



Policy Brief

What the Third International Mathematics and Science Study (TIMSS) Means for Systemic School Improvement

The National Institute on Educational Governance, Finance,
Policymaking, and Management
Office of Educational Research and Improvement
U.S. Department of Education

POLICY BRIEF

**What the Third International Mathematics and Science Study (TIMSS) Means for
Systemic School Improvement**

National Institute on Educational Governance, Finance, Policymaking, and Management
Office of Educational Research and Improvement
U.S. Department of Education

and

Consortium for Policy Research in Education

Foreword

The Third International Mathematics and Science Study (TIMSS), the most extensive international educational comparison ever done, contains a wealth of information about how U.S. students perform in these two critical areas and how various aspects of our educational system affect their performance. Whether the messages from TIMSS spur genuine education reform will depend in large part on the actions of policymakers and educators at the local, state, and national levels.

To explore the policy implications of TIMSS, the National Institute on Educational Governance, Finance, Policymaking, and Management of the Office of Educational Research and Improvement (OERI) in the U.S. Department of Education, and the Consortium for Policy Research in Education (CPRE), held a TIMSS Policy Forum in Washington, D.C. on October 6 and 7, 1997. This forum brought together a diverse group of TIMSS researchers, state and local policymakers, teachers, school administrators, mathematics and science specialists, business representatives, university educators, and federal officials. Participants discussed what TIMSS means for systemic education reform and identified several future directions for policy decisions and policy research.

This policy brief synthesizes the presentations and discussions that occurred during the TIMSS Policy Forum. It also draws from the TIMSS reports of the National Center for Education Statistics (NCES) and from papers prepared by researchers involved in TIMSS.

Contents

Foreword	iii
Executive Summary.....	1
For Further Information.....	5
What TIMSS Says about Student Achievement	6
Using TIMSS to Benchmark State and Local Achievement.....	9
What TIMSS Says about Equity Issues.....	13
Using TIMSS Data to Inform Equity Issues.....	14
What TIMSS Says about Curriculum.....	15
Using TIMSS to Inform Changes in Curriculum	17
What TIMSS Says about Instruction and Professional Development.....	19
Using TIMSS to Inform Changes in Instruction and Professional Development	21
TIMSS and Standards-Based Reforms.....	25
Using TIMSS to Build Public Support for School Improvement	29
TIMSS as a Model for Future Policy Research	30
Conclusion	31
References	33
Participants in the Policy Forum	35

Executive Summary

The Third International Mathematics and Science Study is a rigorous research effort that examined student performance in 41 nations at 3 grade levels. Unlike past international studies, TIMSS not only shows *how well* U.S. students perform compared with their international peers, but also analyzes curriculum and teaching practices in various countries to explain *why* our students perform as they do. TIMSS includes the following messages about education reform and future education policy:

- ***We need to make dramatic improvements in student achievement, especially in middle schools and high schools.*** U.S. students don't start out behind; they fall behind. U.S. fourth-grade students are among the very best in the world in science and above the international average in math. But by high school our students score near the bottom of TIMSS nations in both subjects--alarming news for a nation that wants to remain a world economic leader.
- ***The U.S. math and science curriculum lacks rigor, focus, and coherence.*** According to TIMSS, many middle school students in the United States are still doing elementary arithmetic and introductory science while their international counterparts are studying algebra, geometry, physics, and chemistry. By the senior year of high school, many of our students have stopped taking math and science altogether. And while other countries focus on a limited number of critical topics, curriculum in the United States emphasizes breadth over depth, providing students with superficial exposure to many topics but mastery of none.

- ***U.S. teachers demand less high-level thought of their math students than teachers in Germany and Japan.*** TIMSS videotapes of real classrooms show how differently math is taught in these three countries. The main goal in the United States is to teach students how to do various procedures, while in both Japan and Germany teachers are much more likely to *develop* concepts and procedures rather than merely *state* them, resulting in what experts describe as higher quality teaching in those countries than in the United States.
- ***Unlike in Japan, in the United States education reformers try to change teaching through indirect means, rather than by focusing on improving the quality of classroom lessons.*** In Japan, classroom lessons are highly valued, crafted with greater care, and seldom interrupted. Japanese teachers use collaborative study groups to critique and continually improve their classroom lessons. To make time for this formal planning, Japanese teachers have much larger classes. Most American teachers do not have these built-in opportunities for collaborative analysis and improvement.

TIMSS provides the impetus for states and school districts to think about where their students stand according to international, not just state or local benchmarks, and to identify subjects, grades, and areas where they need to do better. This process can be eye-opening for complacent communities that think their students are “good enough,” or for low-performing schools that have a long way to go to prepare their students for a global labor market. The TIMSS Policy Forum participants came to the following conclusions about the possible implications of TIMSS for American education.

- ***All levels of government have important roles in systemic school improvement.*** Even TIMSS countries with highly centralized systems use both "top down" and "bottom up" strategies to achieve educational goals--for example, by setting curriculum goals nationally and determining teaching methods locally. In the decentralized U.S. system, school improvement can best be accomplished through combined actions at the local, state, and national levels. Local educators could determine the most effective instructional strategies for their students and design systems that hold people accountable for high student performance. States could lead reforms of curriculum and teacher preparation. The federal government could disseminate information about a range of effective strategies for reform, which states and districts could adopt or adapt.
- ***New approaches to curriculum reform are needed.*** Copying the curriculum of other countries is not the answer. Rather, the United States must develop curricular alternatives that fit our pluralistic system and tradition of dispersed control. We could develop new structures for better coordinating curriculum across states and districts and for making difficult choices about which topics are most important and which can be eliminated.
- ***TIMSS can be a tool for professional development.*** American teachers can use TIMSS to analyze and improve their own practices--for example, by watching videos of other teachers teaching or by taking TIMSS test items themselves. TIMSS has also identified interesting models for professional development, such as the Japanese lesson study groups. Changing instruction in the United States not only will require new methods of teacher preparation and professional development, but also may require new approaches to school organization, time, and teacher duties.

- ***TIMSS can help rally public support for school improvement.*** Presenting TIMSS findings in a clear and engaging format can help policymakers, parents, educators, and the public understand why our students need to do better in science and math. For example, showing parents sample items from the TIMSS test or screening classroom videotapes can illustrate quite vividly what U.S. students know and how U.S. teachers teach compared with their counterparts in other countries.
- ***Most TIMSS countries have reached a national consensus about standards for curriculum and instruction.*** The move toward state standards in the United States is a promising way to give more focus to education reform while still respecting our decentralized system. But standards-based reform could compound our curricular problems if states adopt a wide range of content standards that are not high by international benchmarks, or are not concrete enough to be useful to teachers. States and local standards may be more effective if they identify a limited number of topics to be taught and provide guidance on the best ways to teach them.
- ***TIMSS offers a productive way of comparing the U.S. education system with those of other countries.*** TIMSS has gone well beyond "horse-race" comparisons. It provides us with the basis for examining our assumptions about education and the social and cultural underpinnings of those assumptions. By reflecting on these international comparisons, it will allow those trying to improve education a better grasp of the problems to be overcome.

- ***A global economy demands international benchmarks.*** The U.S. Department of Education has developed a “toolkit” of written and video materials that communities can use to measure their students’ achievement against the international benchmarks of TIMSS and learn more about curriculum and instructional practices associated with high performance (see below).

For Further Information

Additional information, sample test items, and other TIMSS resources are available from two of the Department of Education’s web sites at <http://nces.ed.gov/TIMSS/> and <http://www.ed.gov/inits/Math> or from the Eisenhower National Clearinghouse Web site at <http://timss.enc.org>. New information is being added to these sites on a ongoing basis so they are an excellent source of the latest developments relating to TIMSS. Additionally, the Department has a TIMSS Customer Service Hotline to answer questions at (202) 219-1333. Interested persons can obtain copies of the TIMSS reports and a TIMSS Resource Kit through these Web sites, or from the Superintendent of Documents, P.O. Box 371954, Pittsburgh, PA 15250; telephone (202) 512-1800, fax (202) 512-2250. Detailed technical information about the TIMSS data and methodology is available from the TIMSS International Study Center, CSTECP, School of Education, Boston College, Chestnut Hill MA 02167; telephone (617) 552-4521, fax (617) 552-8419, Internet <http://www.csteep.bc.edu>. Contact information for participants in the National Institute’s Policy Forum can be found at the end of this report.

What TIMSS Says About Student Achievement

The Third International Mathematics and Science Study (TIMSS) is the largest, most comprehensive, and most rigorous international comparison of education ever done. In 1995, TIMSS tested the mathematics and science knowledge of half a million students at 3 different grade levels in 41 countries. Participating countries had to follow strict quality control procedures to ensure that samples of students accurately represented each nation's student population, that tests were properly monitored, and that comparisons were fair. The 8th-grade results were released in November 1996, the 4th-grade results in June 1997, and the 12th-grade results in February 1998.

TIMSS represents a great stride forward in international comparative research. It gives a more complete and accurate picture than any previous study of how American students perform compared with their peers in Japan, Germany, England, Canada, Russia, Korea, and many other countries. Through a close examination of curriculum and instruction, TIMSS also provides new information about *why* our students perform as they do. To understand the curriculum as it is *intended* to be taught, TIMSS researchers studied curricular frameworks, textbooks, and other materials of participating nations and surveyed their educational authorities. To understand the curriculum as it is *actually* taught, TIMSS researchers surveyed teachers and students and conducted classroom observations. An integral part of TIMSS research was a videotape study of real teachers teaching eighth-grade mathematics in the United States, Japan, and Germany. These 230 hours of videotaped lessons offer an unprecedented inside look at instructional practices in different countries.

In short, TIMSS has collected a wealth of information with significant implications for educators, policymakers, and anyone else interested in school reform. TIMSS can help the United States address fundamental issues about what our students are learning, what our teachers teach and how they teach relative to other countries, and how we can improve instruction for all our students.

In order to use TIMSS to guide improvement at the state and local level, policymakers and educators must first understand what TIMSS says about science and mathematics achievement for the whole nation. The general story of TIMSS is that U.S. students start out performing at high levels in fourth grade, but by the time they approach high school graduation they are performing at unacceptably low levels in both science and math.

The best performance for U.S. students came from our fourth graders. On the fourth-grade TIMSS test administered in 26 countries, U.S. students scored near the top in science and above the international average in mathematics (NCES, 1997). In science, our fourth-graders were outperformed only by Korea, while 5 other countries scored in the same statistical range as the United States. In mathematics, 7 countries scored above our fourth-graders, while 6 were in the same range, and 12 scored below.

On the eighth-grade TIMSS assessment, U.S. students scored somewhat above the international average in science and somewhat below average in mathematics (Peak, 1996). Our eighth-grade science scores are in the same range as those of Germany, England, Canada, and Russia. In eighth-grade math, the United States outperformed only seven other countries, none of which is a major economic competitor. Five high-performing nations--Singapore, Japan, Korea,

the Czech Republic, and Hungary--did better than the U.S. eighth-graders in both science and math.

On the 12th-grade TIMSS assessment, U.S. students performed among the lowest of the 21 participating nations on tests of general knowledge in science and math (NCES, 1998). Only Cyprus and South Africa scored significantly lower than the United States on these general tests. In science, our 12th-grade scores were in the same overall range as those of seven other countries, including Italy, Germany, France, and Russia, while in math our scores were about the same as those of four other countries, among them Italy and Russia. (The Asian nations chose not to participate in the 12th-grade study.) The TIMSS study also gave tests in physics and advanced mathematics (pre-calculus and beyond) to students who had studied these subjects. On these tests, the advanced U.S. 12th-graders performed among the lowest of the TIMSS nations.

TIMSS also shows how well students perform in various content areas of math and science (Peak, 1996; NCES, 1998; NCES, 1997). For example, U.S. students do relatively well in earth science and life science at both the 4th- and 8th-grade level (the 12th-grade test did not break-out results in this way). U.S. eighth-graders are lagging in geometry, which most of them have not studied yet, and in chemistry and physics topics. And the results of the advanced 12th-grade tests suggest that even our best students are weak in physics and in calculus, geometry, equations, and functions.

Two main messages emerge from these results. First, U.S. students don't start out behind; they fall behind (Schmidt et al, 1996). U.S. fourth-graders do well in math, and in science they are close to achieving the Governors' goal of being first in the world--the National Education Goal many people thought would be hardest to reach when it was set in 1989. But then

achievement drops considerably across the grades in both subjects, so by 12th-grade our students are performing near the lowest performing nations. In fact, the United States is the only TIMSS nation that went from above average in math in fourth grade to below average in eighth grade.

The second key message of TIMSS is that by the time our students complete their formal secondary schooling, they are not achieving at the international standards demanded by a global labor market. Our relatively poor performance in 12th-grade math and science is particularly disappointing for a country that aims to be a world economic leader.

The TIMSS student achievement data can help give national urgency and focus to U.S. school improvement efforts. We clearly need to make dramatic, rapid, and fundamental improvements in math and science education, particularly in our middle schools and high schools.

We need to have higher expectations of our students and better prepare them to meet these expectations. These and other policy implications are explored in more detail below.

Using TIMSS to Benchmark State and Local Achievement

Although comparisons give an indispensable global dimension to performance information, they are by no means the only, or even the most important, contribution of TIMSS. If state and local educators focus only on the "horse race" aspect of the TIMSS data, they will miss its true value for benchmarking--namely, what states and districts do with the information. Countries like Japan that rank near the top of the TIMSS nations use international data to find areas for further improvement, rather than basking in their accomplishments. The U.S. educational system does not have this same concept of continuous improvement, partly because we lack a sense of what

the gradations of quality are and partly because we do not have incentives and support structures that encourage students, teachers, and schools to become better and better.

One basic use of TIMSS at the state and local level is as a benchmark. States and school districts can use TIMSS to help determine how well their own schools and students perform by international standards in mathematics and science and where they need to improve.

Some states and school districts have gone the full route of administering actual TIMSS tests to representative samples of their students, at their own expense and by arrangement with the U.S. Department of Education. The state of Missouri, for example, has administered the TIMSS tests to a sample of 3,000 students. The First in the World Consortium, a group of 20 high-achieving school districts in Illinois, has undertaken an ambitious benchmarking effort by giving the TIMSS tests to a statistically valid sample of Consortium student (Paul Kimmelman, West Northfield School District #31, Illinois; David Kroeze, Northbrook School District #27, Illinois, TIMSS Policy Forum).

After receiving the preliminary achievement data, the First in the World Consortium decided to expand the project by contracting for curriculum and videotape analyses patterned after the TIMSS model. The districts hope to use these data to identify areas where they need to do better and to engage educators, parents, students, and community in a process of continuous improvement. The process designed by the Consortium has four basic components: (1) developing a common, coherent vision; (2) using data as a vehicle for inquiry and self-reflection; (3) encouraging professional mastery and sustained professional development; and (4) collaborating with other school districts, universities, research centers, and additional partners.

A district does not have to go through the full route of testing and analysis in order to use TIMSS as a catalyst for local introspection. The Department of Education has developed a five-part TIMSS "toolkit" to help districts through the benchmarking and improvement process (OERI, 1997). The "Guide" to the kit outlines its contents and how they might be used. The "Education" module of the kit is intended to provide educators, parents, policymakers, and concerned citizens with an overview of the TIMSS findings. It includes summaries of the main TIMSS reports, an introductory video, and a guide for local discussion. It is meant to serve as a kind of "medical alert" that will wake up complacent communities who think their schools are good enough just because they're as good as the neighboring district's (Lois Peak, NCES; and Kroeze, TIMSS Policy Forum).

The "Achievement" module of the toolkit takes users through a "diagnosis" phase. It is intended to help educators use TIMSS items, in conjunction with their own assessment results, to reflect on how their students measure up by international standards (Peak, TIMSS Policy Forum). This module includes in-depth student achievement data from the TIMSS nations, as well as actual items from the TIMSS tests. For communities and parents that are accustomed to seeing their children score above average on norm-referenced tests, this type of analysis can spur them to think in terms of what their children do and do not know, and what they can and cannot do.

The final two modules of the TIMSS toolkit help educators research "treatment options" by making them aware of instructional and curriculum practices in other countries. The "Teaching" module includes videotaped excerpts of lessons from the TIMSS video study; transcripts, notes, and discussion guides to accompany the taped lessons; and excerpts from mathematics standards produced by the National Council of Teachers of Mathematics (NCTM) in

the United States and the Japanese Ministry of Education. The “Curricula” module contains a guidebook to help people analyze their own curriculum and compare it with other international and national approaches, so they can decide directions for curricular and instructional reform.

Some states and districts are already engaged in data-driven benchmarking and introspection, similar to the process laid out in the TIMSS toolkit. When educators deeply analyze student achievement data, they may come across surprises, both positive and negative. Vermont, for example, was compelled to ask hard questions about the impact of its new portfolio assessment program when data showed that less than one-fourth of the participating students had met the state standards for problem solving (Marjorie Petit, Vermont Department of Education, TIMSS Policy Forum). In some schools with fully implemented portfolios, no student met the standard. These results were particularly disconcerting because strengthening students' problem-solving was a primary goal of the state reforms. When state officials delved into samples of student work from the portfolios they found that many teachers did not fully understand what a substantive problem was, and therefore were not adequately teaching problem solving or higher math. Many Vermont schools and districts have responded by providing teachers with professional development to build a deeper understanding of problem-solving in math. Thus, although the Vermont data did not show positive results, something positive came out of it: state and local educators identified a problem with teachers' knowledge and took steps to address it.

TIMSS can be an effective tool for using achievement data to benchmark progress and drive school improvement. Several large urban systems are using a model of a data-driven improvement process for urban schools which includes the following five components: (1) student outcome data (such as disaggregated student test scores and graduation rates) and staff data (such

as staff attendance, involvement in professional development, and preparation in various content areas); (2) information on parents' involvement in activities that support education; (3) information about special programs currently operating at the school level; (4) school-community partnerships; and (5) information on grant writing (Sharon Johnson-Lewis, Council of the Great City Schools, TIMSS Policy Forum). The TIMSS findings can provide an international benchmark for these schools.

What TIMSS Says About Equity Issues

The TIMSS data suggest that the enormous diversity of the U.S. educational system is *not* a reason or excuse for our low average performance. However, currently available analyses do raise questions about the practice of tracking students. TIMSS data will eventually enable us to address other equity issues. More detailed analyses relating to critical equity issues have yet to be done on the TIMSS data, including analyses that will shed light on comparative distributions of achievement and the relationship between poverty and achievement in the TIMSS nations.

The United States is unique, for the first eight grades of schooling, in its practice of tracking students into different math and science courses or content based on past performance or perceptions of ability. Even a nation like Germany, which formally tracks students into different kinds of schools, still requires all students to study basically the same content; the main difference is how deeply they go into various subjects.

By 12th grade, many US. students have stopped taking mathematics and science altogether. Only 66 percent of U.S. graduating seniors were currently taking math, compared with an average of 79 percent in other TIMSS nations; the same trend was also apparent in

science. Consequently, many U.S. students--including college-bound students--graduate without four years of secondary level math and science, and without exposure to physics, calculus, and other rigorous content that their international peers are getting (NCES, 1998).

Using TIMSS Data to Inform Equity Issues

Low-performing students in the United States are often concentrated in cities and other high-poverty districts. Many urban districts have a high proportion of students performing below the United States average and well below the level demanded in a global economy. In the Boston Public Schools, for example, 81 percent of eighth-graders scored at the lowest level on a national standardized achievement test, meaning that they had not reached the basic level of performance (Maryellen Donahue, Boston Public Schools, TIMSS Policy Forum). TIMSS international comparisons are nevertheless relevant in these very diverse districts, because their students must compete in the same global labor market as other students, and educational quality is a key factor that companies consider when deciding where to locate.

The TIMSS data, along with urban districts' own benchmarking data, suggest the need for special improvement strategies for high-poverty districts with below-average achievement levels. Major improvements in student achievement in these districts are unlikely to happen without investments in capacity building (Donahue, TIMSS Policy Forum). Generally, these systems enroll very high percentages of low-income children and children with limited English proficiency, have high student mobility, and have difficulty hiring and keeping good teachers. Additional resources are needed to recruit appropriately certified teachers, provide professional development, and do other kinds of capacity building.

What TIMSS Says About Curriculum

What happens in schools really does matter. That is the key message from the TIMSS studies of curriculum and instruction (Schmidt et al, 1996; Schmidt and Valverde, 1997).

TIMSS found no single "magic bullet" that is universally associated with high or low achievement across the countries. Rather, differences in national performance appear to be connected to what schools teach and how they teach it. What schools teach is delineated in standards, curricula, and textbooks. Fortunately, these are areas that can be influenced by educational policy.

TIMSS performance data contradict the common public perception that American students have not grasped "basics" like arithmetic. The problem instead is that many are not learning the kinds of advanced mathematics and science that form the new basics in the global arena. The TIMSS studies point to the need for all U.S. students to learn more demanding content, taught in depth as they progress through the grades.

Curriculum in American schools lacks coherence, focus, and rigor compared with that of other countries. The United States has a notably fragmented vision of math and science, while other countries have a clear and unitary voice about what is expected.

Most state curriculum frameworks in the United States demand breadth over depth, calling for many more topics and skills to be taught in a given year than other countries expect. At each of the first eight grades, our states have more topics in the average mathematics framework than three-quarters of the TIMSS countries. In science the problem is not as extreme, although we still try to cover numerous topics. Teachers respond to these demands by giving superficial coverage to as many topics as they can, seldom spending enough time on any one topic to allow students to achieve mastery.

The United States also has a problem with packaging and sequencing curriculum. Subjects are taught in an episodic way, skipping from topic to topic with little relationship from week to week or year to year. Topics like basic arithmetic skills are repeated in grade after grade, well into the middle school grades. In most other countries, by contrast, instruction is concentrated on a few related topics for a meaningful period of time. Then instruction moves in sequence to a set of different topics that builds on what students presumably have already learned. As an illustration of this difference, the top-achieving TIMSS countries introduce an average of 20 new math or science topics with an intense focus between fourth and eighth grade; U.S. schools seldom give this kind of focus to any new math or science topic in these years (Schmidt and Valverde, 1997).

The U.S. curriculum is also less demanding than that of other TIMSS countries, particularly in the "wasteland" of the middle school years. While most U.S. middle school students are revisiting elementary school arithmetic and introductory science, their counterparts in the rest of the world are moving into algebra and geometry, and physics and chemistry. By eighth grade, U.S. students are out of synch with other nations, and there we stay through high school. The TIMSS video study videotaped eighth-grade math lessons three countries. The average level of mathematics actually taught varied significantly: in the United States it approximated a mid-seventh-grade level; in Germany it was a mid-eighth-grade level; and in Japan a beginning ninth-grade level (James Hiebert, University of Delaware, TIMSS Policy Forum).

The content in the TIMSS 12th-grade general knowledge tests in mathematics and science is introduced later in the United States than it is in other TIMSS countries, on average (NCES, 1998). American 12th-graders spend limited time studying advanced mathematics, even when it

is available. A relatively low proportion of U.S. students receive at least 5 hours per week of advanced math instruction. The countries with higher proportions of students receiving advanced instruction tended to do better on TIMSS (NCES, 1998).

U.S. textbooks mirror the incoherence, fragmentation, and lack of rigor found in the curriculum. No other country includes as many topics in its textbooks as we do. The textbooks used by American fourth-graders run an average of 530 pages in math and 397 pages in science, compared with an international average of 170 pages in math and 125 pages in science (Schmidt and Valverde, 1997).

How did we arrive at this kind of curriculum? Control of curriculum in the United States is dispersed among local, state, and federal governments and other less official actors. Anyone with a strong idea and a political constituency can successfully add something to the curriculum. Textbooks and packaged curricular materials constitute yet another "hidden" curriculum that is not accountable to the public. The problem is that no one has the courage to take anything away or keep out counterproductive ideas (Elmore, 1997).

Using TIMSS to Inform Changes in Curriculum

The findings suggest directions for changes in U.S. curriculum policies. As a starting point, we need more focus, coherence, and rigor in our curriculum. Higher performance expectations for students and by students, especially during the middle school and high school years, are needed to alter course-taking patterns.

We cannot solve our curriculum problems by simply moving upper-grade courses or advanced topics to lower grades or merely insisting that students take more of what is currently

available. Rather, curriculum reform means redefining content, grade by grade, to ensure coherent transitions from simple to more complex content and skills. It may also mean organizing topics in a different way.

Every national context is different, and similar curricula play out in different ways in various nations, states, and classrooms. Merely emulating the curriculum practices of other countries is not an effective strategy for the United States. Rather, we should apply the new knowledge about curriculum from TIMSS to develop new alternatives that fit our own context. The process of reforming curriculum should include strategies for monitoring how students and parents respond to innovations. In the past, some educators have pushed curriculum innovations like whole language to an extreme before we really knew which aspects worked and which did not. And they lost public support in the process (Richard Elmore, Harvard University, TIMSS Policy Forum).

Implementing these changes will require policymakers to make some difficult choices, such as deciding which topics are most important, which can be eliminated, and what should be done for students who do not learn the content at the time it is offered. Making these changes will necessitate discussions across grades and across levels of government. Right now the United States does not have professional or policy structures that encourage these kinds of conversations--suggesting we may have to create them.

While focus, coherence, and rigor in curriculum appear to be necessary conditions for raising math and science achievement, these characteristics alone cannot guarantee high achievement. The curriculum sets the stage for good instruction (Schmidt and Valverde, 1997).

What TIMSS Says about Instruction and Professional Development

Most of the rich details about comparative instructional practices come from the TIMSS videotape studies. This research shows how differently eighth-grade math is taught in the United States, Germany, and Japan (Stigler and Hiebert, 1997a; 1997b). In the United States, and to some extent in Germany, teachers usually teach a math lesson by explaining a topic and demonstrating a procedure. Then the students practice solving problems while the teacher goes around the room helping those who are having trouble. Problems that are not done by the end of class are often assigned as homework.

In Japan, students spend less time practicing routine procedures and more time analyzing and proving than their American counterparts. A typical lesson in Japan focuses on just one or two carefully selected problems, and forms a complete story with a beginning, middle, and end. Students work on a challenging problem for part of the class period, then share their solutions. The Japanese teacher uses lecture and demonstration to highlight important aspects of the students' solutions or to show another solution. Before the end of class, the teacher summarizes the main point of the day.

This is not to suggest that Japanese teachers are less active or directive than United States or German teachers. After allowing students time to struggle with challenging problems, they often follow up with direct explanations and summaries of what students have learned. Seventy-one percent of Japanese lessons contained at least some lecturing, compared with only about 15 percent of U.S. and German lessons. Japanese teachers also control the direction of the lesson in subtle ways, such as selecting an opening problem that can be solved by modifying a method developed during the previous lesson (Stigler and Hiebert, 1997b).

The TIMSS studies also asked content experts to judge the instructional quality of the taped lessons. None of the U.S. videotaped lessons were judged to be of high quality. For example, one indicator of quality is whether students engage in mathematical reasoning, such as doing mathematical proofs. Sixty-two percent of the Japanese lessons, 21 percent of the German lessons, and none of the American lessons contained instances of this type of reasoning.

Another gauge of quality is whether key mathematical concepts and procedures are *developed* through examples, demonstrations, and discussions, or whether they are simply *stated* by the teacher. For example, one teacher might state that the area of a right triangle is calculated by following a specific formula, $\frac{1}{2}$ base x height. Another might develop this procedure by showing how the formula can be derived by combining two triangles to form a rectangle. In German and Japanese lessons, math concepts and procedures were generally well developed; in U.S. lessons, these concepts were usually just stated and not developed.

How do Japanese teachers generate high-quality lessons? First, the Japanese educational system takes a very different approach to teacher professional roles than the U.S. system. In Japan, the lesson is admired and respected, like a finely crafted work of art. In the United States, we do not place very high value on the concept of a great lesson. As one indicator, 31 percent of teachers' lessons in the United States are interrupted by outside distractions, such as a loudspeaker announcement or a visitor at the door, which can disrupt the coherence of the lesson (Hiebert, TIMSS Policy Forum). This finding speaks to the need for local policies that respect instructional time and make the classroom a haven for teaching and learning.

Second, Japanese education reforms tend to focus directly on classroom teaching-- specifically, on well-designed and well-taught lessons. In the United States, we have a history of

trying to improve the quality of teaching through such indirect levers as certification requirements, teacher accountability standards, and management practices, rather than focusing on teaching itself (Stigler and Hiebert, 1997a). Japanese teachers develop and refine their lessons through a structured, collaborative process characterized by "lesson study groups"--small groups of teachers who meet weekly to design, critique, revise, and try out lessons. A group may take a year to refine three 45-minute lessons. New approaches are tested, with colleagues providing feedback. Critiques focus on student learning: what did this student say, how much did that student understand? When the revisions are finished, the lessons are shared with other teachers, and the best are distributed nationally. From this experience, the teachers not only produce better lessons, but they also learn more about pedagogy and effective practice, and engage in rich intellectual conversations. The U.S. system does not have any similar built-in mechanisms that encourage teachers to study and continuously improve the quality of their instruction.

Using TIMSS to Inform Changes in Instruction and Professional Development

New approaches to instruction will require new ways of supporting, preparing, and strengthening teachers. Holding teachers accountable to curriculum and performance standards is not enough; we must also change the processes that lead to classroom learning. The TIMSS research offers some innovative ways of thinking about teaching and professional development.

The TIMSS reports can themselves be valuable tools for professional development. Positive things can happen when groups of teachers review the TIMSS results, test items, and videotapes, and discuss the implications for their own practice. When teachers review test items, they gain a better appreciation of the difference between surface-level knowledge and high-level

knowledge. When teachers see videos of students in other classrooms solving complex problems or doing things they cannot get their own students to do, they may think critically about their own core beliefs and their own teaching practice. When teachers watch good teachers at work, they pick up nuances that are hard to capture in other ways.

Collective analysis is perhaps the most valuable part of this process. When teachers analyze the TIMSS videos as a group, they not only see vivid examples of good instruction, but they also are inspired to discuss new strategies, share their knowledge, and open their minds to new perspectives. The TIMSS experience can also stimulate teachers to do their own videotaping and critiquing. Watching oneself teach on video can be a daunting but rewarding experience if done sensitively with a peer group.

The Japanese-style lesson study group provides some interesting ideas for redesigning professional development in American schools. Bringing together teachers to develop higher quality lessons could be a logical building block for improving instructional focus and coherence in the United States and exposing more students to advanced math and science content. As with curriculum, the TIMSS researchers do not advocate importing the Japanese or any other system wholesale into U.S. culture. Instead they suggest we develop our own approaches for studying and continuously improving the teaching process.

The First in the World Consortium, for example, has established "teacher learner networks"--long-term structures to provide teacher support and school improvement. These networks, which consist of teachers from across the 20 districts, are studying specific instructional issues, discussing lessons, and analyzing and revising curriculum. Network members will then share their knowledge throughout their own districts.

One key issue is use of time in the school day. TIMSS data indicate that U.S. eighth-grade teachers teach 26 periods per week, compared with 24 in Germany and 16 in Japan. U.S. teachers spend most of their time in the classroom, with few extended periods to meet with colleagues to refine lessons or study pedagogy or content. In Germany, school is finished by early afternoon, and teachers use the rest of the day much as college faculty do, preparing for future lessons. In Japan, time for group planning is built into the school schedule. But Japanese classrooms have much higher student-teacher ratios than American schools. In fact, if one multiplies the number of periods times the number of students per period, Japanese and American teachers actually teach about the same number of students during the year. So there is a tradeoff, which raises the issue of whether American teachers would accept larger class sizes in exchange for more planning time.

To create blocks of time for planning and study, some of the First in the World Consortium districts are considering such options as employing permanent substitutes who would oversee classrooms while the regular teachers engage in professional interaction; this approach would be less expensive than hiring more teachers. Increasing professional interaction among teachers will require administrators, as well as teachers, to think in new ways. Administrators may have to make some unpopular policy decisions.

Other changes in professional development, teacher preparation, certification, and induction may be needed to move instruction in the directions suggested by TIMSS. Many teachers at all grades need more content knowledge in math and science. And in secondary school, many more teachers must be prepared to teach advanced math and science subjects. Some Forum participants suggested alternative certification as a means of bringing people with

math and science expertise into the classroom, while others cautioned that alternate pathways should be designed carefully to provide content experts with structured opportunities to develop and practice teaching skills.

Other policy options include setting more meaningful standards for what teachers should know and be able to do at both the preparation and certification stages; providing longer periods of learning and mentoring for new teachers; and raising teacher salaries to attract the best candidates. TIMSS researchers note that our lack of serious attention to teacher development is closely linked to the issue of teacher professionalism (Stigler and Hiebert, 1997b). In many other countries, teachers enjoy higher professional status and see themselves as having expertise they can contribute to improving their profession.

The TIMSS Forum participants agreed that implementing changes in instruction on a larger scale will require compatible changes in content standards, school organization, and professionalism of teachers, as well as significant funding investments.

TIMSS and Standards-Based Reforms

One reason other countries are successful in teaching a rigorous curriculum is because all the major elements of their educational systems are aligned, with uniform, high standards at the core. Not all high-achieving countries have "national" standards in the sense of being mandated from a central government. But they have reached a national consensus about what constitutes the basics, grade by grade, and this consensus guides their educational systems (Schmidt and

Valverde, 1997). This contrasts markedly with the U.S. approach wherein each locality essentially defines its own standards.

In this sense, the TIMSS findings can be seen as supporting the movement in the United States toward standards-based, systemic reform--although not necessarily in its current forms. Standards can help the United States realize many of the curricular and instructional improvements implied by TIMSS. They can keep distractors out of the curriculum. They can set high goals for students and teachers to strive for. Standards can guide compatible changes in other elements of the educational system. And they can serve as a rationale for all major policy decisions; for example, policymakers can evaluate a school finance plan or a categorical program in terms of whether it will help students achieve standards.

The TIMSS research shows that high standards are not at odds with the goal of increased educational equity. Many high-achieving TIMSS nations have implemented a "high standards for all" approach that applies to low-income and educationally disadvantaged children, as well as the more advantaged (Schmidt and Valverde, 1997). In these nations, high standards are commonly used to justify policies to provide disadvantaged children with the resources necessary to meet the standards.

Most U.S. states are already implementing systemic, standards-based reforms, but the task is complex, given our pluralistic governance system. Some versions of standards may actually be compounding the problem of curriculum overload (Elmore, TIMSS Policy Forum). States, districts, and professional organizations are producing an abundance of standards--often with real teeth--but they are not always grappling with the hard choices of which topics to cover, how to build coherence over time, or how to translate broad expectations into effective instruction.

There is a gap in the United States between what educators say--most report being well aware of current reforms in their field--and what they actually do in the classroom, as illustrated on the TIMSS videotapes. Math teachers have changed some aspects of their teaching in response to NCTM professional standards, such as using real-world problems, manipulatives, or cooperative learning, but they apparently have not changed the more fundamental scripts from which they work.

Bridging this gap will require more than just telling teachers what the expectations are. The real challenge is how to develop standards that are concrete enough to be useful to teachers--in other words, that identify a limited number of topics to be taught at each grade with advice on how to teach them. Standards should also be high by international standards. Right now this isn't the case; in several states, a student who performs at the "proficient" level on the state math assessment may actually be performing well below international standards of proficiency.

Accountability is another critical element of standards-based reform. A major policy challenge for the United States is how to encourage educators and policymakers to adopt more effective approaches to teaching and learning that include accountability for high student performance. Incentives to develop such approaches--which are central in systems like Japan's--exist only in a haphazard way in the United States.

Many states and school districts are already implementing data-based accountability. The Boston Public Schools, for example, have set up an accountability system based on test data and other performance measures, and is also developing indicators for other important school variables such as parent involvement. Schools that show improvement by these measures are eligible for awards of up to \$25,000 (Donahue, TIMSS Policy Forum).

Community District 2 in New York City is taking another approach to accountability by holding all the professionals in the system--including the principals--accountable for effective instruction, and ultimately for student performance. As this strategy recognizes, new forms of curriculum and instruction often demand new kinds of leadership. Just as teachers are being asked to take risks as part of standards-based reform, principals and superintendents must also be willing to take risks; for example, by giving teachers permission to do things differently (Elmore, TIMSS Policy Forum).

TIMSS also can help spur discussion about the respective responsibilities of the local, state, and national levels in standards-based reform. Often in the United States debate about these issues becomes hung up on the merits of "top down" versus "bottom up" reform. In many other TIMSS countries, even those highly centralized systems, curriculum and instruction is a hybrid of both top down and bottom up actions; goals and content are decided nationally, and issues of pedagogy are determined locally. Research indicates, moreover, that having national standards or similar transcendent standards need not preclude innovation; many countries with national standards have implemented effective innovations (Schmidt and Valverde, 1997).

Reforms in the United States are most likely to succeed if they combine top down and bottom up strategies, while still taking into account the decentralized nature of our system. Instead of trying to put responsibility on some other level, educators and policymakers need to focus on what they can accomplish at their own level. For example, reforms in instruction could combine statewide changes in teacher preparation and professional development with school-based changes in teacher interaction.

As the primary movers of standards-based reform, states have a particular role in implementing ideas from TIMSS. States have authority over critical elements of education, such as curriculum, teacher preparation, and funding formulas. In addition, the states are the most logical level for doing the rational, coordinated planning that is necessary to bring cohesion to school improvement.

National testing or national standards is one way to bring focus and rigor to the U.S. curriculum (Martin Orland, NCES, TIMSS Policy Forum). National tests could provide a focus for instruction and a reason for educators, policymakers, and citizens to take school reform more seriously. An alternative strategy is for states to establish collaborative structures for making their policies and standards more uniform across the nation. States could also share curricula or endorse various curricula developed by professional organizations.

The federal government can also encourage improvements that are compatible with TIMSS. In addition to continuing to disseminate information about TIMSS, the Department of Education could serve as a "supermarket of ideas" and curriculum models that are compatible with TIMSS and that states and local districts could adopt or adapt.

Using TIMSS to Build Public Support for School Improvement

A major element of school improvement is to build support for change among policymakers, parents, and the public. This support must be solid and sustained over the long term. Much as Sputnik rallied the American public to support education reforms in the 1950s, the TIMSS data can help galvanize public support for school reform. In Kentucky, for example, the commitment of the state legislature kept that state's massive educational reforms going through a change in governors. Sustaining commitment is becoming more difficult, however, with the adoption of term limits in many states.

Most people in the policy world, including state legislators, are not yet widely aware of TIMSS or its messages. Furthermore, many parents and citizens, including some well-educated people, are not sold on the need for higher standards in math and science. Informed educators, researchers, and policymakers must be prepared to explain to friends and colleagues why we should care about the performance of students in other countries. In a society that still regards children who do not speak English as their first language as a liability, the rationale for international comparisons may not be readily apparent. People in the research community should also be able to provide clear answers to such fundamental questions such as why every student should study algebra and geometry.

Researchers, federal agencies and other organizations involved in publicizing TIMSS could package the TIMSS findings in customized ways to reach different audiences, such as legislators, administrators, or school boards. Busy policymakers, for example, might appreciate a 10-minute excerpt of representative teaching practices from the TIMSS videotapes, accompanied by leading questions about policies that impede or promote various practices.

Since people sometimes base policy decisions on beliefs and emotions rather than rational arguments, TIMSS data should be presented in a way that adds emotional impact to the data. For example, presentations about TIMSS could feature several sample items from the TIMSS tests, along with the percentages of students from the United States and other countries who answered correctly (Steve Leinwand, Connecticut Department of Education, TIMSS Policy Forum). These items should be carefully selected and presented, so that people understand the math and science content behind them and the kinds of instructional practices they imply. Items could be grouped in mini-test booklets that people can take themselves or give in their districts. Emphasizing the relationship between high math and science achievement and economic development is another powerful way to persuade parents, policymakers, and especially business people of the importance of school improvement.

TIMSS as a Model for Future Policy Research

TIMSS offers a promising model of how to do sophisticated policy research while producing findings that are accessible to the people who must implement them. TIMSS represents a major advancement in the literature of comparative assessment by going beyond the horse race aspects in its design and data reporting. By comparing curriculum, textbooks, and instruction in many different countries, TIMSS has opened up dialogue about issues that were once considered risky to discuss. Through its use of video analysis, extensive questionnaires, teacher journals, and classroom time studies, TIMSS has also demonstrated how to apply new research tools to the study of teaching practice. And TIMSS has developed new structures for engaging multiple countries in true, cross-national research, which in turn has stimulated new

cross-national studies of real classroom interactions (Susan Fuhrman, University of Pennsylvania, TIMSS Policy Forum).

TIMSS raises several fascinating avenues for additional research. NCES will soon be releasing international results of a subset of performance-based items included in the TIMSS assessment. Multivariate analyses on the current TIMSS data base have yet to be done. State and local people could benefit from more detailed data about how TIMSS researchers coded the videotaped lessons; this kind of information could help educators develop a common language for analyzing classroom lessons. State and local people also express interest in video studies of fourth-grade lessons that would complement the eighth-grade video project.

TIMSS also points out the need for continuing research about what works--in other words, which practices and policies most effectively support systemic change. Right now state legislators get "mixed messages" about such critical policy issues as the educational value of reducing class size; they would like more definitive answers (Ron Cowell, Pennsylvania State Legislature, TIMSS Policy Forum).

Conclusion

TIMSS contains strong messages about the impact of policies and practices on student achievement in various nations. But these findings will make little difference unless policymakers, educators, and others use them right now to develop strategies for changing education in the United States. This does not mean importing wholesale the educational practices of higher-achieving countries. Rather, U.S. policymakers and educators need to use TIMSS as a benchmark to design their own strategies that take into account our nation's history, culture, and educational structures.

The good news is that many of the issues addressed in the TIMSS study--curriculum, instruction, classroom time, teacher roles, professional development--are already high priorities for local, state, and national policymakers, and are issues that are subject to influence by policy decisions. For example, policymakers involved in standards-based reform could take guidance from TIMSS about the need for rigor, focus, and parsimony in state standards. Those involved in teacher preparation and professional development could consider ways to provide teachers with time, administrative support, and funding to engage in collective analysis of their lessons. Policymakers at all levels of government could use TIMSS to call attention to reform needs in the different grade levels. Other issues where policy reforms could produce meaningful changes include sequencing of curriculum, student tracking, allocation and sanctity of instructional time, and avenues for professional interaction.

TIMSS is a call for immediate action. We know enough from the TIMSS data and other research to design some reasonable solutions now.

References

- Elmore, R. (1997). "Education Policy and Practice in the Aftermath of TIMSS." Paper commissioned by the Board on International Comparative Studies in Education.
- National Center of Education Statistics (1998). *Pursuing Excellence: A Study of U.S. Twelfth-Grade Mathematics and Science Achievement in International Context*. Washington, DC: U.S. Government Printing Office.
- National Center of Education Statistics (1997). *Pursuing Excellence: A Study of U.S. Fourth-Grade Mathematics and Science Achievement in International Context*. Washington, DC: U.S. Government Printing Office.
- Office of Educational Research and Improvement, U.S. Department of Education (1997). *Attaining Excellence: A TIMSS Resource Kit*. Washington, DC: U.S. Government Printing Office.
- Peak, L. (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. Government Printing Office.
- Schmidt, W. H., and Valverde, G. (1997). "Policy Lessons from TIMSS." Paper prepared for the National Governors Association.

Schmidt, W. H., McKnight, C. C., and Raizen, S. A. (1996). *Splintered Vision: An Investigation of U.S. Science and Mathematics Education*. U.S. National Research Center for the Third International Mathematics and Science Study, Michigan State University. Hingham, MA: Kluwer Academic Publishers.

Stigler, J.W., and Hiebert, J. (1997a). "Understanding and Improving Classroom Mathematics Instruction: An Overview of the TIMSS Video Study." *Phi Delta Kappan*, September 1997, Volume 79, Number 1, pages 14-21.

Stigler, J.W., and Hiebert, J. (1997b). "Findings and Implications of the TIMSS Video Study of Teaching." Paper prepared for the National Governors Association

TIMSS Policy Forum (1997). *The Implications of TIMSS for Systemic School Improvement in the United States*. Forum held October 6-7, Washington, D.C., by the National Institute on Educational Governance, Finance, Policymaking, and Management, Office of Educational Research and Improvement, and Consortium for Policy Research in Education.

Participants in the TIMSS Policy Forum

Ron Anson
National Institute on Educational
Governance, Finance, Policymaking, and
Management
555 New Jersey Avenue, NW
Washington, DC 20208

Erling Boe
Center for Research and Evaluation in
Social Policy
University of Pennsylvania
3700 Walnut Street
Philadelphia, PA 19104-6216

Ron Cowell
State Representative
Room 115 South Office Building
Harrisburg, PA 17120-2020

Maryellen Donahue
Director, Research, Assessment, and
Evaluation
Boston Public Schools
26 Court Street
Boston, MA 02108

Richard Elmore
Harvard Graduate School of Education
4th Floor, Gutman Library
6 Appian Way
Cambridge, MA 02138

Susan Fuhrman
Consortium for Policy Research in
Education
3440 Market Street, Ste 560
Philadelphia, PA 19104

John Barth
National Governors' Association
444 North Capitol Street
Washington, DC 20001-1512

Rodger Bybee
Executive Director
National Research Council
2001 Wisconsin Avenue, NW
Washington, DC 20007

Christopher Cross
President
Council for Basic Education
1319 F Street, NW
Washington, DC 20004-4171

James Ellis
20 Bailey Hall
The University of Kansas
Lawrence, KS 66045

Pascal Forgione
Commissioner
National Center for Education Statistics
555 New Jersey Avenue, NW
Washington, DC 20208

Alice Gill
American Federation of Teachers
555 New Jersey Avenue, NW
Washington, DC 20001

Daniel Goroff
Office of Science and Technology Policy
Executive Office of the President
OEOB
Washington, DC 20502

Randy Hitz
Dean, College of Education, Health and
Human Development
250 Reid, Montana State University
Bozeman, MT 59717

Tom Houlihan
Education Advisor to the Governor
Office of the Governor
116 West Jones Street
Raleigh, NC 27603

Sharon Johnson-Lewis
Council of Great City Schools
1301 Pennsylvania Avenue, NW Suite 702
Washington, DC 20004

Nancy Kober
Writer
741 Lexington Avenue
Charlottesville, VA 22902

Joyce Krumtinger
Illinois State Board of Education
100 North First Street
Springfield, IL 62777

James Hiebert
College of Human Resources, Education, &
Public Policy
University of Delaware
Newark, DE 19716-2915

Margaret Hoppe
Consortium for Policy Research in
Education
3440 Market Street, Ste 560
Philadelphia, PA 19104

Deborah Inman
Director, National Institute on Educational
Governance, Finance, Policymaking, and
Management
555 New Jersey Avenue, NW
Washington, DC 20208

Paul Kimmelman
Superintendent
West Northfield SD #31
3131 Techny Road
Northbrook, IL 60062

David Kroeze
Superintendent
Northbrook SD #27
1250 Sanders Road
Northbrook, IL 60062

Gerald Kulm
Program Director-Project 2061
American Association for the Advancement
of Science
1333 H Street, NW
Washington, DC 20005

Karen Larson
National Alliance for Business
1201 New York Avenue, NW
Suite 700
Washington, DC 20005-6143

David V. Lee
Advisor to the Governor's Office on
Education Policy
250 14th Street
Atlanta, GA 30318

Steve Leinwand
Mathematics Supervisor
Connecticut State Dept. of Education
165 Capitol Avenue, Room 203
Hartford, CT 06010

Jack MacDonald
Council of Chief State School Officers
One Massachusetts Avenue, NW
Washington, DC 20001

Margaret McNeely
National Institute on Student Achievement,
Curriculum and Assessment
555 New Jersey Avenue, NW
Washington, DC 20208

Steve Meiring
Mathematics Consultant
Ohio Department of Education
65 South Front Street
Columbus, OH 43215-4183

Martin Orland
Special Assistant to the Commissioner
National Center for Education Statistics
555 New Jersey Avenue, NW
Washington, DC 20208

Lois Peak
TIMSS Project Officer
National Center for Education Statistics
555 New Jersey Avenue, NW
Washington, DC 20208

Marjorie Petit
Deputy Commissioner
Vermont Department of Education
120 State Street
Montpelier, VT 05620

Elle Rustique-Forrester
Coordinator for Policy Development
National Commission on Teaching and
America's Future
New York, NY 10027

William Schmidt
College of Education
463 Erickson Hall
Michigan State University
East Lansing, MI 48824-1034

Susan Shaman
Director of Special Projects
The Institute for Research on Higher
Education
Suite 5A - 4200 Pine Street
Philadelphia, PA 19104

Michael Timpane
Senior Advisor, Education Research
RAND
1333 H Street, NW - 8th floor
Washington, DC 20005

Jeannie Tramill
Institute for Education Research
and Public Service
1 Bailey Hall
The University of Kansas
Lawrence, KS 66045

Gerald Wheeler
Executive Director
National Science Teachers Association
1840 Wilson Blvd.
Arlington, VA 22201-3000

Judith Wurtzel
Director, Mathematics Initiative
U.S. Department of Education
Washington, DC 20202

Susan Zelman
Deputy Commissioner of Education
Missouri State Department of Education
205 Jefferson Street, PO Box 480
Jefferson City, MO 65102