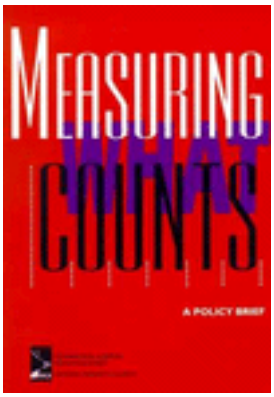


Measuring What Counts: A Policy Brief

Mathematical Sciences Education Board, National Research Council

ISBN: 0-309-59844-3, 32 pages, 7 x 10, (1993)

MEASURING WHAT COUNTS

A POLICY BRIEF

MATHEMATICAL SCIENCES
EDUCATION BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, DC 1993

NATIONAL ACADEMY PRESS 2101 Constitution Avenue, NW Washington, DC 20418

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance. This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.
The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Robert. M. White are chairman and vice chairman, respectively, of the National Research Council.
The Mathematical Sciences Education Board was established in 1985 to provide a continuing national capability to assess the status and quality of education in the mathematical sciences and is concerned with excellence in education for all students at all levels. The Board reports directly to the Governing Board of the National Research Council.

International Standard Book Number 0-309-04986-5

Copies of the policy brief and the complete report may be purchased from the National Academy Press, 2101 Constitution Avenue, NW, Box 285, Washington, DC 20055. Call 800-624-6242 or 202-334-3313 (in the Washington Metropolitan Area).

Printed in the United States of America

CONTENTS

The contents of the entire report, from which this policy brief is extracted, are reproduced on pages iv and v.

PREFACE

Calls for standards in education have been echoing across the nation for several years, especially since political leaders of both parties decided to adopt bipartisan national goals for education. Standards without appropriate means of measuring progress, however, amount to little more than empty rhetoric. To stay the course and achieve the national goals for education, we must measure the things that really count.

Since 1989 mathematics has led the national movement toward establishing standards for curriculum, teaching, and assessment, with publication of several significant reports authored by the National Council of Teachers of Mathematics (NCTM), the National Research Council, and the Mathematical Sciences Education Board (MSEB). *Measuring What Counts* further advances this national discussion by establishing crucial research-based connections between standards and assessment. It demonstrates the importance of three key principles—on content, learning, and equity—for any program of assessment that aims to support the national educational goals.

The intent of *Measuring What Counts* is not to offer immediate practical advice, but to lay out a conceptual framework that will help those who are struggling with the urgent need to develop new assessments that align properly with the national standards for mathematics education. Its message is quite simple, but its implications are profound: Assessment in support of standards must not only measure results, but must also contribute to the educational process itself.

Experience reveals enormous gaps between current assessment practice and new goals for mathematics education. It is clear from the recent history of failed reform that when assessment is

separated from curriculum and instruction, teaching becomes distorted, thus diminishing learning. Experts agree that for education to be effective, curriculum, instruction, and assessment must harmonize for their mutual support. Both internal (teacher-based) and external (district-or state-based) assessment must support improved learning. However, the path from general agreement to specific assessments is far from clear. We are embarking on a new venture, guided by the principles of content, learning, and equity. Success at this venture will require years of exploration by thousands of practitioners working with mathematics education specialists and measurement experts to achieve a more effective balance of assessment in practice. It will also require leadership and support from policymakers at every level of government.

All reform is evolutionary. As society changes, the targets and goals for education change. Assessment is our primary tool for monitoring progress and making midcourse corrections. The principles of assessment set forth in *Measuring What Counts* provide a solid conceptual basis for current efforts to improve assessment and lay the groundwork for more detailed assessment standards to be published by the NCTM.

This policy brief is based on a lengthier report, *Measuring What Counts: A Conceptual Guide for Mathematics Assessment*, which was developed by National Research Council's Study Group on Mathematics Assessment over a period of two years. Members of the study group faced many significant hurdles posed by their differing professional perspectives, by the rapidly changing context of educational assessment, and by the challenges posed by the new *Standards* for school mathematics. We owe each of them a special thanks for persisting in this formidable task to reach consensus on the key principles enunciated in the two reports.

We are indebted as well to authors of the three commissioned papers published with the full volume, to provide additional background on several core issues addressed by the study group.

Like all reports of the National Research Council, the policy brief and the full volume of *Measuring What Counts* were extensively reviewed—first by outside experts in early draft form, then by the MSEB Committee on Policy Studies at several key stages, and, at the final stage, under the careful protocol of the NRC's Report Review

Committee. We thank these many reviewers for their insightful and knowledgeable comments.

Financial support for work of the study group and preparation of *Measuring What Counts* was provided by the U.S. Department of Education and the National Science Foundation. We gratefully acknowledge the support of these organizations.

Hyman Bass, Chairman

Mathematical Sciences Education Board

MEASURING WHAT COUNTS

COMMON ASSESSMENT MYTHS

*The **mathematics for the talented few myth** that only students with special talent can learn the kind of mathematics required to meet new performance standards*. New standards call for *all* students to learn important and challenging mathematics. The prevailing public belief is that only some are capable of meeting this expectation. Yet from elementary school to college, challenging mathematics classes are emerging all across the country that reduce drop-out rates and, more importantly, record significant gains by students from traditionally underrepresented groups. Such programs invalidate the myth that only a talented few can learn important mathematics.

*The **arithmetic myth** that mathematics is synonymous with the facts and procedures of arithmetic*. Mathematics is much more than arithmetic. It is finding, making, and describing patterns. It is constructing mathematical models for both practical and theoretical situations, using technology when appropriate. It is representing and reasoning about quantities and shapes. It is devising and solving challenging problems and justifying and communicating about the solutions. Students today need to learn and be assessed on all of these ways of knowing and doing mathematics.

*The **ground rules myth** that changing methods of assessment is unfair to students who are doing well or showing improvement on the traditional choose-the-right-answer approach to assessment*. In fact, doing well or doing better by old standards is no longer good enough. Current assessment misleads us about how much important mathematics is deeply understood by even high achievers. Current assessment provides little or no information about whether students have developed the skills and concepts they need to live and work in the twenty-first century. Nothing could be *fairer* than to hold students to the new high levels of performance they will need as adults.

*The **high risk myth** that new assessments ignore important, time-honored measurement criteria for evaluating assessments*. Neither mathematics educators nor measurement experts are proceeding alone in assessment reform. In recent years, there has been increased collaboration and cross-fertilization between the two fields on how this might be accomplished. As a result, the educational principles that mathematics educators have articulated and the new technical criteria that measurement experts have proposed are mutually supportive.

*The **expense myth** that new assessments are just too costly to be realistic for a nation such as the United States with its commitment to universal education*. The direct costs will be higher than those associated with traditional testing, but so are the benefits. The "true benefit" of an assessment depends on more than dollars and cents. Time spent on high-quality mathematics assessment is time *well* spent because such assessments function as part of the learning process. The U.S. cannot afford the cost of any assessment that might work against current goals in mathematics education, no matter what its price tag.

*The **magic bullet myth** that mathematics assessment reform, as embodied in three principles, will be a panacea for the many ills that plague education*. Nothing could be further from the truth. Educational reform depends on much more than changing assessment. Efforts to reform school mathematics must proceed along three lines: revitalizing the curriculum, redesigning the professional development of teachers, and reconceptualizing the assessment of learning. Teachers are the central agents for changing school mathematics, and they, in particular, will need continuing support and adequate resources to make the necessary transformations.

POLICY BRIEF

"You can't fatten a hog by weighing it." So said a farmer to a governor at a public hearing in order to explain in plain language the dilemma of educational assessment. To be useful to society, assessment must advance education, not merely record its status.

Assessment is a way of measuring what students know and of expressing what students should learn. As the role of mathematics in society has changed, so mathematics education is changing, based on new national standards for curriculum and instruction. Mathematics assessment must also change to ensure consistency with the goals of education.

Three fundamental educational principles form the foundation of all assessment that supports effective education:

Assessment is a way of measuring what students know and of expressing what students should learn.

The Content Principle

Assessment should reflect the mathematics that is most important for students to learn.

The Learning Principle

Assessment should enhance mathematics learning and support good instructional practice.

The Equity Principle

Assessment should support every student's opportunity to learn important mathematics.

Despite their benign appearance, these principles contain the seeds of revolution. Few assessments given to students in America today reflect any of these vital principles. For educational

reform to succeed, the yardsticks of progress must be rooted in the principles of content, learning, and equity.

MATHEMATICS: THE FLAGSHIP FOR STANDARDS

Measuring What Counts fits naturally into the continuing efforts by the American mathematics community to improve mathematics education. In 1989, the National Council of Teachers of Mathematics released its first standards volume, entitled *Curriculum and Evaluation Standards for School Mathematics*, in a planned release precisely two months after publication by the Mathematical Sciences Education Board (MSEB) of *Everybody Counts*, a "report to the nation on the future of mathematics education." This was followed in 1990 by the MSEB report *Reshaping School Mathematics* and in 1991 by the second NCTM standards volume, *Professional Standards for Teaching Mathematics*. These volumes on curriculum and instruction were followed by *Measuring Up* (MSEB, 1993) in which 13 examples of fourth grade tasks for mathematics assessment offer educators and policymakers a glimpse of how new assessment could support standards-based education.

Measuring What Counts heralds the next major thrust of the standards movement in mathematics education. It sets forth principles upon which assessments must be constructed if they are to serve educational purposes. These principles form the basis for the third volume of the NCTM standards series, on assessment, which will be published in 1995. The mathematics community's standards development work is recognized as the model for reform in other academic disciplines, from history to the sciences.

Measuring What Counts has two primary purposes: to put assessment in the context of broader mathematics reform efforts and to establish educational principles for mathematics assessment. The report was developed by a broad cross section of mathematicians, academics, and educators under the direction of the Mathematical Sciences Education Board of the National Research Council. This policy brief summarizes the salient points of the full report, which is available from the National Academy Press, and describes what changes in assessment reform will mean for a variety of different audiences.

MATHEMATICS IN TODAY'S WORLD

The pressures to change mathematics education reflect society's disappointment with the lack of interest and accomplishment of so many students in today's schools. In the background of public debate is the steady criticism that school mathematics is out of step with today's world and is neither well taught nor well learned.

Unfortunately, these pressures often suggest inconsistent courses of action, with standards-based curriculum and instruction moving in one direction while mandated tests remain aimed in another direction, at an older, more traditional target. Too often, teachers are caught in the middle. To be effective, mathematics education must be rooted in the practice of mathematics, in the art of teaching, and in the needs of society. These pivotal forces drive the current movement to improve mathematics education:

- A more comprehensive view of mathematics and its role in society: mathematics is no longer just a prerequisite subject for prospective scientists and engineers but is a fundamental aspect of literacy for the twenty-first century.
- A recommitment to the traditional wisdom that mathematics must be made meaningful to students if it is to be learned, retained, and used.
- The growing recognition that in this technological era, all students should learn more and better mathematics.

MATHEMATICS FOR LIFE

Because Americans do not know enough mathematics, they are at a disadvantage in a modern world where making many decisions requires mathematical proficiency. We encounter mathematics in every corner of our lives. Graphs, charts, and statistical data appear on television and in newspapers. Results of opinion polls are reported with margins of error. Banks, credit card companies, insurance companies, and other financial services advertise varying rates, both for investment returns and finance charges.

The complexity of the modern workplace requires employees who can think, who can adapt to changing circumstances, and who can cooperate with co-workers to solve unexpected problems. These skills are important if American businesses are to be competitive in a rapidly changing international environment.

The complexity of modern life also requires us to reason with numbers and interpret graphs, maps and diagrams. Although computers perform ordinary tasks, they do not eliminate the need to understand ideas. Any home buyer ought to understand how interest rates operate even if a computer performs the computations. Likewise, anyone building a structure or redecorating a room should understand how to read and prepare scale drawings.

Mathematics also plays a major role in how people make informed social decisions. With so much information today presented in graphics or as survey results, it is critical for Americans to know the correct way to draw accurate conclusions from data.

ASSESSMENT IN TODAY'S WORLD

Assessment is the guidance system of education just as standards are the guidance system of reform. Assessment helps teachers and parents determine what students know and what they need to learn. It can play a powerful role in conveying clearly and directly how well students are learning and how well school systems are responding to the national call for higher education standards.

At its root, assessment is a communication process that tells students, teachers, parents, and policymakers some things—but not

everything—about what students have learned. Assessment provides information that can be used to award grades, to evaluate a curriculum, or to decide whether to review fractions. Internal assessment communicates to teachers critical aspects of their students' performance, helping them to adjust their instructional techniques accordingly. External assessment provides information about mathematics programs to parents, state and local education agencies, funding bodies, and policymakers.

> Assessment can be the engine that propels reform forward, but only if education rather than measurement is the driving force.

Many reformers see assessment as much more than an educational report card. Assessment can be the engine that propels reform forward, but only if we make *education* rather than *measurement* the driving force in the development of new assessments. By setting a public and highly visible target to which all can aspire, assessment can inform students, parents, and teachers about the real performance-based meaning of curriculum guidelines. Assessments not only measure what students know but provide concrete illustrations of the important goals to which students and teachers can aspire.

ASSESSMENT IN THE SERVICE OF EDUCATION

Improved assessment is required to complement and support the changes under way in mathematics education: both in the kinds of mathematics that are taught and in the ways in which they are taught. As such, assessment is an integral part of an interlocking triad of reforms along with curriculum and professional development of teachers. Because assessment is key to determining what students learn and how teachers teach, it must be reshaped in a manner consistent with the new vision of teaching and learning.

Students learn important mathematics when they use mathematics in relevant contexts in ways that require them to apply what they know and extend their thinking. Students think when they are learning and they learn when they are thinking. Good teachers have long recognized that mathematics comes alive for students when it is learned through experiences they find meaningful and valuable. Students learn best and most enduringly by engaging mathematics actively, by reflecting on their experience, and by communicating with others about it. Students want to make sense of the world, and mathematics is a wonderful tool to use in this eternal quest.

Because teamwork is important on the job and in the home, mathematics students learn important lessons when they work in teams, combining their knowledge and discovering new ways of solving problems. Often there is no single right answer, only several possibilities that unfold into new questions. Students need opportunities to advance hypotheses, to construct mathematical models, and to test their inferences by using the mathematics of estimation and uncertainty alongside more traditional techniques of school mathematics. Hand-held graphing calculators allow, for the first time, thorough exploration of complex, real-life problems. Computational impediments need no longer block the development of problem-solving or mathematical modeling skills.

> Educational assessment has been driven largely by practical and technical concerns rather than by educational priorities.

This new vision of learning and teaching is now being tried in some classrooms across the country. Current assessment does not support this vision and often works against it. For decades, educational assessment in the United States has been driven largely by practical and technical concerns rather than by educational priorities. Testing as we know it today arose because very efficient methods were found for assessing large numbers of people at low cost. A premium was placed on assessments that were easily administered and that made frugal use of resources. The constraints of efficiency meant that mathematics assessment tasks could not tap a student's ability to estimate the answer to an arithmetic calculation, construct a geometric figure, use a calculator or ruler, or produce a complex deductive argument.

A narrow focus on technical criteria—primarily reliability— also worked against good assessment. For too long, reliability meant that examinations composed of a small number of complex problems were devalued in favor of tests made up of many short items. Students were asked to perform large numbers of smaller tasks, each eliciting information on one facet of their understanding, rather than to engage in complex problem solving or modeling, the mathematics that is most important.

In the absence of expressly articulated educational principles to guide assessment, technical and practical criteria have become de facto ruling principles. The content, learning, and equity principles are proposed not to challenge the importance of these criteria, but to challenge their dominance and to strike a better balance between educational and measurement concerns. An increased emphasis on

validity—with its attention to fidelity between assessments, high-quality curriculum and instruction, and consequences—is the tool by which the necessary balance can be achieved.

In some ways, test developers do acknowledge the importance of curricular and educational issues. However, their concern is usually about coverage, so they design tests by following check-off lists of mathematical topics (e.g., fractions, single-digit multiplication). This way of determining test content matched fairly well the old vision of mathematics instruction. In this view you could look at little pieces of learning, add them up, and get the big picture of how well someone knew mathematics.

MATHEMATICS AS LANGUAGE

Assessment in writing provides a good analogy for assessment in mathematics. We would not accomplish much if we tried to measure writing skills by testing only grammar, spelling, and vocabulary. To assess writing, teachers must ask their students to actually write: sentences, paragraphs, essays. Moreover, when we test children by having them actually write, teachers tend to focus instruction on the substantive facets of writing rather than only on the mechanics. Similarly, in mathematics, we should assess children by asking them to engage in real mathematics—to solve problems, make conjectures, formulate convincing arguments.

Assessment should reflect the mathematics that is most important for students to learn.

Today we recognize that students must learn to reason, create models, prove theorems, and argue points of view. Assessments must reflect this recognition by adhering to the three principles of content, learning and equity. You cannot get at this kind of deep understanding and use of mathematics by examining little pieces of learning. Assessments that are appropriately rich in breadth and depth provide opportunities for students to demonstrate their deep mathematical understanding. Mathematics education and mathematics assessment must be guided by a common vision.

THE CONTENT PRINCIPLE

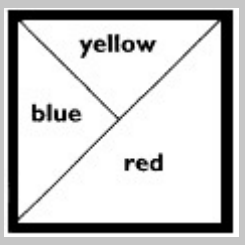
Any assessment of mathematics learning should first and foremost be anchored in important mathematics. Assessment should do much more than test discrete procedural skills so typical of today's topic-by-process frameworks for formal assessments. Many current assessments distort mathematical reality by presenting mathematics as a set of isolated, disconnected fragments, facts, and procedures. The goal ought to be assessment tasks that elicit student work on the meaning, process, and uses of mathematics.

A Mosaic Problem

Tri-Tex Corporation has the following logo. It's going to put a large-scale mosaic of the logo on the side of its corporate headquarters building.

A 2' × 2' scale model of the logo uses:

144 yellow tile pieces
144 blue tile pieces
288 red tile pieces



When the logo is placed on the side of the building it will also have a border that consists of a single band of long black tiles. It takes 40 of the black tiles to make a border for the 2' × 2' scale model.

• How many tiles of each color will the purchasing department have to order if the full-scale mosaic is 4' × 4'? What if it's 6' × 6'?

• Tri-Tex's president says he'll decide later just how big the mosaic will be, but he wants everything set up in advance so the purchasing department can send out the order as soon as he makes up his mind. Your job is to write a two-part memo to the purchasing department.

(a) In part I of the memo, tell how the purchasing department can do a simple calculation to find out how many yellow, blue, red, and black tiles they need to purchase, for any size design. The instructions should be as simple and direct as possible.

(b) In part 2 of the memo, explain how you arrived at the answer you did, so the people in the purchasing department can understand why your instructions give the right answer.

Important mathematics must shape and define the content of assessment. Appropriate tasks emphasize connections within mathematics, embed mathematics in relevant external contexts, require students to communicate clearly their mathematical thinking, and promote facility in solving nonroutine problems. Considerations of connections, communication, and nonroutine problems raise many thorny issues that testmakers and teachers are only beginning to explore. However, these considerations are essential if students are to meet the new expectations of mathematics education standards.

The content principle has profound implications for those who design, score, and use mathematics assessments. Many of the assessments used today, such as standardized multiple-choice tests, have reinforced the view that the mathematics curriculum should be con*structed* from lists of narrow, isolated skills that can be easily disassembled for appraisal. The new vision of school mathematics requires a curriculum and matching assessment that is both broader and more integrated.

Rather than forcing mathematics to fit assessment, assessment must be tailored to the mathematics that is important to learn.

The mathematics in an assessment must never be distorted or trivialized for the convenience of assessment. Assessment should emphasize problem solving, thinking, and reasoning. In assessment as in curriculum activities, students should build models that connect mathematics to complex, real-world situations and regularly formulate problems on their own, not just solve those structured by others. Rather than forcing mathematics to fit assessment, assessment must be tailored to the mathematics that is important to learn.

Implications of the content principle extend as well to the scoring and reporting of assessments. New assessments will require new kinds of scoring guides and ways of reporting student performance that more accurately reflect the richness and diversity of mathematical learning than do the typical single-number scores of today.

THE LEARNING PRINCIPLE

To be effective as part of the educational process, assessment should be seen as an integral part of learning and teaching rather than as the culmination of the process. Time spent on assessment will then contribute to the goal of improving the mathematics learning of all students.

Assessment should enhance mathematics learning and support good instructional practice.

If assessment is going to support learning, then assessment tasks must provide genuine opportunities for all students to learn significant mathematics. Too often a sharp line has been drawn between assessment and instruction. Teachers teach, then instruction stops and assessment occurs. In the past, for example, students' learning was often viewed as a passive process whereby students remember what teachers tell them to remember. Consistent with this view, assessment has often been thought of as the end of learning. The student is assessed on material learned previously to see if her or she remembers it. Earlier conceptions of the mathematics curriculum as a collection of fragmented knowledge led to assessment that reinforced the use of memorization as a principal learning strategy.

Today we recognize that students make their own mathematics learning individually meaningful. Learning is a process of continually restructuring prior knowledge, not just adding to it. Good education provides opportunities for students to connect what is being learned to prior knowledge. Students know mathematics if they have developed the structures and meanings of the content for themselves.

If assessment is going to support good instructional practice, then assessment and instruction must be better integrated than is commonly the case today. Assessment must enable students to construct new knowledge from what they know. The best way to provide opportunities for the construction of mathematical knowledge is through *assessment tasks* that resemble learning tasks in that they promote strategies such as analyzing data, drawing contrasts, and making connections. This can be done, for example, by basing assessment on a portfolio of work that the student has done as part of the regular instructional program, by integrating the use of scoring guides into instruction so that students will begin to internalize the standards against which the work will be evaluated, or by using two-stage testing in which students have an extended opportunity to revise their initial responses to an assessment task.

Assessment tasks must provide genuine opportunities for all students to learn significant mathematics.

Not only should all students learn some mathematics from assessment tasks, but the results should yield information that can be used to improve students' access to subsequent mathematical knowledge. The results must be timely and clearly communicated to students, teachers, and parents. School time is precious. When students are not informed of their errors and misconceptions, let alone helped to correct them, the assessment may both reinforce misunderstandings and waste valuable instructional time.

INTEGRATING ASSESSMENT AND INSTRUCTION

The Pacesetter™ project, currently under development by the College Board, seeks to integrate assessment activities and instruction in a new capstone high school mathematics course. Teachers assist students through case studies of applications of mathematics to problems in industrial design, inventions, economics, demographics, and other areas. Throughout the activities, students answer questions about the models they construct and write them up in a portfolio. Pacesetter also includes methods to certify that students have met the course's goals. Even though Pacesetter has not been fully evaluated, the project shows that assessment can be integrated into instruction in a variety of ways including

- responses to open-ended questions, problems, and tasks;
- projects that take several days or weeks;
- portfolios of mathematical products;
- writing intended to express mathematical ideas;
- demonstrations, including some computer models;
- presentations, discussions, and debates;
- investigations and research-type explorations; and
- models and simulations.

When the line between assessment and instruction is blurred, students can engage in mathematical tasks that not only are meaningful and contribute to learning, but also yield information the

student, the teacher, and perhaps others can use. In fact, an oft-stated goal of reform efforts in mathematics education is that visitors to classrooms will be unable to distinguish instructional activities from assessment activities.

THE EQUITY PRINCIPLE

The idea that some students can learn mathematics and others cannot must end; mathematics is not reserved for the talented few, but is required of all to live and work in the twenty-first century. Assessment should be used to determine what students have learned and what they still need to learn to use mathematics well. It should not be used to filter students out of educational opportunity.

> Assessment should support every student's opportunity to learn important mathematics.

Designing assessments to enhance equity will require conscientious rethinking not just of what we assess and how we do it but also of how different individuals and groups are affected by assessment design and procedures. The challenge posed by the equity principle is to devise tasks with sufficient flexibility to give students a sense of accomplishment, to challenge the upper reaches of every student's mathematical understanding, and to provide a window on each student's mathematical thinking.

Some design strategies are critical to meeting this challenge, particularly permitting students multiple entry and exit points in assessment tasks and allowing students to respond in ways that reflect different levels of mathematics knowledge or sophistication. But there are no guarantees that new assessments will be fairer to every student, that every student will perform better on new assessments, or that differences between ethnic, linguistic, and socioeconomic groups will disappear. While this is the hope of the educational reform community, it seems clear that hope must be balanced by a spirit of empiricism: there is much more to be learned about how changes in assessment will affect longstanding group differences.

Equity implies that every student must have an opportunity to learn the important mathematics that is assessed. Obviously, students who have experience reflecting on the mathematics they are learning, presenting and defending their ideas, or organizing, executing, and reporting on a complex piece of work will have an

advantage when called upon to do so in an assessment situation. Especially when assessments are used to make high-stakes decisions on matters such as graduation and promotion, the equity principle requires that students be guaranteed certain basic safeguards. Students cannot be assessed fairly on mathematics content that they have not had an opportunity to learn.

Assessments can contribute to students' opportunities to learn important mathematics only if they are based on standards that reflect high expectations for all students. There can be no equity in assessment as long as excellence is not demanded of all. If we want excellence, the level of expectation must be set high enough so that, with effort and good instruction, every student will learn important mathematics.

We have much to learn about how to maintain uniformly high performance standards while allowing for assessment approaches that are tailored to diverse backgrounds. Uniform application of standards to a diverse set of tasks and responses poses an enormous challenge that we do not yet know how to do fairly and effectively. Nonetheless, the challenge is surely worth accepting.

Students cannot be assessed fairly on mathematics content that they have not had an opportunity to learn.

OBSTACLES AND CHALLENGES

The boldness of our vision for mathematics assessment should not blind us to either the obstacles educators will face or the limitations on resources we possess for making it come about. Even if new assessments were to magically appear and be implemented across the nation, many substantial problems will remain. Examples of important, unresolved issues abound:

- Open-ended problems are not necessarily better than well-defined tasks. The mere labels "performance assessment" and "open ended" do not guarantee that a task meets sound educational principles. For example, open-ended problems can be interesting and engaging but mathematically trivial. Performance tasks can be realistic and mathematically appropriate but out of harmony with certain students' cultural backgrounds.
- The equity principle implies that students must be provided an opportunity to learn the mathematics that is assessed and that schools must be held to "school delivery

standards" to ensure that students are provided with appropriate preparation, particularly for any high-stakes assessment. However, many would argue that past remedies designed to improve schools often failed precisely because the emphasis was placed on the resources schools should provide rather than the outcomes that schools should achieve.

- The equity principle also requires some consideration of consequences for schools of the way assessments are used. Fair inferences can be drawn and comparisons can be made only when assessment data include information on the nature of the students served by the school, students' opportunities to learn the mathematics assessed, and the adequacy of resources available to the school. Assessments based only on partial data—typically outcome scores on basic skills—can seriously mislead the public about how schools are performing and how to improve them.
- On the job and in the real world, knowledge is frequently constructed and validated in group settings rather than through individual exploration. Mathematics is no exception: learning and performance are frequently improved in group settings. Hence assessment of learning must reflect the value of group interaction. The challenge of fairly appraising an individual's contribution to group efforts is immense, posing unresolved problems both for industry and education.

Our Vision For Mathematics Assessment Should Not Blind Us To The Obstacles That Must Be Overcome.

- New performance-based assessments introduce significant challenges both for the mathematical expertise of those who score assessments and for the guidelines used in scoring. Problem solving legitimately may involve some false starts or blind alleys; students whose work includes such things are doing important mathematics and their grades need to communicate this in an appropriate fashion. All graders must be alert to the unconventional, unexpected answer that, in fact, may contain insights that the assessor had not anticipated. Of course, the greater the chances of unanticipated responses, the greater the mathematical sophistication needed by those grading the tasks.

- As assessments become more complex and more connected to real-world tasks, there is a greater chance that the underlying assumptions and points of view may not apply equally to all students, particularly when differences in background and instructional histories are involved. Despite good intentions and best efforts to make new assessments fairer to all students than traditional forms of testing, preliminary research does not confirm the corollary expectation that group differences in achievement will diminish. Indeed, recent studies suggest that differences may be magnified when performance assessment tasks are used.
- Teachers are a fundamental key to assessment reform. As evaluation of student achievement moves away from short-answer recall of facts and algorithms, teachers will have to become skilled in using and interpreting new forms of assessment. As a result, teachers' professional development—at both the preservice and inservice levels—will become increasingly important.
- To the extent that communication is a part of mathematics, differences in communication skill must be seen as differences in mathematical power. To what extent are differences in ability to communicate to be considered legitimate differences in mathematical power?
- Current assessment frameworks, derived as they were from a measurement-based tradition largely divorced from mathematics itself, rarely conform to the principles of content, learning, and equity. Today's mathematics reveals the paramount importance of interconnections among mathematical topics and of connections between mathematics and other domains. Much assessment tradition, however, is based on an atomistic approach that hides connections both within mathematics and among mathematical and other domains.

BENEFITS FROM IMPROVED ASSESSMENT

Assessment based on the principles of content, learning, and equity are already being tested in numerous schools and jurisdictions

in the United States. It is clear already that despite obstacles and challenges, many benefits accrue even beyond the central goal of improved assessment.

> Assessments represent an unparalleled tool for communicating the goals and substance of mathematics education reform to various stakeholders.

Assessments represent an unparalleled tool for communicating the goals and substance of mathematics education reform to various stakeholders. Assessments make the goals for mathematics learning real to students, teachers, parents, policymakers, and the general public, all of whom need to understand clearly where mathematics reform will take America's children and why they should support the effort. Assessments can be enormously helpful in this re-education campaign, especially if the context and rationale for various tasks are explained in terms that the public can understand.

Improved assessment can lead to improved instruction. Assessment can play a key role in exemplifying the new types of mathematics learning students must achieve. Assessments can indicate to students not only what they should learn but also the criteria that will be used in judging their performance. For example, a classroom discussion of an assessment in which students grade some (perhaps fictional) work provides a purely instructional use of an assessment device. The goal is not to teach answers to questions that are likely to arise, but to engage students in thinking about performance expectations.

Assessment can also be a powerful tool for professional development as teachers work together to understand new expectations and synchronize their expectations and grades. Teachers are rich sources of information about their students. With training on methods of scoring new assessments, teachers can become even better judges of student performance.

LOOKING TO TOMORROW

Improved assessment is not a panacea for the problems in mathematics education. Our findings neither diminish nor reject important, time-honored measurement criteria for evaluating assessment; nor do they suggest that changes in assessment alone will bring about education reform. Clearly, they will not.

What we can say with assurance is that if old assessments remain in use, new curriculum and teaching methods will have little

impact. Moreover, if new assessments are used as inappropriately as some old assessments, little good will come of changes in assessment.

> Although The Necessary Change In Mathematics Assessment Will Be Neither Swift Nor Straightforward, We Cannot Afford To Wait Until All Questions Are Resolved.

It will take courage and vision to stay the course. As changes in curriculum and assessment begin to infiltrate the many jurisdictions of the U.S. educational system, these changes will at the outset increase the likelihood of mismatches among the key components of education: curriculum, teaching, and assessment. It is not unlikely that performance will decline initially if assessment reform is not tightly aligned with reform in curriculum and teaching.

Mathematics education is entering a period of transition in which there will be considerable exploration. Inevitably there will be both successes and failures. No one can determine in advance the full shape of the emerging assessments. Mathematics education is in this respect an experimental science, in which careful observers learn as much from failure—and from the unexpected—as from anticipated success. The necessary change will be neither swift nor straightforward. Nevertheless, we cannot afford to wait until all questions are resolved. It is time to put educational principles at the forefront of mathematics assessment.

NEW ROLES

Like any change in the fundamental way something gets done, the initial adjustment will require work. In no time, people will grow accustomed to their new roles. Soon they will wonder how they got by in the past.

... for Students

Students will learn to work in teams, to solve problems both individually and collectively, and to approach tasks with creativity and precision. In short, they will learn to think, and they will expect to demonstrate what they have learned on tests and assessment tasks. By learning more and understanding better, students will have an easier time showing their teachers what they have mastered and where they need more assistance.

... for Parents

Parents are the most influential advocates for improved mathematics assessment. It is their job to monitor the improvements their children make and then let school boards, administrators, and legislators know their feelings. Part of their challenge will be to relearn mathematics with their children. Mathematics is not just about calculations; but that is what was drilled into the heads of today's adults. It will take perseverance and patience, because parents may no longer recognize their children's homework.

... for Teachers

Teachers will have to carry the major burden, because there is simply no way to reduce the time it takes to analyze each child's unique approach to solving a mathematical problem. Teachers will have to devote more time to assessment as part of improved education. Their roles will change from the authoritarian master who transmits knowledge to an intellectual coach encouraging active exploration. In this new role teachers will exercise greater discretion in selecting, administering, and interpreting assessments.

... for Administrators

School administrators are on the front line when any educational change is introduced. They hear praise and complaints from all directions. Their role is pivotal, in communicating the ideas embodied in the educational principles of assessment—content, learning, and equity—to parents, teachers, and students. Administrators also bear responsibility for implementing policy and legislative mandates. They can help legislators involved in school reform understand the implications of proposed changes in assessment for the quality of learning and teaching. They also can ensure that the assessments they select meet appropriate criteria.

... for Test Developers

Like teachers, those who develop assessments will have more responsibilities. Although multiple-choice tests will not disappear, other types of assessments will require different approaches to development and evaluation. In particular, developers will need to know much more mathematics and will need to use more creativity in developing assessment tasks.

... for Policymakers

Legislators and other policy leaders often set the constraints within which teachers, assessment developers and school administrators work in matters of assessment. They play a very critical role in ensuring that assessment-based accountability programs do not undermine the goals of educational reform. After all, legislators decreed that students would need to meet minimum competency standards for graduation and promotion. Legislators and other policy leaders make decisions about the levels of resources that will be put into assessment programs and other aspects of reform. They need to understand the essential linkage of standards for content, for teaching, and for assessment.

STUDY GROUP ON GUIDELINES FOR MATHEMATICS ASSESSMENT

Jeremy Kilpatrick, Chair
Regents Professor of Mathematics Education
University of Georgia
Athens, GA

Janice Arceneaux
Magnet Specialist
Houston Independent School District
Houston, TX

Lloyd Bond
Professor of Educational Research
University of North Carolina
Greensboro, NC

Felix Browder
Professor of Mathematics
Rutgers University
New Brunswick, NJ

Philip C. Curtis, Jr.
Professor of Mathematics
University of California at Los Angeles
Los Angeles, CA

Jane D. Gawronski
Superintendent
Escondido Union High School District
Escondido, CA

Robert L. Linn
Professor of Education
University of Colorado
Boulder, CO and
Co-Director, Center for Research on Evaluation, Standards, and Student Testing at UCLA

Sue Ann McGraw
Mathematics Teacher
Lake Oswego High School
Lake Oswego, OR

Robert J. Mislevy
Principal Research Scientist
Educational Testing Service
Princeton, NJ

Alice Morgan-Brown
Statewide Director for Academic Champions of Excellence Program
Morgan State University
Baltimore, MD

Andrew Porter
Director, Wisconsin Center for Education Research
Professor of Education Psychology,
University of Wisconsin
Madison, WI

Marilyn Rindfuss
National Mathematics Consultant
The Psychological Corporation
San Antonio, TX

Edward Roeber
Director, Student Assessment Programs
Council of Chief State School Officers
Washington, DC

Maria Santos
Mathematics & Science Supervisor
San Francisco Unified School District
San Francisco, CA

Cathy Seeley
Director of Pre-college Programs
Charles A. Dana Center for Mathematics and Science Education
University of Texas
Austin, TX

Edward A. Silver
Senior Scientist, Learning Research & Development Center
Professor of Education
University of Pittsburgh
Pittsburgh, PA

MSEB MEMBERS

*** Executive Committee**

Hyman Bass, Chair * (I 996)
Professor of Mathematics
Columbia University
New York, NY

Mary M. Lindquist,
Vice Chair * (I 994)
Callaway Professor of Mathematics Education
Columbus College
Columbus, GA

Maria A. Berriozabal (I 993)
Former City Councilwoman
San Antonio, TX

Sylvia T. Bozeman (I 995)
Chair, Department of Mathematics
Spelman College
Atlanta, GA

Sadie C. Bragg (I 996)
Acting Dean of Academic Affairs
Borough of Manhattan Community College
The City University of New York
New York, NY

Diane J. Briars (I 995)
Director, Division of Mathematics
Pittsburgh Public Schools
Pittsburgh, PA

Patricia Chavez (I 996)
Statewide Executive Director
New Mexico Mathematics, Engineering, Science Achievement
Albuquerque, NM

Loring Coes, III * (1993)
Chair, Department of Mathematics
Rocky Hill School
East Greenwich, RI

Nancy S. Cole * (I 994)
Executive Vice President
Educational Testing Service
Princeton, NJ

Rita R. Colwell (I 993)
President, Maryland Biotechnology Institute
University of Maryland
College Park, MD

Shari Anne Wilson Coston (I 996)
Executive Director, Arkansas Education Renewal Consortium
Conway, AK

David Crippens (I 993)
Senior Vice President
Educational Enterprises
KCET-TV, PBS
Los Angeles, CA

Gilbert J. Cuevas * (I 995)
Professor of Education
University of Miami
Coral Gables, FL

Philip A. Daro * (I 993)
Executive Director
University of California
California Mathematics Project
Oakland, CA

Daniel T. Dolan (I 993)
Associate Director, PIMMS Project
Wesleyan University
Middletown, CT

Sue T. Dolezal (I 995)
Mathematics Instructor
Sentinel High School
Missoula, MT

Dale Ewen (I 994)
Vice President
Academic and Student Services
Parkland College
Champaign, IL

Joseph A. Fernandez (I 994)
Chancellor
New York City Board of Education
Brooklyn, NY

Rol Fessenden (I 996)
Director of Inventory Control
L.L. Bean, Inc.
Freeport, ME

Gary Froelich (I 994)
Secondary Schools Project Manager
Consortium for Mathematics and Its Applications
Lexington, MA

John Gage (I 994)
Director, Science Office
Sun Microsystems
Mountain View, CA

Ramesh 1. Gangolli (I 994)
Professor of Mathematics
University of Washington
Seattle, WA

Linda M. Gojak * (I 995)
Chair (K-8) Mathematics Department
Hawken School
Lyndhurst, OH

Jackie Goldberg (I 993)
City Councilwoman
Los Angeles, CA

Jacqueline E. Goodloe * (I 995)
Elementary Mathematics Resource Teacher
Burrville Elementary School
Washington, DC

Ronald L. Graham (I 995)
Adjunct Director
Information Sciences Division
AT&T Bell Laboratories
Murray Hill, NJ

Norbert Hill (I 993)
Executive Director, American Indian Science and Engineering Society
Boulder, CO

Patricia S. Henry (I 994)
President, The National PTA
Lawton, OK

Elizabeth M. Jones (I 995)
K- 12 Mathematics Consultant
Mathematics Office
Lansing School District
L a n s i n g , M I

James S. Kahn * (I 994
President and CEO
Museum of Science and Industry
Chicago, IL

Harvey B. Keynes (I 993)
Professor of Mathematics and Director of Special Projects
University of Minnesota
Minneapolis, MN

Glenda T. Lappan * (I 996)
Professor of Mathematics
Michigan State University
East Lansing , Ml

Paul G. LeMahieu (I 996)
Director, Delaware Education Research & Development Center
University of Delaware
Newark, DE

Sue Ann McGraw (I 994)
Mathematics Instructor
Lake Oswego High School
Lake Oswego, OR

Ruth R. McMullin * (I 993)
Acting President & CEO
Publishing Group
Harvard Business School
Boston, MA

Fernando Oaxaca (I 994)
President, Coronado Communications
Los Angeles, CA

Jack Stanley Price (I 996)
Co-Director, Center for Science and Mathematics Education
California State Polytechnic University
Pomona, CA

C.R. Richmond (I 996)
Director, Science Education Programs and External Relations
Oak Ridge National Laboratory
Oak Ridge, TN

William E. Spooner (I 996)
Chief Consultant, Science Division
North Carolina Department of Public Instruction
Raleigh, NC

Lee V. Stiff (I 996)
Associate Professor of Mathematics
North Carolina State University
Raleigh, NC

Robert A. Strickland (I 996)
Senior Consultant and Director for Information Technology
Harvard Business School
Boston, MA

Stephanie Sullivan (I 996)
Executive Director
Rhode Island Mathematical Sciences Education Coalition
Providence, RI

Richard A. Tapia (I 995)
Professor of Mathematical Sciences
Department of Computational and Applied Mathematics
Rice University
Houston, TX

Daniel J. Teague (I 994)
Mathematics Instructor
North Carolina School of Mathematics and Science
Durham, NC

John S. Toll, (I 993)
President, Universities Research Association, Inc.
Washington, DC, and Chancellor Emeritus and Professor of Physics
University of Maryland
College Park, MD

Philip Uri Treisman (I 993)
Professor of Mathematics
University of Texas at Austin
Austin, TX

Alvin W. Trivelpiece * (Chair 1990-l 993)
Director, Oak Ridge National Laboratory
Oak Ridge, TN

Larry DeVan Williams (I 993)
Mathematics Instructor
Eastwood Middle School
Tuscaloosa, AL

James J. Wynne (I 993)
Manager, Biology and Molecular Science
IBM T.J. Watson Research Center
Yorktown Heights, NY