STATE TESTING
AND GRADUATION REQUIREMENTS:
WHAT CAN MASSACHUSETTS
LEARN FROM OTHER STATES?

Faculty Leader:
Dr. Kathryn A. McDermott
Associate Director, Center for Education Policy

Guest Presenters:
Dr. Catherine Clark
Director for Research and Policy
Charles A. Dana Center, University of Texas, Austin

Dr. Herman “Bud” Meyers
Deputy Commissioner for Standards and Assessment
Vermont Department of Education

Ms. Roseanne DeFabio
Assistant Commissioner for Curriculum, Instruction, and Assessment
New York State Education Department

Dr. Betty Sternberg
Associate Commissioner for the Division of Teaching and Learning
Connecticut State Department of Education

Center for Education Policy
School of Education
University of Massachusetts, Amherst
CEP Paper 2001-02
Acknowledgements

Dr. McDermott and the Center for Education Policy would like to thank a number of people for their many contributions to the success of this conference and to the policy analysis that provided background for it.

Funding for the conference was provided by the Noyce Foundation and Verizon. Bailey Jackson and Andrew Effrat provided leadership support in developing the conference. Catherine Clark, Roseanne DeFabio, Herman “Bud” Meyers, and Betty Sternberg presented the approaches of Texas, New York, Vermont, and Connecticut, respectively. Joseph Berger, John Carey, Lenore Carlisle, Andrew Churchill, and Andrew Effrat helped facilitate the conference. Students Susan Bowles, Maria Calkins, Carolyn Lee-Davis, Jay Rideout, and Debbie Smith, and staff member Catherine Brooks, provided conference staffing and support.

Students Jay Rideout and Carolyn Lee-Davis provided research assistance on the policy analysis. Andrew Churchill assisted with the conceptual framework and revision of the policy analysis, managed the organization of the conference, and edited the proceedings.

Joseph Berger, Catherine Clark, John Carey, Roseanne DeFabio, Andrew Effrat, Bailey Jackson, Penny Noyce, Herman “Bud” Meyers, Marjorie Petit, Stephen Sireci, and Hariharan Swaminathan provided helpful feedback on drafts of the policy analysis.

About the Center for Education Policy

The Center for Education Policy in the School of Education of the University of Massachusetts Amherst was developed to connect the resources of the University to the major education policy issues of the day, with the dual goals of improving public decision-making and enriching scholarly activity. The Center works in an independent and non-partisan way to fulfill the University’s land-grant mission of public service and outreach, with a particular focus on state-level issues of importance in Massachusetts and nationally.

Through conferences and seminars, policy studies and research projects, and the dissemination of information throughout the policy community, the Center focuses attention on the important education policy questions facing decision-makers and helps to achieve greater impact and cohesion in the search for effective policies. Policymakers interested in expert assistance, and faculty and graduate students interested in conducting education policy studies, are encouraged to contact the Center for Education Policy.

For additional information, contact:

Andrew Churchill, Research Coordinator
Phone: 413-545-0958 Fax: 413-545-3855 Email: achurchill@educ.umass.edu

Center for Education Policy
250 Hills South, University of Massachusetts
813 North Pleasant Street
Amherst, Massachusetts 01003-9308
# Table of Contents

Summary .......................................................... 5

Introduction ......................................................... 7

The Analytical Framework ......................................... 8

Massachusetts in National Perspective ......................... 9

Texas .................................................................. 11

Vermont ................................................................. 13

New York ............................................................... 14

Connecticut ............................................................ 16

Implications for Massachusetts ................................. 17

General Discussion .................................................. 19
Table of Contents (continued)

Supplemental Materials 24

General 25
• Conference Agenda
• List of Registrants
• High School Testing and Graduation Requirements Analytical Framework
  Kathryn McDermott

Texas 33
• Texas Testing and Graduation Requirements: The Framework
  Catherine Clark
• The Evolution of High-Stakes Testing in Texas
  Catherine Clark
• Critiques and Commendations for the Texas System
  Catherine Clark

Vermont 41
• Vermont State Board of Education Policy on Testing and Accountability
• Vermont Comprehensive Assessment System (table)
• Part of a Larger System
• Overview of Vermont’s Accountability System (chart)
• Design Decisions

New York 48
• State Testing and Graduation Requirements – How New York Fits Into the Framework
  Roseanne DeFabio

Connecticut 50
• State Testing and Graduation Requirements – Connecticut
  Betty Sternberg
• Position Statement on Measuring Success
  Connecticut State Board of Education
Summary

On February 2nd, 2001, a group of Massachusetts education policymakers and stakeholders gathered at a conference organized by the Center for Education Policy of the School of Education at the University of Massachusetts Amherst. The conference focused on testing policies and graduation requirements in Massachusetts and four other states: Connecticut, New York, Texas, and Vermont.

The conference introduced a new Center for Education Policy study of testing and graduation requirements, based on the idea that Massachusetts needs to look beyond a simple “MCAS: Yes or No” formulation and examine the multiple, concrete policy areas in which state testing approaches can and do vary. The five policy areas examined in the study are as follows:

- **Standards.** What do we want students to know and be able to do? Basic skills, higher-order thinking skills, or mastery of content areas?

- **Assessments.** What kind of assessments do we use to determine whether students are meeting the standards? Who determines competency—state or local authorities? How are special populations treated?

- **Uses and Consequences of Assessments.** Is an assessment “high-stakes” for students (e.g., a graduation requirement) or “low-stakes” (e.g., assessment results used only for diagnostic purposes)? What are the stakes for schools, teachers, and administrators?

- **Supports for Learning.** How are students helped to learn what they need to know and be able to do? What special interventions are targeted toward students at risk of failure? What retest opportunities exist?

- **Implementation Strategies.** How are standards, assessments, consequences, and supports established, sequenced, and phased in? How have states changed their implementation timelines, and why?

Attendees heard from senior-level representatives of four states that combine these components in different ways. For example, Connecticut has relatively high standards, but its assessments are not high-stakes in the sense of students being required to pass a state test in order to graduate (state test results are printed on each student’s transcript, however). Texas, on the other hand, initially set its standards relatively low, but required all high school students to pass its exit test to graduate. New York, like Massachusetts, is moving to a high-standards, high-stakes approach for all students. Vermont has a combination of state and local assessments, including a portfolio of student work over time, none of which must be passed before a student may graduate.

The discussion raised a number of key policy issues:

- Communication about the implementation of Education Reform has not been sufficient in Massachusetts. Participants who play a variety of roles in the state educational community, and who disagree on many of the key issues, generally seemed to agree on this point.

- The states represented had all phased in their reforms over a period of several years or more. Connecticut began with standardized tests in grades 4, 6, and 8 for almost a decade before adding its 10th-grade test. Texas began with a relatively easy test, then made it more challenging when it was aligned to new, more rigorous learning standards. New York built its Regents-for-all approach on the Regents Examinations’ long history.
• Lack of a graduation requirement doesn’t necessarily mean that a state’s assessment lacks high stakes. Several presenters mentioned the pressure schools and districts feel due to media reporting of test results, regardless of whether students are required to pass a test to graduate.

• Local districts and teachers are involved in important ways in state testing strategies. In Vermont, the state’s role is shared with a locally designed assessment component. New York uses local teachers to grade its tests. Several participants noted that a state test, like all state education efforts, depends on local people for its implementation.

• States are developing accountability measures for both children and the adults who teach them. There was significant discussion about what kinds of assessments should be used for each, what the local versus state roles should be in each area, and how these accountability efforts should be sequenced.

• In implementing standards-based assessments, timelines generally have been longer than expected. Participants mentioned the need to get curriculum and instruction aligned with standards and then give students time to absorb the revised instruction before tests will have validity as high-stakes instruments.

Despite ongoing controversy within the state over MCAS as a graduation requirement, presenters from several of the states spoke highly of Massachusetts’ standards and test, saying that they were models of non-multiple-choice assessment. But they also warned of the difficulty of implementing an assessment system if local teachers and administrators are not supportive.

Ensuring that all students, regardless of race, gender, ethnicity, or parental socioeconomic status, attain high levels of academic achievement is a crucial but novel task. Early in the twentieth century, only a minority of Americans completed high school. More recently, high school graduation became the norm, but only the college-bound students were expected to attain high academic standards.

Massachusetts is now striving to bring all students to desired levels of knowledge. Several conference participants suggested that this may mean reconsidering implementation timetables and/or looking again at the difficulty of the tests. However, these are matters for further discussion. By breaking the monolithic question of “MCAS: Yes or No” down into more specific tradeoffs between the areas of standards, assessments, consequences, supports, and implementation strategies, conference participants began utilizing a shared vocabulary for that discussion.
Introduction

These are the proceedings of a policy leadership conference, sponsored by the Center for Education Policy of the School of Education at the University of Massachusetts Amherst, which took place on February 2, 2001, in Needham, MA. In response to policymakers’ and education stakeholders’ interest in learning more about how state educational testing issues are being approached in other states, the Center organized an event bringing together (1) leading Massachusetts policymakers and stakeholders, (2) senior-level representatives from four states with diverse approaches to testing and graduation requirements (Connecticut, New York, Texas, and Vermont), and (3) an analytic framework that would enable participants to consider Massachusetts’ testing and graduation issues in new ways. The purpose was to provide an opportunity for key policy leaders to discuss and lay the groundwork for future collaboration on these issues within the Commonwealth.

Dr. Kathryn McDermott, a faculty member at the University of Massachusetts Amherst and Associate Director of the Center for Education Policy, conducted a new, comparative study analyzing state testing policies across the nation. This study provided the analytic foundation for the conference and examined how various states use a number of different components in their reform strategies, including the following elements: standards (e.g., content mastery or basic skills); assessments (i.e., how attainment of standards is measured); consequences (e.g., high or low, school or student); support programs (i.e., how students are helped to learn and to pass the tests); and implementation strategies (e.g., how tests and stakes may be phased in).

The conference began with opening remarks from Dean Bailey Jackson of the University of Massachusetts, Amherst, School of Education and from Professor Andrew Effrat, Chair of the Department of Educational Policy, Research, and Administration and Director of the Center for Education Policy in the School of Education.

The remainder of the conference had three stages. First, all participants heard Dr. Kathryn McDermott’s presentation of her study’s analytical framework for understanding controversies about testing, graduation requirements, and state education reform. Continuing the opening session, panelists from Texas, Vermont, New York, and Connecticut gave brief presentations on their states’ policies. In the second session, participants divided into four smaller groups for a more in-depth discussion of how the analytic framework illustrated testing policy choices in a particular state. In the final session, the entire group reconvened for reports from the small groups and general discussion of the implications of the analytic framework and the different approaches of other states for Massachusetts.

In these Proceedings, a summary of the analytical framework and how Massachusetts fits into it is drawn from the Executive Summary of Dr. McDermott’s paper, State Testing and Graduation Requirements: A Framework for Analysis. A summary of the information presented at the conference about each state follows, compiled from the panelists’ remarks and the notes from the small-group sessions. The final section reports on the general discussion that followed the small-group work, along with Dr. McDermott’s concluding remarks on major themes and challenges.
The Analytical Framework

Dr. Kathryn McDermott, Associate Director of the Center for Education Policy, presented her framework for analyzing how state policies on testing and graduation requirements vary and how Massachusetts fits into that framework.

The goal of the analytical framework and of the conference is to examine various dimensions of state testing policy separately, with reference to the experiences of other states that have assembled all or some policy elements differently from Massachusetts in their pursuit of standards-based reform. The MCAS test and its high-stakes use reflect decisions made along five policy dimensions:

1. **Standards.** What are the state’s expectations in terms of what students should know and be able to do? Is the focus on basic skills, higher order thinking skills, or mastery of content areas? Which content areas are covered by standards?

2. **Assessments.** How do we know if students are meeting the standards? What kinds of test format and test items are used? Are tests connected to specific courses (like Algebra) or more general subject areas (like Mathematics)? Is there a single test or a range of assessments including portfolio and performance elements? Who determines competency—state or local authorities? What alternate assessments and accommodations are used for special education students and limited English proficient students? How are students in vocational-technical programs assessed?

3. **Uses and Consequences of Assessments.** What are the stakes, and for whom? Is an assessment high stakes for students (e.g., a graduation requirement) or low stakes (e.g., assessment results used only for diagnostic purposes)? Are assessment results used to measure student achievement, school performance, or both? What are the stakes for schools, teachers, and administrators? Do standards and assessments lead to a single diploma or multiple credentials?

4. **Supports for Learning.** How much progress has been made in implementing standards-based reform and teaching students the material and skills in the standards? What types of programs exist to help students learn what they need to know and be able to do? What special interventions are targeted toward students at risk of failure? What retest opportunities exist?

5. **Implementation Strategies.** How are standards, assessments, consequences, and supports established, sequenced, and phased in? What is the time lag between establishment of standards and enactment of assessments, and between implementation of the tests and imposition of consequences? Is there an initial focus on testing in a limited number of disciplines, which may be expanded later? Is the passing score for the assessment started low and ratcheted up? What types of changes have states made in their implementation timelines, and for what reasons?

This framework is intended to provide policymakers and stakeholders with a logical and usable vocabulary that goes beyond all-or-nothing, pro or anti positions on standards and testing. By breaking these policy questions down into five components, stakeholders may find that they are able to discuss reform issues in a more precise and nuanced way, and possibly find new areas on which they can agree and move forward.
Massachusetts in National Perspective

Standards. Massachusetts stakeholders who disagree on many other issues agree that it is crucial to educate all students to a high level. Controversy arises, however, over the standards to which students should be held as a criterion for graduation.

In the early days of standards-based reform, many states enacted graduation tests that assessed students’ grasp of basic skills. In some cases, students could take and pass these tests as early as eighth or ninth grade. More recently, the trend has been toward requiring students to reach higher levels of performance, such as mastery of material and skills identified as being at a tenth-grade level. Of the eighteen states that currently have graduation tests, only five identify the test as being at a tenth-grade level. By 2008, twenty-four states (including Massachusetts) will have graduation tests, of which twenty-one, including MCAS, will be at the tenth-grade level.

“Tenth-grade level” is a relative rather than an absolute standard. Based on comparisons of state graduation test pass rates and National Assessment of Educational Progress (NAEP) scores, the material tested by the MCAS appears to be more demanding than current tenth-grade level tests in other states.

Assessments. Backed by federal laws such as the 1994 reauthorization of the Elementary and Secondary Education Act, the Individuals with Disabilities Education Act, and the Perkins Vocational and Technical Education Act, all states include students with disabilities, students whose first language is not English, and students in vocational-technical programs in their assessments. State assessment systems vary along several dimensions:

1) The balance between machine-scorable test items and “constructed response” items that require students to write an essay or show how they solved a problem;
2) Whether or not the examinations are connected to, and given at the end of, specific courses;
3) Whether or not a portfolio or other form of “authentic assessment” is included;
4) Whether the way in which students’ mastery of curriculum frameworks will be assessed is determined only by state authorities or by both state and local authorities;
5) The way in which the test is administered to special education students, students whose proficiency in English is limited, and students in vocational and technical programs.

In its format, MCAS is consistent with most other states’ testing programs. A few states currently use portfolio or performance assessments, but not as high-stakes tests for students because of concerns about the consistency of how they are scored. Accommodations on the MCAS and the alternate MCAS for special education students adhere more closely to the state curriculum frameworks than similar measures in some other states, but not unusually so.

The format of the MCAS, the inclusion policies, and the alternate assessment for special education students are broadly consistent with what is done in other states.

Uses and Consequences of Assessments. States currently use their assessments for a variety of purposes, including evaluating students, holding schools accountable, and sometimes rewarding or sanctioning individual teachers and administrators.

Nearly all states test their high-school students’ mastery of skills and academic subjects, but at this point only about half require—or plan to require—their students to pass a test in order to graduate.

Controversy over high-school graduation tests has become more intense as more states have enacted such tests and as the material on those tests has become more difficult.

States with graduation tests generally award two different high school credentials: a diploma for students who pass the test, and a “certificate of completion” for students who complete all the other
requirements but do not pass the test. A few states provide alternate means by which students who do not pass the tests may nonetheless earn a diploma.

States are experimenting with a few kinds of rewards for high test scores, either in addition to or instead of attaching sanctions to failure. These include scholarships to state universities and “honors diplomas.” However, distributing benefits such as financial aid for higher education, on the basis of test scores is likely to produce results that will strike many people as discriminatory, since students in different racial, gender, and economic groups have not been equally likely to score well on the tests.

**Supports for Learning.** All states with graduation tests allow students several opportunities to retake the tests prior to their scheduled graduation dates, and most enable students to continue retaking the test if necessary after they leave twelfth grade.

Critics of MCAS as a graduation requirement often claim that it is too soon to hold students responsible for mastering the material in the Curriculum Frameworks because schools have not fully aligned their curricula and instructional practices with the frameworks.

Massachusetts is not the only place in which this criticism has been made. The complete vision of standards-based school reform includes changes in what is taught and how it is taught, not just implementation of tests that assess mastery of the standards. In the long run, the tests are meant to drive a fundamental improvement in the quality of public schools in all communities. Nationwide, states have tended to implement testing requirements before fully implementing the other standards-based reforms, rather than the other way around. The idea is that the test requirement will create an impetus for the other changes.

In addition to more fully implementing standards-based reform, states have enacted a variety of shorter-term support programs intended to help students pass the tests. The most successful involve certified educators providing extensions or supplements to the regular school program, rather than volunteer tutors. Charges that some schools have improved their students’ apparent performance levels by teaching low-level skills and retaining students who are likely to fail the exam until they reach the legal dropout age remain controversial.

**Implementation Strategies.** Many stakeholders in Massachusetts have argued that the state should delay the use of MCAS as a graduation requirement in order to give schools more time to implement standards-based reforms. Several states have adjusted their schedules or changed their assessment plans in the course of implementation, often in order to give the other components of standards-based reform time to work. Massachusetts itself has already deviated from the schedule initially set out in the Education Reform Act of 1993.

Changes in the plan for implementing a policy are a sign neither of surrender nor of bad faith. It is often necessary to be flexible as a policy is put into effect, in particular when implementation involves a number of actors at state and local levels.

Separating the various dimensions of testing and graduation policies can help identify what kinds of compromises are consistent with the broad consensus that high standards are desirable for all students. Some states, like Connecticut, typically do well in national comparisons despite not having a high-stakes test. Texas and New York provide examples of different kinds of high stakes tests. To date, the Texas Assessment of Academic Skills (TAAS) at the exit level has been a high-stakes test of basic skills. This will change in 2003 with implementation of the TAAS-II exit-level test in eleventh grade. The new test will cover more subjects and be aligned with higher academic standards. New York is transforming the Regents Examinations, recognized as fairly challenging tests of content mastery, from tests taken only by the college-bound into high-stakes tests for all students. The timetable according to which this is happening is different, and in some sense slower, than that of Massachusetts. Vermont’s long experience with a system of different kinds of low-stakes assessment
provides perspective on the decision in Massachusetts to use only a standardized test for school and student accountability.

## Texas

Information about testing and state policy in Texas was presented by Dr. Catherine Clark, the Director for Research and Policy of the Charles A. Dana Center at the University of Texas, Austin.

**Standards.** Initially, the goals of education reform in Texas were to ensure that students were mastering basic skills and that gaps in this mastery between students in different racial and ethnic groups were narrowing. Original exit-level TAAS tests covered eighth and ninth grade skills and content established in the “essential elements.” Recently, Texas established new academic standards called the Texas Essential Knowledge and Skills (TEKS), and the Texas Education Agency has linked TAAS to the TEKS. As a result, the tests are now more challenging. A further challenge awaits high school students in 2003 when the TAAS-II exit examination is administered to students in eleventh grade. The test will include language arts, writing, algebra, geometry, biology, and integrated chemistry and physics.

Local districts decide on curriculum and instruction that will help students master TEKS, and they may select textbooks that the state has identified as “conforming” to TEKS. If the district selects a conforming text, the state pays the full cost. If the district selects a “nonconforming” textbook (defined as one that covers 50 percent or more of the TEKS) then the state pays for part of the textbook. Districts may elect to purchase other textbooks or materials using their own operating funds.

**Assessments.** TAAS is a battery of tests given to students in grades three through eight and again in high school. Reading and math are administered in every grade and writing is administered in grades four, eight, and ten. Science and social studies tests are administered to eight grade students. TAAS is designed to measure whether or not students have mastered the TEKS. The form of the examinations is mostly multiple-choice. The writing test has an essay component. In addition to serving as assessments of individual student performance, the TAAS tests are part of the Texas accountability system that results in ratings for school districts and individual campuses.

Texas also administers end-of-course examinations to high school students in Algebra I, Biology, American History, and English III. Students do not need to pass these tests to graduate or to get course credit, and these tests are not used to establish accountability ratings.

Early in its implementation, some critics claimed that TAAS included obscure information or was part of an “outcomes-based education” program that was violating student and family privacy. To stem this controversy, the state began releasing each year’s tests to the public, even though the practice increases the overall cost of the assessment program.

Students identified for special education take TAAS unless the IEP exempts the student. Roughly seven percent of students are exempted for reasons of special education. In 2003, Texas will have implemented an alternative assessment for special education students who do not take the regular TAAS examinations.

**Uses and Consequences of Assessments.** Texas was among the first states to require high school students to pass standardized tests in order to earn a diploma. Students must pass either the
exit-level TAAS or three of the four end-of-course examinations. Students have seven chances to pass the TAAS for graduation.

TAAS results are also used for district and campus accountability. The accountability system considers average aggregate TAAS results as well as average performance of individual student groups (Anglo, Hispanic, African American, low-income) at both the district and the campus levels. Ratings are based on all subgroups performing at or above standards on TAAS, attendance, and dropout rate measures. Schools and districts are rated “exemplary,” “recognized,” “acceptable,” or “low-performing.”

Low-performing schools receive direction from the state education agency regarding improvement plans and reports, and the district is expected to seek additional technical assistance to improve performance. Low-performing districts may be assigned a master or board of managers and may also seek technical assistance. If a school or district remains in the low-performing category for three consecutive years, the commissioner of education can impose further sanctions. In the case of schools, the commissioner of education can appoint a board of managers or even close the school or order it to be reorganized. In the case of a school district, the commissioner may appoint a board of managers to run the district or annex the district to another district.

Principals’ and superintendents’ performance is measured, in part, by student performance results on TAAS.

Supports for Learning. Testing and accountability are components of a broader standards-based reform effort, in which instruction for all students is to be aligned with the same standards. Some schools participate in the Texas School Improvement Initiative, in which teams of educators trained by the state education agency spend a week observing in a school and advising staff members about strategies for improvement.

Texas offers financial rewards for improvement and performance through the Texas Successful Schools Awards System. The financial rewards are very modest and educators acknowledge that they are too small to serve as an incentive. The stronger incentive is the pressure from the public and the media to perform well and continue improving. Press coverage of successes is highly desirable, and schools and districts seek to avoid low-performing status because of the negative publicity that it generates.

Support programs for students include accelerated instruction in elementary reading, the Ninth Grade Success Initiative for failing ninth grade students, an optional extended-year program, and TAAS remediation at the school or at home with web-based materials for parents.

TAAS scores have improved over time, and the gaps between racial and ethnic groups have narrowed. Advocates point to these trends as evidence that standards-based reform and testing have increased the attention paid to students who were once ignored by the system or held to different, lower standards. The effects, intended and unintended, of standards-based reform in Texas have been at the center of a controversy. Some researchers believe that the TAAS exit-test requirement has increased the proportion of students who are held back and who drop out of school. The state’s four-year dropout rate for the class of 1999 is 8.5 percent, but the Intercultural Development Research Association estimates the actual dropout figure at 20 percent or higher. Critics of TAAS also claim that teachers are spending too much time on rote memorization and test preparation rather than higher-level thinking skills in all academic subjects.

Implementation Strategies. Implementation of increasing accountability standards has followed a gradual timeline. The individual student passing rate of 70 percent has been consistent over seven
years, but the requirements for a rating of “acceptable” or higher have increased steadily. Put simply, each year, more students (in all subgroups) are required to pass TAAS for the school to maintain an acceptable or higher rating. Texas is revising all the TAAS tests for all grades and subjects tested. The revised tests are referred to, generally, as TAAS-II. The exit-level TAAS, as mentioned earlier, will test more subjects and be administered in eleventh grade. Also, ninth and tenth grade will take TAAS tests in language arts and mathematics. New tests are scheduled for implementation in 2003 and at that time the high school end-of-course tests will be phased out.

Vermont

Information on Vermont policies was presented by Dr. Herman “Bud” Meyers, Deputy Commissioner for Standards and Assessment of the Vermont Department of Education.

Standards. The Vermont Framework of Standards and Learning Opportunities outlines high standards for students, based more on skills than on content. The standards are based in part on the work of the nationwide New Standards project, which is aligned with international benchmarks.

Assessments. A balance between state and local roles is a major priority of assessment policy. The Vermont Comprehensive Assessment System is intended to strike a balance between state- and locally-determined assessments that address the Framework of Standards and Learning Opportunities. In addition, at least two types of measures are included in each assessment component in order to allow all students a full and fair opportunity to demonstrate what they have learned. Vermont is one of only a few states to incorporate portfolio assessment into its state assessment system.

The state component of the assessment system includes writing portfolios in grades five and eight and mathematics portfolios in grades four, eight, and ten. Standardized assessments include Early Reading in grade two, English and mathematics in grades four, eight, and ten, science in grades six and eleven, and history in grade nine.

The local component of the assessment system consists of locally-evaluated writing and mathematics portfolios at the same grade levels as the state-evaluated portfolios, norm-referenced standardized tests of English and mathematics in grades five, nine, and eleven, other norm-referenced tests at local option, and other locally-developed assessments.

Vermont’s policies on inclusion and alternate assessments for students in special education have been cited as exemplary by the United States Department of Education.

Uses and Consequences of Assessments. Vermont does not have a high-stakes test for students. One reason is concern over measurement error and validity in standardized assessments. But because the tests are not high-stakes for students, there is also some concern over whether students’ performance on them reflects the best work of which they are capable.

The main goal of the Vermont system is accountability and improvement for schools, as opposed to accountability for individual students. Assessment results are not used to evaluate individual teachers because of recognition that cohorts of students vary greatly, especially in small schools: a school may have an exceptionally able group of fourth graders in one year, and a much less able group the following year, for reasons unrelated to teachers’ efforts.
Vermont schools are required to engage in an ongoing Action Planning process focused on needs identified by analyzing assessment results. This process must involve both teachers and parents, and may involve students. Schools determine their own sets of incentives and rewards for performance.

A biennial Accountability Review by the state identifies schools that are making insufficient progress in meeting and exceeding the standards, as well as schools that do not meet the requirements of the Vermont School Quality Standards, which include provision of a safe, civil, and orderly learning environment. Schools identified in these categories receive technical assistance from the state.

When technical assistance does not improve conditions or outcomes in an underperforming school, the school is subject to a Commissioner’s Review, which may result in continued technical assistance, adjusted responsibilities or boundaries for the superintendent, state assumption of administrative control to correct identified deficiencies, or ultimately closure of the school. If a school is closed, its district pays students’ tuition to attend another school.

**Supports for Learning.** Vermont’s testing and accountability policies are part of Act 60, which also redistributes education funding among towns in an effort to produce equity. The Vermont School Quality Standards include requirements that schools be safe, civil, and orderly, and the state collects “opportunity to learn” data that allow better analysis of school resources and student achievement. One challenge in using assessment data to support learning is that there is a time lag between when students take the tests and when schools receive their scores. The Action Planning process currently must begin before scores are available.

The state and the local Supervisory Unions share responsibility for educators’ professional development. Curriculum work and professional development have focused on the area of mathematics. Early evidence suggests that teachers are pleased with the direction reform is taking, though some are also overwhelmed by the increased time spent on assessment.

**Implementation Strategies.** Vermont began its portfolio assessment system in the early 1990s. The standardized assessments were added later because of concerns over the validity and reliability of portfolio scores, in particular the likelihood that different scorers would not come up with the same scores for the same work.

When the state legislature passed Act 60 in 1997, the state was already revamping its assessment system to comply with federal Title I requirements, so the two initiatives were combined. The curriculum standards had been in place for two years prior to the first assessments. Timelines have been revised due to challenges in developing the curriculum frameworks.

One issue is the state’s capacity to provide technical support to all underperforming schools. The first year in which such assistance was available, four schools were on the “underperforming” list. All of them improved and were taken off the list. In the second year, forty schools were identified for assistance, but only about half were able to be taken off the list after receiving assistance.

**New York**

Information on New York’s testing and graduation requirements policies was presented by Ms. Roseanne DeFabio, Assistant Commissioner for Curriculum, Instruction, and Assessment, New York State Education Department.
**Standards.** Until 1996, New York had two sets of high-school level standards. One led to a lower-level local diploma, and the other to a more prestigious Regents diploma. In 1996, the state Board of Regents voted to require all students to meet the Regents standards, which represent the level of skills needed for college entry without remediation and for high skilled employment. This higher standard for all is being implemented gradually, as described below.

**Assessments.** When the new standards have been completely implemented, students will be required to pass Regents end-of-course examinations in five subjects. The examinations include both multiple choice and constructed response items. Portfolios are not used in the state assessments, though some alternative high schools have petitioned the state to allow them to substitute their students’ portfolios and demonstrations for the Regents examination. Other tests, such as the Advanced Placement, International Baccalaureate, and SAT II, may be substituted for the Regents examinations.

The state tests are written by the New York State Education Department and are scored locally by teams of New York teachers (state authorities audit this process). These practices are less expensive than working with a testing contractor, and they also have increased acceptance of the tests by teachers and the general public.

As in many other states, testing accommodations that do not alter the construct being measured are allowed for students with Individualized Education Plans.

**Uses and Consequences of Assessments.** The Regents Examinations, which date back to the nineteenth century, were originally intended to be college entrance tests. When the new system is fully implemented, students will have to pass five Regents Examinations in order to earn a high school diploma. Students who perform exceptionally well qualify for an Advanced Regents Diploma. Either the regular or advanced diploma may also carry a Career and Technical endorsement. Some colleges and universities have begun using students’ scores on the Regents Examination in English for course placement decisions.

The examinations are also the basis of New York’s system of accountability for its high schools.

**Supports for Learning.** The new Regents Examination requirement reflects the Board of Regents’ priority of “closing the gap” between advantaged and disadvantaged students and eliminating the two-tiered expectations for high-school students. Students may take Regents examinations whenever they complete the coursework that the exams cover, and may continue to re-take them throughout high school. They are offered three times per year.

The State Education Department provides core curriculum guides and coordinates networks for professional development. The examinations are scored quickly, which provides immediate feedback to students and schools. The state outlines the Academic Intervention Services that are required for students at risk of failing the tests. Results from the first high-school class for whom the English Regents Examination is a graduation requirement were encouraging, with over 90% passing. However, the state recognizes the need to develop programs to prevent students from dropping out, either because of the high-stakes tests or before they even reach the point of taking the tests.

**Implementation Strategies.** Standards-based tests are central to New York’s education reform. The standard-setting effort began in 1991, and the standards themselves went into effect in 1996. The high-school level Regents Examinations, however, were already over one hundred years old at this point. With the advent of standards-based reform, they were aligned with the standards and began gradually changing from tests taken only by the college-bound few to tests that all students must take.
This evolution is a gradual process. The first class required to pass one Regents examination graduated in 2000. The full Regents requirement will take effect for the Class of 2005. During the phase-in period, the passing score on the examinations has been lowered, and the Board of Regents will consider extending the phase-in period to permit continuation of the lower passing score. There is also a “safety net” system in place for students in special education.

During this period, the State Education Department has been soliciting information on effective practices from the state’s teachers and publishing them. State authorities are also evaluating the impact of the new assessments on teaching and learning.

**Connecticut**

Dr. Betty J. Sternberg, Associate Commissioner for the Division of Teaching and Learning in the Connecticut State Department of Education, presented information on Connecticut’s testing and graduation requirement policies.

**Standards.** Connecticut’s standards, articulated in the state’s Common Core of Learning, are generally recognized as demanding ones. Connecticut students consistently score very well on the National Assessment of Educational Progress (NAEP). State authorities’ goal is to produce standards that are “challenging, rigorous, and reasonable” for all students.

**Assessments.** Connecticut has two state assessments, the Connecticut Mastery Test (CMT) for 4th, 6th, and 8th grades and the Connecticut Academic Performance Test (CAPT) in 10th grade. The CMT assesses student knowledge and performance in reading, writing, and mathematics. Many local educators and business people were involved in writing the assessments. The tests include some items that are similar to the ones included in the NAEP. The first generation of CAPT contained language arts, mathematics, science, and interdisciplinary sections. The second generation of the CAPT, implemented in May 2001, contains sections on mathematics, science, and reading and writing across the disciplines. The first-generation CAPT included laboratory activities, including a science lab experiment that is conducted by all the students in the state, and group work. The second generation CAPT maintains the laboratory activities, but the group work has been dropped.

**Uses and Consequences of Assessments.** For students, the CAPT and CMT are “low stakes,” in that neither is used as a criterion for promotion or graduation. The CAPT scores are included on their transcripts, however, which can increase the pressure on students to do well. Students who meet the mastery standard on any CAPT component earn a Certificate of Mastery for that component. In some local school districts, students who master all sections of the CAPT are exempt from their locally developed tests required for graduation. Some employers reward students with higher starting salaries, based on their performance on the CAPT.

Test data is used to identify “low performing” schools. Strategic School Profiles report the percentage of students who score at or above the mastery standard on the assessments, by school and by district.

Despite the fact that the state’s tests do not have direct consequences regarding promotion or graduation for individual students, schools and districts feel a great deal of pressure for achievement due to the media’s reporting of test scores. Even though the state Board of Education has publicly stated that measures other than test performance must be taken into account in evaluating schools and districts, media outlets such as the Hartford Courant produce “ranking” tables, and there is public
outcry if districts’ performance appears to be lagging or slipping. Schools receive their test results four to six weeks before statewide results are released to the media. State law requires each local board of education to hold an annual public discussion on student-related information from its Strategic School District Profile, which includes test score data. Overall, the state emphasizes the importance of communication between state and local authorities and between school districts and the public.

Supports for Learning. The state is required to identify the lowest-performing elementary schools. The state has chosen to use the CMT as the primary indicator for this purpose. Schools so identified receive more money (approximately $200,000 per school) and additional technical assistance from the state department of education.

Some observers have criticized the CMT tests for narrowing the curriculum; since the CMT focuses solely on reading, writing, and math, there has been a tendency to devote less emphasis to other areas of the curriculum. Observers have also criticized what they see as schools turning over the month of September to CMT preparation rather than instruction (the CMT tests are given in late September and early October). On the other hand, the 10th-grade CAPT has been viewed as helping broaden the high school curriculum by encouraging critical thinking and interdisciplinary emphasis across the curriculum.

Even though the CAPT is not a graduation requirement for students, there are re-test opportunities in eleventh and twelfth grades for students who do not meet the mastery standard as tenth-graders. Because the test is not a graduation requirement, it appears that the test has not had a direct effect on dropout rates. Sternberg also suggests that it has raised teachers’ expectations of all students and has had a particularly positive effect in urban settings. Because CAPT includes science laboratory work, one beneficial consequence of the test has been that some schools that did not have the necessary equipment have received it.

Implementation Strategies. Connecticut’s new generation of standards-based tests was implemented in the 4th grade first. In the late 1970s, the state required a 9th-grade competency test which at the time was set at a relatively low standard. Subsequent legislation instituted mastery tests in 4th, 6th, and 8th grades (the CMT) and no longer required the 9th-grade proficiency test. After a number of years of implementation of the 4th, 6th, and 8th grade test, a new requirement for the 10th-grade CAPT test was implemented, almost a decade after the first 4th-grade test was first administered.

Implications for Massachusetts

Texas. In some ways, it is difficult to compare state testing policies in Texas and Massachusetts. Because the TAAS is a less difficult test than the MCAS, its political and educational consequences have been different. In addition, a portion of special education students do not take TAAS or, if they are tested, their scores may not be included in the accountability rating system, whereas they have been included in the MCAS from its inception. Finally, to date there has been more emphasis in Texas than in Massachusetts on intervention and support for schools and school systems.

At the same time, Texas’ longer experience with high-stakes tests, and the broad similarities between the states’ policies, suggest issues that Massachusetts policy makers should consider carefully. The first is the notion of “easing into” high-stakes testing. Texas did this by starting with a high-stakes test that was relatively easy to pass, with the intention of gradually increasing the test’s difficulty once alignment with more challenging academic standards was completed and students, parents, and
educators were used to the practice of test-taking. This is in marked contrast to the approach taken in Massachusetts, in which a difficult test and high stakes have been combined from the beginning.

The second issue concerns the difficulty of establishing the “real story” about the effects of reform. Some analysts believe the TAAS has had beneficial results; others say it has narrowed the curriculum and increased dropout rates. A broader issue is the unintended consequences of reform. TAAS scores have affected real estate markets as well as public opinion. Reform efforts may contribute to a shortage of principals for lower-performing schools because many principals chose to move to positions in higher-performing schools and districts.

**Vermont.** Vermont’s use of assessment data for school improvement is similar to that of Massachusetts. However, the states’ policies are quite different in other ways. Despite concerns about whether students try their hardest on a test that is not high-stakes for them, Vermont does not plan to attach student-level consequences to their assessments.

Vermont’s blend of state and local roles in assessment is quite different from the dominant state role in Massachusetts. Assessment is seen as a partnership between the state and local schools, with local districts having a significant role in determining the mix of assessments by which their students will be tested and their schools held accountable.

Both Vermont and Massachusetts have language in their statutes calling for multiple forms of assessment, but Vermont’s mix of standardized and portfolio assessment goes much further in this direction than does the MCAS. Vermont’s relatively small student population makes portfolio assessment logistically easier than it would be in a larger state. And despite strong support for the principle of authentic assessment, Vermont has added standardized assessment to increase the system’s reliability and validity.

**New York.** New York is an especially relevant case to compare with Massachusetts because the two states are moving in the same general direction, but employing different implementation strategies. The Regents Examinations have a much longer history than the MCAS, and in some respects New York’s transition to requiring Regents Examinations for all students is proceeding more gradually than the implementation of the Competency Determination in Massachusetts. In New York, more time elapsed between when the standards were put in place and when the Regents English examination became a graduation requirement than has been the case with the Massachusetts Curriculum Frameworks and MCAS.

Current Massachusetts policy calls for only one state-endorsed diploma beginning with the Class of 2003. Some local educators have called upon the state to allow school districts to continue issuing their own diplomas to students who satisfy all local graduation requirements but do not pass MCAS. In effect, this is the system New York is moving away from. Formerly, New York students received either a local diploma or a higher-level Regents Diploma, depending upon which state tests they attempted and passed.

Although the New York Regents Examinations are at a level comparable to that of MCAS, the Regents requirement has been less controversial thus far than the MCAS. One reason may be that the Regents Examinations are much more familiar to the public given their long history. Another possibility is that there is a greater sense of local ownership of the tests, since scoring is done by local teachers rather than by outside contractors. Local scoring also allows the results to be given to students and teachers much more promptly than has been the case for the MCAS.

Pass rates on the English Regents Examination (the only one that was high-stakes for students in the Class of 2000) were much higher than MCAS pass rates thus far. One reason is that the New York
results were for students who had taken the exam multiple times, compared with Massachusetts results for students who had only taken MCAS once. Another is that the Regents pass rates are stated as a percentage of twelfth-grade students, compared with MCAS pass rates as a percentage of tenth-grade students. Tenth-graders have not yet reached the age at which they may legally leave school, but the twelfth-grade population has been shrunk by students dropping out.

New York state officials are closely monitoring the effect of the Regents Examination requirement on the dropout rate, especially in New York City and the state’s other large urban districts. The dropout rate is just one indicator that can be used to assess the test’s “consequential validity,” or effects on classroom teaching and learning. As in other states, the unintended effects of the Regents requirement are as important as the intended ones.

**Connecticut.** Like Massachusetts, Connecticut is a state where public schools function quite well in many districts, but where poor students are concentrated in urban districts and often do not benefit as much from the system.

The major policy contrast between Connecticut and Massachusetts is in the uses of assessment results. Despite some legislators’ support for the idea, Connecticut has no plans to make the CAPT a high-school graduation requirement for students. Connecticut’s experience demonstrates that even low-stakes tests can be high-pressure for students and educators because of the attention paid to the scores by the media, employers, and parents.

The political environment in which education policy is made also differs between the two states. The Connecticut State Board of Education opposes making tests high-stakes for students, where the Massachusetts Board of Education supports it. The process of developing the Common Core of Learning in Connecticut did not generate the same controversy as the Curriculum Frameworks did in Massachusetts, and the Connecticut standards have not been revised as often as the Massachusetts Frameworks. In general, the absence of high stakes for students probably reduces the overall level of controversy about testing.

**General Discussion**

In the concluding large-group session, each small group presented a summary of its discussion of how one of the four states fits into the analytical framework and the benefits, challenges, and implications for Massachusetts. Then the floor was opened for general discussion. The following comments and questions are arranged mostly in the order in which they were posed, but in some cases related points made at different times are grouped together.

**College Admissions.** A conference participant asked the panelists from New York and Texas whether there were formal links between test scores and college admissions in their states.

Dr. Clark replied that Texas college students must pass the Texas Academic Skills Program (TASP) test to remain enrolled in a Texas public four-year college or university. Students who have high scores on the TAAS exit-level test may be exempted from the TASP requirement.

Ms. DeFabio said that there is no systematic connection in New York. Students can go to a community college without a high school diploma. Additionally, some colleges, at their discretion, use test results for placement.
**Potential Changes in Federal Policy.** An attendee asked all the panelists whether President Bush’s proposed changes to federal policy on accountability and testing [requiring testing of all students in both reading and math annually in grades 3 through 8] would have an effect on their states’ testing systems.

According to Dr. Meyers, Vermont authorities have calculated that implementing the Bush plan would cost $9 million the first year, compared with the state’s current testing budget of $5 million. The federal proposal would only provide funding for the cost of developing the additional tests, not of implementing them.

Dr. Clark noted that, unsurprisingly, Texas policy is already in line with the Bush administration proposal.

Ms. DeFabio seconded Dr. Meyers’s concerns about cost. She also noted that New York residents are not eager for still more state testing. She expressed the hope that as the bill goes through Congress, it will be modified to allow states to select the tests they will use.

Dr. Sternberg also said the budget implications of the Bush proposal are an area of concern. She noted that the emphasis on testing is not new, and that it is hard to come to agreement on what is tested and the standards for student performance within a state, let alone for the whole nation.

**Inappropriate Use of Test Results.** A participant called attention to the potential for intentional and unintentional use of testing information for purposes other than the ones for which tests are designed; for example, to influence real estate prices or to make political points. This misuse may have negative consequences.

**Links Between Student and School Accountability.** Another attendee asked about the impact of holding students accountable before schools are held accountable, and said it was a question that should be researched. Are student high stakes needed to drive schools to change? Is “blaming and shaming” of schools enough, or do students need to be held accountable as well?

Ms. DeFabio replied that the high-stakes discussion needs to be expanded to include the negative impact of not having known performance information for all students.

Dr. Meyers said that the answer to the question depends on how it is framed. The way in which policies are formulated depends on the initial assumptions and phrasing of questions. The effect of state testing also depends upon the design of the tests. The effect of local tests depends on the relationship between teachers and students.

**Goals of Public Education.** A member of the group said that nobody at the conference had yet asked the question of whether academics is the sole goal of public education. Schools also serve their students’ spiritual and emotional growth, this participant noted, and they are being asked more and more to deal with values issues. The developmental nature of education is not being considered, and different types of learning are not being measured. These issues are missing from today’s discussion. People who work with kids haven’t been involved enough.

Dr. Meyers agreed that he would add “purpose of assessment” to the analytical framework, to encourage states to make clear the goals of their education system so that the assessment system can be evaluated according to overall educational goals, not as an end in itself.
**Teacher Preparation.** A participant asked what the implications of each state’s approach to standards and testing have been for professional development, and for teacher training institutions. In particular, have there been changes to the approach, program, and curriculum?

Dr. Sternberg replied that Connecticut took a comprehensive approach, with several strands. One strand was attracting and retaining high quality teachers, through the Education Enhancement Act. The Act raised standards for admission to teacher education and for certification, and created institutes for teaching and learning. It also raised teachers’ salaries and changed the state’s program for continuing professional development. Teachers now have to qualify for continued certification every five years. A Common Core of Teaching was linked to the Common Core of Learning for students.

Ms. DeFabio referred interested people to the New York State Education Department website (www.nysed.gov) for information about their revised teacher development policies, which are similar to Connecticut’s. Teacher education programs must have 80% of their graduates pass a performance evaluation in order to remain in operation.

Dr. Clark said that teacher preparation has become standards-based and is linked to the TEKS. As a result, more programs to prepare teachers have a field-based component. Teacher certification tests are also linked to state standards.

Dr. Meyers said that Vermont teacher education programs teach the state’s standards and how to assess them. Vermont has the highest required scores in the Northeast on the PRAXIS I and II tests for aspiring teachers.

**Communication.** One participant comment reiterated the need to stop thinking in pro- and anti-MCAS terms, and to think in terms of the five components listed in the Center for Education Policy’s analytic framework. However, she challenged the attendees to think about whether they were still communicating well enough to do this together. Are stakeholder groups no longer talking/listening to each other? Is joint progress still possible, or is it all power politics now? Can activities like today’s conference help us move toward this?

Dr. Meyers praised the analytical framework for helping make clear where the enterprise started. Thinking in this way, he said, can help get people to talk to each other rather than retreat into camps.

A different participant from the one who posed the question originally commented that the divisiveness over the testing issue has been extremely harmful to education in Massachusetts. According to this speaker, there is a disconnect between higher education and K-12, and between policymakers and implementers. In 1993, when the education reform law was passed, teachers were excited. Now, they are totally demoralized. There has been too much too soon, without careful implementation and evaluation. The system will be threatened if some changes to the current process don't happen.

Joining the discussion, two other attendees asked how other states encourage communication, and how we can move towards at least agreeing on what we don’t agree about.

Ms. DeFabio said that the New York State Education Department used the media. As one example, the Commissioner appeared on a call-in program on public television. Teachers and parents called in and had their voices heard.

Dr. Sternberg replied that Connecticut state authorities had used education coalitions—collections of the major education organizations in the state, including teachers unions, principals, superintendents,
school board members, and the Department of Education—to go together to the legislature. For the most part, they reached consensus. However, she said, there are still some parts of the state’s policies that will not please some groups. But at some point you have to make a decision to move ahead with a generally satisfactory approach.

According to Dr. Meyers, Vermont also used education coalitions, which meet regularly with the Commissioner. Action planning is required to include representative groups.

In Texas, according to Dr. Clark, the absence of a strong teachers’ union meant that teacher groups were not united in supporting or opposing accountability. A business group took a leadership place at the table and was greatly involved in promoting standards, testing, and accountability. Otherwise, most groups had representatives in the state policy discussion. However, given the size of the groups, many people still felt left out or that having only one representative had not been enough.

A conference participant pointed out that this sort of coalition effort has begun in Massachusetts, where a group of professional education associations called Mass Partners now meets regularly with the Commissioner.

One participant repeated the question of how higher education could be more involved.

Continuing the theme of communication, another person pointed out that school committees also feel left out of the discussion. This participant felt that most of the major decisions are being made by gubernatorial appointees. For school committees to attract the attention of the state Department of Education or legislature, he said, they first must attract the attention of the Board of Education, which seems to them to be a circuitous route.

Closing Comments. Dr. McDermott’s closing comments pointed to common themes. Everybody present at the conference is engaged with the issue of standards-based reform and is struggling to figure out what it really means. How do tests work as motivators? What is the role of tests as a source of information? What do they provide incentives for people to do? What are the intended and unintended consequences of policy? She also referred back to the point about the purposes of public education that was made in the final discussion. Controversy over the aims of Education Reform is forcing us to think anew about what we expect public schools to do.

The discussion raised a number of key policy issues:

- Communication about the implementation of Education Reform has not been sufficient in Massachusetts. Participants who play a variety of roles in the state educational community, and who disagree on many of the key issues, generally seemed to agree on this point.

- The states represented had all phased in their reforms over a period of several years or more. Connecticut began with standardized tests in grades 4, 6, and 8 for several years before adding its 10th-grade test. Texas began with a relatively easy test, then made it more challenging when it was aligned to new, more rigorous learning standards. New York built its Regents-for-all approach on its long history of test-taking in the state.

- Lack of a graduation requirement doesn’t necessarily mean that a state’s assessment lacks high stakes. Several presenters mentioned the pressure schools and districts feel due to media reporting of test results, regardless of whether students are required to pass a test to graduate.

- Local districts and teachers are involved in important ways in state testing strategies. In Vermont, the state’s role is shared with a locally designed assessment component. New York
uses local teachers to grade its tests. Several participants noted that a state test, like all state education efforts, depends on local people for its implementation.

- States are developing accountability measures for both children and the adults who teach them. There was significant discussion about what kinds of assessments should be used for each, what the local versus state roles should be in each area, and how these accountability efforts should be sequenced.

- In implementing standards-based assessments, timelines generally have been longer than expected. Participants mentioned the need to get curriculum and instruction aligned with standards and then give students time to absorb the revised instruction before tests will have validity as high-stakes instruments.

In reflecting on the day, Dr. McDermott noted that even though testing is based upon a large body of technical knowledge, science alone cannot yield the “correct” answers for testing policy. Psychometricians seek the best solutions to test design problems, but policy makers often must work incrementally and choose which of several options is “good enough” given time and resource constraints. For example, even though several associations of professionals involved in test construction have warned against using a single test for multiple purposes, policy makers strive to get as much use as possible out of single tests because they work with limited budgets.

In closing, Dr. McDermott observed that testing and accountability policies for public education will always be controversial because in order to hold students or schools accountable for results, we must first decide which results are the important ones. Debate over the purposes of public education is as old as public education itself, and will probably never be resolved. Until quite recently, most Americans were satisfied with a system that produced a wide range of outcomes for different students. Standards-based reform attempts to replace this state of affairs with one in which all students master higher-level skills and more advanced material. How to make this new vision a reality in a society that is also as diverse as that of the contemporary United States will be the major educational challenge of the new century.
Supplemental Materials

General
- Conference Agenda
- List of Registrants
- High School Testing and Graduation Requirements Analytical Framework
  Kathryn McDermott

Texas
- Texas Testing and Graduation Requirements: The Framework
  Catherine Clark
- The Evolution of High-Stakes Testing in Texas
  Catherine Clark
- Critiques and Commendations for the Texas System
  Catherine Clark

Vermont
- Vermont State Board of Education Policy on Testing and Accountability
- Vermont Comprehensive Assessment System (table)
- Part of a Larger System
- Overview of Vermont’s Accountability System (chart)
- Design Decisions

New York
- State Testing and Graduation Requirements—How New York Fits Into the Framework
  Roseanne De Fabio

Connecticut
- State Testing and Graduation Requirements—Connecticut
  Betty Sternberg
- Position Statement on Measuring Success
  Connecticut State Board of Education
Supplemental Materials

General

- Conference Agenda

- List of Registrants

- High School Testing and Graduation Requirements Analytical Framework
  - Kathryn McDermott
STATE TESTING AND GRADUATION REQUIREMENTS:
WHAT CAN MASSACHUSETTS LEARN FROM OTHER STATES?
—A Policy Leadership Conference—

Friday, February 2, 2000
Sheraton Needham Hotel
7:30am – 12:30pm

7:30 – Continental Breakfast and Registration

8:00 – Welcome and Introduction to the Conference
Andrew Effrat, Director, Center for Education Policy, University of Massachusetts-Amherst
Bailey Jackson, Dean, School of Education, University of Massachusetts-Amherst

8:15 – Presentation of State Testing Policy Framework
Kathryn McDermott, Associate Director, Center for Education Policy

8:30 – State Perspectives
Connecticut – Betty Sternberg, Associate Commissioner, Connecticut Department of Education
New York – Roseanne DeFabio, Assistant Commissioner, New York Department of Education
Texas – Catherine Clark, Program Director, Charles A. Dana Center, University of Texas-Austin
Vermont – Herman Meyers, Deputy Commissioner, Vermont Department of Education

9:45 – Breakout Groups
Attendees will be divided into four groups, each using the state testing policy framework to examine the approach of one of the presenting states—strengths, challenges, and possible implications for Massachusetts.

11:00 – Coming Together and Looking Forward
Presenters and attendees will compare the approaches of the presenting states and explore possible implications for Massachusetts.

12:20 – Concluding Remarks
Dr. McDermott and presenters

12:30 - Adjourn
Dr. McDermott and presenters will be available for further discussion with attendees

This conference is sponsored by the Center for Education Policy of the School of Education at the University of Massachusetts Amherst, with support from the Noyce Foundation and Verizon
# State Testing and Graduation Requirements: What Can Massachusetts Learn From Other States?

**A Policy Leadership Conference**

**February 2, 2001**

## Registrant List

<table>
<thead>
<tr>
<th>Dana Ansel</th>
<th>Marguerite Clarke</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Director, MassInc.</td>
<td>Associate Director, National Board on Educational Testing and Public Policy, Lynch Graduate School of Education, Boston College</td>
</tr>
<tr>
<td>Robert A. Antonioni</td>
<td></td>
</tr>
<tr>
<td>Senate Chair, Committee on Education, Arts and Humanities</td>
<td></td>
</tr>
<tr>
<td>Maura Banta</td>
<td></td>
</tr>
<tr>
<td>Manager, Corporate Community Relations, IBM</td>
<td></td>
</tr>
<tr>
<td>Laura Barrett</td>
<td></td>
</tr>
<tr>
<td>Division of Communications, Massachusetts Teachers Association</td>
<td></td>
</tr>
<tr>
<td>Geoffrey Beckwith</td>
<td></td>
</tr>
<tr>
<td>Executive Director, Massachusetts Municipal Association</td>
<td></td>
</tr>
<tr>
<td>Joseph Berger</td>
<td></td>
</tr>
<tr>
<td>Associate Director, Center for Education Policy, UMass/Amherst</td>
<td></td>
</tr>
<tr>
<td>Noah Berger</td>
<td></td>
</tr>
<tr>
<td>Aide to Senate President Birmingham</td>
<td></td>
</tr>
<tr>
<td>Sandra Berkowitz</td>
<td></td>
</tr>
<tr>
<td>Chair, Amherst School Committee</td>
<td></td>
</tr>
<tr>
<td>Mary Brabeck</td>
<td></td>
</tr>
<tr>
<td>Dean, Lynch Graduate School of Education, Boston College</td>
<td></td>
</tr>
<tr>
<td>Lenore R. Carlisle</td>
<td></td>
</tr>
<tr>
<td>Assistant Professor, School of Education, UMass/Amherst</td>
<td></td>
</tr>
<tr>
<td>Andrew Churchill</td>
<td></td>
</tr>
<tr>
<td>Research Coordinator, Center for Education Policy, UMass/Amherst</td>
<td></td>
</tr>
<tr>
<td>Catherine Clark</td>
<td></td>
</tr>
<tr>
<td>Director for Research and Policy, Charles A. Dana Center, Univ. of Texas/Austin</td>
<td></td>
</tr>
<tr>
<td>Patricia Crosson</td>
<td></td>
</tr>
<tr>
<td>Professor, School of Education, UMass/Amherst</td>
<td></td>
</tr>
</tbody>
</table>

**Kevin Courtney**
Superintendent, Pioneer Valley Regional School District

**Ronald R. Cromwell**
Dean, School of Education & Allied Studies, Bridgewater State College

**Gerald Croteau**
Superintendent, Taunton Public Schools

**Patricia A. Crutchfield**
Massachusetts Board of Education

**William L. Dandridge**
Dean, Graduate School of Education, Lesley College

**Roberta Deering**
School Committee Member

**Roseanne DeFabio**
Assistant Commissioner for Curriculum, Instruction and Assessment, New York State Board of Education

**Philip Devaux**
Superintendent, Marblehead Public Schools

**Kathleen Devlin**
Research Analyst, Joint Committee on Education, Arts and Humanities, Massachusetts

**Ralph Devlin**
Professional Development Consultant, Massachusetts Teachers Association

**Andrew Effrat**
Director, Center for Education Policy, UMass/Amherst
Alan McDonald
Executive Director,
Massachusetts Business Roundtable

Katherine S. McHugh
Director of Donor Services,
Jesse B. Cox Charitable Trust

H. W. “Bud” Meyers
Deputy Commissioner for Standards and Assessment,
Vermont Department of Education

Beverly Miyares
Professional Development,
Massachusetts Teachers Association

Jeffrey Nellhaus
Associate Commissioner for Student Testing,
Massachusetts Department of Education

Pendred E. Noyce
The Noyce Foundation

Vera Ossen
Director of Education Certification Programs,
UMass/Lowell

Donald E. Pierson
Dean, College of Education, UMass/Lowell

John Portz
Political Science Department,
Northeastern University

Jill Reynolds
Deputy Education Advisor,
Governor’s Office

Connie Rizoli
Research Director,
House Committee on Education,

Mark Roosevelt
President, CEO & Treasurer,
Massachusetts Biomedical Initiative

Gus Sayer
Superintendent, Amherst-Pelham Regional Schools

Paul Schlichtman
School Committee Member

Michael Sentence
Education Advisor, Governor’s Office

Dennis Shirley
Associate Dean, Lynch School of Education,
Boston College

Katherine Skinner
Director, Center for Education Quality and
Professional Development,
Massachusetts Teachers Association

Sylvia Smith
Director of Education Policy & Legislation,
Senator Robert A. Antonioni’s Office

Betty J. Sternberg
Associate Commissioner, Division of Teaching and
Learning, Connecticut State Department of Education

Sandra Stotsky
Deputy Commissioner for Academic Affairs and
Planning, Massachusetts Department of Education

Patricia Sweitzer
MassPartners

Elaine Taber
Aide to Speaker Thomas M. Finneran

Philip Veysey
Director of Education Policy,
Massachusetts Federation of Teachers

Katherine Viator
Director of Student Testing,
Massachusetts Department of Education

Ann Walsh
Massachusetts PTA

Melanie S. Winklosky
Special Assistant to the Board of Education,
Massachusetts Department of Education

Jennifer Wood
Donahue Institute, University of Massachusetts

Laurie Wright
Assessment Specialist,
Massachusetts Department of Education

Jodie Zalk
Data Specialist for MCAS,
Massachusetts Department of Education
State Testing and Graduation Requirements:  
A Framework for Analysis

<table>
<thead>
<tr>
<th>Standards</th>
<th>• How advanced are the skills and material on which students are assessed?</th>
</tr>
</thead>
</table>
| Assessments                                                              | • What is the balance between state and local authority?  
• What kinds of test items?  
• Standardized test only, or portfolio/performance component?  
• What accommodations or alternate tests are used? |
| Uses and Consequences of Assessments                                     | • Is there a graduation test?  
• Are tests high-stakes for schools or individual staff members?  
• Does the state award multiple high-school credentials or diplomas? |
| Supports for Learning                                                    | • How many chances do students have to pass a high-stakes test?  
• What supports exist for students at risk of failure?  
• Is the test or assessment part of a standards-based reform initiative?  
• Has instruction been aligned with the same standards/curriculum as the test? |
| Implementation Strategies                                               | • How much time between alignment and assessment?  
• Will passing standard rise over time?  
• Are consequences of test phased in?  
• Have changes been made in the implementation timeline? |
## State Testing and Graduation Requirements: How Massachusetts Fits Into the Framework

| Standards                                                                 | • Material in 10\textsuperscript{th} grade frameworks is at high level  
|                                                                          | • Compared with most other states’ 10\textsuperscript{th} grade level tests, MCAS is almost certainly more difficult  
|                                                                          | • Focused retest  
| Assessments                                                              | • MCAS has more constructed-response questions than many state tests  
|                                                                          | • No portfolio assessment in regular MCAS  
|                                                                          | • Special education alternate assessments based on curriculum frameworks  
|                                                                          | • Assessment of vocational-technical school students is a more pressing issue here than in many states  
| Uses and Consequences of Assessments                                     | • MCAS scheduled to become graduation test for class of 2003  
|                                                                          | • MCAS also used for school level accountability and for diagnostic purposes  
|                                                                          | • Certificate of completion has been recommended  
|                                                                          | • Certificate of Mastery and Certificate of Occupational Proficiency  
| Supports for Learning                                                     | • Five opportunities to pass test  
|                                                                          | • Support programs underway  
|                                                                          | • Test is part of comprehensive standards-based reform  
|                                                                          | • State has been criticized for implementing graduation test before alignment of curricula was complete  
| Implementation Strategies                                                | • Competency determination being phased in  
|                                                                          | • Some support for delaying high-stakes use of MCAS  
|                                                                          | • Some alterations already made to implementation timeline  

| Standards | • 47 states have standards-based assessments  
• Trend is towards mastery rather than basic skills  
• Of 18 existing graduation tests, 5 are considered to be at a 10<sup>th</sup> grade level  
• By 2008, 21 states will have graduation tests at a 10<sup>th</sup> grade level |
| --- | --- |
| Assessments | • Few examples of portfolio assessment at high school level in regular education  
• Inclusion (with accommodations) is federal mandate  
• Special education alternate assessments are portfolio/performance based  
• Most alternate assessments are based on curriculum frameworks |
| Uses and Consequences of Assessments | • Tests used for multiple purposes  
• 24 states have or plan to have graduation tests  
• 26 states neither have graduation tests nor plan to implement them  
• States with graduation tests award certificate of completion  
• A few states have multiple pathways to diploma  
• Honors diplomas gaining popularity |
| Supports for Learning | • Multiple opportunities to pass test  
• Support programs with certified educators  
• Tests often in place before other standards-based reforms |
| Implementation Strategies | • Phasing in of more demanding assessments  
• Midcourse corrections are common |
Supplemental Materials

Texas

- Texas Testing and Graduation Requirements: The Framework  
  - Catherine Clark

- The Evolution of High-Stakes Testing in Texas  
  - Catherine Clark

- Critiques and Commendations for the Texas System  
  - Catherine Clark
## Texas Testing and Graduation Requirements:
The Framework

<table>
<thead>
<tr>
<th>Standards</th>
<th>• How advanced are the skills and material on which students are assessed?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Texas tests are aligned with the Texas learning standards: Texas Essential Knowledge and Skills (TEKS). Taken together, TEKS represent high standards for learning. The Texas Assessment of Academic Skills (TAAS) measures whether students meet the minimum expectation level for the standards. Textbooks and professional development are also aligned with TEKS.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assessments</th>
<th>• Balance between state and local authority?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The state establishes the learning standards and TAAS. Local districts determine the curriculum and professional development. Local districts can select any texts, but the state will pay some of the costs of texts that cover more than half the TEKS and all of the cost of texts that cover all the TEKS. TEKS presents districts with opportunities to use a wide variety of materials and approaches.</td>
</tr>
<tr>
<td></td>
<td>• What kinds of test items?</td>
</tr>
<tr>
<td></td>
<td>Multiple choice. The writing test, given in grades 4, 8 and 10, has an essay component.</td>
</tr>
<tr>
<td></td>
<td>• Package of assessments: standardized test only, portfolio/performance component, others?</td>
</tr>
<tr>
<td></td>
<td>Standardized tests, including reading and math in grades 3-8 and 10, writing in 4, 8, and 10, and end-of-course exams in Algebra I, Biology, American History, and English III.</td>
</tr>
<tr>
<td></td>
<td>• What accommodations or alternate tests are used?</td>
</tr>
<tr>
<td></td>
<td>Texas will implement an alternative assessment for some special education students**. Students may take as much time as they need for each TAAS section. A Spanish version of TAAS is available for certain students in grades 3-6.</td>
</tr>
</tbody>
</table>
| **Uses and Consequences of Assessments** | • **High stakes for students—graduation test?**  
Yes. Students must pass TAAS (or a combination of end-of-course tests) to graduate.  
• **High stakes for adults—schools and/or staff?**  
Yes. Principal and superintendent evaluation is measured, in part, by student performance. State sanctions (including takeover) for low-performance 3 years in a row.  
• **Does the state award multiple high-school credentials or diplomas?**  
Regular diploma (22 credits), recommended high school program (24 credits), and distinguished achievement program (24 credits). |
| **Supports for Learning** | • **How many chances do students have to pass a high-stakes test?**  
Seven  
• **What supports exist for students at risk of failure?**  
Accelerated instruction in elementary school, Ninth Grade Success Initiative for students in grades 8 and 9, optional extended year program for grades K-8, TAAS remediation at the school or at home with web-based materials for parents.  
• **Is the test or assessment part of a standards-based reform initiative?**  
Yes  
• **Has instruction been aligned with the same standards/curriculum as the test?**  
It is intended to be; implementation may vary. |
| **Implementation Strategies** | • **How much time between alignment and assessment?**  
Alignment should be complete in 2000; new tests to be developed for 2003 and beyond (still aligned to TEKS).  
• **Will passing standard rise over time?**  
Passing standard is 70%. Accountability standard has increased every year for previous 6 years; anticipated increase in standards or test difficulty in next 5 years.  
• **Are consequences of test phased in?**  
Texas has used sanctions since the mid-1990s. For individual schools and districts, there are consequences (and technical assistance) for low performance the first year and the second (subsequent) year. After three years of low performance, the school is subject to being reconstituted, the principal is likely to be removed.*** For districts, the third year consequence could be state takeover. Commissioner of education has several options. |
| **Have changes been made in the implementation timeline?**
| No. But new tests are under development. Texas will have 9th and 10th grade TAAS. End-of-course tests will be phased out and replaced by an 11th grade exit test (covering biology, chemistry and physics, algebra, geometry, English III, writing, and American History).

** Some special education students take regular TAAS. Some will take the alternative assessment. Some students will be exempted from testing. The IEP guides the determination of whether to test and the type of test administered.

*** There are cases when the principal might not be removed. For example, if the school is experiencing a third year of low performance but the principal had been there only one year.
The Evolution of High-Stakes Testing in Texas

Catherine Clark
Charles A. Dana Center
February 2001

Texas Chronology

Early 1980s: Minimum competency testing (Texas Assessment of Basic Skills—TABS)
Late 1980s: Criterion-referenced testing to measure “essential elements of the curriculum” in reading and mathematics (Texas Educational Assessment Measurement System—TEAMS)
1991: Criterion-referenced test revised and made more challenging (Texas Assessment of Academic Skills—TAAS). Passing TAAS required for graduation
1991-1992: Accountability system developed for Texas
1993: Academic Excellence Indicators System (AEIS) and public reports of TAAS results for districts and schools
1994: Accountability ratings for Texas public schools and districts
1996: Policy to release TAAS tests to the public soon after the test
1998: Implementation of more challenging learning standards (Texas Essential Knowledge and Skills—TEKS)
1998-2000: Alignment of TAAS to TEKS (while accountability performance criteria continued to rise)
2003 (spring): Revised and expanded TAAS (including alternative TAAS for special education)

Scaling up to a Full Accountability System

- Establish learning standards or curricula describing what students should know and be able to do.
- Implement a data reporting system that includes a wide variety of performance indicators for public education.
- Develop or select a test aligned with what students should know and be able to do, and establish passing and mastery standards for performance.
- Assess all students, each year, in most grades.
- Establish an accountability system that incorporates key public education indicators. Incorporate performance standards that most people can understand and establish a time frame for reporting. Make sure all subgroups of students and schools are required to meet the accountability standards. Establish consequences and sanctions (and means of providing technical support) for low performance. Implement the system over a period of time during which the standards for accountability increase.
- Explain the system to educators before applying it and continue to explain the system as it evolves.
- Benchmark student performance before applying accountability ratings.
- Report results of assessments at the state, region, district, school, and grade-within-school level in the data or indicators system.
• Disaggregate student performance at all levels (by race/ethnicity, income level, and special program participation).
• Report accountability ratings.
• Follow up on consequences and sanctions.
• Provide access to performance information on all indicators through the Internet and through reports to parents and the public. Provide information about the tests, preferably access to all tests within a few weeks of test administration.

TAAS Mathematics Performance. All Texas Students in Grades 3-8 and 10 (all grades tested)

<table>
<thead>
<tr>
<th>Student Groups</th>
<th>Percent Passing</th>
</tr>
</thead>
<tbody>
<tr>
<td>All students</td>
<td>58</td>
</tr>
<tr>
<td>African American</td>
<td>36</td>
</tr>
<tr>
<td>Hispanic</td>
<td>45</td>
</tr>
<tr>
<td>White</td>
<td>70</td>
</tr>
<tr>
<td>Low-income</td>
<td>43</td>
</tr>
</tbody>
</table>

TAAS Reading Performance. All Texas Students in Grades 3-8 and 10* (all grades tested)

<table>
<thead>
<tr>
<th>Student Groups</th>
<th>Percentage Passing</th>
</tr>
</thead>
<tbody>
<tr>
<td>All students</td>
<td>74</td>
</tr>
<tr>
<td>African American</td>
<td>58</td>
</tr>
<tr>
<td>Hispanic</td>
<td>63</td>
</tr>
<tr>
<td>White</td>
<td>85</td>
</tr>
<tr>
<td>Low Income</td>
<td>61</td>
</tr>
</tbody>
</table>

Texas Accountability Ratings History---Campus Ratings. Percent of Campuses with Various Ratings, 1995 through 2000

<table>
<thead>
<tr>
<th>Rating</th>
<th>Percent of Campuses in Rating Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exemplary</td>
<td>4.1</td>
</tr>
<tr>
<td>Recognized</td>
<td>16.1</td>
</tr>
<tr>
<td>Acceptable</td>
<td>69.9</td>
</tr>
<tr>
<td>Low-performing</td>
<td>4.3</td>
</tr>
<tr>
<td>Alternative</td>
<td>4.1</td>
</tr>
<tr>
<td>Other</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Critiques and Commendations for the Texas System

- Tests are (or have become) too easy, are too narrow, and are faulty because they use a multiple-choice format.
- TAAS and the accountability system drive students out of school (dropouts).
- Most Texas gains come from dumping the regular curriculum and instituting TAAS practice and “getting rid of” low-performing students who would have driven results lower.
- Teaching to the test and other test preparation activities take up a disproportionate amount of instructional time, sometimes several months.
- TAAS and accountability are detrimental to student learning: the teaching of math and language arts is perverted and other subjects are short-changed.
- Having to pass Exit-TAAS for high school graduation is unfair and inappropriate.
- Student performance at and near the top is not well measured.
- TAAS “works” because of blame and shame, not because of any positive incentives.

- Criticisms that will be addressed by new tests in 2003
  - High school tests do not measure algebra and geometry.
  - High school tests ignore other important subjects.
  - There is no 9th grade test.
  - The Exit-TAAS does not indicate what a student knows as a result of 4 years of high school.

- TAAS and the rating system are a response to the public’s demand for accountability.
- TAAS and the accountability system continue to improve and evolve.
- TAAS and ratings have focused educator attention on reading and math when other reforms failed to do so.
- People other than educators pay attention to TAAS and accountability. When the results come out, it is “big news.”
- There is evidence that Texas students are improving on national tests, as reflected in improved performance on various NAEP tests.
- TAAS has produced the greatest results for the most needy and underserved students.
- Requiring all student subgroups to do well on TAAS means all students must be taught and low scores of poor and minority students cannot be hidden in school and district averages.
- TAAS, textbooks, and teacher professional development are all aligned to learning standards adopted for use in all Texas schools. Texas has devised a coherent system.
Supplemental Materials

Vermont

- Vermont State Board of Education Policy on Testing and Accountability
- Vermont Comprehensive Assessment System (table)
- Part of a Larger System
- Overview of Vermont’s Accountability System (chart)
- Design Decisions
X The Vermont Comprehensive Assessment System has state and local components which together address Vermont’s Framework of Standards and Learning Opportunities. The system components are presented on the reverse of this sheet. Content areas, grade levels and mandated participation levels are presented for each element.

X The system reflects a balance of assessment approaches and formats. At least two types of measures are included in each component to allow students full and fair opportunity to demonstrate learning.

X Each assessment tool has been selected because it provides a useful measure of (important clusters of) Vermont’s Standards.

X Each assessment tool provides valid and reliable information at the reporting level(s) it is primarily used to address.

X As required by current Federal regulations (e.g., Title I) to ensure that high standards are learned by ALL students, demographic data on students and schools will be collected to enable the report of achievement for sub-populations within Vermont. (e.g., student data: gender, ethnicity, English proficiency, disability status, migrant status, economic status; school data: enrollment, child counts)

X In order to ensure that learning opportunities which address the standards are provided to ALL Vermont students, opportunity-to-learn data from students and schools will be collected to enable the analysis of achievement.

X The Data Council, recommended as part of the Department’s Management Information System study, will be responsible for determining how to collect the demographic and opportunity-to-learn data.

X Support for professional development needed to implement the instruction/assessment of Vermont’s standards will be built into each component. Professional development provision is the shared responsibility of the State and Supervisory Unions.

<table>
<thead>
<tr>
<th>Title I Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessments must:</td>
</tr>
<tr>
<td>X assess at least Reading/Language Arts and Mathematics;</td>
</tr>
<tr>
<td>X be administered at some time during grades 3-5, 6-9, AND 10-12;</td>
</tr>
<tr>
<td>X include multiple, up-to-date, measures (further defined as different approaches and formats) [At least two types of measurements to show proficiency on standards for showing substantial progress (multiple-choice vs extended response; on-demand vs over-time; snap-shot vs portfolio; one right answer vs open-ended). All measures must address proficiency on the standards.];</td>
</tr>
<tr>
<td>X include the measurement of higher-order thinking skills and understanding of challenging content;</td>
</tr>
<tr>
<td>X provide reasonable adaptations and accommodations for children with diverse learning needs;</td>
</tr>
<tr>
<td>X include Limited English Proficient students, and assess in the language most likely to yield accurate and reliable information in subjects other than English;</td>
</tr>
<tr>
<td>X provide individual student interpretive and descriptive reports that include individual scores, or other information on the attainment of student performance standards; and</td>
</tr>
<tr>
<td>X enable results to be reported within each State, LEA and school by gender, each major racial or ethnic group, English proficiency status, migrant status, students with disabilities as compared to those without disabilities, and economically disadvantaged students as compared to those who are not economically disadvantaged.</td>
</tr>
</tbody>
</table>
## VERMONT COMPREHENSIVE ASSESSMENT SYSTEM

<table>
<thead>
<tr>
<th>Tool</th>
<th>Content Area</th>
<th>Grade Levels</th>
<th>Participation Level</th>
<th>Reporting Level(s)</th>
<th>Frequency</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State Component</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Portfolios</td>
<td>Writing, Mathematics</td>
<td>5, 8</td>
<td>A stratified random</td>
<td>State</td>
<td></td>
<td>Mandated (If selected, a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4, 8, 10</td>
<td>sample of all students is centrally scored with the Vermont rubrics</td>
<td>Writing/ Math: Alternate Years</td>
<td></td>
<td>student’s portfolio or comparable work must be submitted)</td>
</tr>
<tr>
<td>Standards Based Tests/Assessments</td>
<td>Early Reading, English/LA, Mathematics</td>
<td>2, 4, 8, 10</td>
<td>All students: E/LA and Math are scored externally</td>
<td>Student; School; SU; State</td>
<td>Annually</td>
<td>Mandated</td>
</tr>
<tr>
<td></td>
<td>Science, History/ Social Studies/ Geography</td>
<td>6, 11, 6, 9, 11</td>
<td>Student; School; SU; State</td>
<td>Sci/SS: Alternate Years</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Local Component               |                               |              |                     |                    |           |                          |
| Student Portfolios            | Writing, Mathematics          | 5, 8         | All students’ portfolios are scored locally | Student; School; SU (Includes index of scoring consistency) | Annually | Strongly Recommended (Mandated for Title I Schools*) |
| Norm-referenced Standardized | English/LA Mathematics         | 5, 9, 11     | All students (who can take the test without adaptations) | Student; School; SU | Annually | Strongly Recommended     |
| (Recommended by State Panel)  | Others locally determined (Must be linked to Vermont Standards) |                  |          |                    |           |                          |
| Locally developed or selected assessments | Determined Locally (To address Vermont Standards not assessed elsewhere in the system) | Determine Locally | All students | Student; School; SU | Annually | Strongly Recommended     |

*Schools which are not already participating in the Vermont Portfolio Assessment Program will be required to adopt that program OR demonstrate (with technical assistance from the VTDOE) that their existing program is a comparable measure for Adequate Yearly Progress.

Portfolios offer measures of student work which comes from classroom instruction and represents each student’s best effort. Tests/assessments that are standardized usually offer measures of student work which may not be directly related to current classroom instruction and represent each student’s level of effort at one point in time.

Assessment of portfolios using the Vermont rubrics is strongly recommended, although other rubrics which address the same elements of mathematics problem solving and writing, when part of an existing cohesive local instruction/assessment plan may be used.
Part of a larger system

The Commissioner of Education and the State Board of Education are required by law, 16 V.S.A. § 165 (b), to implement a school accountability system that determines whether schools have made sufficient improvement in student performance in relation to Vermont’s Framework of Standards and Learning Opportunities. The law also specifies that schools are held accountable for demonstrating such improvement and that the Commissioner is responsible for providing technical assistance to or taking other actions in those schools making insufficient progress.

The Accountability System Based on Student Performance was developed as a result of a three-year process with the statewide Adequate Yearly Progress (AYP) Committee. Consultation with the National Center for the Improvement of Educational Assessment (NCIEA) was arranged to obtain technical support to ensure that the system was legally and technically sound. The NCIEA consultant worked with the AYP Committee through the last year of development.

Vermont’s Accountability System Based on Student Performance is part of the more comprehensive Vermont School Quality Standards, established in law and by rule and implemented by the Department of Education and local schools throughout the state. The School Quality Standards outline a broad framework that includes:

- development and implementation of standards and learning opportunities;
- use of action plans by schools and districts to focus their improvement efforts;
- assessment of students in a state comprehensive assessment system;
- development of local comprehensive assessment systems;
- reporting of results;
- professional development;
- staff evaluation; and,
- access to technical education.

In addition, the School Quality Standards cover “conditions, practices and resources,” which include school leadership, staff qualifications and requirements, support and health services, graduation requirements, instructional practices, curriculum, resources, facilities and a safe, civil and orderly learning environment.

The relationships of these parts of the larger system are illustrated in the overview that follows.
Overview of Vermont’s Accountability System

**Public Reporting** (16 V.S.A. § 165 (a)(2))
- ✔ Provides a vehicle for schools to present a broad range of information.
- ✔ Increases confidence in schools when they are clear and public about performance.

**Vermont School Report**
- 34 indicators

**School Improvement Support Guide**
- Graphing site for assessment results

**Press Conferences**
- Annual release of assessment results

**Local School Reports**
- Student performance and other indicators of health and well being.

**Action Planning** (16 V.S.A. § 165 (a)(1))
- ✔ Provides a vehicle for in-depth analysis of student learning and for focusing on needs identified in the analysis.

**Equity and Excellence through Action**

**Implementation of School Quality Standards** (16 V.S.A. § 165 (b))
- ✔ Provides a vehicle to focus on those conditions, practices, and resources that support the implementation of standards.

**Biennial Accountability Review of All Schools to Identify Schools for Technical Assistance** (16 V.S.A. § 165 (b))
- Provides an important safety net for schools:
  - ✔ making insufficient progress in students meeting and exceeding the standards; or
  - ✔ not meeting the requirements of the Vermont School Quality Standards, including safe, civil, and orderly schools.

**Two Years Of Technical Assistance For Identified Schools**

**Biennial Accountability Review Of Schools Receiving Technical Assistance**

**Commissioner’s Review** is an additional analysis of the implementation of the School Quality Standards, as required by Rule 2123.2 (b); after this review the Commissioner will recommend to the State Board:

**Other actions:**
- ✔ adjust supervisory union boundaries or responsibilities of the superintendency, or
- ✔ assume administrative control only to the extent necessary to correct deficiencies, or
- ✔ close the school and require that the school district pay tuition to another public school or an approved independent school pursuant to Chapter 21 of this title.

**Continued Technical Assistance**

**OR**
Design Decisions

The purpose of the accountability system is to help students become well educated and productive individuals and citizens, and to help schools do an effective job of educating Vermont’s children. There is a strong commitment to equity — that *every child counts*. There is recognition that school accountability can provide the right balance of pressure and support to achieve these goals. The Department and Board are strongly committed to accountability as a means to help schools improve. Accountability is used not to be punitive but rather to strategically provide supportive, positive assistance. The Department’s record of working with schools identified during the transitional accountability system supports this view. At the same time, the Department is deeply committed to supporting schools and districts proactively, seeking the educational welfare of children. The accountability system is fundamentally about improving the education of children.

The accountability system is designed to achieve these purposes and to be technically sound and operationally feasible. This *Manual* reflects the balance of these design specifications.

**The components of this system include**

- *Data from student performance and those indicators most closely associated with student performance* - This aspect of the school accountability system is based on measurable student performance. “Input” and “process” variables are addressed in the accountability system based on *School Quality Standards*.

- *Data from statewide testing* - Vermont has an operational state assessment program which incorporates statewide testing data. The accountability system is structured to include other test instruments, and may be changed as new assessments or data become available.

- *Testing data from cohorts/classes at the same grades each year* - The cross-sectional design gathers results from the same grade each year. For example, to gauge school performance over time, results from one class of grade 4 students are compared with the results from another class of grade 4 students the next and following years.

- *Accounts for all students* - The system accounts for all students enrolled in the school. Students who do not have valid assessments or who are not assessed and do not have a valid excuse or exemption will receive the point value for the lowest performance level, Little Evidence of Achievement, because there is no evidence of their performance.

- *Data from local assessments* - The system encourages the development and use of local assessments by providing schools the option of including student performance data from such assessments in the school accountability system.

- *Identification of schools for technical assistance* - This system identifies schools by applying quantitative formulas to the test results. It minimizes the staff resources required for identification and review, so that those resources may be devoted to assistance and support.

- *Status and Change Indexes* - The system holds schools accountable for their absolute performance at a point in time (status) and requires that performance improve over time (change). The focus on absolute performance means that each school is evaluated in relation to a standard, not to other schools. The focus on improvement means that each school is reviewed based on its required growth, rather than on how other schools change.

- *Flexibility* - This system is designed to incorporate modifications to the state’s student assessment system, expanded local assessments, and refinements of growth expectations for schools and other necessary modifications.
• **Safeguards, Reviews and System Supports** - The system includes the following safeguards:

- An appeals process
- A review of other local data will be applied during Cycle 1
- A further review of data from identified small schools
- A review of the School Quality Standards prior to the Commissioner recommending other actions beyond technical assistance
- Allows for adjustments by the Commissioner, if necessary, if Growth Targets or other aspects of the system are determined to be systematically unreasonable
- Recommendations for inclusion of newly implemented state assessments in the Accountability System by the Technical Advisory Panel
- Establishment of criteria for inclusion of local assessments by the Vermont Technical Advisory Panel
- Independent external audit to determine the effectiveness of the accountability system based on student performance
Supplemental Materials

New York

- State Testing and Graduation Requirements—How New York Fits Into the Framework
  Roseanne De Fabio
State Testing and Graduation Requirements:
How New York Fits into the Framework

<table>
<thead>
<tr>
<th>Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Commencement level of standards represents skills needed for college entry without remediation and for high-skilled employment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>• State Regents examinations required in 5 subjects. All Regents are standards-based and criterion referenced.</td>
</tr>
<tr>
<td>• Each test includes multiple choice and constructed response items</td>
</tr>
<tr>
<td>• Portfolios are not used for state assessment.</td>
</tr>
<tr>
<td>• Accommodations that do not alter construct being measured allowed for students with IEPs.</td>
</tr>
<tr>
<td>• Approved alternative tests are allowed for each Regents (e.g., AP, IB, Cambridge International, SAT II)</td>
</tr>
<tr>
<td>• Component Retesting</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uses and Consequences of Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Students must pass five Regents exams to graduate</td>
</tr>
<tr>
<td>• Regents exam performance is basis of high school accountability system</td>
</tr>
<tr>
<td>• Two levels of high school diploma — Regents diploma (5 exams) and Advanced Regents diploma. Either may have Career &amp; Technical endorsement.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Supports for Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Students make continue to retake Regents exams.</td>
</tr>
<tr>
<td>• Offered 3 times per year.</td>
</tr>
<tr>
<td>• Academic Intervention Services are required for at risk students.</td>
</tr>
<tr>
<td>• Standards-based tests are central to New York’s education reform.</td>
</tr>
<tr>
<td>• Core curriculum guides provided by State for aligning instruction to standards and assessments. Research study as well as test results indicate varying degrees of alignment.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Implementation Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Low-pass score permitted during phase-in period.</td>
</tr>
<tr>
<td>• Board of Regents will consider extending phase-in period to permit continuation of low-pass score.</td>
</tr>
</tbody>
</table>
Supplemental Materials

Connecticut

- State Testing and Graduation Requirements—Connecticut
  - Betty Sternberg

- Position Statement on Measuring Success
  - Connecticut State Board of Education
CONNECTICUT ACADEMIC PERFORMANCE TEST (CAPT)

Standards

- Materials in Common Core of Learning (CCL) and Frameworks are high level
- Knowledge, skills and critical thinking abilities tested represent a demanding level of achievement, reasonable to expect of students in the spring of their 10th grade—NOT BASIC
- Low stakes if definition means tied to graduation
- High stakes if definition means scores are reported for individual students, not sample testing

Assessments

MATHEMATICS
- 8 open-ended items—students show work and explain reasoning
- 24 grid-in items
- No multiple choice
- All contextually based
- Calculators allowed for entire test

READING AND WRITING ACROSS THE DISCIPLINES

Reading Across the Disciplines
1. Response to literature—demonstrate initial understanding, interpretation, connection and critical stance by reading and responding to four open ended questions about a short story
2. Reading for information—demonstrate developing an interpretation and a critical stance by reading three pieces of authentic non-fiction material and responding to 12 multiple choice and 6 short-answer questions

Writing Across the Disciplines
1. Interdisciplinary writing
   - Assesses students' abilities to think critically, solve problems, make decisions and communicate their ideas in a realistic and interdisciplinary context
   - Two test items each with three short, non-fiction informational and persuasive reading pieces
   - One extended response to assess how well students write to persuade
2. Editing and Revising
   - Assess students' editing and composing/revising skills
   - Three short reading passages with embedded errors
   - Eighteen multiple-choice questions

SCIENCE
1. Experimentation
   - Performance task precedes written test
   - Four open-ended and four multiple-choice items directly related to performance task
2. Life Science
   - One open-ended and 16 multiple-choice items
3. Physical Science
   - One open-ended and 16 multiple-choice items
4. Earth Science
   - One open-ended and 16 multiple-choice items
Uses and Consequences of Assessments

- Scores produced for individual students
- Scores reported on individual students' transcripts
- Scores NOT tied to graduation
- Scores NOT tied to a formalized accountability system 1985-1999
- Scores tied to a formalized accountability system 1999

Supports for Learning

- Retest option in 11th and 12th grade
- Fostered interdisciplinary collaboration of teachers
- Fostered modification of instructional techniques and use of time

Implementation Strategies

- 9th grade proficiency test (low level, low stakes) 1980-1985
- Criterion Referenced Tests in 4, 6 and 8th grade since 1985
- Connecticut Academic Performance Test (CAPT) in 10th grade since 1995
The State Board of Education defines successful students as those who can read, write, compute, think creatively, solve problems and use technology. All students should also enjoy and perform in the arts and athletics and understand history, science and other cultures and languages. Each student must be responsible for his or her learning and behavior, work well with and be helpful to others, and contribute to the community. Every student must graduate from high school and be prepared to move on to productive work and further study and to function in the global economy. Ultimately, students must become active citizens and lifelong learners who lead healthy lives.

We can meet our responsibility of providing an educational experience that achieves these goals only if we regularly and consistently assess how well our students are doing and make the changes that are necessary in order to improve all facets of our schools.

**SETTING STANDARDS.** The most critical set of responsibilities for a local board of education is to articulate clearly what success means in its district; establish standards of performance; measure performance against those standards; regularly make this information available to the public; and ensure that this information is used to make good decisions which support student success. Defining standards for success and continually monitoring progress enable schools and school districts to make informed decisions about allocation of resources, curricular priorities and new initiatives that will directly enhance the success of all their students. Schools must constantly build on their accomplishments, while also addressing areas in need of improvement.

**MEASURING SUCCESS.** Because success is multifaceted, it must be assessed using multiple measures: academic achievement over an extended period of time; student achievements that are other than academic; unique local indicators that represent community values; and the extent to which the performance gaps between various groups of students (by gender, race, economic status, etc.) are being reduced. The responsibility of measuring success is one shared by local boards of education and the State Board of Education.

There are several sources of data that a local board of education may use to measure the success of its students. The Connecticut State Department of Education (SDE) provides statewide, district and regional data on the critical aspects of education that contribute to and measure student performance. These include student test scores; dropout rates; graduate follow-up data; student participation in various courses and programs; expenditures; class size; and number of academic computers per student.

This information is published on the SDE web page and in publications such as the Strategic School Profiles (SSP), Condition of Education, Special Education in Connecticut, and CMT and CAPT Interpretive Guides. Each district should also use local
data such as local assessments and teacher observations; measures of opportunities for students to learn with students from diverse backgrounds; community service work; participation in extracurricular activities; professional development acquired by teachers; parent involvement; and other indicators unique to the environment of the school and community.

CAUTIONS. Connecticut Mastery Test (CMT) and Connecticut Academic Performance Test (CAPT) results provide important information about student performance on a selected set of skills and competencies in mathematics, reading and writing in Grades 4, 6 and 8, and also science in Grade 10. However, these results do not provide a comprehensive picture of student accomplishments. There is a danger that overemphasizing state test scores to evaluate a student’s, a school’s or a district’s performance can result in an inappropriate narrowing of the curriculum and inappropriate classroom instructional practices. Focused preparation for state tests should be a small fraction of a year-long comprehensive curriculum that balances the competencies assessed on the state tests with other critical skills and objectives. Teaching isolated skills for test preparation or using repetitive tasks that go far beyond reasonable practice do not represent good instruction. In addition, no one assessment—state or local—should be the sole basis for promotion, graduation or other important decisions in the education of a student.

REPORTING RESULTS. Each local board of education is required by law to annually hold a public discussion with the community on relevant student-related information from its Strategic School Profile. There should also be ongoing communication with the community on the successes and needs of its schools. This can include additional information about students, curriculum, staff, new initiatives and programs, and evaluations of programs’ effectiveness. When and where appropriate, the use of technology (e.g., e-mail, school and district web pages, closed-circuit broadcasting) should be used to provide the community with greater access to critical information on and understanding of the district’s performance. We encourage local boards of education to communicate frequently and openly with the communities they serve, using every effective means available.

IMPROVING INSTRUCTION. Defining standards, measuring success and reporting the results are important steps in the process of improving education. However, the most important step is using this information to make good decisions about adjusting curriculums, improving teaching, designing new programs and providing more specific and more effective instruction for each student.