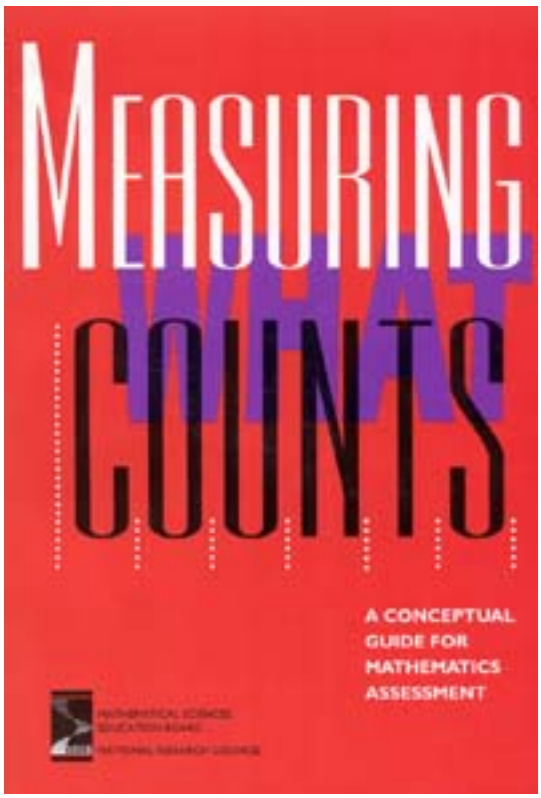
cover next page >



title :	Measuring What Counts : A Conceptual Guide for Mathematics Assessment
author :	Bass, Hyman
publisher :	National Academies Press
isbn10 | asin :	0309049814
print isbn13 :	9780309049818
ebook isbn13 :	9780585020969
language :	English
subject	Mathematics--Study and teaching--United States--Evaluation, Mathematical ability--Testing.
publication date :	1993
lcc :	QA13.N37 1993eb
ddc :	510/.71/2
subject :	Mathematics--Study and teaching--United States--Evaluation, Mathematical ability--Testing.

cover next page >

Measuring What Counts

A Conceptual Guide For Mathematics Assessment

MATHEMATICAL SCIENCES
EDUCATION BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, DC 1993

< previous page | cover-0 | next page >

< previous page | page_ii | next page >

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.

Library of Congress Catalog Card Number 93-85917
International Standard Book Number 0-309-04981-4

Copies of the 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

B122

< previous page | page_ii | next page >

Contents

< previous page | page_iii | next page >

< previous page | page_v | next page >

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.

Standards take many forms and appear under many guises. Curriculum (or content) standards tell what students should learn. Teaching (or pedagogical) standards tell how students learn and how teachers should teach. Delivery (or opportunity-to-learn) standards tell what is necessary of schools so that students can learn and teachers can teach. Assessment (or performance) standards tell what students should know and be able to do as well as how evaluators can judge levels of performance.

Since 1989 mathematics has led the national movement towards standards with *Everybody Counts* (National Research Council, 1989), *Curriculum and Evaluation Standards for School Mathematics* (National Council of Teachers of Mathematics (NCTM), 1989), and *Professional Standards for Teaching Mathematics* (NCTM, 1991). In April 1991, the Mathematical Sciences Education Board (MSEB) convened a national summit on assessment, which led to *For Good Measure* (1991), a concise statement of goals and objectives for mathematics assessment. To move the national discussion from generalities to specifics, the MSEB then published *Measuring Up* (1993), which provided prototype assessment tasks for fourth grade mathematics that illustrate in concrete terms the goals of the NCTM *Standards*.

< previous page | page_v | next page >

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 principleson content, learning, and equityfor any program of assessment that is intended to support the national educational goals. The message of *Measuring What Counts* 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.

The analyses and recommendations in this report were developed by a study group formed by the NRC in 1991 to develop conceptual guidelines that would make content the driving force in the reform of mathematics assessment and that would explore a variety of related measurement and policy issues. 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 assessment that align properly with the national standards for mathematics education.

The three principles on content, learning, and equity articulated in *Measuring What Counts* are necessary but not sufficient criteria for effective assessment. They set forth fundamental conditions that form a foundation on which to build new approaches to traditional technical testing issues such as reliability and validity. Assessment should foremost reflect important mathematics, support good instructional practice, and enhance every student's opportunity to learn.

Although these principles are rooted in both informed practice and extensive research, it is fair to say that there remain many open questions. Research shows clearly that the task of assessing mathematical learning is far more subtle than previously believed; 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. Exploration of this new world of alternative assessment will take years of work from thousands of practitioners working with mathematics education specialists and measurement experts to achieve a more effective balance of assessment in practice.

Although many of the issues raised in this report apply to all disciplines, it is mathematics education that provides the primary motivation for the study, the background of the authors, the source of examples, the domain of research, and the field of practice on which the conclusions and recommendations are based. *Measuring What Counts* seeks to address issues in assessment that are important to the discipline of mathematics and about which the expertise of mathematics educators can make a singular contribution. Content, learning, and equity emerged as fundamental principles for assessment because they are fundamental concerns of mathematics education.

One consequence of *Measuring What Counts* should be a new wave of research on assessment, on learning, and on instruction. Since much that is in this report is based on expert conjecture rather than firm evidence, it opens scores of potential areas for further research. Indeed, the changing practice of mathematics itselfthe increased focus on computer-enhanced work, on group problem solving, on modeling complex problemschallenges researchers in assessment and learning with issues rarely before considered. The resulting iteration of practice and research will provide an effective guidance system to keep assessment reform aligned with curricular objectives and principles of learning.

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.

When the stakes for improved education are so high, when our children's futures are at stake, we must ensure that assessment supports standards-based education by adhering to fundamental principles of content, learning, and equity.

Acknowledgments

Measuring What Counts was prepared by the National Research Council's Study Group on Mathematics Assessment, which was chaired by Jeremy Kilpatrick, Regents Professor of Mathematics Education at the University of Georgia. The group met over a period of two years to develop drafts of this report. Members of the study group, in addition to the chair, were Janice Arceneaux, Magnet Specialist in the Houston Independent School District; Lloyd Bond, Professor of Educational Research Methodology at the University of North Carolina-Greensboro; Felix Browder, Professor of Mathematics at Rutgers University; Philip C. Curtis, Jr., Professor Mathematics at the University of California at Los Angeles; Jane D. Gawronski, Superintendent of the Escondido Union High School District; Robert L. Linn, Professor of Education at the University of Colorado-Boulder; Sue Ann McGraw, Mathematics Teacher at Lake Oswego High School; Robert J. Mislevy, Principal Research Scientist at Educational Testing Service; Alice Morgan-Brown, Statewide Director for Academic Champions of Excellence Program at Morgan State University; Andrew Porter, Director, Wisconsin Center for Education Research at the University of Wisconsin-Madison; Marilyn Rindfuss, National Mathematics Consultant at The Psychological Corporation; Edward Roeber, Director, Student Assessment Programs at the Council of Chief State School Officers; Maria Santos, Mathematics and Science Supervisor in the San Francisco Unified School District; Cathy Seeley, Director of Pre-college Programs, Charles A. Dana Center for Mathematics and Science Education at the University of Texas-Austin; and Edward A. Silver, Professor and Senior Scientist, University of Pittsburgh.

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 this report. Particular thanks are due Jeremy Kilpatrick not only for his able leadership as chairman of the Study Group, but also for the substantial contribution he made in writing and editing the various parts of the report.

The work of the study group was enriched by three commissioned papers that appear at the end of this volume. These

papers provided useful background for deliberations of the study group and constituted valuable additions to the research literature on assessment. We are particularly grateful to their authorsLynn Hancock and Jeremy Kilpatrick, Stephen B. Dunbar and Elizabeth A. Witt, and Diana C. Pullinfor their contributions to this endeavor. We wish also to thank Linda Dager Wilson and Lynn Hancock for supplying other background information of value to this study.

Like all reports of the National Research Council, *Measuring What Counts* has been extensively reviewedfirst 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. Special thanks are due Nancy Cole, Chair of the MSEB Committee on Policy Studies, for providing consistent and wise counsel as the report worked its way through various drafts. *Measuring What Counts* is much stronger as a result of the input and advice of these outside reviewers.

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.

Staff work in support of the study group was provided by the MSEB Office of Policy Studies directed by Linda Peller Rosen, Associate Executive Director of the MSEB. Edward T. Esty and Patricia A. Butler deserve special thanks for managing this complex project with unfailing perseverance and tireless energy. Thanks are due also to Anuradha Sapru Kohls, Ramona Robertson, and Altoria Bell, who undertook diverse tasks without which this report would never have been completed. The staff of the National Academy Press deserves special mention for their efforts on our behalf.

Hyman Bass, Chairman
Mathematical Sciences Education Board

< previous page | page_1 | next page >

Measuring
What
Counts

Executive
Summary

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

"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:

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 American today reflect any of these vital principles. For educational

< previous page | page_1 | next page >

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

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.

Assessment is the guidance system of education just as standards are the guidance system of reform.

Assessment in Today's World

Assessment is the guidance system of education just as standards are the guidance system of reform. It helps teachers and parents determine what students know and what they need to learn. Assessment 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.

< previous page | page_3 | next page >

At its root, assessment is a communication process that tells students, teachers, parents, and policymakers some thingsbut not everythingabout 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

< previous page | page_3 | next page >

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.

> For decades, educational assessment in the United States 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 criteriaprimarily reliabilityalso 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 validitywith its attention to fidelity between assessments, high-quality curriculum and instruction, and consequencesis 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.

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

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

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.

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,

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

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 constructed 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.

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.

> Assessment should enhance mathematics learning and support good instructional practice

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.

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

< previous page | page_7 | next page >

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.

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

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.

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.

< previous page | page_7 | next page >

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 oftstated 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

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

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.

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 assessment 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,

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

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.

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