



Achieve, Inc.

Measuring Up

A Report on
Education Assessments for

MINNESOTA

ACHIEVE'S
BENCHMARKING
INITIATIVE



About Achieve, Inc.

Achieve is an independent, bipartisan, nonprofit organization created by governors and corporate leaders to help states and the private sector raise standards and performance in America's schools. Achieve was founded at the 1996 National Education Summit and subsequently sponsored another Summit in the fall of 1999 that brought together over 100 governors, business leaders and education officials from around the nation.

Achieve helps states raise academic standards, measure performance against those standards, establish clear accountability for results and strengthen public confidence in our education system. To do this, we:

- help states benchmark their standards, assessments and accountability systems against the best in the country and the world;
- provide sustained public leadership and advocacy for the movement to raise standards and improve student performance;
- build partnerships that allow states to work together to improve teaching and learning and raise student achievement; and
- serve as a national clearinghouse on education standards and school reform.

MEASURING UP

A BENCHMARKING STUDY OF THE MINNESOTA COMPREHENSIVE ASSESSMENTS

Prepared by Achieve, Inc. for

Commissioner Christine Jax
Minnesota Department of Children, Families, and Learning

April 2001

TABLE OF CONTENTS

Executive Summary	3
Introduction and Methodology	7
Results for Minnesota	13
Recommendations for Moving Forward.....	19
Appendix.....	21

EXECUTIVE SUMMARY

Achieve, Inc. was created in 1996 by governors and business leaders to serve as a clearinghouse and resource center on education standards, assessment and accountability. As part of its mission, Achieve provides states with candid feedback on the quality of their academic standards, assessments and accountability policies. At the request of the Department of Children, Families, and Learning (CFL), Achieve and the Council for Basic Education (CBE) reviewed the Profile of Learning and related policies, resulting in the November 2000 report, *Aiming Higher: A Report on Education Standards and Policy for Minnesota*.

In late 2000, Achieve also conducted an intensive study of the quality, rigor and alignment of the Minnesota Comprehensive Assessments (MCAs) to the Profile of Learning standards for reading, writing and mathematics. *Measuring Up: A Benchmarking Study of the Minnesota Comprehensive Assessments* presents our conclusions and observations for further strengthening the state's education reforms. *Measuring Up* answers a central question for Minnesota policymakers: How well do the Minnesota Comprehensive Assessments measure the knowledge and skills laid out in the Profile of Learning standards?

RESULTS FOR MINNESOTA

In *Aiming Higher*, Achieve and CBE reported that our principal concern is with the quality and rigor of Minnesota's Profile of Learning content standards and the decentralized methods for assessing those standards. We recommended revising the Profile standards and then using the MCAs as the state's key tool for measuring student achievement and holding students and schools accountable for progress toward meeting the revised standards. We noted that the results of our study of the MCAs could be useful to state policymakers as they consider our recommendations, particularly if we were to find that the MCAs are a starting point for building assessments in the middle and high school grades.

Achieve's review of the MCAs in reading and mathematics in grades 3 and 5 found that the tests are a solid foundation on which the state can build. We also found a number of areas that should be addressed if the tests are to provide the state with strong information on how well students are attaining challenging standards:

- **Minnesota has commendably chosen to develop its own assessments.** Our experience has shown that such tests are more likely than tests bought off the shelf to measure the state's own standards. By investing in the MCAs, the state has shown its commitment to measure student and system performance against standards.
- **Our review found that, on the whole, the tests are consistent with the content and skills outlined in the Profile of Learning standards. And the tests contain a number of strong elements.** Nearly every test question on the MCAs matches

academic content and skills found in the reading and math standards. In the few cases where the questions do not have a home in the standards, the content these items measure is important and appropriate. Also, virtually all of the test questions on the mathematics and reading tests are technically sound. This is an issue with which other states have had trouble — but not Minnesota.

- **Supplementary documents developed to clarify the standards may be confusing to schools and students trying to prepare for the MCAs.** While every sound testing program bases state tests on “test blueprints,” we are concerned that various documents that were developed to clarify the standards are not aligned with each other. And such documents, even if closely aligned, are no substitute for a strong set of rigorous and explicit standards. In order for students, teachers, parents, administrators and test-makers to have shared expectations for performance, the statewide standards must be specific enough to convey what the state considers essential.
- **Because the Profile of Learning standards are broad and general, the tests are unbalanced and skew toward low-level skills.** An on-demand test is unlikely to measure all of a state’s standards; states generally make choices about which standards are most important to measure on a test and which can be assessed in other ways. In our view, the broad and vague language of the Profile of Learning content standards has made these choices more difficult and ambiguous than they should be. In the absence of sufficiently clear and specific standards, the hard work of deciding what standards to measure at grades 3 and 5 has been left to the test-makers. Some of the choices are understandable, but for others, the reasons for measuring certain objectives and not measuring others are not altogether clear.
- **The tests’ level of rigor is a serious concern.** In both reading and mathematics, the tests are not as rigorous as Achieve’s reviewers expected them to be. Important content is sometimes underemphasized in favor of lower-level skills, and the test questions often fail to tap more challenging concepts and skills that students at grades 3 and 5 should be able to master. Too often the tests have a preponderance of “easy” or “superficial” items, instead of including a range of difficult items.

RECOMMENDATIONS FOR MOVING FORWARD

As Minnesota develops the next generation of MCAs, Achieve recommends that the state:

- ✓ **Revise the standards to clearly and consistently communicate to teachers, students and test-makers what the state considers most essential for students to learn.** While the tests match the standards at a basic level, the standards do not provide enough guidance to test-makers about what is most important to measure at each grade level. Thus, the tests are somewhat haphazard and unbalanced, unfortunately in the direction of easier content and skills. Strengthening the content

standards, as we recommended in our first report, is essential to help ensure that the tests emphasize the most important standards.

- ✓ **Revise the assessments to ensure that they measure a broader range of standards, particularly the more challenging standards.** Our review found that in too many cases the tests place too heavy an emphasis on certain standards and leave other worthwhile content areas unassessed. The test results may provide a misleading picture of student achievement: They show only how students performed on certain standards. While not all standards can be assessed in an on-demand test, the tests should be better balanced than they currently are.
- ✓ **Increase the level of rigor throughout the tests.** Just as the MCAs could be better balanced in their content, they also could be better balanced in the level of challenge they pose to students. Relatively easy items dominate the tests. While not all items should be difficult for students, the tests should include more items that challenge students to demonstrate what they have learned.

INTRODUCTION AND METHODOLOGY

Since the early 1990s, 49 states have developed academic standards for their students and 48 are putting in place assessments to measure those standards. By stating clearly the knowledge and skills students are expected to gain as a result of their schooling, reformers hope that students will better understand what is expected of them, schools will improve their programs to help students achieve those expectations, and low-performing districts and schools in particular will be challenged to raise the level of teaching and learning. Because states have made substantial investments in new standards and tests — and many states are beginning to hold students and schools accountable for performance — policymakers and the public want to know about the quality of standards and tests. Achieve, Inc. was created precisely to address this issue.

Born out of the 1996 National Education Summit, Achieve helps states ensure that they have in place standards that compare favorably with the academic expectations of other states and high-performing nations and assessments that accurately measure student achievement against those standards. An independent, bipartisan, nonprofit organization overseen by a board of directors composed of governors and corporate CEOs, Achieve serves as a clearinghouse and resource center on education standards, testing and accountability, working primarily with states to support their work in these areas.

Central to the success of standards-based reform is high-quality information about student performance against standards. Statewide tests are among the best measurement tools available to see how well schools and students are meeting standards. However, if the assessments states are using do not closely match the expectations set out in the standards, then a state is sending conflicting messages about what is most important for students to learn. Students and schools who have used the state standards to lead curriculum and instruction should not be surprised by what they find on the tests. Moreover, such a mismatch means that the information students, parents, policymakers and the public receive about student performance is misleading: It may suggest that students are attaining standards when this may not be the case. As states move to implement assessments, they want to be sure that they measure what they expect students to know and be able to do.

Benchmarking is a highly respected practice in the business world. It is an activity that looks outward to find best practice and high performance and then measures actual business operations against those goals. Benchmarking in education follows the same principle. It is appropriate at a time when state education reforms are focused on raising student and school performance, as states want and need an external yardstick to gauge their efforts.

Achieve is involved in benchmarking for another important reason. States traditionally have had limited access to high-quality, trustworthy information about education standards. This is partly due to the fact that the standards movement in education is relatively young. But it is also a result of the disparate nature of much of the work that

has been done to date. While the standards reviews and “report cards” issued by other organizations have helped to focus national attention on the quality of standards, their judgments often have been in conflict and their tone has not always been constructive. States increasingly are looking for independent, credible advice on these issues.

Achieve’s benchmarking efforts are not designed to grade or rank states. Instead, the analysis of state standards and assessments is diagnostic in nature; one that yields detailed, reliable information that provides a roadmap for continuous improvement of standards and tests. In addition, our focus on the alignment between standards and assessments allows us to truly determine what the state expects all its students to know and be able to do and whether the state’s standards and tests are a strong enough foundation for the state’s efforts to improve education performance.

EDUCATION STANDARDS, TESTING AND POLICY IN MINNESOTA

During the development of the Basic Requirements and the Profile of Learning, Minnesota set some specific goals for the state’s education system:

- that students be prepared for postsecondary education and the world of work and be competitive in a global economy;
- that all learners are provided opportunities that recognize their uniqueness and maximize their achievement;
- that parents and local communities maintain the right to design curriculum, instruction and school opportunities they believe will best serve their students;
- that learning experiences be comprehensive;
- that student achievement be recorded and reported meaningfully; and
- overall, that there be stronger accountability for learning standards, higher achievement and better preparation of Minnesota students.¹

These goals resulted in a multifaceted set of education reforms that include the Profile of Learning, local performance assessments, the Basic Skills Tests, the MCAs, and accountability policies for schools, districts, and students. In *Aiming Higher: A Report on Education Standards and Policy for Minnesota*, our November 2000 report to the Minnesota Department of Children, Families, and Learning (CFL), Achieve and the Council for Basic Education (CBE) analyzed the quality, clarity, specificity, and rigor of the Profile of Learning content standards and the strength of state laws, rules, curriculum guides and support documents that support schools’ and districts’ efforts to ensure common high standards for all students. The analysis was aimed at making judgments about the usefulness of the Profile for schools, families, colleges, employers and other Minnesota citizens to raise standards and achievement.

¹ Minnesota State Board of Education, Statement of Need and Reasonableness in the Matter of the Proposed Permanent Rules Relating to Graduation Rule, Profile of Learning: Chapter 3501 (3501.0300 to 3501.0469), pp. 15–19

Aiming Higher identified a number of strengths of the Minnesota system:

- State officials have taken a strong stand on standards-based reform and remain committed to using standards to improve teaching and learning.
- The emphasis on applying knowledge to real-world problems and measuring achievement through performance tasks is commendable.
- The Minnesota assessment system includes multiple ways for students to demonstrate what they know and can do.

However, the report also urged the state to address significant shortcomings in its system of standards, assessments and accountability:

- The emphasis on applied learning in the Profile of Learning has resulted in too little emphasis on content knowledge in the academic disciplines; a more balanced approach is needed.
- The standards should be revised to include more clarity, specificity, depth and focus so that they can be applied rigorously throughout the state. Crosscutting skills and processes should be integrated into the content standards.
- The comprehensiveness of the Profile undermines the standards' focus and manageability. The state should focus on the core academic content areas as it moves ahead to improve the standards.
- The process for holding students accountable for learning is too decentralized to guarantee comparability and equity across the state. Statewide testing is needed particularly in the high school years to ensure that all students are meeting high standards and will be prepared for higher education, employment and citizenship.
- The legislative decision to make the Profile standards a local option does not advance the state's goal of high standards for all students. All schools and districts should implement content standards in the core content areas once they are revised.
- The state should develop a more comprehensive accountability system that includes assistance and interventions for consistently low-performing schools and districts and rewards for high-performing school systems. A statewide effort to build and sustain support for education reform also is needed.

In 2000, Minnesota policymakers also asked Achieve to conduct a review of the alignment between the state's High Standards and the MCAs. *Measuring Up* is aimed at providing detailed information to the state about the quality, alignment and rigor of the MCAs.

ALIGNING STANDARDS TO TESTS

As the statewide tests are designed to measure the extent to which schools and districts are implementing the Profile of Learning content standards, it is important that the MCAs are as closely aligned to the standards as possible. Beginning in 1998, Achieve developed and refined a methodology to judge the degree of alignment between state assessments and standards to ensure that the most important elements are fully captured. Achieve now has worked with nine states to analyze alignment. Achieve’s assessment-to-standards comparison has been developed to address the following questions:

- Does each assessment measure only content and skills reflected in the standards? In other words, can everything on the test be found in the state standards?
- Does each assessment fairly and effectively sample the important knowledge and skills in the standards? In other words, to what extent does each assessment measure the breadth and depth of the key content and skills for a grade level?
- Overall, is each assessment sufficiently challenging? Do the assessments as a whole grow more challenging from grade to grade?

Alignment is not a “yes or no” question; it is the degree to which standards and assessments are in agreement and serve in conjunction with one another to guide and support student learning. Alignment can be improved by strengthening the standards, the assessments or both. Consequently, responding to the above questions requires a systematic procedure to probe the different factors that contribute to alignment.

Achieve’s protocol for determining the alignment of assessments to standards is based on five criteria. Some criteria apply to individual items and result in individual ratings, while others apply to the entire set of items related to a given standard, such as “measurement” or “literary analysis,” and result in written judgments. Application of the protocol thus provides rich information about alignment of tests and standards that is typically unavailable to states. The alignment criteria Achieve considers are:

- **The accuracy of a test blueprint.** Reviewers check to see that each test question corresponds to at least one state standard. If no test blueprint is provided or the blueprint supplied by the test developer does not stand up to scrutiny (i.e., a significant number of items mapped to one standard or objective are found to be more closely related to a different one), then reviewers construct a new blueprint. Because test blueprints are typically the basis for state score reports it is important that they be as accurate as possible.
- **Content centrality.** This criterion examines the quality of the match between the content of each test item and the content of the related standard. Reviewers determine how closely the content of the item aligns with that of the related standard and then assign the item to one of four categories based on the degree of alignment: clearly

consistent with the standard, somewhat consistent because the item only matches part of the content in the standard, somewhat consistent because the standard is broad or vague, or not consistent with the standard.

- **Performance centrality.** Each item places a certain type of cognitive demand on a student (e.g., a student is asked to “identify” or “analyze”). If an item simply asks a student to “identify” and the corresponding standard requires a student to “analyze,” then there is a mismatch between the two performances. Reviewers assign each item to one of four categories based on the degree of alignment: clearly consistent with the standard, somewhat consistent because the item only matches one of several performances in the standard, somewhat consistent because the standard is broad or vague, or not consistent with the standard.
- **Challenge.** Two factors are considered in judging challenge. Reviewers determine if the *source of challenge* in an item comes from content in the standards that students must know to answer the item correctly or from some extraneous factor such as the language of the item or the way the item is constructed. Reviewers then determine the *level of challenge* by comparing the kinds of performances required by a set of items to those called for by the related standard and supporting materials and also judge whether the set of test items has a span of difficulty appropriate for students at the grade level tested.
- **Balance and range.** Balance compares the extent to which the knowledge and skills delineated in the standards receive the same emphasis on the assessment and if that emphasis is appropriate for the given grade level. Though it is very difficult for one assessment to measure the full complement of knowledge and skills required by state standards, range examines the degree to which the assessments sample the standards. Evaluating balance and range provides both qualitative and quantitative information about the choices states or test developers have made.

Despite this detailed protocol, there is no mathematical formula for matching a test to standards. Rather, the alignment analysis is a process that relies on experienced, knowledgeable educators who bring their knowledge to bear in applying the criteria for judging alignment. The judges selected by Achieve to analyze the alignment between assessments and standards are a deliberate mix of classroom teachers, curriculum specialists and subject-matter experts, with extensive expertise in content and assessment design. They often have experience in large-scale assessments and/or standards development and represent diverse viewpoints. This diversity of backgrounds has proven invaluable in arriving at considered judgments (for example, determining the appropriateness of a test item for a particular grade level).

Teams of reviewers, facilitated by an experienced group leader, meet in small groups to discuss their judgments. The findings and recommendations contained in this report are derived from the teams’ analyses of actual test items’ and secure test forms’ alignment to the Profile of Learning standards. These detailed analyses of secure test forms are included in the supporting technical documentation delivered to CFL.

RESULTS FOR MINNESOTA

The MCAs are the statewide tests currently given in grades 3 and 5 in reading and math and in writing at grade 5. We understand that the state plans to introduce MCAs for middle and high school students as well. These tests are part of the state's system for raising standards and assessing achievement; while they provide scores for individual students, they are intended to measure how well schools are implementing the Profile of Learning High Standards. In this way, the scores are not meant to be used for student accountability, but rather for system accountability.

Achieve's review of the MCAs is intended to help the state build and sustain high quality standards and a system of assessments to measure those standards that begins in the primary grades. As we described in *Aiming Higher*, we believe that clear and explicit high standards should be assessed regularly with common, standards-based assessments. We offer the following conclusions regarding the Profile of Learning and the MCAs so that Minnesota can revise the Profile of Learning with much-needed clarity, specificity and focus; strengthen the primary level MCAs; and develop MCAs for English language arts, mathematics, science and the social sciences for middle and high school students.

STRENGTHENING THE MINNESOTA COMPREHENSIVE ASSESSMENTS

The Achieve review of the Minnesota assessments in reading and mathematics in grades 3 and 5 found some strengths that can provide a solid foundation on which the state can build. Minnesota has commendably chosen to develop its own assessments, and our experience has shown that such tests are more likely than tests bought off the shelf to measure the state's standards. In fact, our review found that, on the whole, the tests measure content and skills found in the standards. The review also found a number of well-crafted items.

However, a number of areas should be addressed if the tests are to provide the state with strong information on how well students are attaining challenging standards. The tests tend to concentrate on relatively low-level standards and, as a result, they are not as rigorous as they could be.

1. The tests for the most part measure content and skills outlined in the Profile of Learning standards for reading, writing and mathematics.

Overall, the MCAs are clearly consistent or somewhat consistent with the content and skills in the state's standards. Virtually all the mathematics items in grades 3 and 5 match the standards in those grades, and the overwhelming majority of reading items match the standards in that subject. In some cases, however, our reviewers could find no matching standard for an item, but nevertheless felt that the content assessed by the item is appropriate and important at the grade level. In these cases, we recommend including that content in any subsequent revision of the standards.

It is important to note that this finding does not suggest that the alignment between the Minnesota standards and tests is exemplary. It does not say that the tests measure all of the standards well. Because the Profile of Learning is written at a high level of generality and does not always specify important content, the state cannot be certain that schools and students can successfully use the standards to guide teaching and learning toward high performance on the assessments.

Nevertheless, this finding is significant. It suggests that the MCAs measure academic content and skills, some of which are outlined in the Profile standards and others of which are not. This finding should counter the concerns of some critics in the state who may argue that the MCAs are unreliable tools for measuring student achievement against standards because at a basic, fundamental level the MCAs match the standards.

2. The tests include a number of well-crafted components.

All of the tests include a number of well-crafted items and some sections that are particularly strong. The reading passages for the most part are commendable. The vocabulary and syntax level are generally appropriate for the grade levels, and the passages represent an appropriate mix of various genres. Particularly commendable are two successful attempts that ask students to answer questions based on two passages simultaneously. The open-ended questions associated with these passages represent good opportunities to demonstrate levels of comprehension.

However, some of the passages are unnecessarily contrived and may not hold reader interest well. In addition, the range of genres, length and quality is too extreme. For example, some of the passages on the grade 3 test we reviewed are either childlike or unnecessarily complex for the questions that follow them; still other passages contain misleading illustrations, awkward presentations and poor organization. These problems are easily fixed, and our reviewers could offer specific suggestions for selecting exemplary reading passages.

In mathematics, nearly all the test questions earn high marks for technical quality. In all of these items, mathematics is the main challenge students have to meet in order to give a correct response. None of these items are written in a way that could cause students to answer the item incorrectly for other reasons besides the mathematics, such as the format of the item, confusing language or a particularly inviting distracter. This is commendable; in other states in which Achieve has worked, a high percentage of items have been overly tricky for students.

3. The state has developed several documents to support the standards, such as test blueprints, that are not consistent with each other; this may be confusing to schools and students trying to prepare for the MCAs.

Currently the state is sending mixed messages. The standards, the “Internal Test Specifications” that include test blueprints, the “Content Standards and Item

Classifications,” the curriculum frameworks and the sample Minnesota Comprehensive Assessment student reports each convey different messages.

This problem is particularly acute in the grade 5 writing assessment. The state’s scoring criteria delineate one set of expectations for Composition, Style, Sentence Formation, Usage and Grammar and Mechanics and Spelling that differ in significant ways from the expectations in the standards. The criteria appear to be related to the broadly-worded objectives outlined in the standards, yet it is unclear to reviewers whether students, teachers, and test-makers are all operating with the same set of expectations. It is unfair to hold students to standards for grammar, mechanics, spelling, and style if these expectations are not clearly outlined in the standards.

In order for students, teachers, parents, administrators and test-makers to have shared expectations for performance, the standards must be specific enough to convey what the state considers essential. While we applaud state officials for developing various documents to communicate expectations for students and schools, using supplementary documents to clarify the standards is no substitute for well-developed, detailed standards.

4. Because the Profile of Learning standards are broad and general, the tests are unbalanced and skew toward low-level skills. A number of important objectives are not assessed.

An on-demand test is unlikely to measure all of a state’s standards; states generally make choices about which standards are most important to measure on a test and which can be assessed in other ways. In our view, the broad and vague language of the Profile of Learning content standards has made these choices more difficult and ambiguous than they should be. In the absence of sufficiently clear and specific standards, the hard work of deciding what standards to measure at grades 3 and 5 has been left to the test-makers. Some of the choices are understandable, but for others, the reasons for measuring certain objectives and not measuring others are not altogether clear.

In some cases, the test items represent an imperfect measure of the state’s standards. This is largely because of the standards themselves; often they include a number of related concepts or skills that no single test item could possibly measure comprehensively. This is an issue Achieve has uncovered in many states, but in Minnesota it is particularly problematic. Some standards are over-sampled at the expense of other important content implied by the standards as well as important content that is not in the standards at all.

For example, the grade 5 standard for Literal Comprehension includes the following objective:

A student shall demonstrate comprehension of literal meaning by reading, listening, and viewing of nonfiction and fiction selections to identify main ideas and supporting details, retell main events or ideas in sequence, pronounce new words using phonics, demonstrate techniques of improving and expanding vocabulary, and demonstrate an age-appropriate reading rate.

This single standard includes multiple, distinct concepts which are important enough to be measured on the state tests. If the test included items that measured each of those objectives (all of them were listed as separate objectives for grade 3), then the test could represent a fair measure of the standards. But this is not the case. Test items for grade 5 that measure this standard place far too great an emphasis on some skills more than others, and those that are over-emphasized happen to be relatively low-level for grade 5 students.

While the comparatively few items that measure the Interpretation and Evaluation standards are balanced, the Profile of Learning standards omit important concepts that should be taught and assessed at this grade level. In particular, literary analysis skills (e.g., the identification and analysis of literary forms) and knowledge of literary elements (e.g., plot, setting, character and style) are omitted from the grade 5 standards and, thus, are not assessed on the MCAs.

The same patterns are true in mathematics. In grade 3, too many of the items mapped to the second objective under primary level Number Sense measure the same concept at the expense of others included in the objective. In grade 5, the assessment measures a reasonable proportion of objectives under the three standards Number Sense; Shape, Space, and Measurement; and Chance and Data Handling. However, the items are not evenly distributed among all of the concepts, procedures, and skills students are expected to know as stated in the standards. For example, a significant majority of the regular items corresponding to the grade 5 Number Sense standard match only some of the less complex objectives at the expense of more complex objectives under this standard.

5. The tests' level of rigor is a serious concern.

In both English and mathematics, the tests are not as challenging as they could be. The reading tests for grades 3 and 5 do not exhibit a broad enough range of difficulty. In part, this relatively low level of challenge reflects the overemphasis on low-level objectives. But in addition, the test items themselves often fail to tap more challenging concepts and skills that students at grades 3 and 5 should be able to master.

In mathematics, very few of the items require students to show they understand the connections among two or more mathematical ideas. Even the questions requiring students to write short responses test only basic knowledge and require students to do only very simple activities. In grade 5, for example, the items measuring Number Sense are skewed towards items requiring a direct application of operations. And, knowledge and understanding of important concepts, such as fractions, are underassessed.

In reading, meanwhile, the grade 3 test has a preponderance of “superficial” items, especially with respect to the “main idea/detail” standard (Literal Comprehension Standard Objective #1), which focuses too much on recalling details. The grade 5 reading test as a whole tends to overemphasize lower-level skills at the expense of more challenging ones. Too many test questions measure the Literal Comprehension standard,

while too few measure the Interpretation and Evaluation standard. And Achieve’s reviewers found that most of the items measuring Literal Comprehension are “easy,” while the items measuring Interpretation and Evaluation are more appropriately challenging for grade 5 students. Although Interpretation and Evaluation may be more cognitively demanding skills than those of Literal Comprehension, the state nonetheless should ensure that there is a range of difficulty represented for both strands.

RECOMMENDATIONS FOR MOVING FORWARD

The MCAs represent an integral element in the standards-based system the state is putting in place. As we noted in our previous report, these assessments have laid the groundwork for a future expansion of statewide testing in Minnesota, which we strongly recommend.

In the meantime, the state can build on its solid foundation to ensure that the tests provide high-quality information about student progress toward high standards. Minnesota has commendably chosen to build its own tests to measure its standards, rather than buy an “off-the-shelf” version that was not designed to measure Minnesota’s standards. And the content of the tests reflects that decision; the tests generally measure knowledge and skills in the standards.

Moreover, the high quality of many of the test items, particularly the mathematics items and well-chosen reading passages, demonstrates that Minnesota test-makers are skillful — we are confident that the next generation of tests can represent a substantial improvement. As Minnesota develops this next generation, Achieve recommends that the state:

- ✓ **Revise the standards to clearly and consistently communicate to teachers, students and test-makers what the state considers most essential for students to learn.**

Our findings suggest that while the tests match the standards at a basic level, the standards do not provide enough guidance to test-makers about what is most important to measure at each grade level. As a result, the tests are somewhat haphazard and unbalanced, unfortunately in the direction of easier content and skills. Strengthening the standards to provide greater detail and depth and to specify what content is most important at particular grade levels will help ensure that the tests emphasize the most important standards.

- ✓ **Revise the assessments to ensure that they measure a broader range of standards, particularly the more challenging standards.**

Our review found that in too many cases the tests place too heavy an emphasis on certain standards and leave other worthwhile content areas unassessed. As a result, the test results may provide a misleading picture of student achievement: They show only how students performed on certain standards. While not all standards can be assessed in an on-demand test, the tests can be more balanced than they currently are.

One way to improve the balance in mathematics would be to replace some of the items that are redundant — those that measure objectives measured by other test items — and replace them with ones that measure important content currently not included. In reading, the balance could be improved by placing somewhat less emphasis on literal

understanding and more on interpretation and evaluation of literary ideas and informational text.

✓ **Increase the level of rigor throughout the tests.**

Just as the MCAs could be better balanced in their content, they also could be better balanced in the level of challenge posed to students. The tests contain a preponderance of relatively easy items. While not all items should be difficult for students — on the contrary; tests should include a mix of easy, moderate and hard items — the tests should include more items that challenge students to demonstrate what they have learned.

In making this recommendation we are not suggesting that the state include more open-ended items and fewer multiple-choice items. That may be a good idea, but the decision to do so should stem from the content the state is trying to assess. Multiple-choice does not mean easy; nor does open-ended mean difficult. In fact, our review found that both types of items could be made more challenging.

APPENDIX

ABOUT THE MINNESOTA COMPREHENSIVE ASSESSMENTS

MCAs in reading for grades 3 and 5 are untimed. Both tests are administered in four separate sessions over a two-day period; testing sessions are usually 25–35 minutes. Each of the four sections of the tests has approximately 15 questions. Each section contains one open-ended question. Students take test sessions one and two on day one and sessions three and four on day two. Each test includes embedded field test items, so that each test contains approximately 60 questions total with nine to 11 total reading passages (50 base items and 10 field test items). Students record their answers on a separate answer sheet.

The writing test, administered in grade 5 only, is a matrix-sample; four prompts are administered, one representing each of the four modes of discourse (narrative, descriptive, problem/solution, clarification), and each student responds to one randomly assigned prompt. A testing period of 60–90 minutes is suggested for the writing tests, but, like the reading test, it is untimed. Two raters on each of the following domains score each student response: composing, style, sentence formation, usage/grammar and mechanics/spelling. The Achieve reviewers examined only two of the four prompts and some student anchor papers.

The grade 3 mathematics test consists of four parts, two of which were administered on the first day and two on the second day of the assessment. The first part, on day one, consists of 11 multiple-choice items and one open-ended item. Students were not permitted to use a calculator on part one. They could use a calculator on the second part given on day one, which consists of 14 multiple-choice items and one open-ended item. The third part consists of 10 multiple-choice items and one open-ended item. Students were not permitted to use a calculator on part three, but could use one on part four. The fourth and final part of the test consists of 15 multiple-choice items and one open-ended item.

The grade 5 mathematics test also consists of four parts. The first part consists of 18 multiple-choice items and one two-part open-ended item. The second part consists of 17 multiple-choice items and one open-ended item. Part three, the first part given on the second day, and the only part on which students are not permitted to use a calculator, consists of 21 multiple-choice items and one open-ended item. On the fourth part of the test, consisting of 10 multiple-choice items and one open-ended item, students are asked to explain their reasoning as to why an answer is wrong.

ACHIEVE STAFF AND EXPERT REVIEW TEAMS

Achieve relied on the expertise of nationally respected experts in academic content, standards, curriculum and assessment design to inform and conduct the standards benchmarking and alignment of assessments to standards.

English language arts

- Sheila Byrd, education consultant (team leader)
- Vernon Gettone, Ph.D., California Teachers' Association
- Laura McGiffert, Senior Project Associate, Achieve, Inc.
- Melanie Pritchett, Ed.D., Just for the Kids

Mathematics

- Kaye Forgione, Ed.D., University of Texas-Austin
- Lucio Calzada, Corpus Christi Public Schools, Texas
- Lynn Raith, Pittsburgh Public Schools
- Norman Webb, Ph.D., University of Wisconsin-Madison (team leader)

The following Achieve staff and senior consultants led the analysis and report development for Minnesota:

- Ellen Clark, Benchmarking Consultant
- Matt Gandal, Vice President, Achieve
- Lauren Resnick, Ed.D., Director, Learning Research and Development Center, University of Pittsburgh
- Robert Rothman, Senior Project Associate, Achieve
- Jean Slattery, Ed.D., Senior Benchmarking Consultant
- Jennifer Vranek, Director, Benchmarking and State Services, Achieve
- T. Jason Weedon, Project Associate, Achieve

Board of Directors, Achieve, Inc.

Co-Chairs

Louis V. Gerstner, Jr., Chairman and CEO
IBM Corporation

John Engler, Governor
State of Michigan

Co-Vice Chairs

Gray Davis, Governor
State of California

Philip M. Condit, Chairman and CEO
The Boeing Company

Board Members

Keith E. Bailey, Chairman, President and CEO
Williams

Roy E. Barnes, Governor
State of Georgia

Craig R. Barrett, President and CEO
Intel Corporation

Frank Keating, Governor
State of Oklahoma

Gary Locke, Governor
State of Washington

Edward B. Rust, Jr., Chairman and CEO
State Farm Insurance

Arthur F. Ryan, Chairman and CEO
Prudential

Bob Taft, Governor
State of Ohio

President

Robert B. Schwartz



Achieve, Inc.
www.achieve.org

400 North Capitol Street, NW
Suite 351
Washington, DC 20001
Phone: (202) 624-1460
Fax: (202) 624-1468

8 Story Street
1st Floor
Cambridge, MA 02138
Phone: (617) 496-6300
Fax: (617) 496-6361