

04 - 0201

esota
Department
of Education

**RECOMMENDATIONS
MINNESOTA'S HIGH
ACADEMIC STANDARDS**

**REPORT TO THE
LEGISLATURE**

February 2004

**As required by
Minnesota**

2003 Regular Session Laws

Chapter 129

Article 2

Section 2



COMMISSIONER
CHERI PIERSON YECKE, Ph.D.

DEPUTY COMMISSIONER
CHAS ANDERSON

FOR MORE INFORMATION CONTACT:

Pat Olson, Assistant Commissioner
Minnesota Department of Education
Office of Accountability and Improvement
T: (651) 582-8856

FAX: (651) 582-8874

E-MAIL: pat.olson@state.mn.us

1500 Highway 36 West
Roseville, MN 55113-4266
(651) 582-8200 TTY: (651) 582-8201

Upon request, this report can be made available in alternative formats.

RECOMMENDATIONS

**MINNESOTA'S HIGH
ACADEMIC
STANDARDS**

**REPORT TO THE
LEGISLATURE**

February 2004

As required by
Minnesota

**2003 Regular Session
Laws**

Chapter 129
Article 2
Section 2

This report required the collection of information that the Department of Education does not collect as part of its normal business functions. It was therefore necessary to gather and analyze information in order to prepare this report. The cost of preparing this report includes estimates of the Department of Education information collection costs as well as the estimated costs of the providers of the information. The cost of preparing this report was \$4,582.00.



DESCRIPTION OF REPORT

The Commissioner of MDE, after consulting with stakeholders, must provide written recommendations including the following:

1. Identification of cut-scores on reading and math assessments that indicate remedial instruction in college is unneeded;
2. Alternative assessments;
3. Whether students must pass a state end-of-course graduation exam;
4. Feasibility of including state percentile rankings and a national comparison; and
5. A method for using the Grade 8 math and language arts tests to satisfy basic skills requirements.
6. Attachment A: Special Education Alternate Assessment
7. Attachment B: Special Education Functional Skills Alternate Assessment

Making use of the high school MCA results for higher education – options and recommendations

Executive Summary

Minnesota is developing new high school Minnesota Comprehensive Assessments, as part of its statewide accountability system and in accordance with provisions of the federal legislation "No Child Left Behind". These tests are in transition, as they are aligned with the new academic standards. Minnesota Session Laws 2003, Chapter 129, Article 2, Sec 2 required that recommendations be delivered to the Legislature by February 2004 to *Identify the cut-scores on high school reading and mathematics assessments indicating that remedial instruction in the state's two-year higher education institutions is unneeded.*

Minnesota Statute 120B.365 established an Assessment Advisory Committee. This Committee was comprised of school superintendents, teachers, representatives of higher education, and members from the community at large. The Committee began meeting in September 2003 to analyze these issues.

The Commissioner of Education agrees with the Committee's belief that it is important to ensure that these new MCA-II tests provide information that is useful and relevant for college placement purposes. Making this happen requires collaboration across institutions involved in P-16 education. The Commissioner recommends that the existing collaboration through the Minnesota P-16 partnerships that is already making progress should be supported.

Background

Students will benefit when higher education clearly articulates its expectations for students, and when higher education and pre-K-12 work to assure that students are appropriately prepared to take advantage of their post-secondary education. MCA-II test scores should in principle be able contribute to a better understanding of the student's likelihood of success in a college-level math or reading program through showing that students appear to be "on-track" to success at the time they take the MCA-II tests.

The high school Minnesota Comprehensive Assessments (MCAs) were first given in production versions in 2003. These students have yet to enter college¹. The tests are in transition, as they are being aligned with the new academic standards. The reading test has limited writing requirements. There is also a writing test.

Some generalizations of the skills shown by typical students at different score levels are possible (and is part of the standards-setting procedures), but these typical patterns are only approximately predictive for individuals.

A project called "Standards for Success" examined the alignment of the tests by using expert raters to code the apparent alignment of items against the requirements of the Knowledge and Skills for University Success (KSUS) standards. [KSUS standards were developed by about 400 faculty members across twenty universities]. The results of this project suggested that the MCA high school mathematics test has potential to provide useful information related to post-secondary readiness in three of four areas where the content is covered (Algebra, mathematical reasoning and geometry). The challenges for utilizing the MCA reading tests may be greater, as the MCA reading test (grade 10) was seen by the project as having limited potential to provide information related to post-secondary readiness.

¹ There was a full length pilot version of the high school MCAs in 2002. The P-16 partnership working group expects to complete a data analysis based on these results and some of these students' achievements in first semester college-level study during 2004.

Key Issues:

A test result is only *one* piece of information, and it is a single piece of information at a single time. The MCA-II mathematics test is planned for spring of grade 11. College is over twelve months later. So the result can not represent what *will* definitely happen, that a student will or will not need particular instruction later on. The test result can represent whether the student appears to be "on-track" for success, giving students, parents and teachers some information about what still needs to be done. The result can also be used as one of several factors to be taken into account in making college-level placement decisions.

To establish connections between the high school tests and college readiness (with the implication that therefore remedial instruction will not be needed), it is necessary to identify entry-level expectations for mathematics and reading for colleges and universities. Then these expectations would need to be compared to identify the extent and nature of any common ground. Next, the alignment of the common expectations with the grade 10 reading and grade 11 mathematics tests would need to be determined, as well as any changes/enhancements required to these tests to bring them into a suitable alignment.

Currently, an analysis of college-level mathematics placement tests used in Minnesota is being facilitated through a working group set up by the P-16 Partnership. This analysis seeks to determine the nature and extent of differences and common ground, both amongst these tests and between the tests and the high school mathematics standards. The next step in this work is to identify the nature and extent of any required enhancement to the anticipated high school mathematics MCA-II. After reviewing the placement of the tests, MDE would identify changes/enhancements to the tests needed to improve the alignment and determine results that report student achievement in relevant ways.

In order for these changes to be helpful, relevant scores must be available to higher education in some effective and efficient way such as including the scores on transcripts. Colleges and universities would then have to develop procedures to make effective use of this information. Data must be generated, collated, and studied to allow review of the validity of the information as used in this manner.

Tests given at different times and for different purposes may not be universally applicable to meet all needs. For example, a test designed to determine whether all students have met a certain basic standard in 10th grade may not be sufficiently discriminating at higher levels to make it useful in college-level placement (as much as two years later).

Developing the high school mathematics MCA-II to provide information relevant to college-level placements may have important potential implications for students, parents and teachers. For example, if a mathematics test given in grade 11 shows that a student needed further study before being ready for college-level courses, it would allow time for the student to take appropriate further study in grade 12.

Some results of the initial stages of the work of the P-16 Partnership in this area are expected in January/February 2004.

Recommendations

The Commissioner of Education concurs with the Assessment Advisory Committee in acknowledging the valuable work of the Minnesota P-16 partnership on establishing the links between the high school Math MCA and the requirements of college-level courses and recommends that the work of the P-16 partnership be continued and supported by the partnership's sponsors.

The Commissioner recommends that

- development of the capacity of the high school MCA-II tests to give information of use in placement in college-level courses be continued
- potential users of such information be asked to develop ways to use such information
- schools be asked to record information about students' results on the high school MCA-II on student transcripts, and
- the information about a student's achievement on the high school MCA-II tests be *one* of several pieces used to make college-level placements, including remedial course placements.

Additionally, the Commissioner would like to see the development of the links between high school tests and college-level placements be extended, if possible, to the high school reading test after the work in mathematics is completed, and will consider the potential relevance of the high school tests to college-level placements in decisions about the most appropriate grade-placement for these tests.

1 Appendices

1.1 Overview of advisory structure for MDE statewide assessment program

- a) MDE's assessment program is advised by:
- i) the Assessment Advisory Committee, formally constituted, responsible for advice to the Commissioner about principles and guidelines for a coherent and effective assessment system
 - ii) Subject Advisory Committees (reading, mathematics and science), formally constituted, responsible for advice to the Office of Academic Excellence on principles and guidelines needed to ensure a coherent K-12 perspective within each subject area. As at December 2003, these committees have yet to be constituted.
 - iii) Teacher/community panels, constituted on an *ad hoc* basis as required to do the detailed work (test specifications, alignment studies, item review, standards setting) needed to provide the Division of Assessment & Testing with the advice required to implement the Commissioner's approved policies and guidelines
 - iv) the National Technical Advisory Committee, with members who provide expert advice on technical issues associated with assessment and testing. Present National TAC members are
 - (1) Dr Roger Trent, former state assessment director Ohio
 - (2) Professor Mark Davison, Minnesota Office of Educational Accountability
 - (3) Emeritus Professor William Schafer, University of Maryland
 - (4) Dr Susan Phillips, legal and psychometric consultant
 - v) the Local Technical Advisory Committee with members drawn from district testing coordinators, OEA, and other people with knowledge and expertise in assessment.

1.2 Assessment Advisory Committee – membership

- vi) two superintendents, selected after consultation with MASA
- Roger Giroux, Superintendent Anoka-Hennepin *
 - Rick Ellingworth Superintendent Redwood Falls *
- vii) two teachers, selected after consultation with Education Minnesota
- Neil Witikko, Hermantown High School *
 - Rose Gundacker, Rosemount High School

viii) two members from higher education

- Sandra G. Johnson, Assoc. Professor, St. Cloud University
- Dr. Craig Swan, Vice Provost for Undergraduate Education

ix) five additional members

- Dr. Valerie Pace, State Mgr. IBM Corporate Community Relations
- Glenn Dorfman, Chief Operating Officer, Minnesota Assoc. of Realtors
- Yusef Mgeni, Director, Office of Educational Equity, SPPS
- Virginia Richardson, Director, Parent Training, PACER
- Jerry Hanson, Principal & Title I Dir., Roosevelt Elementary School (selected after consultation with MESPA and MASSP)

1.3 Functions of the Assessment Advisory Committee

Under MN 120B.365, the Assessment Advisory Committee reviews statewide assessments and the Commissioner will consider the recommendations of the committee before finalizing a statewide assessment.

The functions of the Assessment Advisory Committee are to support the validity and technical adequacy of statewide assessments by

- a) responding to requests for advice from the Commissioner about assessment and testing issues
- b) advising the Commissioner about principles and guidelines for
 - (1) developing test specifications
 - (2) reviewing tests against test specifications
 - (3) aligning tests and test results with the state's high academic standards
 - (4) reporting test results to parents and teachers in ways that indicate students' educational needs.

Alternate Assessments

Introduction*

Alternate assessments are designed to measure the performance of students who are unable to participate in general large-scale assessments used by districts and states. This is due, in part, to the reference to "alternate assessment" in the 1997 Federal Reauthorization of the Individuals with Disabilities Education Act (IDEA), which called for states to have alternate assessments in place by July 1, 2000.

The primary purpose for alternate assessments in state assessment systems is to increase the capacity of large-scale accountability systems to create information about how a school, district, or state is doing in terms of overall student performance. Gathering data on the performance of students through alternate assessments requires rethinking traditional assessment methods. An alternate assessment is neither a traditional large-scale assessment nor an individualized diagnostic assessment. Alternate assessments are designed for a small number of students who are unable to participate in general large-scale assessments. Only a small percentage of students with disabilities should participate in alternate assessments; most students participate in the general state or district assessment with accommodations as needed. For students with disabilities, alternate assessments can be administered to students who have a unique array of educational goals and experiences, and who differ greatly in their ability to respond to stimuli, solve problems, and provide responses.

Because the expectations that educators and parents hold for students with disabilities vary, it is critical that clear guidelines be established to decide who participates in alternate assessment systems. This decision should **not** be based on which students are expected to perform poorly on the general education assessment.

Alternate assessments need to be aligned with the general curriculum academic content standards set for all students, and should cover the broad content areas (such as reading, mathematics, writing) assessed in the general assessment.

* This section is an excerpt of a report from a leading national resource for alternate assessments — **Measuring Academic Achievement of Students with Significant Cognitive Disabilities: Building Understanding of Alternate Assessment Scoring Criteria**, National Center for Educational Outcomes, University of Minnesota, Synthesis Report 50 / 2003.

Title I Regulation on Alternate Achievement Standards

- Alternate assessment based on alternate achievement standards must be aligned with the State's academic content standards, promote access to the general curriculum, and reflect professional judgment of the highest achievement standards possible.
- Each State using alternate achievement standards must establish clear and appropriate guidelines for individualized education plan (IEP) teams to apply in determining when a child's disability justifies an alternative assessment.
- When measuring Adequate Yearly progress (AYP), State and school districts have the flexibility to count the "proficient" and "advanced" scores of students who take the alternate assessment as long as the number of those proficient and advanced scores does not exceed one percent of all students in the grades tested.
- Under IDEA, IEP teams do *not* have complete discretion regarding the assessment of students with disabilities. The team decides *how* a student participates, not *whether* the student participates in the assessment. States must develop and disseminate guidelines to inform IEP teams about how students may be assessed properly.
- State guidelines for the use of the alternate assessment are communicated to local school districts.

Minnesota's Alternate Assessment System

The Individuals with Disabilities Act (IDEA, 1997) requires that children with disabilities be included in general state and district-wide assessment programs. The law requires states to evaluate students with an alternate assessment when a student is exempted from a statewide accountability test. 34 C.F.R. § 300.138 (b). When a student with an Individualized Education Plan (IEP) or Section 504 Plan is exempted from Minnesota statewide accountability testing, the state-designated alternate assessment **must** be used and the data **must** be submitted to the Minnesota Department of Education.

Minnesota's system for alternate assessments was developed in response to the reauthorization of IDEA in 1997. During the school year 1998-1999, a group of stakeholders including representatives from general and special education, school psychologists, the Minnesota Association of School Psychologists, Education Minnesota, and the National Center for Educational Outcomes at the University of Minnesota, met to determine what approach to take in developing the system. The stakeholder group recommended building upon a framework already in place in select districts to measure student progress, the continuous progress scales. These scales were modified and became the foundation for the alternate assessment. Priority was given to the development of a system of accountability that would have a minimum cost to the district, an ease of administration and scoring at the district level, and efficiently measure the achievement of a small number of the total group of students in special education. In the spring of 1999, the alternate assessment was piloted statewide. The pilot involved hundreds of general and special education teachers working together to examine

the scales for reading and math in order to gain confidence that the scales adequately covered grade-level content and skills. Following the pilot, the alternate assessment scales were modified. Beginning in the 1999-2000 school year, the alternate assessment was available online for use in all districts. It has been part of the statewide system of accountability since that time and is part of a Federally required reporting system covering a variety of aspects of the special education program in each state.

In 2000, another scale, the Functional Skills Rating Scale, was developed in response to the need to assess students with severe cognitive impairments. These scales were developed in conjunction with the Special Education Regional Low Incidence Facilitators, a statewide team of experts in the less frequently occurring disabilities including Mild to Moderate Mental Impairments.

In addition to previous alternate assessment requirements, the Reauthorization of ESEA 2001 (No Child Left Behind Act) requires the collection of some new information and a close accounting for results related to reading and math at grades 3, 5, 7, 10 and 11. Under the No Child Left Behind Act, districts are allowed to count 1% of their total scores as "proficient" using the alternate assessment. If more than 1% of the student population happens to take the alternate assessment, the number reported as "proficient" is limited to a cap of 1%; additional scores must be reported as "not proficient."

The decision to exempt students rests solely with the student's IEP or 504 team. The parameters for these exemption decisions are defined by Minn. R. 3501.0090, Subpart 1(a)(3):

An exemption from the statewide standard shall be granted to a special needs student when the student cannot demonstrate the required degree of learning with appropriate accommodation or modifications if:

- (a) the student's IEP or section 504 accommodation plan does not and never has included the requirements on which the tests are based; or
- (b) the student is enrolled in special education classes for the subject matter included in the test, but the student's IEP or section 504 accommodation plan does not include a majority of concepts tested.

A decision scale has been developed by the Department of Education (Special Education Policy Division) to assist district staff in the appropriate use of the alternate assessment with individual students.

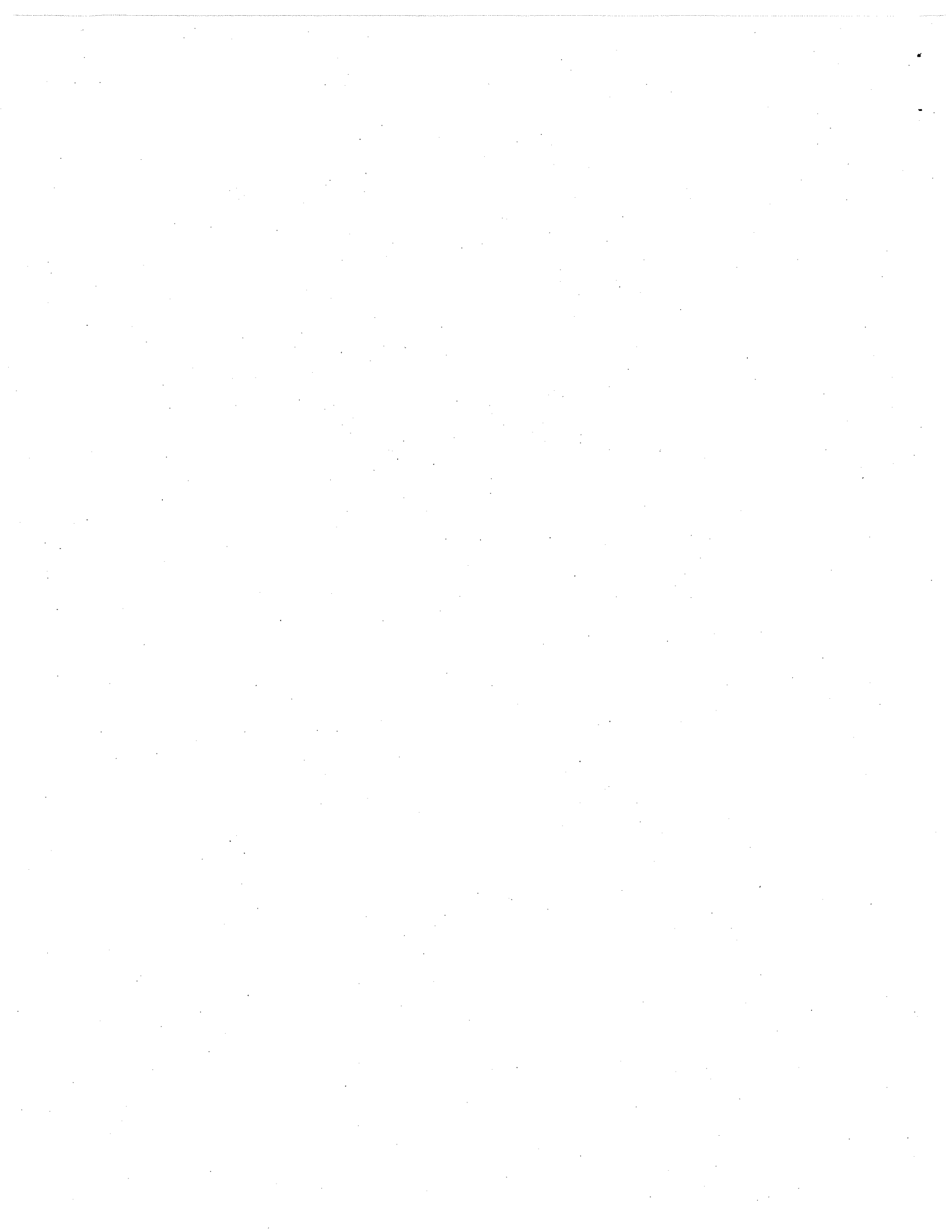
When a student is exempted from a specific content area on a state test because of a gap between daily academic skill focus and the content of the state test, progress in relevant skill area is assessed using a seven-point rating scale. The scale is used to rate progress on a continuum ranging from awareness to understanding to application of grade-level academic skills. Points are assigned to rate the student's level of functioning based on teacher observation. Separate scores are given in the areas of reading, mathematics and writing at various grade levels. See Attachment A for an example of a content area alternate assessment. The scores achieved on this alternate assessment have been norm-referenced and correspond to the proficiency levels 1-5 currently used in reporting student performance on the statewide assessments.

Some students who are exempted from the state tests are working on functional living skills rather than developmental academic skills. These students tend to be those with the most significant cognitive disabilities. For state accountability purposes, their progress towards acquiring functional living skills is evaluated using the Functional Skills Rating Scale. They are also assessed by using a seven-point scale in the areas of home living, recreation and leisure, community participation, jobs and training, social skills, communication, and academics. See Attachment B for an example of the Functional Skills Rating Scale.

Recommendations

- The Minnesota Department of Education should continue to implement the alternate assessment for students with disabilities who are determined to be unable to participate in the statewide assessments.
- Minnesota's system for the alternate assessment should continue to align with the content areas and proficiency levels of the Minnesota statewide assessments.
- The Minnesota Department of Education should integrate the alternate assessment data collection system into the system for data collection and reporting of progress for statewide assessments.
- The Special Education Policy Division of the Minnesota Department of Education should continue to provide training and technical assistance to local school districts for the correct implementation of the alternate assessment; including compliance with the 1% rule of the No Child Left Behind Act.

END-OF-COURSE EXAMINATIONS



Executive Summary

Minnesota currently requires public school students to pass the Basic Skills Tests (BST) in mathematics, reading, and writing composition before graduation from high school. Reading and mathematics tests are administered in eighth grade, and the written composition test is administered in tenth grade. Students who fail to pass the test the first time are afforded multiple opportunities to re-take the test.

Minnesota Session Laws 2003, Chapter 129, Article 2 Section 2 required that recommendations be delivered to the Legislature by February 2004 to recommend whether students must pass state *end-of-course* examinations as a requirement of graduation. (As a point of clarification, end-of-course exams are different than the currently mandated BST in that they are administered upon completion of a specific course and cover information learned only in that course.) In the course of preparing this report, the Department also examined the issue of lack of student motivation on tests that lack serious consequences, such as the MCA's.

The Minnesota Department of Education recommends that in the area of testing, the State stay the course at this time in its goal of implementing a statewide accountability system based on academic achievement. At a later date, consideration of adoption of legislation of end-of-course exams should occur, as this type of exam creates a closer connection for the student between the courses taught and subsequent testing, which will undoubtedly improve their motivation to prepare well for the test. On the issue of motivation to do well on lower-stakes tests, such as the MCAs, the Department urges the Legislature consider implementing the Governor's Seal of Honor and Governor's Scholar of Distinction programs, which would recognize and reward participating students for educational excellence. The importance of the Governor's Seals will be promoted to businesses and higher education institutions, and they will be urged to make receipt of such seals one criterion for scholarships they might award. Districts may also want to consider placing MCA scores on transcripts or incorporate high school MCA scores into a course grade.

Context and Purpose of this Report

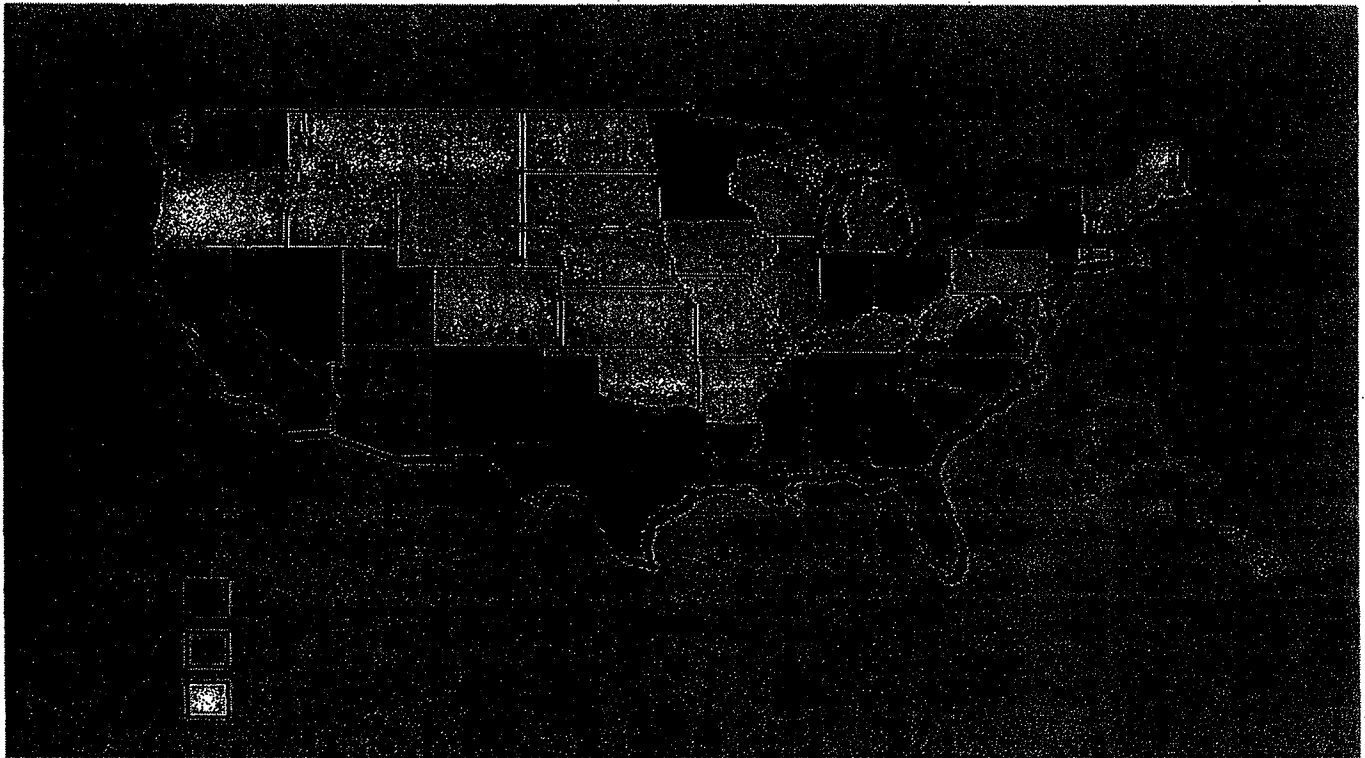
Minnesota Session Laws 2003 Chapter 129, Article 2, Section 2 required the Commissioner of Education to recommend whether students must pass state end-of-course examinations as a requirement for high school graduation. This report presents background information, key issues, options, and recommendations.

Background

According to a recent report released by the Center on Education Policy, as of 2003, 19 states have legislated mandatory end of course or exit exams, and five more states plan to phase in the tests by 2008.

States Requiring End-of-Course Exams

(Source: Center on Education Policy, based on information collected from state departments of education, June 2003)



Map legend:

Dark shading: States with mandatory exit exams in 2003:

AL, FL, GA, IN, LA, MD, MA, MN, MS, NV, NJ, NM, NY, NC, OH, SC, TN, TX, VA (19)

Medium shading: States phasing in exit exams by 2008 but not yet withholding diplomas:

AK, AZ, CA, UT, WA (5)

Light shading: States with no mandatory exit exams

In Minnesota's public schools, students are required to pass the BST before graduation. The BST ensures that students graduate from Minnesota public high schools with essential competencies in reading, mathematics and writing. Students have multiple opportunities to pass these tests before graduation.

All public school eighth grade students must take the reading and math tests, and all public school tenth grade students must take the written composition test. Older students who have not yet passed a BST may retake the test with first time test takers, and retake opportunities are also offered in the summer for all students and in the spring for seniors. Students in special education who have an Individualized Learning Plans (IEP) or 504 Plan may have their scores modified for a Pass Individual (PI) rather than a Pass State (PS) earned by student in regular education. Eighth grade students with limited English proficiency who have been in the U.S. for less than a year are waived from taking the eighth grade tests. Tenth grade students with limited English proficiency who have been in the U.S. for less than a year are waived from taking the tenth grade test. For students in grades 9-12 who have not been enrolled in a school for three or less years in which the primary language of instruction is English, the temporary 36-month exemption for the BST graduation requirement also remains.

To pass the reading test, students must be able to demonstrate a reasonable understanding of factual information. Students must read a passage and be able to:

- Identify the main idea
- Recognize supporting information
- Identify the meaning of words and phrases
- Recognize the author's point of view
- Draw logical conclusions
- Distinguish between fact and opinion

To pass the mathematics test, students must be able to solve problems involving numbers, shapes and symbols commonly used in adult life. The test covers material introduced to students before the sixth grade.

Students must be able to:

- Solve problems involving whole numbers, fractions, decimals and integers
- Solve problems involving percents, rates, ratios and proportions
- Use concepts of number sense, place value and number relationships to compare, order and determine equivalence
- Use estimation in problem solving
- Apply measurement concepts
- Read, interpret and use tables, charts, maps and graphs to analyze data, identify patterns and make predictions
- Use elementary concepts of probability and statistics
- Apply geometric and spatial relationships

To pass the written composition test students must write a clear message in English for an adult reader that is well organized and has only minor mechanical or spelling errors. Successful compositions must have:

- A clearly stated message
- Supporting ideas related to the message
- A clear connection among ideas
- A logical beginning, middle and end
- Correct use of grammar, spelling, and punctuation

Passing scores on the BST in reading and mathematics are defined in terms of a scale score. Scale scores allow for comparisons when test questions have varying difficulty levels and when the questions change from year to year. This type of scoring system ensures scoring fairness and consistency from year to year. The passing score is 600. Students must earn a score of 3 or above on a scale of 0 to 6 to pass the written composition portion of the test.

**Percentage of Minnesota students passing an exit exam on the first try
(All students and subgroups)**

Student Subgroups	MN Math 2003	MN Reading 2003
All	72%	81%
White	78%	87%
Black	33%	49%
Hispanic	43%	55%
Asian	61%	62%
ELL	34%	35%
Free/reduced lunch	49%	60%
Students with disabilities	30%	42%

There has been considerable debate over whether high-stakes testing has a positive motivating effect on students. One recent study,² performed by Mr. Braun of the Educational Testing Service, examined this issue and found mixed results. When he looked at overall changes in previously identified high stakes states, he found an overall academic gain in 4th and 8th grade test scores. However, when Mr. Braun looked at how cohorts of students fared over time, the improvements in academic achievement were greater – albeit to a lesser degree – in states with low-pressure testing systems.³

³

Key Issues

There are numerous issues to consider when contemplating the use of end-of-course exams. The table below details the positive and negative aspects of this type of exam.⁴

Pros	Cons
<p>Proponents of end-of-course exams say:</p> <ul style="list-style-type: none"> ➤ Students know what is expected and that the test really counts, so they work harder. ➤ Schools identify and can address student weaknesses early. ➤ Schools discover areas of overall weakness, prompting them to refocus resources where they are most needed. ➤ Education across the state is more consistent, eliminating situations where schools in some districts are superior to others. ➤ The public sees gains from year to year and regains confidence in public schools. ➤ Exit exams appear to be changing curriculum and instruction in ways that have led to greater internal curriculum alignment and focus within districts and schools. 	<p>Critics of end-of-course exams say:</p> <ul style="list-style-type: none"> ➤ States have tried to do much too soon without the proper preparation and support for everyone involved. ➤ The tests are sometimes too hard, lead teachers to teach to the test, take time away from instruction, and are expensive. ➤ Teachers may be unprepared to teach to the standards, and students may claim they're being tested unfairly on content they haven't yet had. ➤ Exit exams force teachers to squeeze out content not covered by the tests, and cover more contents superficially rather than fewer topics in-depth. ➤ Some studies offer a moderate degree of evidence that exit exams are associated with higher dropout rates. Other studies have found no such link. ➤ Cost ➤ Issue of "too much testing"

¹Braun, H. (2004, January 5). Reconsidering the impact of high-stakes testing, *Education Policy Analysis Archives*, 12(1).

²Viadero, Debra, *Study Offers Mixed Results on Impact of High-Stake Tests*. *Education Week*. January 28, 2004.

³Sources: *State High School Exit Exams Put to the Test*, Center on Education Policy, August 2003
Education Commission of the States Web Site <http://www.ecs.org>, Issue area: Assessment, Sub-issue: High Stakes/Competency

Options and Recommendations

For standards-based accountability to work, clearly written, comprehensive, and rigorous statewide standards must be first implemented. The standards must detail what students should know and be able to do in core subjects. Minnesota, in the adoption of its new standards, is well on its way to fulfilling that charge.

Implementation of end-of-course exams may be more difficult without this framework. For example, some states that have instituted end-of-course exams have had to re-trace their steps and change their laws due to opposition and controversy created after end-of-course testing commenced. Some states created waivers, special exemptions, or alternative routes to a diploma. Other states delayed or suspended implementation, or lowered cut scores. One state went so far as to void test results. However, some states have held fast in the face of criticism.⁵

The Department recommends that in the area of testing, the State stay the course at this time in its goal of implementing a statewide accountability system based on academic achievement. At a later date, consideration of adoption of legislation of end-of-course exams should occur, as this type of exam creates a closer connection for the student between the courses taught and subsequent testing, which will undoubtedly improve their motivation to prepare well for the test. On the issue of motivation to do well on lower-stakes tests, such as the MCAs, the Department urges the Legislature consider implementing the Governor's Seal of Honor and Governor's Scholar of Distinction programs, which would recognize and reward participating students for educational excellence. The importance of the Governor's Seals will be promoted to businesses and higher education institutions, and they will be urged to make receipt of such seals one criterion for scholarships they might award. Districts may also want to consider placing MCA scores on transcripts or incorporate high school MCA scores into a course grade.

⁵ *State High School Exit Exams Put to the Test*, Center on Education Policy, August 2003

State percentiles, national and international comparisons – options and recommendations

Executive Summary

The tests required for Minnesota's statewide accountability system and to meet the requirements of the federal legislation "No Child Left Behind" must be aligned with the state's academic standards and provide "standards-based" achievement levels. They must also provide teachers, parents and students with information about students' individual achievements. Minnesota Session Laws 2003, Chapter 129, Article 2, Section 2 required that recommendations be delivered to the Legislature by February 2004 *evaluating the feasibility of including state percentile rankings and a national comparison for these state tests.*

Minnesota Statute 120B.365 established an Assessment Advisory Committee. This Committee was comprised of school superintendents, teachers, representatives of higher education, and members from the community at large. The Committee began meeting in September 2003, and considered four options regarding state percentile ratings and the feasibility of a national comparison.

After reviewing the deliberations of the Assessment Advisory Committee, the Commissioner of Education recommends the use of concordance tables based on NAEP data and local district testing data (Option 4), as being the approach that is the most cost-effective, while still providing approximate information for parents about relative performance of their student compared with others.

The Committee also suggested that the Minnesota Department of Education (MDE) further explore options for benchmarking state performance in comparison with other countries.

Background

The tests required for Minnesota's statewide accountability system and to meet the requirements of the federal legislation "No Child Left Behind" must be aligned with the state's academic standards and provide "standards-based" achievement levels⁶. Additionally, they must provide some form of "value-added" or "growth" information by 2006, and provide teachers, parents and students with information about students' individual achievements.

These requirements mean that the state tests must answer questions such as "what learning has this student shown?", and "how much improvement/growth has this student shown since the last test?" It is reasonable for parents, teachers and the community to ask how a particular student's achievement compares with that of other students in Minnesota. This would be done using state percentiles⁷.

We may also want to know how Minnesota students' achievements compare with those of students elsewhere in the United States, by means of national comparisons. Finally, in our global economy, Minnesotans will want to know how our students' achievements compare with those of students in other countries. The value of making international comparisons is already affirmed by MN Statute 120B.30(d)(4), which refers to the state benchmarking "its performance against the nation and other states, and, where possible, against other countries." MDE issued a request for proposals (RFP) in February 2003 in relation to the new MCA-II in grades 4, 6, and 8. This RFP included request for proposals for national comparisons.

Key Issues include:

- can state percentile rankings be included in reporting of results on statewide assessments?
- what level of precision is required for national comparisons – student level, building level, district level or state level data? There are significant technical differences from being able to state: "this student's performance compares nationally ..." to "the performance of students in this grade compares nationally ..."
- what are feasible methods for benchmarking state performance against other countries?

⁶ Minnesota has five achievement levels, 1, 2, 3, 4, and 5. Level 3 represents "proficiency" for NCLB purposes. Levels of achievement are "standards-based" in the sense that each level is intended to represent a particular standard of achievement, not an achievement relative to that of other students.

⁷ A percentile is a number that represents a student's score in comparison with the scores of other students. A student is in the 80th percentile, for example, if that student's score is only exceeded by twenty per cent of the group of students taking that test. So if 20 percent of students get a perfect score, the student with one wrong will be placed in the 80th percentile.

State Percentiles: State percentile rankings can be derived from the existing data and included in reports. This is not a straightforward process. There are several technical challenges that need to be considered and are difficult to communicate to the public:

- Percentile rankings provide accurate comparisons of standards between years **only if the overall performance remains the same from one year to the next** – if all students in fact achieve more highly, then percentile rankings will *falsely suggest no improvement*.
- Percentile rankings provide comparisons of how much better a student does in one grade in comparison with the previous **only if the progress that all students make is the progress they “ought” to make** – if the group as whole makes *less* progress than it ought to, or even slips backwards, *the percentile ranking of a student who does not slip backwards will rise*.
- Percentile rankings **only provide detailed information if the test is constructed to provide suitable discrimination** – if there are many students at a restricted number of score points the percentile rankings will not be stable.
- **Percentile rankings lose the scale properties of the calibrated test scores** – the same difference in test score at different parts of the scale can be represented by very different changes in percentile rank
- **Percentile rankings may not be stable enough at the extremes to be useful**
- **Percentile rankings depend on the base population** – including different groups of students changes the rankings.

National Comparisons:

Comparisons of the performance of our students with those of students in other states could be made by: having all students, or a sample of students, take a commercial norm-referenced⁸ test (NRT) separately; including sufficient items taken from a norm-referenced test in all tests to give individual student level data; using embedded NAEP items (assuming that NAEP items with data will be available); or, developing concordance tables using basic NAEP data, or using data from district tests that include national comparisons, to establish approximate concordance tables for the statewide data.

The choice among these options depends in part on the level of precision that is required. It is most difficult to determine accurate student-level information, measuring how an individual student's achievement compares with national norms. Less precision would be required in order to determine accurate building level information (how the achievement of a group of students compares with national norms). Finally, the most readily available comparison would be one that requires accurate district level information.

Other issues affecting the choice are costs – which include both financial and student time commitments.

⁸ A norm-referenced test gives a student's results in terms of “norms.” Norms are often derived by giving the items to random samples of students. Different tests have different norms. A criterion-referenced test seeks to give results in terms of standards rather than norms – where some people must be above a norm and others must be below it (that is what makes it a norm), it is possible that all students will meet a standard.

The Committee considered the following options:

	Testing Implications	Policy Implications	Financial Implications
<p>Option One:</p> <p>A sample of students takes a national Norm Referenced Test (NRT)</p>	<ul style="list-style-type: none"> - Extra testing for some students - No direct student information for those not in the sample - Test is outside the state assessments 	<ul style="list-style-type: none"> - Requires the choice of a particular NRT (Many districts have strong preference for one NRT over another) 	
<p>Option Two:</p> <p>Embed NRT items in state assessments</p>	<ul style="list-style-type: none"> - Lengthens the test, perhaps prohibitively if accurate student level information is required (10 items needed for building level data, 20-30 items for individual student level information) 	<ul style="list-style-type: none"> - Requires selecting a particular NRT- different districts currently use different NRTs - Raises issues about ownership of embedded items and release of tests 	<p>Cost appears to be about \$100,000 per grade- more if more items are required</p>
<p>Option Three:</p> <p>Embed NAEP items in state assessments</p>	<ul style="list-style-type: none"> - Lengthens the test - If items are well-aligned with the state tests, this could provide student level information 	<ul style="list-style-type: none"> - It could be very difficult and expensive to obtain permission to use the NAEP items and items statistics 	<p>Because of the cost, it is likely that it could be done for limited grades. The Committee recommended grades 4 and 8 if this option were selected</p>
<p>Option Four:</p> <p>Concordance Tables</p>	<ul style="list-style-type: none"> - Provides very approximate information at the student level- shows that a student with this result is "likely" to have this range of result nationally - Can be based either on NAEP data, or on local district data, if these test results can be made available to the state 	<ul style="list-style-type: none"> - Responds to parental interest in getting a sense of how their child is doing in comparison with others - Responds to interest by school districts in being able to offer approximate comparative information 	<p>Very little additional cost</p>

The first three options for national comparisons can give student level data with different levels of precision. We achieve more precision when we have more test items included. It is also important to have the most accurate and relevant⁹ national norm information possible about the items, in order to achieve precision of the student level information.

⁹ The extent to which the items being used for the national comparisons align with Minnesota's academic standards is clearly central to the usefulness of the information – if there is a weak relationship between what is in the national test and what is in the Minnesota tests the norming information is of less meaning.

The fourth option, developing concordance tables, will show the relationship of results on different tests – students with this *range of results* on one test have *results in this other range* on the other test. Concordance tables reflect the fact that tests are different, rather than claim that two different tests are essentially equivalent.

International Comparisons:

Benchmarking state performance against other countries could provide a valuable perspective on the world-class competitiveness of standards and outcomes in Minnesota. One possible option is to use existing international comparison programs (e.g. PISA¹⁰ or TIMSS¹¹) with appropriate participation at the state level. Standard practice in these programs is to provide only state level data with no individual or school level data – a (small) random sample of students takes the tests. No information is given back to students or schools. State level participation in these established programs is available in some years at a price usually around \$350,000.

It may be possible to benchmark state performance against other countries by using items drawn from an established international program to give school level information to participating schools/districts as desired and to give enough information linked at the student level to generate reasonable concordance tables at the state level for participating grades and comparisons/illustrations of the skills shown by typical students at these grades. There is, for example, the program of International School Assessments (ISA) that offers tests at grade 3 and grade 9/10 in reading and mathematics at around \$20 per student. Participation by a limited number of students (a few thousand) at a limited number of schools could provide those schools (and their students) with useful information and the state with the capacity to report on the concordance of the state assessments with these internationally calibrated scales.

Recommendations:

After reviewing the options, the Commissioner recommends the following:

- State percentile rankings are feasible but care is needed in their interpretation.
- The use of concordance tables based on NAEP data and local district testing data (Option 4) is the approach that is the most cost-effective, while still providing approximate information for parents about relative performance of their student compared with others.
- International comparisons, benchmarking performance at the state level, have a powerful potential to inform debate about what is achieved and what could or should be achieved in education in Minnesota. Informed debate about these issues makes a major contribution to building and maintaining a world-class system.

¹⁰ PISA – program for international student assessment. This is OECD program providing information about reading, mathematics and science literacy amongst 15-year-olds in the principal industrialized countries.

¹¹ TIMSS – Third International Math and Science Study.

1 Appendices

1.1 Overview of advisory structure for MDE statewide assessment program

- b) MDE's assessment program is advised by:
- i) the Assessment Advisory Committee, formally constituted, responsible for advice to the Commissioner about principles and guidelines for a coherent and effective assessment system
 - ii) Subject Advisory Committees (reading, mathematics and science), formally constituted, responsible for advice to the Office of Academic Excellence on principles and guidelines needed to ensure a coherent K-12 perspective within each subject area. As at December 2003, these committees have yet to be constituted.
 - iii) Teacher/community panels, constituted on an *ad hoc* basis as required to do the detailed work (test specifications, alignment studies, item review, standards setting) needed to provide the Division of Assessment & Testing with the advice required to implement the Commissioner's approved policies and guidelines
 - iv) the National Technical Advisory Committee, with members who provide expert advice on technical issues associated with assessment and testing. Present National TAC members are
 - (1) Dr Roger Trent, former state assessment director Ohio
 - (2) Professor Mark Davison, Minnesota Office of Educational Accountability
 - (3) Emeritus Professor William Schafer, University of Maryland
 - (4) Dr Susan Phillips, legal and psychometric consultant
 - v) the Local Technical Advisory Committee with members drawn from district testing coordinators, OEA, and other people with knowledge and expertise in assessment.

1.2 Assessment Advisory Committee – membership

- vi) two superintendents, selected after consultation with MASA
 - Roger Giroux, Superintendent Anoka-Hennepin *
 - Rick Ellingworth Superintendent Redwood Falls *
- vii) two teachers, selected after consultation with Education Minnesota
 - Neil Witikko, Hermantown High School *
 - Rose Gundacker, Rosemount High School
- viii) two members from higher education
 - Sandra G. Johnson, Assoc. Professor, St. Cloud University
 - Dr. Craig Swan, Vice Provost for Undergraduate Education

ix) five additional members

- Dr. Valerie Pace, State Mgr. IBM Corporate Community Relations
- Glenn Dorfman, Chief Operating Officer, Minnesota Assoc. of Realtors
- Yusef Mgeni, Director, Office of Educational Equity, SPPS
- Virginia Richardson, Director, Parent Training, PACER
- Jerry Hanson Principal & Title I Dir., Roosevelt Elementary School (selected after consultation with MESPA and MASSP)

1.3 Functions of the Assessment Advisory Committee

Under MN 120B.365, the Assessment Advisory Committee reviews statewide assessments and the Commissioner will consider the recommendations of the committee before finalizing a statewide assessment.

The functions of the Assessment Advisory Committee are to support the validity and technical adequacy of statewide assessments by

- a) responding to requests for advice from the Commissioner about assessment and testing issues
- b) advising the Commissioner about principles and guidelines for:
 - 1) developing test specifications
 - 2) reviewing tests against test specifications
 - 3) aligning tests and test results with the state's high academic standards
 - 4) reporting test results to parents and teachers in ways that indicate students' educational needs.

MDE Assessment Advisory Committee

Impact of the grade 8 MCA-II on the
Basic Skills Test – options and
recommendations

Executive Summary

As Minnesota complies with the federal 2001 Elementary and Secondary Education Act known as No Child Left Behind (NCLB), there are several policy considerations that must be addressed. Minnesota Session Laws 2003 (Chapter 129, Article 2, Sec.2) required that recommendations be delivered to the Legislature by February 2004 to "*establish a method for using the grade 8 language arts and math tests to satisfy basic skills requirements*".

By the spring of 2006, Minnesota will have in place new Minnesota Comprehensive Assessments series II (MCA-II). These will be aligned with state academic standards in Language Arts and Mathematics. Currently, Minnesota 8th graders are required to take Basic Skills tests (BST) in Reading and Math. If there are no changes made, grade 8 students would have to take two tests of reading and mathematics in a single year.

Minnesota Statute 120B.365 established an Assessment Advisory Committee. This Committee was comprised of school superintendents, teachers, representatives of higher education and members from the community at large. The committee began meeting in September 2003, and considered five options regarding the Grade 8 MCA-II and the BST.

The Committee's work was helpful in assisting the Commissioner of Education to develop recommendations regarding the grade 8 MCA-II. The Commissioner agrees that the MCA-II should not contain the same test, nor be the same test as the BST. In response to concerns about double-testing of 8th graders, as well as with respect to fiscal considerations, the Commissioner recommends that the first administration of the BST be moved to 9th grade. The Commissioner believes Information Technology (IT) has real potential to provide improved service and cost-effectiveness and agrees that the Department of Education should examine this possibility further.



Background

No later than spring 2006, Minnesota will have in place fully operational grade 8 statewide assessments (Minnesota Comprehensive Assessments series II, MCA-II) aligned with state academic standards in Language Arts and in Mathematics. The MCA-II will be used to provide students, parents, teachers and schools with information about students' achievements in terms of the state's high academic standards, and will meet requirements of No Child Left Behind (NCLB).

Current high school graduation requirements provide that students¹² must pass the Basic Skills Tests (BST) in reading, mathematics and writing. The BST in reading and mathematics is first taken in grade 8. The BST in writing is first taken in grade 10.

BSTs are offered twice per year – February and July, with an additional April testing period for 12th grade students only. Students therefore have up to eleven opportunities to pass the reading and mathematics components and seven opportunities to pass the writing component.

If the Legislature does not change the law, beginning in 2006 grade 8 students will have to take two tests of reading and mathematics in a single year. This report examines issues and options for changes to the grade 8 testing arrangements that seek to:

- maintain effective basic skills requirements in reading, mathematics and writing
- provide technically sound grade 8 statewide assessments in reading and mathematics
- eliminate a double testing requirement in grade 8
- deliver a cost-effective approach.

Linking the BST and the MCA-II

There are several options for linking the BST and the MCA-II. One approach is to embed the BST in the MCA-II – which in effect would be two tests taken at the same time. Another approach is to ensure that the new MCA-II in grade 8 in reading and math would provide students with the opportunity to show that they clearly satisfied basic skills requirements in reading and math (not writing). Those who did not demonstrate this would need to take a Basic Skills test later.

There are complex technical challenges involved in linking the BST and the MCA-II. **The BST is designed as a test for an individual to meet a graduation requirement.** This purpose is markedly different from the purpose of the grade 8 MCA which is based on high academic standards at grade 8. The MCA provides information about the proportion of students at “grade-level”, as well as information about higher

achievement. There are five levels of achievement on the MCA-IIs. Levels 1 and 2 are below "grade-level, level 3 represents "grade level" and levels 4 and 5 provide information about the knowledge and skills of higher achieving students.

Conversely, the purpose of the BST is reflected in its design – enough items with the right level of difficulties to provide information about students near the cut-off. This type of test emphasizes providing good information at the cut-off at the cost of providing little useful information about the higher achieving students. Since the BST is high stakes for students, it is important that diplomas are denied only to those who clearly do not meet the basic skill requirements. The grade 8 MCA-II cannot be constructed so that it provides this information as well as the information needed for assessing students in terms of the state's high academic standards. On the 2003 BST reading test, for example, amongst the over sixty thousand grade 8 students taking it, the most common score was one wrong, and twenty-four thousand grade eight students had none, one, or two wrong.

In order to compare scores on the various versions of the BST, linking with previous year cut-offs is done for current Basic Skills Tests by using items from the previous years' field tests, and occasionally items from operational tests. If we were to administer the BST in its present form only to those who did not "pass" the grade 8 MCA, the students taking the field test items would be a very limited segment of the total population. By eliminating the top 40 to 60 percent of the population, year-to-year equating would be highly problematic. The differences between the BST versions students take are dealt with through a technical "equating" process intended to make scores comparable from year-to-year. Raw scores are not directly comparable across different versions.

The approach of finding a score on the grade 8 MCA-II that justifies students not having to take the BST is importantly different from finding a score on the grade 8 MCA-II that is the same as the BST passing score of 600. It would be very difficult to be sure that a score on an MCA-II is an exact (or even good) equivalent of a score on another, different test. It is possible, however, to find a score on an MCA-II that we can be very sure identifies students who would have passed the BST. Students who scored below this cut-score *might* well have passed the BST, but we can't be confident enough to determine they don't need to take the BST.

This is something like a situation where there are two kinds of sporting tests – a person who does very well cycling will almost certainly be a *basic* runner – if we set the cycling test cut-off high enough we can be sure enough that we don't need to ask these people to take the basic running test. This doesn't mean we are saying that those who don't reach the cycling cut-off *can't* run – it means we can't be sure enough that we don't need to ask them to show their skills.

One challenge with this transition period, is that in order to keep the passing standard the same, the present BST relies on having *all* grade 8 students take the test with *sufficient field test* items included to

¹² There are alternative assessments for "the very few students with disabilities for whom statewide assessments are inappropriate and for students with limited English proficiency". For the BST, the alternative for a limited English proficiency student applies where the student has been in the United States for fewer than three years.

generate the base test forms for the following year. There are three BST sessions per year: This requires three different sets of test items – a total of about 220 different math items per year.

Under the present system, results from the full range of Minnesota students on the field test item are needed to provide the “linking” or “equating” data – the information needed by the processes that ensure that students taking different forms of the test (in the same year and in different years) are held to the same standard – the equivalent cut-score. The process works because one set of students (the prior year) takes both the base¹³ items (for that year) and the field test items (that will be the base items the next year). If the field test items are not taken by the full range of Minnesota students¹⁴ or not taken by Minnesota students at all (for example, if the items were bought from another state), the technical process used to equate the scores would not give a sufficiently solid and fair equivalence of cut-scores for this high-stakes test.

Why not reuse test forms from previous years? This would give an unfair advantage to students who happened to have come across and practiced on these items.¹⁵ Why not reuse items from old tests? First, the year-to-year linking or equating is thought to be best and fairest when it is done using items taken the prior year – item statistics have a “shelf-life”. Second, the less the variety in the test forms within a year and between years, the greater the chance that a group of students will have an unfair advantage/disadvantage because there has been access to sets of items. Having new items in each test form and using item information from the previous year gives the best chance of being fair to all – the more that items are reused, the longer the time gap between field test and base test use, the more uncertainty that students are being held to the same standard.

There is always some level of risk associated with BST type of testing – risks to security, integrity and equivalence of results. The present procedures (calibrate the items used in one year with the entire grade 8 population, using items taken from recent years) represent a certain (low) level of risk. As significant changes are made to the procedures it will be necessary to show not that there *is* a level of risk (there is always some risk) but that the level is kept about the same as in previous years.

Continuing to reuse old items indefinitely would eventually lead to the situation where it would be likely that the students who were just passing on a set of old items would not pass if they were given a set of new items – that their scores reflected familiarity with the items rather than possession of the underlying skills.

The technical challenges of linking the BST and the MCA-II are accompanied with policy challenges:

- 1) Simply embedding BST items in the grade 8 MCA-II will lengthen the test substantially and this may not solve the technical equating issues.
- 2) Using the MCA-II items to identify a “pass” score that shows that students **clearly do** meet basic skill requirements will be seen as stating that students below this score **do not** meet those requirements. As

¹³ The “base” test is the set of items that count, the ones on which a student’s result is based. The field-test items are included in the test form but they do not affect a student’s test score. The students are not told which items are base test items and which items are field test items.

¹⁴ A given student takes only some of the field test items. The test booklets are given out in such a way that enough students from across the full range take each item. Because there are only nine field test items on each test form, no student has to take a lot of items and the risk of losing security on next year’s test forms is minimized.

¹⁵ The security on the BST test forms is reasonably high, but no security system is perfect.

explained above, this will not necessarily be true. There will be students below this score who probably can meet the requirements, and, may have passed a BST, but who do not show this clearly on the challenges of a grade 8 MCA-II. This complexity arises because of the different purposes of the two tests; the BST in effect denies a diploma to a student who does not pass, and so the items and cut-scores are set so that diplomas are denied only to those who **clearly do not** meet the basic skill requirements. The grade 8 MCA-II cannot be constructed so that it equally well provides both this sort of information and the information needed for assessing students in terms of the state's high academic standards.

- 3) The "pass" score and the cut-scores for showing that students meet grade-level requirements will not, in principle, align exactly – there is no reason to suppose that *all and only* those students who meet the high academic standards for grade 8 in reading/math also clearly satisfy basic reading and math skill requirements.¹⁶
- 4) The MCA-II in grade 8 in spring 2006 will be taken by all students and will count for statewide accountability (AYP) purposes. Should there be one year in which students take both the BST and the MCA-II, or should the new arrangements apply, if possible, beginning in 2006? A full pilot of the grade 8 MCA-II is planned for the spring of 2005.¹⁷

Implications for the BST of using grade 8 MCA-II scores to meet BST requirements

The central implication for the BST of using grade 8 MCA-II scores to meet BST requirements is that the existing approach to ensuring a fair equating of cut-scores across different BST test forms will become progressively less effective. That is, an unintended consequence of this process is that it will not be possible to continue the BST indefinitely in its present form.

The use of grade 8 MCA-II so that many students will not need to take the BST would suggest a significant reduction in the cost of the BST. Under the present model of three BST sessions per year, however, this change will not make a dramatic difference to the cost of the BST. That is, if the technical challenges of embedding the BST in the grade 8 MCA-II could be overcome it would still be necessary to provide students with further opportunities to take the BST so that they could meet graduation requirements.

It may be possible to develop a computer-administered version of the BST for the students who do not "pass" in grade 8, to be taken in approved test centers¹⁸ under appropriate security/integrity conditions. This might include the BST writing requirement – there are possibilities for computer aided scoring of the sort of writing required for the BST.

The development of these possibilities will require funding (for the test and any required infrastructure developments) while the existing BST is still being operated. That is, although doing this will in the longer term save a lot of money it will cost more money in the short term.

¹⁶ This is what "exactly" would require: students above the cut-score on first test *did* meet the requirements on the second test and students below the cut-score on the first test *did not* meet the requirements on the second. Realistically, with different tests designed for different purposes some appreciable uncertainty is unavoidable.

¹⁷ A full pilot requires a test that is as long as the real test and is taken by nearly all students.

¹⁸ For example, schools may be approved test centers.

The recommendations in this report have been arrived at after examination of the following options:

	Testing and Policy Implications	Financial Considerations
<p>Option One: Make the Grade 8 MCA-II the equivalent of the BST, and require the students who "fail" to retake it each year at the same time</p>	<ul style="list-style-type: none"> - The BST and the grade 8 MCA-II have different purposes - The BST is not designed to align with Minnesota's high academic standards - The BST <i>is</i> designed to provide information about pass/fail in terms of basic skills requirements. It is <i>not</i> designed to (and does not) provide information about higher levels of achievement (levels 4 and 5 in the MCAs) - There would be only four re-take opportunities (Spring of grades 9, 10, 11, and 12). - Could not continue indefinitely in its present form, given the impact on field testing of items and consequent test equating through not having full grade 8 participation. 	
<p>Option Two: Embed the BST in the grade 8 MCA-II and maintain the remaining BST sessions for students in grades 9-12</p>	<ul style="list-style-type: none"> - This lengthens significantly (by 20-30 items) the grade 8 MCA-II 	<p>This increases the costs of the grade 8 MCA-II without a major reduction in the total cost of the BST</p>
<p>Option Three: Establish a cut-score on the grade 8 MCA-II that shows a student has demonstrated skills to the extent that it is not necessary to take the BST and maintain the remaining BST sessions for students in grades 9-12.</p>	<ul style="list-style-type: none"> - This reduces the number of students taking the reading and math BST - This presents significant field-testing and equating difficulties for the BST, as the BST sessions would be taken by a restricted segment of the population (those who had not reached the cut-score on the grade 8 MCA-II). - Insufficient items would be field-tested in the grade 8 MCA-II to maintain having three different BST forms per year using new items - This has the unintended consequence that the BST could not continue in its present form indefinitely; creating new forms by recycling old items is possible in the short-term, but would eventually introduce an appreciable risk of unfair advantage to a group of students who had obtained access to particular sets of items. 	<ul style="list-style-type: none"> - There is no major impact on BST costs
<p>Option Four: Maintain the BST as it is (pencil-and-paper, two main sessions per year) but move the first opportunity to take the BST to grade 9</p>	<ul style="list-style-type: none"> - Eliminates the double testing in grade 8 - Maintains the BST in its present form indefinitely- items are calibrated on the entire year cohort - Reduces the number of opportunities to pass the BST by two (some may see this as an advantage, others as a disadvantage) - Requires a transition process to manage the situation that most of the grade 9 students in the first year of operation will already have passed the BST in the previous year - Requires some technical study to take of the difference between the two cohorts to ensure that the cut-scores remain comparable 	<p>Maintains existing BST cost structure</p>
<p>Option Five: Establish a cut-score on the grade 8 MCA-II that shows a student has demonstrated skills to the extent that it is not necessary to take the BST (in reading and math) and replace the BST sessions for students in 9-12 with an IT-delivered and scored BST in reading, math and writing</p>	<ul style="list-style-type: none"> - Eliminates the double-testing option in grade 8 - First step towards use of IT in statewide assessment - Results available to students sooner - Multiple, but limited, re-take opportunities - Requires appropriate level of infra-structure (computers and support) to be available in districts - Would need to be developed in stages, "practice test", "pilot administration", and "full implementation". 	<ul style="list-style-type: none"> - Eventually reduces costs and enhances flexibility

There are some important technical challenges with Option 4, to ensure that the cut-scores remain comparable with the cut-scores set during the years when the tests were taken by grade 8 students – items may look easier when taken only by all grade 9 students than when they were taken by all grade 8 students, simply because of the extra year rather than because the items are in fact easier. Some form of equating study¹⁹ will be needed to demonstrate that the cut-scores under the new arrangements are equivalent in standard to those set under the old arrangements.

Options for the transition process include:

- Requiring all grade 9 students to re-take the BST in that first year and using the information from that to recalibrate the item pool to take into account the effect of the change in the age of the cohort (this assumes that there will be no systematic and large distortion to the results attributable to differences in motivation).
- Making the first year of operation one where all sessions are taken only by students who did not pass in grade 8 and conducting equating studies with samples of students.

Considerations for Information Technology (IT) delivered tests:

If Minnesota were to move to an Information Technology (IT) delivered and scored BST in reading, math and writing the student would take the test(s) on a computer as required as available under secure conditions. MDE would need a large item bank for reading, math and writing prompts. The test each individual student would see would be a selection from the large item bank. The multiple-choice items would be pre-equated so that the difficulty level of each test would be about the same. Once the student finished, the responses would be analyzed by computer and a result given to the student and notified officially to the school district.

On the writing portion of the test, the student would enter a response to the prompt on the computer, using the keyboard.²⁰ The response would be analyzed by computer and compared on a wide range of key characteristics with a large number of existing responses that have been scored by human scorers. Where this process showed that the response matched the characteristics of responses given a result at BST “pass”, or better, the student would be given the result and feedback. The school district would then be notified officially. Where this process does not indicate a result at BST “pass”, or better, the student would receive feedback, and the response would be routed to a human scorer for review. After review the school district and student would be notified officially of the final result.

There would be appropriate accommodated versions available for students who require them. A limited number of BST-style items would be included in the grade MCA-II field test items to refresh the item-

¹⁹ “Equating studies” are used to provide quantitative estimates of the effect of any impact on cut-scores of these sorts of changes. Typically, equating studies involve administering tests to representative samples of students and analyzing the results.

²⁰ The use of keyboards for entering written text is widespread and allows the assessment of students’ skills in producing and correcting written text. The mechanical skill of handwriting may soon be assessable on computer using tablet and stylus. The use of keyboards for input of text (a mode of input in use as a testing accommodation) will raise questions about the extent to which of the skill of using a pen to fashion letters is one of the basic skills that must be included as a graduation requirement.

bank. The number of opportunities to take the test would be limited to reduce the possibility of extensive "teaching to the test", or of students repeatedly taking the test on many occasions in a short period of time until they score above the cut-off more due to chance than by having the required knowledge and skills.

This option requires the use of existing technologies. To put it in place for Minnesota might best be achieved by development and pilot use in schools of a practice test version of an IT-delivered and scored BST in reading, mathematics and writing. The practice test would provide information to students and teachers and provide information to the system about the demands on infrastructure in districts and schools. It would also require analyses of grade 8 MCA-II scores and grade 8 BST scores for the same students, and development of options for schools to ensure secure administration of an IT-delivered and scored BST. Development of infrastructure options for schools and districts would also be necessary, as would pilot studies to show that the two forms (old pencil-and-paper and IT-delivered) were equivalent. It would also be important to provide appropriate notice to students.

It would be important to develop this option in stages – "practice test", "pilot administrations", "full implementation" - rather than in one step. At the "practice test" stage, schools would have access to a practice version. From this, it would be possible to learn about the capacity of and need for infrastructure (computers, web-access, technical support) in districts and schools). It would also help identify many of the likely problems in a real test situation (crashes, access difficulties, overloads, errors in returning data to the server), as well as the acceptability to educators and the public of the soundness of the proposed model for scoring the writing component (human scoring as a validation of computer-scoring).

The IT-delivered and scored option responds to the key challenge of a fair BST test in grade 9-12 by ensuring that each student has an individual selection of items. A class of students may have practiced on a particular set of BST items. If those items appeared on a real BST pencil-and-paper test form, then all those who had practiced that set of items would have a systematic advantage over those who had not (who had perhaps practiced some other set of items).

The individualized nature of the IT-selection provides security, and minimizes the risk of any systematic advantage/disadvantage. It also would provide for the continuation into the future, as it allows for the inclusion of one or two items being calibrated.

The inclusion of a limited number of BST items in the grade 8 MCA-II field and base tests would allow the BST item-bank to be refreshed. The individualized nature of the IT-selected test forms makes best use of this information in keeping the BST scores comparable and the cut-scores for passing equivalent across the different forms.²¹

As well as a large number of multiple-choice items, there would also be a need for a large number of writing prompts- each student would be assigned one of these. The large number means that security and integrity of administration of the writing test in this form is comparable with that for the present writing test.

²¹ Of course, this is not the *post-equating* used in the present system. It should give appropriately comparable results because each test form is individualized.

Conclusions and Recommendations:

The replacement of the grade 8 BST with the grade 8 MCA-II would mean that the BST could not continue indefinitely in its present form, given the impact on field testing of items and consequent test equating through not having full grade 8 participation. Moving the first opportunity to take the BST in reading and mathematics to grade 9 will enable the BST to continue in its present pencil-and-paper form.

Students must have adequate notice of this change, making 2006 the last year when students must take both the grade 8 MCA-II and the BST in reading and mathematics. The BST administrations in the year immediately following the last administration of the BST to grade 8 students will be needed only by students re-taking the BST.

Information Technology (IT) has real potential to provide improved service and cost-effectiveness. The Department needs to examine further this possibility.

- i) The Commissioner recommends that students' first opportunity to take the BST in mathematics and reading be in grade 9 beginning in 2006-2007. Further, the Department should review the timing and number of BST test administrations each year, and be given rule-making authority to make changes to MN Rule 3501 required to implement these recommendations. Finally, the Department should examine the possibility of delivering and scoring of the BST through Information Technology (IT).

1. Appendices

1.1. Overview of advisory structure for MDE statewide assessment program

- a) MDE's assessment program is advised by:
 - i) the Assessment Advisory Committee, formally constituted, responsible for advice to the Commissioner about principles and guidelines for a coherent and effective assessment system
 - ii) Subject Advisory Committees (reading, mathematics and science), formally constituted, responsible for advice to the Office of Academic Excellence on principles and guidelines needed to ensure a coherent K-12 perspective within each subject area. As at December 2003, these committees have yet to be constituted.
 - iii) Teacher/community panels, constituted on an *ad hoc* basis as required to do the detailed work (test specifications, alignment studies, item review, standards setting) needed to provide the Division of Assessment & Testing with the advice required to implement the Commissioner's approved policies and guidelines
 - iv) the National Technical Advisory Committee, with members who provide expert advice on technical issues associated with assessment and testing. Present National TAC members are
 - (1) Dr Roger Trent, former state assessment director Ohio
 - (2) Professor Mark Davison, Minnesota Office of Educational Accountability
 - (3) Emeritus Professor William Schafer, University of Maryland
 - (4) Dr Susan Phillips, legal and psychometric consultant
 - v) the Local Technical Advisory Committee with members drawn from district testing coordinators, OEA, and other people with knowledge and expertise in assessment.

1.2. Assessment Advisory Committee – membership

- vi) two superintendents, selected after consultation with MASA
 - Roger Giroux, Superintendent Anoka-Hennepin *
 - Rick Ellingworth Superintendent Redwood Falls *
- vii) two teachers, selected after consultation with Education Minnesota
 - Neil Witikko, Hermantown High School *
 - Rose Gundacker, Rosemount High School
- viii) two members from higher education
 - Sandra G. Johnson, Assoc. Professor, St. Cloud University

- Dr. Craig Swan, Vice Provost for Undergraduate Education

ix) five additional members

- Dr. Valerie Pace, State Mgr. IBM Corporate Community Relations
- Glenn Dorfman, Chief Operating Officer, Minnesota Assoc. of Realtors
- Yusef Mgeni, Director, Office of Educational Equity, SPPS
- Virginia Richardson, Director, Parent Training, PACER
- Jerry Hanson Principal & Title I Dir., Roosevelt Elementary School (selected after consultation with MESPA and MASSP)

1.3. Functions of the Assessment Advisory Committee

Under MN 120B.365, the Assessment Advisory Committee reviews statewide assessments and the Commissioner will consider the recommendations of the committee before finalizing a statewide assessment. The functions of the Assessment Advisory Committee are to support the validity and technical adequacy of statewide assessments by:

- c) responding to requests for advice from the Commissioner about assessment and testing issues
- d) advising the Commissioner about principles and guidelines for
 - (1) developing test specifications
 - (2) reviewing tests against test specifications
 - (3) aligning tests and test results with the state's high academic standards

SPECIAL EDUCATION ALTERNATE ASSESSMENT MATH - 5th GRADE

1-----2
Awareness

3-----4-----5
Understanding

6-----7
Application

Exemption Decision: (1) Compare the student's math level to 5th grade math material, and (2) determine a number rating for the following two attributes: **Shape and Space, and Fractions and Decimals**. See the back of this sheet for the attributes and their indicators. Average the attribute ratings. **Average Rating = _____**. If you rate the student a 1 or 2, it is suggested you exempt the student and complete the alternate assessment. However, the IEP team has primary responsibility and may use other data/information to support their decision.

NOTE: If you exempt a student it is key that you complete the AA by rating the student on each attribute based on what is currently developmentally appropriate for the student.

1-7 Scale: Use the following definitions of awareness, understanding and application to make your ratings.

Awareness (1 to 2): Overall, the student demonstrates little or no ability to perform on this attribute.

Understanding (3 to 5): Overall, the student performs inconsistently and needs substantial teacher assistance.

Application (6 to 7): Overall, the student performs consistently well and independently on this attribute.

Basic Conceptual Understanding

- Counts orally to 30 with proper number sequencing
- Orally repeats a number pattern composed of three numbers up to 10 (e.g. 5+9+8)
- Understands concepts of more, less, same and grouping
- Given a set of partial and complete objects, can identify parts of objects that belong with the whole
- Can divide an object into halves, thirds and fourths

Basic Procedural Skills

- Using manipulatives can add with sums to 20
- Using manipulatives can subtract with differences less than 10
- Given a set of objects up to 100, can group them by 2's, 3's, 5's and 10's and state the dividend
- Recognizes common symbols for addition, subtraction, multiplication, division and equality

Basic Shape, Size and Measurement

- Can distinguish between the notions of round, flat, and straight
- Names seasons
- Names months
- Names directions (North, South, East, West)
- Can tell time to the hour
- Can distinguish between a circle, a square and a triangle

Basic Organizing Skills

- Can sort objects on the basis of one major characteristic (e.g. color, size, shape, etc.)
- Can imitate or reproduce rhythms, rhymes and visual patterns
- Recognizes and understands the concepts of before, after, and between, using concrete objects

Shape and Space

- Sketches or builds common 2 and 3 dimensional shapes
- Identifies and describes common geometric shapes and relationships (e.g. circle, cube, triangles, parallel, vertical)
- Can follow simple directions using common terms of directions using or directionality (e.g. left/right, north/south, up/down)
- Identifies locations on maps and other coordinate graphs

Computation With Whole Numbers

- Can add: 2 digits + 2 digits
- Can subtract: 2 digits – 2 digits
- Can multiply: 2–3 digits x 1 digit
- Can divide: 2 digits ÷ 1 digit (with no remainders)

Conceptual Understanding: Number Sense

- Counts by twos, fives and tens
- Builds 3-digit whole numbers with concrete materials (e.g. 243)



- Can accurately identify number or shape patterns
- Compares and orders two numbers within a given set of whole numbers, simple fractions (denominators of 2, 3, 4, 5, 6, 8, 10), money amounts
- Can identify the fraction represented in a visual display



Measurement

- Given a set of bills and coins, can total the amount of money (e.g. 2 dimes, 1 quarter and 3 pennies = \$0.48)
- Can accurately use a measurement tools (e.g. clocks and calendar) to record and interpret information
- Can match appropriate tools with the measurement function (e.g. clocks/time; gallons/liquid; temperature/thermometer; height/ruler)

Fractions and Decimals

- Adds and subtracts decimals to two decimal places with or without a calculator
- Multiplies and divides decimals using a calculator and rounds answers, if appropriate

Data and Chance (For Grades 5 and higher)

- Given questions and directions, can carry out a simple interview or survey
- Given a format and directions, can collect and organize simple data
- Can describe simple data from graphs, charts, tables, and pictures
- Given graphical format and directions, can record simple data

SPECIAL EDUCATION FUNCTIONAL SKILLS ALTERNATE ASSESSMENT

1-----2
No Participation/Full Support

3-----4-----5
Moderate Participation/Moderate Support

6-----7
Full Participation/No Support

On the Functional Skills Attributes, use the degree of student participation and student support required to make your ratings

1-7 Scale: Use the following definitions of awareness, understanding and application to make your ratings.

No Participation/Full Support: Overall, student shows little or no ability to participate and requires full support.

Moderate Participation/Moderate Support: Overall, the student shows some ability to participate, but requires some support.

Full Participation/No Support: Overall, the student performs well and independently on this attribute.

HOME LIVING

- Eating (uses conventional eating utensils)
- Grooming (dresses, toilets, cares for personal appearance)
- Kitchen skills (uses utensils and appliances, plans and prepares food)
- Chores (sweeps, dusts, vacuums, does laundry and dishes)
- Safety (is aware of appropriate dress and safety procedures related to electricity, tools, fire, and emergencies)

RECREATION AND LEISURE

- Safety (follows safety rules or guidelines, uses materials appropriately)
- Chooses or selects activities
- Uses recreational equipment (simple technology, television, games, and toys)
- Accesses activities (finds transportation, locates game, uses telephone)
- Engages in activities with others

COMMUNITY PARTICIPATION

- Functions safely in the community
- Accesses community resources (bank, library, grocery store)
- Manages personal belongings (wallet, ID)
- Demonstrates appropriate public behavior
- Uses mobility skills to access buildings and public transportation

JOBS AND TRAINING

- Performs assigned task (follows schedule of work activities, does enough work)
- Attends to task (focuses, works independently, transitions to next task or activity)
- Works with others (cooperates, accepts help, gives assistance)
- Understands safety procedures (recognizes and responds to emergency alarm, uses protective equipment)
- Is productive

SOCIAL SKILLS

- Initiates or sustains social interactions
- Maintains socially acceptable behavior
- Has positive relationships with peers
- Shares and takes turns
- Responds/asks questions

COMMUNICATION

- Responds to environmental stimuli
- Indicates continuation or refusal of interaction
- Makes requests
- Follows instructions
- Describes events or objects

ACADEMIC

- Foundations (matches, knows symbols, sorts, manipulates objects)
- Math (tells time, counts money, measures)
- Reading (knows survival words, reads name, enjoys being read to)
- Writing (stamps or writes name, writes letters, writes sentences, keyboards)