Taking Stock: What Have We Learned About Making Education Standards Internationally Competitive? Summary of a Workshop

ISBN
978-0-309-05944-2

48 pages
8.5 x 11
PAPERBACK (1997)

Alexandra Beatty, Editor; National Research Council

Taking Stock

What Have We Learned About Making Education Standards Internationally Competitive?

Summary of a Workshop

Alexandra Beatty, *Editor*

Board on International Comparative Studies in Education
Commission on Behavioral and Social Sciences and Education
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1997

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.

This report was supported by Cooperative Grant No. OSR-9355774 from the National Science Foundation, which includes funds from the National Center for Education Statistics of the U.S. Department of Education. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for this project.

International Standard Book Number 0-309-05944-5

Additional copies of this report are available from:
National Academy Press
2101 Constitution Avenue, N.W.
Lock Box 285
Washington, D.C. 20055
Call 800-624-6242 or 202-334-3313 (in the Washington Metropolitan Area).

This report is also available on line at **http://www.nap.edu**

Printed in the United States of America

BOARD ON INTERNATIONAL COMPARATIVE STUDIES IN EDUCATION

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of 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. William A. Wulf are chairman and vice chairman, respectively, of the National Research Council.

Contents

Taking Stock

What Have We Learned About Making Education Standards Internationally Competitive?

INTRODUCTION

The past 15 years have seen a growing concern in the United States over the quality of the nation's educational system. *A Nation at Risk*, released in 1983, warned of "a rising tide of mediocrity that threatens our very future as a nation and as a people" (National Commission on Excellence in Education, 1983:5). In this report and in the discussion that it inspired, the assumption that the United States seemed to be doing so poorly at educating its children was based in part on comparison with other nations. International studies of educational achievement have repeatedly indicated that American students do not excel in comparison to their counterparts in other countries (Elley, 1992; International Association for the Evaluation of Educational Achievement, 1996a and 1996b; Lapointe, 1992). At the 1989 education summit, the President of the United States and most of the nation's governors for the first time articulated a set of education goals that included an explicit call for international competition, that American students should be "first in the world in mathematics and science achievement by the year 2000" (National Education Goals Panel, 1995:12). In response, vigorous efforts have been initiated at the national level and within many states to define new content and performance standards in these and other subjects. In providing federal support for these efforts to define standards, the U.S. Congress called for standards that would be "internationally competitive and among the best in the world." These policy directives did not, however, settle the question of how any set of standards might be made internationally competitive.

The drive for education standards that can help American students to compete with the top students around the world grows out of a shift in the United States' view of its place in the world. Other calls for education reform—such as those following the Soviet launch of Sputnik, for example—have centered around the need to maintain military preparedness. With the end of the Cold War, however, improved schooling has come to be seen as a way of fending off external economic threats. Innovations in technology and communications have drastically changed the nature of work and the economic relationships among nations. It is in this context that human capital, and hence the quality of education, have moved high on the American political agenda.

Businesses have developed increasingly sophisticated ways of comparing their processes and products to others around the world. One of those, the concept of benchmarking—measuring one's own practices against those of others—has had a substantial impact on education policy makers. Benchmarking academic achievement and setting high standards for students have come to be viewed as prime elements in the reform of U.S. schools. Many people believe that holding U.S. students to internationally competitive standards of performance can ensure that they will grow into workers who are capable of mastering new technologies and who can help the nation

prosper. The movement to define and impose such internationally competitive standards has been welcomed by many constituencies concerned about the failings of American schools.

Recent national polls have found wide support for the idea that students and schools should be held to high academic standards even if that means that some students will fail to meet them, and even if serious consequences (such as not graduating from high school) are attached to failure (LeMahieu and Bickel, 1996:9-10). Vast numbers of scholars, educators, and interested citizens have participated in the process of developing national and state standards for most of the subjects taught in kindergarten through high school. The national standards are voluntary—states are invited to adopt or modify them as they choose, and many have done so. To date the vast majority of the 50 states have adopted their own specific academic standards or are working to establish them, and a number of studies and reports have documented these efforts (*Education Week*, 1997; American Federation of Teachers, 1995b). The President recently called for voluntary national tests in reading at grade four and mathematics at grade eight, to be based on the frameworks that guide the National Assessment of Educational Progress. The mathematics test is to be linked to the Third International Mathematics and Science Study so that the performance of American students can be benchmarked to that of students around the world (see box on p. 3).

These circumstances might be interpreted as evidence that the education community in the United States has a fairly robust notion of what internationally competitive standards for students and schools might be, but such is not the case. Many of the national standards documents have provoked continuing vigorous discussion and disagreement. In some subjects competing standards have been developed by different groups. Numerous constituencies have criticized the standards that have been developed for a dizzying variety of reasons. Indeed, the first version of the standards for U.S. history was formally repudiated by the U.S. Senate. Moreover, the content and performance standards that have been adopted by the states vary significantly in purpose, form, content, and rigor.

> . . . the notion of international competitiveness is a dynamic sort of notion. It's not a static thing; it's not criterion referenced. Countries can get better.
>
> *Andrew Porter*

State education communities are struggling to define and implement standards that are appropriately challenging and to address a host of related issues. One issue is that what people mean when they talk about standards can vary significantly. Others include: How might formal academic standards fit into an education system? Just how high should performance targets be? How can one tell how high they are? How have others set standards, and what are they? By what means should the United States—or a particular state—compare itself to others? To whom should the comparison be made?

Both local communities and scholars have been struggling with these kinds of questions for some time, and many valuable insights

International Comparisons—
The Third International Mathematics and Science Study (TIMSS)

Based on presentation by Mary Lindquist

The TIMSS study, which provides mathematics and science achievement data for students at three age levels from around the world, is the largest and most complex international comparative study ever undertaken. The study was designed to produce not only data on student achievement, but also a variety of information about the contexts in which student learning takes place. Background questionnaires completed by students, teachers, and administrators were designed to produce data about educational practices, attitudes, school characteristics, and other factors. These data can be correlated with achievement data so that possible relationships can be identified. A monumental study of curricula and textbooks used around the world was also part of TIMSS. In addition, two studies sponsored by the United States explored contextual factors in the United States, Germany, and Japan in greater detail. One of these was a videotape study of classroom practice; the other was a set of ethnographic case studies that explored some of the attitudes and experiences of students at the middle school level.*

While the analysis of the results of TIMSS is still under way, the results that have been released have played an important role in the ongoing discussion of internationally competitive standards. TIMSS, Lindquist noted, will make it possible not only to identify countries with high-achieving students, but also to use information from the teacher questionnaires, the curriculum study, and the two qualitative studies to develop a picture of educational practice in those countries. She reminded the group that one significant outcome of the Second International Mathematics Study (SIMS) was its influence on the widely respected National Council of Teachers of Mathematics standards. SIMS suggested that clear standards would be beneficial, and this was one impetus for the development of the new mathematics standards. It also provided the basis for many of the specific recommendations that went into those standards.

For Lindquist the value of TIMSS lies less in the international benchmarks it provides than in the opportunity it provides to generate more questions. Because TIMSS was based on a framework negotiated among approximately 40 countries, it is not particularly well aligned to the curriculum of any one of them. Consequently, the student achievement benchmarks it provides are somewhat vague. Lindquist noted further that the achievement rankings alone offer no guidance on ways to improve teacher practice and student learning. Her hope is that the contextual information about teachers' lives, instructional practices, curricula, and other issues will provide clues for reform that leads not just to further research, but also to focused reflection and improvement among teachers.

*For more information about TIMSS and the reports that are available, see the following sites on the World Wide Web: http://www.ed.gov/NCES/index.html, http://ustimss.msu.edu./, http://wwwcsteep.bc.edu/timss, and http://uttou2.to.utwente.nl/, or write or call the International Study Center: CSTEEP, Campion Hall 323, Boston College, Chestnut Hill, MA 02167, (617) 552-4526, or the National Center for Education Statistics, TIMSS Project, 555 New Jersey Ave. NW, Suite 02A, Washington, DC 20208, (202) 219-1333.

have emerged. The Board on International Comparative Studies of the National Research Council sponsored a workshop on November 6, 1996, to take stock of what has been learned about making education standards internationally competitive and to examine why the criterion of international competitiveness has been so difficult to articulate and to apply to education standards. The purpose of the workshop was to bring together a variety of people, as both presenters and observers, who have thought creatively about bringing the idea of international competitiveness to school reform to share their perspectives and to discuss ways of keeping the process moving forward. (See Appendix A for the list of participants.) The agenda of the 1-day workshop was structured around presentations that focused on different perspectives and on the emerging body of empirical evidence on internationally competitive standards. (See Appendix B for the workshop agenda.)

The board commissioned three papers for the workshop to aid in sorting out some of the conceptual issues, and these were the focus of the first half of the workshop. One paper (by Paul LeMahieu and William Bickel) explored what people mean when they call for internationally competitive standards. A second (by Alan Ruby) provided an international perspective by reviewing key aspects of the Australian experience with standards. The third (by Robert Floden) tackled the question of implementing standards by reviewing recent work on teacher practice with regard to standards. (See Appendix C for a list of the papers presented at the workshop.)

In the second half of the workshop two panels explored perspectives on particular international benchmarking efforts. One focused on the experiences of three states that have incorporated an international context into their standard-setting efforts: Colorado, Illinois, and Maryland. The other provided a forum for discussion of the Third International Mathematics and Science Study, the international benchmarks developed by the New Standards Project, and the business community's approach to benchmarking. The panels were designed to permit considerable discussion by participants and presenters alike, some of which has been woven into the main sections of this summary. In addition, boxes interspersed through the text provide synopses of the information conveyed by each of the panelists.

Approximately 80 people, including representatives from national organizations with a specific interest in education standards, university researchers, and education policy makers attended the workshop. Their participation was a key component of its success, in particular because thoughtful questions asked throughout the day helped to link some of the major themes. The purpose of this document is to provide a summary of the issues that were explored at the workshop.

WHAT ARE INTERNATIONALLY COMPETITIVE EDUCATION STANDARDS?

The evidence that the concept of high standards for education currently enjoys wide support in the United States is unambiguous. Between 80 and 90 percent of respondents to national opinion surveys support the idea that students and schools should be expected to meet specific standards in basic subjects. The support drops only to 70 percent when respondents are asked whether they support standards that are connected to serious consequences for failure. The support is strong across demographic groups and among both parents and nonparents. Other surveys have documented that the quality of public education is a major national concern and that it was on the minds of the voters during the 1996 presidential election (LeMahieu and Bickel, 1996:9-19). Education standards have also been widely discussed in the mainstream press (Toch et al., 1996; *U.S. News*, 1996; Gagnon, 1995).

LeMahieu and Bickel provided a valuable grounding for the workshop by exploring both some of the details of the public opinion data and some of the nuances of the views and expectations about internationally competitive standards held by a group of key leaders. They collected opinion data from a variety of sources and conducted a small-scale interview protocol with a few political, academic, and education policy leaders. Their goal was to illuminate the discussion by clarifying what people with various perspectives mean when they speak of internationally competitive standards.

One of their major findings is that the urge for standards that are internationally competitive seems to grow primarily out of economic anxiety—a sense that U.S. students are not being adequately prepared to compete in a global market. Pervasive fears during the 1980s about economic decline were clearly linked to the sense that U.S. schools were failing; people became accustomed to looking to the nation's chief economic rivals not just for insights into economic productivity, but also for examples of success in education.

Though the survey and interview data do not probe deeply into the public's understanding of the content of standards, one interesting concern did emerge. Numerous surveys showed that what the public really wants is "basics first" and that they were likely to be disappointed with standards that did not seem to reflect this priority. As LeMahieu and Bickel noted, this is "a point of potential conflict between reformers and the public" because current thinking about classroom practice has moved away from the language of "basics first" to an approach in which basic skills are integrated with other skills and material (LeMahieu and Bickel, 1996:13-14).[1]

The focus on international comparisons reflects a shift in thinking

[1]A study done by the Public Agenda Foundation that explored public opinion on this point suggests that there may not be a conflict (see Johnson and Immerwahr, 1994).

about what can be expected from schools. As Alan Ruby noted in his paper, the emphasis in the United States during the first half of this century was on the uniformity of public schools, which were expected to transcend class and provide opportunity for every child. In the 1960s an influential body of research by James Coleman and others held that socioeconomic factors accounted for most of the variation in student achievement (see Coleman et al., 1966). A response to that work emerged in the 1980s, known as the "effective schools movement." (Steller, 1988; Lezotte, 1986; Rossmiller and Holcomb, 1993) This newer paradigm holds that differences among schools do make a difference and that what takes place in effective schools can be studied and profitably imitated. Clearly those in favor of standards are in sympathy with the notion of effective schools, and the leaders consulted by LeMahieu and Bickel are no exception.

While the group whose views were collected by Bickel and LeMahieu generally shared a belief that standards can be an important vehicle for improvement, there were a few differences within the group. Perhaps the most striking difference in perspective existed between representatives of the business community and those from academic and education communities. The business leaders tended to see the actual development and implementation of standards as a largely technical matter and to see their value as that of a guidepost, a source of motivation. Those from the education and policy communities, in contrast, were far more aware that both developing and using standards are political and social processes, in which decisions among alternative methods of setting and enforcing standards have profound implications.

A related difference lay in what standards-based reform brings to the separate worlds of business and education. For business, which had traditionally focused most of its attention on such outcomes as the quality of the product and its profitability, the application of standards encouraged a focus on process—an examination of alternate methods of production in order to find the most efficient and successful one. Standards in education, however, accomplish the reverse—attempts to improve learning have long focused almost exclusively on matters of process, such as teaching practices, funding, and school structuring. With discussion of standards came a focus on outcomes, a reminder that the primary goal is student learning. LeMahieu and Bickel explained that many of the leaders they consulted believe that specific expectations for performance—beyond statements of goals—are critical to the success of standards. The authors also cautioned that the standards movement brings with it a risk that some will focus exclusively on measurable outcomes and ignore the importance of process. In short, LeMahieu and Bickel concluded that successful standards would need to focus on both process and outcomes.

> Without a . . . statement requiring some agreed-upon performance, content standards alone are . . . likely to give rise to more attention to process: more courses, more seat time, etc.
>
> Respondent in LeMahieu and Bickel

Another distinction that emerged both in Bickel and LeMahieu's work and in discussion at the workshop was between "world-class" and "internationally competitive" standards. When some people speak of a world-class standard, they mean a level reached only by a few, a model that serves primarily as an inspiration. A different notion of a standard is what many now signify with the term "internationally competitive standard"—a high standard that all students can hope to achieve and should be urged to achieve. Implicit in both of these models, of course, is the notion that the standard is defined through benchmarking to high-achieving students in other countries. Also implicit in both models is the belief that the standard itself should be continually reevaluated and raised to stimulate progress in student achievement over time. Most of the business leaders consulted thought of standards in the world-class sense, as goals for a few; in contrast, the educators and policy makers described a firm commitment to holding all students to high standards.

This distinction is significant in part because any group attempting to establish standards must choose one or the other model before specifying the requirements for meeting them. But perhaps more important is the philosophical issue lurking behind the distinction. While representatives from the education and policy communities speak of high standards for all, comparisons with other countries have revealed low expectations for U.S. students. A variety of studies have suggested that U.S. standards of performance are lower than those in many high-achieving countries, that average U.S. student achievement in many subjects is below that of students in many other nations, and that the conditions in which public education takes place in the United States—and, consequently, the opportunity to learn—are in many respects inadequate in comparison to those in other places. What has also emerged from such studies is that there are few systemic incentives for students in the United States to meet high standards or penalties for those who do not do so (Resnick et al., 1995:439; Ravitch, 1995; American Federation of Teachers, 1995b; American Academy of Arts and Sciences, 1995). As many have observed, it is not the case that U.S. children are any less capable of learning than are children in other countries (American Federation of Teachers, 1995a). Case studies, anecdotal evidence, and common sense all suggest that students will achieve at higher levels when they are expected to do so and given the opportunity. Consensus on this point, however, does not fully resolve the tension in pursuing standards that are both extremely demanding on the one hand, and realistic and fair to all students on the other. This tension was explored at a number of points during the day.

When LeMahieu and Bickel asked business and education leaders why they were so confident that standards would improve student learning, they generally did not cite data. As Paul LeMahieu put it: "the belief was much stronger than the evidence was deep." Nevertheless, both the public opinion data and the leaders' comments show

Benchmarking in Business

Based on presentation by Anne Miller

The education community borrowed the concept of benchmarking from the world of business. The term, as LeMahieu and Bickel noted in their paper, comes from an engineering process for measuring and testing a device by putting it on a test bench. For business it has meant a willingness to look closely at the methods and processes others use and to allow others to look at theirs.

It has only been in the past few decades that American businesses have espoused the concept of benchmarking. In the face of crippling competition from Japan in the 1970s and 1980s, the U.S. auto industry recognized that, in order to compete, it had to benchmark—to identify and meet the specific standards of design and production set by the Japanese. The automakers' success using the benchmark process served as a model for other businesses, and it is cited as the impetus for a new quality movement in which many businesses have chosen to adopt new quality standards.

Eastman Kodak has adopted benchmarking enthusiastically, and the company currently benchmarks 44 key processes—though Miller noted that they began with overkill, measuring everything down to the size of the muffins in their cafeterias. The benchmarking is done to companies of all sorts, both in and out of the United States, not only ones that are similar to Kodak. She noted, for example, that many companies look to L.L. Bean for benchmarks in the distribution process, since its is widely recognized as one of the most efficient in the world.

Miller's description of Kodak's "more than casual" interest in seeing the education community in the United States adopt tougher standards in order to better prepare its future workers is very much in tune with the views of the business leaders presented by Bickel and LeMahieu. She noted that Kodak made a commitment at the National Education Summit to begin looking at the high school transcripts of potential employees, although it has not traditionally done so "because . . . they haven't given us the information that's important or relevant for the jobs for which these people are being hired." Kodak administers its own mathematics and reading tests to applicants and will continue to do so despite the desire to send a message about the importance of a high school diploma.

For Miller the positive impact benchmarking has had on American business is clear. She acknowledged that the education community confronts different problems in attempting to benchmark—noting that film canisters are easier to measure than is student learning. Her message, however, was that education faces a need to learn from the experiences of others just as urgent as the one that has faced business.

I contend that standards would have the net effect of reducing inequities—a system without expectations invariably accepts inequities.

Respondent in LeMahieu and Bickel

that expectations for standards are quite high. Poll respondents cited as reasons for their desire for higher education standards a "pervasive moral decay in American culture," a "widespread economic anxiety," and the view that "leadership is out of touch with the concerns of average Americans," although, of course, these are not problems that are addressed in any way by academic standards. Much of the discussion at the workshop strongly suggested that curing these ills is an unrealistic and inappropriate goal for education standards.

STANDARD-SETTING AS A POLITICAL PROCESS

Every educational system has standards for content, performance, and the opportunity to learn, whether they are formally developed and intentional or implicit. A system that is not providing all of its students with opportunities that would permit equivalent accomplishment is the result of social and political choices about expectations for population groups, even though these decisions may not have been deliberate. Alan Ruby raised this issue by describing a debate that occurred in Australia when that country decided to increase its target graduation rates for secondary students. The debate was over what percentage of students ought to be expected to graduate, with one faction arguing that a target over 70 or 80 percent was unrealistic, and others arguing that to declare that any percentage was not expected to graduate was an injustice to some students. (The parties settled on a goal of 95 percent.)

Ruby's response to this tension between high standards for all and high standards for the few was that there is no one correct balance. He maintained that a specific purpose for any set of standards must be explicitly worked out in a public political forum and that this purpose will determine how high the bar should be set in that context. In the Australian example, the purpose of the standard was to help the nation produce workers who could compete for a particular category of jobs; hence, the percentage of expected nongraduates could not exceed the projected percentage of jobs for which high school graduation would not be necessary. A similar calculation of the specific purpose of having a set of standards, and public discussion of the implications of the established purpose, is a necessary step, in Ruby's view, for any community that wants standards that are both valid and fair. A part of this calculation is to plan not just for groups expected to meet or exceed the standard (the intended outcome for those who pass might be grade promotion, employment, or college entrance, for example), but also to plan effectively for those who may not. A plan for the latter group might be based on a determination that the failing group will be brought to a passing level, and allocate

resources for remedial work or other supports. Alternatively, an educational pathway for students not meeting a certain level of achievement by a certain point might be developed, and standards developed for it as well.

Addressing the equity concerns that many have raised about standards, Ruby said in his presentation:

> I would like to get right down to what we are saying. Are we saying that there are children in our community that because of their race, their gender, their social class will learn less, should learn less, or are likely to learn less than others? The problem is not in the individual, the problem is in the system of delivery, and that's where our standards are. Then it becomes an argument about resources and affirmative action . . . .

Though the workshop did not explicitly address opportunity-to-learn standards, Ruby and others stressed that it does not make sense to think about performance or content standards by themselves. If the same set of standards applied in different contexts would be likely to produce different results, then the problem is not to find the "best" or the "highest" standards, but to find those that will enable the students for whom they are intended to reach the highest level of achievement they can. An understanding of the content and performance standards that are in place for high-achieving students around the world should surely inform a local search for standards, but simply imitating standards that have worked in one place would clearly not be an effective strategy because context is so important. Ruby joked that the term "internationally competitive" might be understood to refer to standards that are in competition with one another as documents. His point was that an exclusive focus on comparing the form and content of standards could obscure the need to consider the entire network of factors that affect student achievement.

Despite his warnings, Ruby did identify several strategies that he believes characterize effective standard-setting, which were amplified by the comments of other participants.

Design standards that are context-specific. Effective standards are linked to well-defined objectives and priorities that the community—whether a nation, state, or district—that is adopting the standards has chosen.

Synthesize goals and performance. To be effective, standards must address the gap between what is expected and what students actually do. This means that content standards and resource standards are as important as, and must fit with, outcome standards.

Focus on the educational core. Standards should be related to central and enduring parts of the education system. Although some people may criticize this approach as one that will encourage teaching to the test (assuming the standards are linked to assessments), others would argue that if the test reflects thoughtful priorities, preparing for it is not a bad thing.

Link standards closely to schools. Standards should focus on things that can be affected by changes within the school system: setting standards for aspects of students' experiences that are outside the influence of education policy is pointless.

Make the standards clear and transparent. Standards must be easily and widely understood if they are to be widely espoused and effectively implemented.

Adopt a systemic approach. Standards can only work if they are reinforced by other elements of the education system, and if they, in turn, reinforce the system's goals.

As several participants pointed out in the course of the workshop, the United States seems to be stuck at a plateau with regard to standards-based reform. Although much work has been done in developing many sets of standards, and in building consensus that standards are desirable, the nation has so far not committed itself to any one vision of standards. The reasons for this situation primarily lie in the U.S. federal political system, under which the funding, management, and policy making for primary and secondary schools are largely the responsibility of state and local officials. In this context, political resistance to any suggestion of a national curriculum is high, although such a curriculum is not a necessary element of a standards-based system. Moreover, Ruby argued, there is an ambiguity about who is responsible to whom for success or failure in education, and this ambiguity could and should be resolved through the process of developing standards. His recommendation is that the United States needs to have a discussion of its economic and social goals in order to clarify the decisions that need to made about education standards. "There seems to be a lack of political agreement about the purpose of schooling" in the United States, he explained. Clearly, such a discussion could be illuminated by an understanding of the goals for education, and the means for achieving them, that have emerged in other countries.

> We seem to be stuck at a point right now where we're trying to develop a perfect standard that will drive the system to all of these wonderful goals . . . We are sort of standing at the precipice and can't quite jump.
>
> *Maggie McNeely*

Ruby noted that there is no regulatory framework in the U.S. education system, though standards in other countries are generally attached to regulatory or compliance mechanisms. This issue was raised a number of times during the day, and the experiences of some of the states that have moved toward incentives were cited. In Maryland, for example, a system of assessments has recently been put in place for grades 3, 5, and 8; the state's goal is to make passing the grade twelve assessment a requirement for graduation by 2004 (see the box on page 12). Ruby remarked that national standards for the content to be mastered by licensed teachers would be fundamental to the implementation of academic standards with teeth for students. No disagreement with the idea that incentives are crucial to making stan-

The State Experience—Maryland
Based on presentation by Robert Rice

The state of Maryland has found an idiosyncratic way of obtaining international comparisons: it administered its state assessment to students in both Baden-Wurttembourg, Germany, and Taiwan, Republic of China. While these two international collaborators were identified more by coincidence than by design, the testing has provided useful input into an ongoing debate in Maryland about whether the state's standards are sufficiently challenging and whether the state assessment program is adequately measuring student progress toward meeting the standards. Specifically, the international effort was designed to identify both levels of achievement and means of assessing it that are respected across national boundaries. It was also intended to provide opportunities for cross-national collaborations in test design, development, and analysis that would have lasting benefits for Maryland.

The Maryland curriculum is not particularly well aligned with those of either Baden-Wurttembourg or Taiwan, and the state faced a number of stiff technical challenges in the course of the project. Not surprisingly, translation was a major issue. However, the joint assessments have yielded some useful insights into the age-appropriateness of certain kinds of items and material, as well as on the relative achievement of Maryland students. In general, results showed that Maryland students lag somewhat behind both the Germans and the Taiwanese, particularly in mathematics.

dards effective was voiced at the workshop, though participants did not specifically address the implications of this point for assessment.

One participant pursued one of the political issues inherent in setting standards. He noted that a seemingly unbridgeable gap exists between those who favor a systemic approach to standards, in which the various components of the system are aligned with a centrally developed set of standards, and those who view such a system as a fundamental threat to the autonomy of teachers. Ruby, however, dismissed the notion that this gap is unbridgeable. He described the system currently in place in New South Wales, Australia, in which assessments that are closely aligned to standards and are used as high-stakes exit exams, are devised by teams of disciplinary experts from universities and exemplary teachers. In this system, teachers are not only given considerable influence on the standards and assessments, but are also given an excellent opportunity to become thoroughly familiar with the details and spirit of the standards.

Noting that the United States had so far shied away from mandatory high-stakes assessments at the national level (though many exist at the state level), Ruby pointed out that in the absence of a formal credential such as passage of a test, those in the business of choosing

among people for various purposes will use a proxy; in most cases the proxy is far more likely to disadvantage particular groups in an unfair way than is standards-based assessment that is open to public scrutiny. He suggested that high school diplomas offer no firm basis for selection and that potential employers might consciously or unconsciously use race, gender, zip code, or some other means of selecting from among a large group.

With regard to employment, Ruby also called for the establishment of standards for adult learning. Given that a primary reason for establishing challenging standards for students is to prepare them for employment, he pointed out that life-long learning will be crucial if these workers are to maintain their competitive edge. Consequently, clear thinking about continuing education, and standards by which progress can be regulated and monitored, would be a valuable component of the standards movement in the United States.

Ruby remarked that "one of the great fascinations about the United States is that [whenever] anyone gets a good idea, someone has got another one." The debates, discussions, and research about education reform in this country seem almost self-perpetuating, he explained, but, "the actual delivery of a policy solution just doesn't happen." His view was that a balance must be struck. Once the large policy questions about the purpose of imposing standards and the role they will play are addressed, it makes sense to move forward without reaching a state of scientific certainty about which, precisely, would be the best way to proceed. While he cautioned against "overselling" the value of standards, he acknowledged that much can be learned from the standards that have been used in other education systems.

> . . . the notion of internationally competitive standards implies a belief that there is just one set, that all nations across the world should have this same set of standards, that it's a universal set. Now I think that that's simply not so.
>
> *Alan Ruby*

The key for the United States is to adapt its own standards to its own circumstances and its own goals. In response to a question, Ruby explained that Australia was able to make the leap of committing to standards because of a combination of factors that sound strikingly familiar to American ears: economic pressure to prepare workers to compete in a service economy, pressure from the business community to increase the competence of secondary school graduates, and a political climate amenable to reform. Australia adopted content and performance standards that are context-specific, and so, he argued, should the United States.

A key insight that emerged from Ruby's discussion was a possible answer to the question of why internationally competitive standards have seemed such an elusive goal in the United States. His message was that while educators and others in the United States clearly need to learn from the experiences of others, it is a mistake to think of internationally competitive standards as a platonic ideal that, once realized, will transform American schools. Ruby described internationally competitive standards as "the Holy Grail of education

reform." Robert Floden, who looked at the role standards actually play in American mathematics classrooms, reinforced this point in the next discussion.

IMPLEMENTING STANDARDS

Permeating the discussion throughout the workshop was a sense of the importance of teachers. Several participants pointed out that teachers are cast in a variety of sometimes conflicting roles when standards are discussed. Critics have cast them as both targets who are accountable for students' failure to achieve high standards of performance and obstacles to the kinds of reforms that high standards will require. Others have described teachers as coaches who can help students achieve goals that are established outside the school walls. Many have commented on the vital importance of including teachers in the process of developing standards, not only so that the standards will benefit from teachers' wisdom, but also so that teachers will truly understand and support the standards. But most important is the recognition that teachers are vital to the successful implementation of any education standards.

Robert Floden examined that point in detail. He argued that regardless of the quality and content of any standards document that is adopted, the standards actually in use are those that exist in teachers' minds (Floden, 1996). Floden reviewed a variety of research from the past 20 years that has explored the relationship between the practices of mathematics teachers and the materials and goals that are meant to guide them. Researchers in one of the earliest studies he discussed, conducted in the 1970s, identified four basic decisions that teachers make about what they will teach. Their point was that these decisions effectively constitute the implementation of any set of content and performance standards for students:

What topics will I include?
How much time will I spend on each topic?
To whom will I teach each of these topics?
What level of mastery of each topic will I expect from my students?

These four decisions are influenced by the content of standards documents, assessments, textbooks, parental expectations, a teacher's own background and experience, the students' prior instructional experiences, as well as other factors, but they are decisions that teachers must make, whether or not they receive clear guidance on how to make them. Floden's point was that standards alone, even standards that are closely aligned with assessments and other aspects of the education system, cannot do the job of reforming educational practice. For anyone involved in attempting to raise content and performance standards, an understanding of the standards in use, and of how their use is determined, is crucial.

In general, Floden reported, mathematics teachers were not teaching to formally established standards 20 years ago, and, by and large, they are not doing so today. He noted that the impetus behind one study of this issue conducted in the 1970s was a concern that children were being held to *de facto* standards established by assessment companies and textbook manufacturers. The hypothesis this study explored was that the content of elementary mathematics classrooms was being routinized by the existence of norms established without public scrutiny or input. The study found that the reverse was true: the content of tests and texts was by no means consistent. Teachers were receiving conflicting messages from various sources about what content ought to be covered or emphasized, and they were making their own judgments. The result was that even in a subject that many lay people would consider straightforward—elementary school mathematics—there were dramatic differences in the topics covered, the time spent on topics, and even in the total time spent on mathematics instruction.

More current findings look extremely similar, Floden reported. The National Council for Teachers of Mathematics (NCTM) has developed coherent standards in mathematics that have been widely espoused. The movement for systemic reform has encouraged people to focus on the alignment of assessments with the NCTM standards and the use of textbooks that are similarly aligned. Nevertheless, even recent studies of classroom practice show that teacher behavior is frequently not in line with the goals of the NCTM standards, though it is moving in that direction. Floden was quick to point out that the explanation for the gap does not lie in a resistance on the part of teachers to change. He cited a variety of data indicating that teachers are willing to follow guidance if the guidance is consistent, and that they are very open to incorporating new content and strategies into their practice.

One problem, however, is that teachers are generally not inclined to discard anything from their existing curricula in order to make room for the new. The result—and this conclusion was generally reinforced by the findings of the curriculum study that was part of The Third International Mathematics and Science Study (TIMSS) (see box on p. 3)—is that teachers have too many topics and themes crammed into their schedules. Choices still need to be made about allocating time to various topics and about reorganizing instruction, and standards documents generally provide very little guidance for these choices. A criticism of many standards documents has been that they are long and detailed enough to serve as textbooks. To have an impact on teachers' choices, they may need to incorporate more decisions about priorities.[2]

[2]The Council for Basic Education has addressed this issue by undertaking to organize the standards for each of the major disciplines in a consistent manner and to cross-reference them. This shorter version of the standards will be published in a single volume.

> Many of the teachers whom we observed did change their practice in response to the new policy, but the frame for those changes was the pedagogy that had been pressed by the older policies. New wine was poured, but only into old bottles.
>
> *Robert Floden*

A related problem is that the NCTM standards are often not completely understood either by the teachers who are attempting to incorporate them or by the administrators who are advocating them. As LeMahieu and Bickel found in their study, educators tend to show relatively less support than do other groups for implementing new standards, not because they question the value of standards, but because they believe high standards are already in place. Ninety-five percent of mathematics teachers surveyed in TIMSS reported that they were familiar with the NCTM standards. In general, the teachers surveyed believed that their practice was in line with those standards, but these same teachers' responses to specific questions about their practice do not support their beliefs (National Center of Education Statistics, 1996:4-5). The disjuncture suggested by these reports, and by Floden's observations may be explained by the fact that the NCTM standards are not simply a new prioritization of existing mathematics topics. Rather, they reflect a fairly profound rethinking of mathematics education. Consequently, a teacher who has not had the opportunity to study and truly digest them may not recognize the ways in which they are at odds with many traditional methods. An often-cited example is that of problem solving. The NCTM standards ask teachers to incorporate the view of mathematics as a tool for real-world problem solving throughout their practice. Floden has found, however, that many teachers have not absorbed what is conceptually new in this familiar-sounding language, and they believe they can meet that standard by assigning more story problems.

> . . . you would say, we want you now to teach for understanding, and they said, "you've got to be kidding, like I was teaching for misunderstanding before?"
>
> *Robert Floden*

Teachers' decisions about what content to emphasize are also affected by both their perceptions about their students' backgrounds and abilities and their own knowledge of the subject, Floden reported. Naturally enough, teachers were generally found to be less comfortable teaching material they themselves had not studied and were consequently more likely to omit or downplay it. Teachers clearly need the opportunity to learn new content before they can teach it. To the extent that the NCTM standards call for a true rethinking of mathematics instruction, this may be a particularly urgent need right now.

Although Floden's conclusion is that standards by themselves will do little to improve student performance, he had several specific suggestions about ways to enhance their effectiveness. Primarily, he stressed that rather than simply creating financial or other practical incentives for teachers to comply with higher standards, education leaders should find ways to engage them with the content of the standards. He stressed that for standards to be effective, they must be internalized by teachers. In particular, teachers need the opportunity—especially time in their schedules—to really

learn what the standards are about. In many cases, as noted above, they need the opportunity to learn new content, and as one administrator pointed out to Floden, this generally requires more than a weekend workshop. Teachers also need opportunities to interact with other professionals who know the standards well, have used them, and are in a position to share ideas about implementing them.

For all of these things to occur, public support for professional development is crucial. One participant asked whether Floden considered it a good idea to require teachers to pass assessments tied to new, higher standards in order to remain in their jobs. Floden acknowledged the appeal of this idea as a common-sense method of ensuring that teachers have the minimum qualifications to implement new standards, but he pointed out the risk in setting teachers up for a humiliating public failure. A more constructive strategy, he suggested, might be to use the existence of new standards to build public support for providing the time and funds for whatever training teachers need to enable them to implement the high standards the community has adopted.

> . . . if teachers are treated as a target, the likelihood of them really being enthusiastic and bringing about the kinds of changes that are needed is less than if they see themselves as contributing to the establishment of the standards.
>
> *Ruth Hayhoe*

Looking at the long term, Floden also pointed out that the many policies that send messages to teachers about what to teach are not consistent. Picking up on that notion, Andrew Porter made the more general point that although excellent standards documents have been developed for the major academic disciplines in the United States, there are few supports in place to facilitate their implementation. In particular, he noted, there could be assessments, materials (textbooks and the like), and teacher education, both preservice and ongoing professional development, that are all in sync with the standards.

A number of participants also raised the point that incentives can play a key role in implementation of standards-based programs. As noted above, many countries have tied high performance standards to specific and serious consequences at various points in a student's career. Though Floden agreed that penalizing teachers would not be constructive, positive incentives can work. Alan Ruby explained that in Australia, universities have been given incentives to develop courses specifically for practicing teachers who need to update their discipline knowledge. These courses offer the teachers a credential that can be tied to promotion or other benefits; to obtain the credential, teachers must demonstrate not only mastery of the content, but also the ability to present it in a lesson.

Several participants took up this point, noting that schools of education are not necessarily offering the material that will best prepare beginning teachers to make use of the NCTM standards. This is partly because in the United States colleges and universities have little incentive to pay attention to the specific credentials that beginning teachers might need. Floden noted that prestigious universities

are the least likely to be concerned about these credentials. The relatively low status that teaching has traditionally suffered in comparison with other professions is part of the reason. Prestigious universities are perfectly willing to ensure that students are offered the courses that will prepare them for careers in law or medicine, for example. A formal system of licensure and professional development for teachers that is both coherent and rigorous is an obvious solution. Under such a system teachers would be expected to update their knowledge, and, like physicians, would be supported in doing so. The National Board for Professional Teaching Standards has been working since 1987 to develop assessments for experienced teachers that could complement and reinforce academic standards: eventually the board plans to offer certificates in approximately 30 subject areas. The board awarded its first certificates to 81 teachers in 1995. A similar effort has been undertaken by the Interstate New Teacher Assessment and Support Consortium (INTASC) which is working to develop an assessment system that can be used to license beginning teachers.

> What I have a hard time seeing in the United States is how we change higher education both in initial teacher preparation and in continuing professional development.
>
> *Amy Stempel*

The emphasis on making sure teachers have the opportunity for professional development and collaboration was echoed in the panel discussion of the New Standards Project, which works with 17 states and seven large districts to set high academic standards and to develop a system of performance assessments to measure progress toward these standards (see box on p. 19). The New Standards Project has made a significant effort to collect and synthesize information about the standards used in other countries and to benchmark its standards to them.

The paper and presentation prepared by Kate Nolan for the workshop focused on the importance of using consensus and discussion to build a well-grounded notion of standards for a given academic area. Nolan described the dynamic in a typical standard-setting session run by New Standards. Participants would begin by looking at a sample of student work and would be asked to evaluate it. Invariably the discussion would quickly zero in on key questions, such as "what grade is the student in?" "What kinds of work have been done in this student's classroom?" In other words, the participants would find that there was no absolute "high" standard, but that the caliber of the performance could only be judged in the context of a set of specific expectations for students in a particular context.

Moreover, Nolan reported, the teachers who worked together to understand the standards by which they were to judge a particular body of student work came away with a far deeper understanding of both the evaluation criteria and the assumptions about the classrooms that produced the work than they could have developed by simply reading about the standards, or even attempting to apply them on their own. Similar experiences are often reported by teachers who have

Approaches to Benchmarking International Standards

Based on presentation by Katherine J. Nolan

The New Standards Project reversed the usual approach to benchmarking by moving from the global to the specific. Most of the interest in international benchmarks in this country has come from districts, states, and education leaders who have examined what's in place and have sought perspective from other countries as they progressed. The New Standards staff began at the other end. They first collected and thoroughly examined standards and assessments from all over the world. They synthesized this information and found ways of linking systems that were not structurally similar. They have also developed relationships with educators around the world so that they can maintain a kind of ongoing comparative dialogue about standards and ways of assessing them. In this way it is possible for them to benchmark one set of standards against another and, arguably, to provide international benchmarks that can help states and districts discover how rigorous their own standards are. A primary goal of the New Standards Project has been to develop an archive of both standards documents and student work linked to various standards.

For Nolan, the value in international comparisons comes from "the widespread dissemination and discussion of truly excellent student work." By seeing for themselves what students of a given age can do, she argues, educators can expand their expectations and use this understanding to increase student achievement. She summarized her view this way: "[We need examples] where I as a teacher get into thinking 'okay, I don't think the kids can do it,' and . . . other groups of teachers come forward and they say 'yes they can and here's how.' . . . I think the more examples we can get that force us to reflect on our process and challenge us to do better . . . , the better off we all are."

participated in scoring open-ended exercises for assessments. Christopher Cross noted that Maryland teachers who had scored the state assessment not only found it an excellent professional development activity, but also became ambassadors for the program. Having had the opportunity to understand the assessment in some depth, they were able to explain it to others.

Nolan noted that language barriers are a major obstacle to attempting such collaboration internationally. While in a sense this is a technical obstacle, and one that has been addressed by international studies such as TIMSS, it raises the issue of culture and context. The example of TIMSS, in which the United States worked with Germany and Japan in an unprecedented effort to collect data about contextual factors and tie them to international achievement data, suggests that coordinating these different kinds of information is not a straightforward exercise. The study made innovative use of several different

methodologies and has little precedent to follow in bringing together and presenting the results from its different components (see National Research Council, 1997). Similarly, those involved in the New Standards Project have found that truly understanding what a particular standard means in its own setting requires a deep level of knowledge of that setting. This perspective could be seen as further reinforcing Ruby's point that a search for a specific international standard in a given subject may be less fruitful than incorporating international comparisons and perspectives into the process of tailoring standards to a particular context.

States have confronted the need to tailor standards directly since it is a variety of competing demands that have led them to develop standards in the first place. These demands include pressure from business interests to educate skilled workers for local employment as well as the reality that, in a mobile society, the students they are educating now may be employed all over the United States and the world. Local political concerns faced by state education policy makers might range from an industry-based need for particular skills or a governor's devotion to a particular subject to public pressure to align themselves with national discipline standards and national goals for public education.

A few states have made a deliberate effort to incorporate an international perspective into their standard-setting efforts, and one consortium of districts in Illinois has done so in a high-profile way by participating in TIMSS on its own (see box on p. 21). The vast majority of states are working on standards, and many have begun to collaborate and to pool some of their resources. One participant suggested that, as both state and national efforts begin to build on experience, the United States is slowly groping its way toward some national priorities for standards in "bottom-up" fashion. However, the three state representatives at the workshop—from among the very few states that have looked internationally in working on their standards—all stressed the value of the state's controlling the process.

Robert Rice of Maryland explained: "There is a value for us in doing it ourselves because teachers and communities buy into it. If a federal standard is handed to us, we would have a very different selling job." Tom Kerins of Illinois made a very similar point: "Consensus was difficult enough for us as it was. Something from D.C. would be much harder. There is a lot of duplication; it's slow and laborious, but it has the best chance for affecting schools given the structure in this country." Wayne Martin of Colorado echoed this sentiment and added: "It's the auxiliary things we need help with. For example there's a need for a literacy test for ESL [English as a second language] students in their native language. We need to find out if they are literate at all before we work on their knowledge of English."

Another participant made the point that an international benchmark does not necessarily identify the highest standard—or one worthy of serving as a benchmark for others—citing the International

The State Experience—Illinois

*Based on presentation by C. Thomas Kerins**

The state of Illinois was an early convert to the standards movement, having begun the process of both defining content standards and developing its own assessments to track student progress in meeting them in 1985. An international focus emerged later. The state is currently in the midst of a cycle of reevaluation of its standards and assessment programs, and it was primarily this process that motivated them to begin looking at international comparisons. Both the state as a whole and a consortium of school districts near Chicago arranged to participate in TIMSS, and their reasons for doing so reflect points raised at the workshop discussion in interesting ways.

The state's interest in participating grew in part out of its desire to obtain contextual data that could help them in evaluating the many variables that affect educational outcomes, and, in particular, to assist them as they work on revising their content standards. The state has not collected background data through its own assessment program and has been particularly interested in obtaining, through TIMSS, data about attitudes toward mathematics and science, teaching practice, resources, and the like. The state was also interested in providing itself with achievement benchmarks. By sorting TIMSS items to match the state's content standards, state officials have been able to link the TIMSS tests to the state assessments and thus to provide state educators with targeted, relevant comparisons. They have identified a list of top-performing countries against which to compare Illinois's performance.

For the consortium of districts near Chicago (known as the First in the World Consortium because it was formed to meet the national goal of educating students to be first in the world in mathematics and science by the year 2000), the motivation for participating was somewhat different. Businesses in these districts had long complained that the graduates of the district's schools were not sufficiently prepared to compete with workers around the world. Business leaders were unconvinced by gains in scores on the statewide assessments. As Kerins put it, "[a]nother source of information had to be found in order to focus the dialogue with the business community."

As it turned out, the consortium's eighth-grade students performed extremely well on TIMSS, with only Singapore scoring better in mathematics and no nation scoring better in science. According to its report, the consortium attributes its success to the fact that its members were already holding their students to standards higher than the average for the United States. However, like the state, the consortium intends to build on its success by continuing to plumb the TIMSS results for insights about the strategies used in high-performing nations and to continue its international focus in the long term.

*This summary is based in part on the paper and presentation prepared by Thomas Kerins for the workshop and in part on information available on the First in the World consortium web site at http://www.ncrel.org/sdrs/firstwor.htm.

Association for the Evaluation of Educational Achievement (IEA) literacy study in which the United States performed well, despite its students' generally disappointing performance on the reading assessment of the National Assessment for Education Progress (Elley, 1992; Campbell et al., 1996). It seems clear that Colorado, Illinois, and Maryland have sought international comparisons as sources of information to use in the development of standards to fit their own needs, rather than as sources of absolute performance objectives.

SUMMING UP

The dialogues that many in the United States education community have had with international colleagues have revealed an interesting perspective—our obsession with defining and adhering to higher standards is very puzzling to many of them (Resnick et al., 1995:441). The reason seems to be that in many countries both content and performance standards are deeply embedded in the educational system—and tightly coupled with other elements of the system—and are simply not discussed in isolation. Arguably this state of affairs is what standards-based reformers here aspire to, but arriving at it in a more deliberate, albeit slow, manner should be an opportunity for some valuable reflection about the nation's commitment to providing an excellent education for all of its children.

One of the intended goals of the workshop was to consider ways of moving the process of standards-based reform forward, and participants offered both general and specific suggestions to that end. While no attempt was made to establish consensus about any of the suggestions that were made, several themes emerged.

Standards need to be embedded in and integrated into the entire education system. As all the presenters made clear, standards will not work by themselves. In order for standards to be fully integrated into a system, the social and political purposes of having them need to be worked out in a public forum. Bickel and LeMahieu amplified this point with their observation about the unrealistically high expectations many leaders have of standards' capacity to improve education—and society in general—in the United States. Moreover, a community needs to plan for the consequences that standards will have on all the parts of the system and for the changes—in assessments, in textbooks, and in teacher training, for example—that will be needed to accompany new standards. For standards to work, they need to be understood by all who will be affected by them, and they need to be made to matter to those who must meet them and those who must support them.

There is a need for greater and clearer communication about standards than is currently taking place. A variety of different constituencies—students, teachers, parents, employers, the academic

The State Experience—Colorado
Based on presentation by Wayne Martin

The state of Colorado has become extremely committed to an international perspective, although, as Mr. Martin noted, this was not a deliberate goal. The primary catalyst for the international focus was an economic crisis. Colorado experienced an oil boom and then an oil bust. One of its major industries, beef cattle, declined at the same time. The state's response was to look to trade with other countries for a solution to rising unemployment. There was a recognition that all students needed to be well educated in order for Colorado to compete in a global arena.

An additional impetus for the state to investigate international comparisons came from its then governor, Roy Romer, who had served on the National Education Goals Panel and has had a long-standing interest in education. Colorado made three major commitments to international comparisons: participation in International Assessment of Educational Progress (IAEP) studies of math, science, and geography achievement at ages 4, 9, and 13, participation in the New Standards Project, and participation—they were one of a handful of states—in the Third International Mathematics and Science Study (TIMSS). As the TIMSS data are released, Colorado will focus on the background data collected from the top-performing countries, with the goal of identifying effective strategies they might adapt for their own purposes. For Colorado, international comparisons are seen as a way of informing the process of developing its own standards.

community, political leaders, voters—have a stake in education standards, and they have differing priorities, biases, and understandings. In order for the adoption of standards to result in improvements in schools and in student performance, all of these constituencies must understand and support them. They will also need a common understanding of terms, as the difference between business and other leaders' definitions of "world class" illustrates. Discussions of standards that do not distinguish among performance, content, and opportunity-to-learn standards (as occurred occasionally at the workshop) can contribute to misunderstanding about the precise nature and purpose of standards in different contexts and about their potential value.

It is also important that those involved in developing and implementing standards share information and solutions to avoid duplicating effort. Though a good deal of information is available—both in print and electronically—about standards efforts, no central clearinghouse for information and ideas exists. The context of the standards movement is quite different from one state to another, but, as one participant noted, the general momentum developing is in the direction of consensus-building, a sort of "bottom-up" progression toward national standards.

There is a need for greater involvement of teachers in standards-based reform. Floden's argument that, as he put it, "Most of the effect . . . policy makers have on student learning is channeled through those actually teaching children" was extremely compelling to the group, and many participants cited the importance of involving teachers (Floden, 1996:1). Teachers cannot implement standards they do not fully understand and support. Moreover, since it is they who best understand the contexts in which learning takes place, teachers are in a position to make an invaluable contribution to the development of standards.

The support of discipline experts is also important, not only for political reasons, but also because their expertise is fundamental to the quality of standards. If the goal of increasing the professional status of teachers while raising content standards for their licensure is achieved, a corollary benefit should be an increased sense of solidarity between university professors and those who teach younger students.

There is a need for more data. At many points during the workshop, participants remarked on the need for additional information, particularly international comparative data. More knowledge about systems of teacher preparation and development that have been successful in other countries would be helpful as states and others struggle to improve that critical leg of standards-based reform. More information about incentives for meeting high standards that have been developed in other educational systems and about assessments that have been used would be equally helpful, as would data about the extent to which standards already in use have made a difference.

Participants also spoke of the need to gain as much from the Third International Mathematics and Science Study as possible, both by pursuing promising secondary analyses and by examining the effects the results have on practice. Several participants highlighted the need for international comparative data in general, and stressed that international assessments would be more useful if their content, and their schedules, were more consistent and predictable. Clearly, information about what other countries are doing is crucial to the development and maintenance of standards that will help U.S. students to compete internationally.

These themes suggest that there may not be a simple answer to the question of how to define international competitiveness for education standards. While the discussion at the workshop clearly supported the notion that international comparisons are vital, participants agreed that each set of standards—for content or performance—must fit its own circumstances. Countries vary considerably in terms of virtually every aspect of their educational systems. Some have formal written standards; others have unwritten ones. Some have standards that are extremely specific; others have only general goals. The education goals that countries establish for their students range across a broad

continuum. All of these differences make direct comparisons of standards difficult. By understanding that one size does not fit all, as Ruby remarked, those in the business of developing standards will be better able to discover ways of ensuring that they lead to the improvements that are hoped for.

Ruby and other international observers have been struck by the fact that many in the United States are obsessed not just with explicit standards, but with those of a chief economic rival, Japan. The need to compete economically has inspired and driven the urges to reform education and to learn from international comparisons. The workshop discussions served as a reminder that goals for education run deeper than a mere desire for economic parity or hegemony. Gains in understanding of the standards for education in other countries have shown that they grow out of profound social aspirations.

> When you've never seen what's possible, it's hard to raise the standards.
>
> *Mary Lindquist*

Perhaps the primary message from the workshop was a recognition that searching for an operational definition of world-class standards is, in Ruby's words, a search for a holy grail. A community that wants to raise its standards to an internationally competitive level cannot do so simply by writing clear, excellent guidelines for content and performance that are grounded in an understanding of high expectations elsewhere, even though that task itself may be difficult. Standards are not, in themselves, a reform—they are an organizing device through which a system can be reformed. From a variety of perspectives, the participants in the workshop seemed to say clearly that many elements in the education system are equally important to a successful outcome for students and that focusing on any one in isolation is unlikely to be worthwhile.

REFERENCES

American Academy of Arts and Sciences
1995 American education: Still separate still unequal. *Dædalus* 124(4).

American Federation of Teachers
1995a Helping students in the middle: What average students can achieve when standards are high and stakes are clear. *American Educator* 19(4):2. Available: http://www.aft.org//middle.htm [March 24, 1997].
1995b *Making Standards Matter: A Fifty-State Progress Report on Efforts to Raise Academic Standards.* Washington, DC: AFT Educational Issues Department.

Campbell, Jay R., et al.
1996 *NAEP 1994 Reading Report Card for the Nation and the States.* National Center for Education Statistics. Washington, DC: U.S Department of Education.

Coleman, James S., Ernest Q. Campbell, and Carol J. Hobson
1966 *Equality of Educational Opportunity.* Available from the Superintendent of Documents, U.S. Government Printing Office. Washington DC: U.S. Department of Health, Education and Welfare.

Education Week
1997 *Quality Counts: A Report Card on the Condition of Public Education in the 50 States*. XVI:January 22, 1997. A supplement to *Education Week* in collaboration with the Pew Charitable Trusts.

Elley, Warwick B.
1992 How in the World Do Students Read? IEA Study of Reading Literacy. Hamburg: The International Association for the Evaluation of Educational Achievement.

Floden, Robert E.
1996 Teachers' Choices About Content: The Standards in Use. Unpublished paper prepared for November 1996 symposium sponsored by the Board on International Comparative Studies in Education of the National Research Council, Washington, DC.

Gagnon, Paul
1995 What should children learn? *The Atlantic Monthly* 276(6):65-78.

International Association for the Evaluation of Educational Achievement
1996a Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS). Chestnut Hill, MA: Boston College. Available: http://wwwcsteep.bc.edu/timss [June 19, 1997].
1996b Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS). Chestnut Hill, MA: Boston College. Available: http://wwwcsteep.bc.edu/timss [June 19, 1997].

Johnson, Jean, and John Immerwahr
1994 *First Things First: What Americans Expect from the Public Schools.* Report available from the Public Agenda Foundation, New York, NY.

Lapointe, Archie E.
1992 *Learning Mathematics: the Second International Assessment of Educational Progress.* Princeton, NJ: Center for the Assessment of Educational Progress.

LeMahieu, Paul G., and William E. Bickel
1996 The Public's and Policy Makers' Perspectives on Internationally Competitive Standards. Unpublished paper prepared for November 1996 symposium sponsored by the Board on International Comparative Studies in Education of the National Research Council, Washington, DC.

Lezotte, Lawrence W.
1986 School Effectiveness: Reflections and Future Directions. Paper presented at the 1996 Annual Meeting of the American Educational Research Association.

National Center for Education Statistics
1996 *Pursuing Excellence: A Study of U.S. Eighth-Grade Mathematics and Science Teaching, Learning, Curriculum, and Achievement in International Context.* Washington, DC: U.S. Department of Health and Human Services.

National Commission on Excellence in Education
1983 *A Nation at Risk: The Imperative for Educational Reform.* A Report to the Nation and the Secretary of Education, U.S. Department of Education by the National Commission on Excellence in Education. Washington, DC: U.S. Government Printing Office.

National Education Goals Panel
1995 *The National Education Goals Report: Building a Nation of Learners: 1995.* Washington, DC: U.S. Government Printing Office.

National Research Council
1997 *Learning from TIMSS: Results of the Third International Mathematics and Science Study, Summary of a Symposium.* Alexandra Beatty, Editor. Board on International Comparative Studies in Education, Board on Testing and Assessment, Committee on Science Education K-12, and Mathematical Sciences Education Board. Washington, DC: National Academy Press.

Ravitch, Diane, ed.
1995 Debating the Future of American Education: Do We Need National Standards and Assessment? Report of a conference sponsored by the Crown Center on Education Policy at the Brookings Institution. Available from the Brookings Institution, Washington, DC.

Resnick, Lauren B., Katherine J. Nolan, and Daniel P. Resnick
1995 Benchmarking education standards. *Educational Evaluation and Policy Analysis* 17(4):438-461.

Rossmiller, Richard A., and Edie L. Holcomb
1993 The Effective Schools Process for Continuous School Improvement. Paper presented at the 1993 Annual Meeting of the American Educational Research Association. Wisconsin Center for Education Research, Madison.

Ruby, Alan
1996 The Holy Grail of Education Reform: Internationally Competitive Standards. Unpublished paper prepared for November 1996 symposium sponsored by the Board on International Comparative Studies in Education of the National Research Council, Washington, DC.

Stedman, Lawrence C.
1996 International achievement differences: An assessment of a new perspective. *Educational Researcher* 26(3):4-15.

Steller, Arthur W.
1988 *Effective Schools Research: Practice and Promise.* Bloomington, IN: Phi Delta Kappa Educational Foundation.

Toch, Thomas, with Robin M. Bennefield and Amy Bernstein
1996 The case for tough standards. In *U.S. News* Winter 1995-96, [Online]. Available: http://www.usnew.com/usnews.com/usnews/ISSUE/STAND.HTM [March 14, 1997].

U.S. News
1996 The *U.S. News* school standards poll: Most Americans say that school standards should be higher than they currently are, a new *U.S. News* poll reports. In *U.S. News* 3/23/96, [Online]. Available: http://www.usnews.com.usnews.NEWS.STANPOLL.HTM [March 14, 1997].

Appendix A

Workshop Participants

PRESENTERS

William E. Bickel, University of Pittsburgh Learning Research and Development Center, Pittsburgh, Pennsylvania
Christopher T. Cross,* Council for Basic Education, Washington, D.C.
Robert Floden, College of Education, Michigan State University, East Lansing,
C. Thomas Kerins, Illinois State Board of Education, Springfield, Illinois
Michael W. Kirst,* School of Education, Stanford University
Paul G. LeMahieu,* University of Delaware and Delaware Department of Public Instruction, Newark
Mary M. Lindquist,* School of Education, Columbus State University
Wayne H. Martin, State Education Assessment Center, Council of Chief State School Officers, Washington, D.C.
Anne W. Miller, Education Solutions and Service, Eastman Kodak Company, Rochester, New York
Katherine J. Nolan, University of Pittsburgh Learning Research and Development Center, Pittsburgh, Pennsylvania
Jessie H. Pollack, Maryland State Department of Education, Baltimore, Maryland
Andrew C. Porter,* Wisconsin Center for Educational Research, School of Education, University of Wisconsin, Madison
Robert Rice, Maryland Department of Education, Baltimore, Maryland
Alan Ruby, Department of Employment, Education, Training and Youth Affairs, Canberra, Australia

*Member, Board on International Comparative Studies in Education

ATTENDEES

Cynthia Almeida, District of Columbia Public Schools, Washington, D.C.
Gordon M. Ambach,* Council of Chief State School Officers, Washington, D.C.
Jan Anderson, Planning and Evaluation Service, U.S. Department of Education
John Barth, National Governors' Association, Washington, D.C.
Halsey Beemer, China and Mongolia Department, World Bank, Washington, D.C.
Rosemary Bellew, Human Resources Two Africa Region, World Bank, Washington, D.C.
Sue Betka, Planning and Evaluation Service, U.S. Department of Education
Rolf Blank, Council of Chief State School Officers, Washington, D.C.
Wanda Brown, Educational Services, Potomac Electric Power Company, Rockville, Maryland
Ruben Carriedo, San Diego City Schools, San Diego, California
Joseph Conaty, Student Achievement Institute, U.S. Department of Education
John A. Dossey,* Department of Mathematics, Illinois State University, Normal
Irene Eckstrand, National Institute of General Medical Sciences, National Institutes of Health
Willem van Eeghen, Middle East and North Africa Department, World Bank, Washington, D.C.
Larry Feinberg, National Assessment Governing Board, Washington, D.C.
Joan Ferrini-Mundy, Center for Science, Mathematics, and Engineering Education, National Research Council, Washington, D.C.
Anne Fickling, Center for Civic Education, Washington, D.C.
Ray Fields, National Assessment Governing Board, Washington, D.C.
Rita Foy, Office of Educational Research and Improvement, U.S. Department of Education
Matthew Gandal, American Federation of Teachers, Washington, D.C.
Lenore Yaffee Garcia, Planning and Evaluation Service, U.S. Department of Education
Alice Gill, American Federation of Teachers, Washington, D.C.
Aimee Guidera, National Alliance of Business, Washington, D.C.
Ruth E.S. Hayhoe,* The Ontario Institute for Studies in Education, University of Toronto, Canada
Cadelle Hemphill, Education Statistics Service, Washington, D.C.

Joyce Higginbotham, District of Columbia Public Schools, Washington, D.C.
Carrie Chimerine Irvin, New American Schools, Arlington, Virginia
Jean M. Johnson, Division of Science Resource Studies, U.S. National Science Foundation, Arlington, Virginia
Barbara A. Kapinus, Council of Chief State School Officers, Washington, D.C.
Sally B. Kilgore, Hudson Institute, Indianapolis, Indiana
Karen Larson, National Alliance of Business, Washington, D.C.
Frans Lenglet, Human Resources and Poverty Division, World Bank, Washington, D.C.
Marlaine E. Lockheed,* Human Development Department, World Bank, Washington, D.C.
John T. MacDonald, Council of Chief State School Officers, Washington, D.C.
Julia MacMillan, Council of Chief State School Officers, Washington, D.C.
Barbara Marenus, National Center for Education Statistics, U.S. Department of Education
Pamela E. Matthews, Department of Mathematics and Statistics, American University
Marilyn McConachie, member, National Assessment Governing Board, Glenbrook High Schools, Northbrook, Illinois
Maggie McNeely, Student Achievement Institute, U.S. Department of Education
William J. Moloney, member, National Assessment Governing Board, Calvert County Public Schools, Prince Frederick, Maryland
Jay Moskowitz, Pelavin Research, Washington, D.C.
Mari Muri, Association of State Supervisors of Mathematics, Hartford, Connecticut
Ken Nelson, National Education Goals Panel, Washington, D.C.
Lynn W. Paine,* Department of Teacher Education, Michigan State University, East Lansing
Cynthia Pantazis, American Society for Training and Development, Alexandria, Virginia
Lois Peak, National Center for Education Statistics, U.S. Department of Education
Richard Phelps, Pelavin Research, Washington, D.C.
Jeffrey M. Puryear, Inter-American Dialogue, Washington, D.C.
Senta A. Raizen, National Center for Improving Science Education, Washington, D.C.
Francisco O. Ramirez,* School of Education, Stanford University
Gretchen F. Ridgeway, Research and Evaluation, U.S. Department of Defense Education Activity, Arlington, Virginia
Linda P. Rosen, National Council of Teachers of Mathematics, Reston, Virginia

Alcyone Saliba, Latin American and Caribbean Human Development Division, World Bank, Washington, D.C.
Elois Scott, Planning and Evaluation Service, U.S. Department of Education
Sharif Shakrani, National Center for Education Statistics, U.S. Department of Education
Amy Stempel, Council for Basic Education, Washington, D.C.
Stephanie Stullich, Planning and Evaluation Service, U.S. Department of Education
Larry E. Suter, Division of Research, Evaluation, and Dissemination, U.S. National Science Foundation, Arlington, Virginia
Marcia P. Sward, Mathematical Association of America, Washington, D.C.
Stewart Tinsman, International Services, U.S. Department of Education
Valerie Truesdale, Lexington School District 5, Balentine, South Carolina
Laurence Wolff, Principal Operations Office, World Bank, Washington, D.C.

Appendix B

Workshop Agenda

TAKING STOCK: WHAT HAVE WE LEARNED ABOUT MAKING EDUCATION STANDARDS INTERNATIONALLY COMPETITIVE?

Board on International Comparative Studies in Education
National Research Council
2001 Wisconsin Avenue, NW, Washington, D.C.

AGENDA

9:00 a.m. ***OPENING REMARKS***
Michael W. Kirst, Stanford University School of Education, Stanford, California, and Board on International Comparative Studies in Education *(chair)*
Andrew C. Porter, Wisconsin Center for Educational Research, University of Wisconsin, Madison, Wisconsin, and Board on International Comparative Studies in Education

9:20 ***SESSION 1***
Perspectives on Internationally Competitive Standards from the Public and Policy Makers

SPEAKERS
William E. Bickel, University of Pittsburgh Learning Research and Development Center, Pittsburgh, Pennsylvania
Paul G. LeMahieu, University of Delaware and Delaware Department of Public Instruction, Newark, Delaware, and Board on International Comparative Studies in Education

9:45 ***Session 2***
The Holy Grail of Education Reform: Internationally Competitive Standards

Moderator
Michael W. Kirst, Stanford University School of Education, Stanford, California, and Board on International Comparative Studies in Education *(chair)*

Speaker
Alan Ruby, Department of Employment, Education, Training and Youth Affairs, Canberra, Australia

10:45 *Break*

11:00 ***Session 3***
Teachers' Choices About Content: The Standards in Use

Moderator
Andrew C. Porter, Wisconsin Center for Educational Research, University of Wisconsin School of Education, Madison, Wisconsin, and Board on International Comparative Studies in Education
Speaker
Robert Floden, Michigan State University College of Education, East Lansing, Michigan

12:00 p.m. *Lunch in meeting room GREEN 104*

1:00 ***Panel Discussions 1***

GREEN 104

Panel One: Opportunities for International Benchmarking

Moderator
John A. Dossey, Illinois State University Department of Mathematics, Normal, Illinois, and Board on International Comparative Studies in Education

Third International Mathematics and Science Study (TIMSS)
Mary M. Lindquist, Columbus College School of Education, Columbus, Georgia, and Board on International Comparative Studies in Education

New Standards Project
Katherine J. Nolan, University of Pittsburgh Learning Research and Development Center, Pittsburgh Pennsylvania

Eastman Kodak's Experience with Benchmarking
Anne Miller, Eastman Kodak Company, Rochester, New York

GREEN 118

Panel Two: From Theory to Practice: The State Experience

Moderator
Paul G. LeMahieu, University of Delaware and Delaware Department of Public Instruction, Newark, Delaware, and Board on International Comparative Studies in Education

State of Illinois
C. Thomas Kerins, Illinois State Board of Education, Springfield, Illinois

State of Maryland
Robert Rice and Jessie H. Pollack, Maryland Department of Education, Baltimore, Maryland

State of Colorado
Wayne H. Martin, Council of Chief State School Officers State Education Assessment Center, Washington, D.C.

2:15 *Break*

2:30 ***Panel Discussions 2***

GREEN 104

Panel One: Opportunities for International Benchmarking

Moderator
John A. Dossey, Illinois State University Department of Mathematics, Normal, Illinois, and Board on International Comparative Studies in Education

Third International Mathematics and Science Study (TIMSS)
Mary M. Lindquist, Columbus College School of Education, Columbus, Georgia, and Board on International Comparative Studies in Education

New Standards Project
Katherine J. Nolan, University of Pittsburgh Learning Research and Development Center, Pittsburgh Pennsylvania

Eastman Kodak's experience with benchmarking
Anne Miller, Eastman Kodak Company, Rochester, New York

GREEN 118

Panel Two: From Theory to Practice: The State Experience

Moderator
Paul G. LeMahieu, University of Delaware and Delaware Department of Public Instruction, Newark, Delaware, and Board on International Comparative Studies in Education

State of Illinois
C. Thomas Kerins, Illinois State Department of Education, Springfield, Illinois

State of Maryland
Robert Rice and Jessie H. Pollack, Maryland Department of Education, Baltimore, Maryland

State of Colorado
Wayne H. Martin, Council of Chief State School Officers State Education Assessment Center, Washington, D.C.

GREEN 104

3:45 ***Next Steps: Using What We've Learned About Internationally Competitive Standards***

Moderator
Christopher T. Cross, Council for Basic Education, Washington, D.C., and Board on International Comparative Studies in Education

5:00 *Adjournment*

Appendix C

Workshop Papers

The papers listed below were prepared for and presented at the workshop. They are available on the web page of the National Academy of Sciences, at http://www.nas.edu.

Floden, Robert E.
1996 Teachers' Choices About Content: The Standards in Use. Michigan State University.

Kerins, C. Thomas
1996 What Have We Learned About Making Education Standards Internationally Competitive? From Theory to Practice: The State Experience in Illinois. Illinois State Board of Education, Springfield.

LeMahieu, Paul G., and William E. Bickel
1996 The Public's and Policy Makers' Perspectives on Internationally Competitive Standards. University of Delaware.

Miller, Anne
1996 Benchmarking in Business and Applying a Similar Strategy to Education: Summary of Remarks. Eastman Kodak Company, Rochester, N.Y.

Nolan, Katherine J.
1996 World-Class Standards. University of Pittsburgh.

Rice, Robert C., and Jessie H. Pollack
1996 Challenges of Cross-National Studies of Educational Achievement: The Maryland Experience. Maryland Department of Education, Baltimore.

Ruby, Alan
1996 The Holy Grail of Education Reform: Internationally Competitive Standards. Department of Employment, Education, Training and Youth Affairs, Canberra, Australia.