in one of the other three areas in which MCAS high school STE tests are available. As in ELA and Mathematics, there is no time limit or limit on the number of times a student may participate in MCAS testing for the purpose of earning a CD in STE.

### **Performance Appeals**

Taking and earning the required score on the high school ELA, Mathematics, and STE MCAS tests is the primary method by which a student meets the CD requirement. A small number of students, however, are unable to demonstrate that they have the requisite knowledge and skills in a content area through the MCAS tests. These students have the option to demonstrate competency in one or more content areas through a performance appeal.

If after multiple, unsuccessful attempts to earn a CD through the MCAS tests, the student's school and district believe that there is compelling evidence that the student possesses the required level of knowledge and skills to earn a CD, the school and district may submit a performance appeal to the Commissioner of Elementary and Secondary Education on behalf of the student. The performance appeal is not a waiver of the CD requirement, but an alternate means of demonstrating competency. No student who can demonstrate competency in the required content areas will be denied a diploma based solely on performance on the MCAS tests.

More information regarding performance appeals can be found online at <http://www.doe.mass.edu/mcasappeals>.

### **Additional Resource**

Further information regarding the MCAS program, including full details and appendices of interest to psychometricians and data analysts, is available in the *2007 MCAS Technical Report*, which is posted at <http://www.doe.mass.edu/mcas/results.html?yr=2007>.