After the release of the TIMSS 1995 Video Study, many educators concluded that only Japanese teaching methods would produce high achievement. The follow-up study reported here looked at teaching methods in five additional high-achieving countries to determine whether this was indeed the case.

BY TIMSS VIDEO MATHEMATICS RESEARCH GROUP

The TIMSS 1999 Video Study is a follow-up and expansion of the TIMSS 1995 Video Study of mathematics teaching, which was itself a part of the analysis of the Third International Mathematics and Science Study. Larger and more ambitious than the first, the 1999 study investigated science as well as mathematics and expanded the number of countries from three to seven. The countries participating in the mathematics portion of the TIMSS 1999 Video Study included Australia, the Czech Republic, Hong Kong SAR, Japan, the Netherlands, Switzerland, and the United States. In this article, we focus on the mathematics lessons; the science results will be available at a later date.

Stimulated by a summary article that appeared in the Kappan and by other reports, interest in the TIMSS 1995 Video Study focused on its novel methodology and the striking differences in teaching found in the participating countries. In particular, the sample of eighth-grade members of the TIMSS Video Mathematics Research Group

The following individuals are members of the TIMSS Video Mathematics Research Group and joint authors of this article:

James Hiebert, University of Delaware
Ronald Gallimore, UCLA and LessonLab
Helen Gamier, UCLA and LessonLab
Karen Bogard Givvin, LessonLab
Hilary Hollingsworth, LessonLab
Jennifer Jacobs, LessonLab
Angel Miu-Ying Chui, LessonLab
Diana Wearne, University of Delaware
Margaret Smith, Iona College
Nicole Kersting, UCLA
Alfred Manaster, University of California, San Diego
Ellen Tseng, Open University of Hong Kong
Wallace Etterbeek, California State University, Sacramento
Carl Manaster, Christian-Albrechts University, Kiel, Germany
Patrick Gonzales, National Center for Education Statistics
James W. Stigler, UCLA and LessonLab
en, Japan taught mathematics differently from their peers in the U.S. or Germany. Because Japan was a high-achieving country and because the Japanese method of teaching resonated with many U.S. mathematics educators, it was tempting to draw the conclusion that such teaching is essential for high achievement. The TIMSS 1999 Video Study addressed this issue by sampling lessons in more countries whose students performed well on the TIMSS 1995 achievement study, in Europe as well as in Asia.

As shown in Table 1, eighth-graders in all of the countries participating in the TIMSS 1999 Video Study scored significantly higher than U.S. eighth-graders on the TIMSS 1995 achievement test, which was used to select countries for this study.

The mathematics portion of the TIMSS 1999 Video Study included 638 eighth-grade lessons collected from the seven participating countries. The Japanese lessons were the same ones collected in 1995 as part of the earlier study, but they were reanalyzed for the current study. A random sample of lessons was filmed across the school year, one lesson per teacher. Sampling information, videotaping procedures, and other methodological notes are detailed in an appendix to the report Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study, from which this summary is drawn. For a more detailed discussion of the technical aspects of the study, readers should see the companion technical report.

WHAT CAN A VIDEO SURVEY TELL US?

Although many factors inside and outside of school influence students’ level of achievement, the quality of classroom teaching is a key to improving students’ learning. But why undertake an expensive and labor-intensive videotaping and analysis of hundreds of hours of randomly selected classroom lessons from all over the world? Here are a few reasons.

• Surveying national samples of classrooms provides information about students’ common experiences. Teaching is rarely studied at a national level, but education policy is often discussed nationally. It is important to know what teaching looks like, on average, so that national discussions of teaching focus on the typical experiences of students.

• Using video makes possible a detailed examination of the complex act of teaching from different points of view. Video preserves classroom activity so that it can be “slowed down” and viewed multiple times, by many people with different kinds of expertise, making possible detailed descriptions of many classroom lessons.

• Comparing teaching across cultures allows educators to see their own teaching with fresh eyes and reveals new alternatives. Teaching is such a common activity that it can be difficult to notice all its features. Contrasts with unfamiliar methods used in other countries make one’s own methods more visible and open for inspection and provide an expanded repertoire of possible alternatives.

ANALYZING THE VIDEOS ACROSS SEVEN COUNTRIES

As the videotapes from each country arrived in Los Angeles, they were analyzed by an international team of researchers representing all the countries in the study. We started by simply watching and discussing the tapes. All of the lessons, from all countries, were easily recognizable as math lessons — students sitting at desks, a teacher at the front of

<table>
<thead>
<tr>
<th>Country</th>
<th>1995*</th>
<th>1999**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia†</td>
<td>519</td>
<td>525</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>546</td>
<td>520</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>569</td>
<td>582</td>
</tr>
<tr>
<td>Japan</td>
<td>581</td>
<td>579</td>
</tr>
<tr>
<td>Netherlands</td>
<td>529</td>
<td>540</td>
</tr>
<tr>
<td>Switzerland</td>
<td>534</td>
<td>—</td>
</tr>
<tr>
<td>United States</td>
<td>492</td>
<td>502</td>
</tr>
</tbody>
</table>

*In TIMSS 1995, Australia and the Netherlands scored significantly higher than the U.S.; Japan, significantly higher than the Czech Republic, Australia, the Netherlands, Switzerland, and the U.S.; the Czech Republic and Switzerland, significantly higher than Australia and the U.S.

**In TIMSS 1999, Australia and the Netherlands scored significantly higher than the U.S.; Hong Kong and Japan, significantly higher than Australia, the Czech Republic, the Netherlands, and the U.S.

†Nation did not meet international sampling and/or other guidelines in 1995. For details, see Alfred Beaton et al., Mathematics Achievement in the Middle School Years: IEA’s Third International Mathematics and Science Study (Chestnut Hill, Mass.: Boston College, 1996).
the room often working math problems on the chalkboard or overhead projector, some time spent discussing problems as a class, and some time devoted to students’ individual work at their desks. But each researcher also saw classroom practices that differed from those in his or her own country, sometimes strikingly so. The challenging task was to develop a reliable and consistent way of analyzing the lessons that would capture both the similarities and differences, especially those that might influence students’ mathematics learning.

Our decisions about what features to code were informed by several sources, including previous studies on mathematics teaching and learning, suggestions solicited from mathematics education experts and cultural “insiders,” and careful observations of the videotapes themselves. In the end, we coded more than 75 different features of the lessons. These features can be organized around three broad dimensions of classroom practice: the structure and organization of lessons, the mathematical content, and the way in which content is worked on during the lesson.

In order to simplify the coding process, we decided to parse the lessons into meaningful chunks and then apply codes to these smaller units of analysis. The most likely candidate for a meaningful unit of mathematics teaching is a mathematics problem. If we could mark the beginning and ending of every problem, then we could examine the kinds of problems presented, make distinctions between problems of varying levels of challenge, and characterize how they were worked on with the students. This took us beyond the more general features of teaching and into the mathematics of the lesson and the way in which it was developed.

**MAJOR FINDINGS FROM THE TIMSS 1999 VIDEO STUDY**

1. All countries share a number of teaching features. Our suspicion that much of mathematics is taught through working on mathematics problems was confirmed. At least 80% of lesson time, in every country, was spent on problems. Other similarities were also found, often by considering general ways in which the lessons were structured. As noted above, most of the lessons in all countries devoted some time to whole-class discussion and some time to individual student work. In all the countries, teachers did most of the talking, in a ratio of at least eight teacher words to every student word. These kinds of similarities probably can be explained by a convergence of global institutional trends.9 But differences between the countries quickly became evident as well.

2. High-achieving countries teach mathematics in different ways. A lingering question from the 1995 study concerns whether high achievement in eighth-grade mathematics is dependent on using Japanese teaching methods. The clear answer is that it is not. With this larger sample of high-achieving countries and with the more-detailed analyses, Japan retained its distinctive profile. A sense of Japan’s distinctiveness is conveyed

---

**FIGURE 1.**

Average Time Spent per Mathematics Problem

<table>
<thead>
<tr>
<th>Country</th>
<th>Time in minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austl.</td>
<td>3</td>
</tr>
<tr>
<td>Czech Rep.</td>
<td>4</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>4</td>
</tr>
<tr>
<td>Japan</td>
<td>15</td>
</tr>
<tr>
<td>Neth.</td>
<td>2</td>
</tr>
<tr>
<td>Switz.</td>
<td>4</td>
</tr>
<tr>
<td>U.S.</td>
<td>5</td>
</tr>
</tbody>
</table>

*Significant differences: Czech Rep., Hong Kong, and Switz. were significantly different from the Neth.; Japan was significantly different from all other countries. The tests for significance take into account the standard error for the reported differences. Thus a difference between averages of two countries may be significant, while the same difference between two other countries may not be significant.*
Defining what counts as a problem is not that simple, especially when the definition must be reliably applied by a team of coders from seven countries.

in Figure 1, which shows the time spent per problem.

What makes this finding important is that the longer time for each problem allowed Japanese students to engage in different kinds of learning experiences. For example, compared to all of the other countries, a greater percentage of mathematics problems per lesson in Japan involved proving or verifying mathematical statements, and a smaller percentage of mathematics problems per lesson were repetitions of previous problems. In addition, Japanese students spent a greater percentage of individual work time doing something other than repeating procedures they had already been taught. Instead, they spent more time analyzing new problems and developing new solution methods. The pattern of teaching found in the Japanese sample and reported in the Kappan in 1997 is a pattern of introducing a problem for the day and asking students to work on it for some time, discussing the solutions, and then presenting one or two more problems to complete the lesson — was not found in any of the other countries.

But differences among the high-achieving countries also became apparent. And these differences were found along dimensions of classroom practice that are also likely to influence the kinds of mathematics learning in which students might engage. Two examples illustrate the nature of the differences. The first has to do with the extent to which mathematics problems are presented in real-life contexts. The appropriate relationship of school mathematics to life outside the classroom has been discussed for some time, with many mathematics educators advocating the use of real-life contexts in mathematics classrooms.

Figure 2 shows the percentages of mathematics problems per lesson that were presented within real-life contexts. These include story problems and other presentations that made reference to real-life situations. The Netherlands and Japan lie at the opposite ends of the spectrum, with the Netherlands placing considerable importance on real-life connections, and Japan, along with the Czech Republic and Hong Kong SAR, placing less importance on this feature of teaching.

As a second example of differences between high-achieving countries, consider the relative emphasis placed on problems that focused on the procedural versus conceptual aspects. Like the discussion of the appropriate role of real-life situations in mathematics class, the debate between procedural and conceptual emphases has a long history. Although a compelling current view is that both procedures and concepts are critical, with no tradeoffs needed, it is still possible to ask whether classroom teachers emphasize them in different ways.

One method that can be used to determine the procedural versus conceptual emphasis of a lesson is to ask what kinds of mathematics problems are presented. For all the problems in this study that were completed with some public discussion during the lesson, the statement of the problem was classified in-

![Figure 2. Percentage of Problems Presented Using Real-Life Contexts Or Mathematical Symbols Only](image)

*Presented using mathematical symbols only: all countries were significantly different from the Neth.; Japan was also significantly different from Austl., Switz., and the U.S.
**Presented using real-life contexts: Austl. and Switz. were significantly different from Japan; the Neth. was significantly different from Czech Rep., Hong Kong, Japan, and the U.S.
†Percentages may not sum to 100 because some problems were marked as "unknown" and are not included here.
to one of three types: using procedures, making connections, and stating concepts. These describe the processes implied by the statements of the problems. “Using procedures” suggests a procedural emphasis, whereas “making connections” (between ideas, facts, and procedures) suggests a conceptual emphasis. A statement of a using-procedures problem might be, “Solve for \( x \) in the equation \( 2x + 5 = 6 - x \).” A statement of a making-connections problem might be, “Graph the equations \( y = 2x + 3 \), \( 2y = x - 2 \), and \( y = -4x \), and examine the role played by the numbers in determining the position and the slope of the associated lines.” Stating-concepts problems often are definition-like problems and so are not easily classified as conceptual or procedural. A statement of a stating-concepts problem might be, “Show the point (3, 2) on the coordinate plane.”

Figure 3 shows the percentages of problems of each type per lesson. On this dimension, Hong Kong SAR and Japan lie at opposite ends of the spectrum, with Japan emphasizing conceptual problems and Hong Kong SAR emphasizing procedural problems.

What does it mean that high-achieving countries differ, sometimes substantially, on features of teaching that can be considered critical for student learning? Does it mean that any method of mathematics teaching supports high achievement? Further insights into this question can be gained by looking at some additional results. As we will show, the higher-achieving countries shared a small number of features that distinguished them from the U.S. These were found, in part, by looking at how features, such as a procedural emphasis, were actually implemented during instruction.

3. High-achieving countries share a few, potentially important features. A close-up analysis of the kind of mathematics presented and how it was worked
Even though U.S. curricula apparently provided problems with conceptual intent, the conceptual aspect of the problems did not reach the students.

on with the students will begin to reveal a few features that distinguish some of the higher-achieving countries. Except for the Czech Republic, all the higher-achieving countries spent more time working on new content than reviewing old. The U.S. spent about the same amount of time on each. Figure 4 shows the percentages for each country, with time per lesson for new content split into introducing the new content and practicing the new content (e.g., by solving problems using new procedures).

As a second example, return to the nature of the mathematics problems presented during the lesson (Figure 3). Each of these problems was coded twice, the first time according to the way in which the problem was presented to the students and the second time according to the way in which the problem actually was discussed publicly during the lesson. Math teachers know that problems can be presented with one apparent intent (e.g., making connections between ideas, facts, and procedures) but then transformed into something different (e.g., demonstrating and practicing a procedure) — perhaps because students are struggling with the original problem and the teacher perceives that they need additional help. So not all problems retain their intent as they are worked on and discussed with students.

Recall that 17% of the problem statements in the U.S. suggested a focus on mathematical connections or relationships (Figure 3). This percentage is within the range of many higher-achieving countries. What happened to these problems?

Figure 5 shows that virtually none of the making-connections problems in the U.S. were discussed in a way that made the mathematical connections or relationships visible for students. Mostly, they turned into opportunities to apply procedures. Or they became problems in which even less mathematical content was visible — i.e., only the answer was given.

That few opportunities were afforded U.S. students to participate in discussions about mathematical relationships speaks to the importance of teaching as a key activity for defining students’ learning opportunities. Even though U.S. curricula apparently provided problems with conceptual intent, the conceptual aspect of the problems did not reach the students.

A plausible conclusion from these selected results is that teachers in the higher-achieving countries attended more to the conceptual development of the mathematics than teachers in the U.S. Even teachers in Hong Kong SAR, who appeared to focus on procedures when presenting problems (Figure 3), were found to examine conceptual underpinnings in an explicit way (Figure 5).

What would content experts say if they examined these lessons as a whole, without knowing from which country the lessons came? If the results just presented sketch an accurate picture of teaching, then experts should come to

![Figure 5. Percentage of Problems Presented as Making Connections And Solved in Each Way](image)
a similar conclusion when judging the degree to which the mathematics was developed during the lessons. And this is exactly what was found.

In order to supplement the analysis of teaching, feature by feature, a group of four mathematicians and teachers of postsecondary mathematics was asked to review country-blind written records of a random subsample of 20 lessons from each country. This was the same group that had analyzed the TIMSS 1995 video data. Because the group members had previously examined the Japanese sample, they reviewed lessons from all countries except Japan.

One of the codes the content expert group developed for this study was the degree to which mathematical concepts or procedures were developed during the lesson. Development required that mathematical reasons or justifications be given for the mathematical results presented and used. A rating of 1 indicated that a lesson was descriptive or routinely algorithmic, with little mathematical justification provided by the teacher or students for why things work as they do. A rating of 5 was assigned to a lesson in which the concepts and procedures were mathematically motivated, supported, and justified by the teacher or students.

The content expert group placed each lesson, by consensus judgment, into one of five categories: 5) fully developed, 4) substantially developed, 3) moderately developed, 2) partially developed, and 1) undeveloped. Forty percent of the U.S. lessons received a rating of 1, undeveloped; no other country received a rating of 1 on more than 15% of its lessons. Averaging the ratings for each country yielded the following, in order of mathematical development: Hong Kong SAR (3.9), Switzerland (3.4), the Czech Republic (3.3), Australia (3.0), the Netherlands (2.8), and the United States (2.4). Because of the small sample sizes, no statistical comparisons were run. But the relative ratings for the U.S. lessons are what we would have predicted from the results presented earlier.

CONCLUSIONS

The results of this study reveal, once again, the complexity of classroom teaching and renew the challenge of discerning the most effective methods of teaching to support high achievement. After the TIMSS 1995 Video Study, it was tempting to regard the Japanese method of teaching as a standard to emulate. Although the distinctive Japanese method remains worthy of careful examination, students in other countries perform well with quite different methods of teaching. After reviewing the results of the TIMSS 1999 Video Study, the matter of how to use information about teaching in other countries is much less straightforward. Videos from other countries might best be viewed as a source of alternatives to study and consider, not as a source of practices to emulate.

We conclude with two observations based on the results summarized here. The first is that high performance on international tests is not a sufficiently defined learning goal to determine the selection of teaching methods. Even if we could successfully import teaching methods from other countries, the results of this study suggest that, based on achievement levels alone, we could not choose which country's method was the "best." Many different methods are associated with high achievement. More precise and clearly articulated learning goals are needed as a first step toward analyzing the potential benefits of the wide repertoire of teaching strategies that were found in this study.

Second, it is not just the presence or absence of individual features of teaching that defines the nature of classroom practice, but how those features are implemented. Consider the types of mathematics problems presented to students (Figure 3). The U.S. falls in the middle of the distribution of higher-achieving countries with regard to problems posed with an apparent emphasis on applying mathematical procedures versus examining mathematical relationships. But when the problems were worked through with the class, the U.S. moved to the end of the continuum, with little attention paid to the conceptual underpinnings of the mathematics (Figure 5 and the findings of the content expert group). Combine this with the relatively light emphasis on new content (Figure 4), and it is clear that, as implemented, the features of eighth-grade mathematics teaching in the U.S. reinforce a more limited and thinner range of learning experiences for U.S. students than their peers in higher-achieving countries receive.

How can classroom practice in the U.S. be strengthened? After summarizing the results of the TIMSS 1995 Video Study, the same question was posed in rather urgent terms: "Our biggest long-term problem is not how we teach now but that we have no way of getting better." That is, the U.S. had no large-scale mechanism for sustained teacher learning. Although this is still the case, there is a growing sense that long-term, continuing teacher learning is a key to improving practice.

The results of this study suggest that a core of the curriculum for this kind of sustained teacher learning must focus on the whys and hows of implementing particular features of teaching. It is not enough to say that students should be taught to execute and apply procedures or to say that they should be presented with challenging and conceptually rich problems. It is by understanding why these features are important and when and how they can be implemented to achieve specific learning goals that teaching can be enriched.

The TIMSS 1999 Video Study not only highlights the importance of studying how features are implemented in the
classroom, but it also provides a range of concrete examples showing teachers in different countries implementing various features. Public release videotapes, four lessons from each country, are available on CD-ROM and in other media. They include commentary by the teacher and other educators from each country; lesson artifacts such as worksheets, textbook pages, and lesson plan documents; and text tracks of the lesson in English and in the native language of the classroom. These videos provide a rich resource for teachers that will allow them to examine their own practice from new perspectives and to consider why and how teachers in other countries implement features of teaching in different ways.

1. The TIMSS 1999 Video Study was funded by the National Center for Education Statistics and the Office of Educational Research and Improvement of the U.S. Department of Education, as well as the National Science Foundation. It was conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA), based in Amsterdam. Support for the project was also provided by each participating country through the services of a research coordinator, who guided the sampling and recruiting of participating teachers. In addition, Australia and Switzerland contributed direct financial support for data collection and processing of their respective samples of lessons. The views expressed in this article are part of ongoing research and analysis and do not necessarily reflect the views of the IEA or the funding agencies.

2. For convenience, in this report Hong Kong SAR is referred to as a country, although it is a Special Administrative Region (SAR) of the People’s Republic of China.


5. The Japanese mathematics lessons collected for the TIMSS 1995 Video Study were reexamined according to the revised and expanded coding scheme developed for the present study. As noted in reports of the 1995 study, the Japanese sample was filmed over a part of the school year rather than the whole year.


10. Stigler and Hiebert, p. 18.


17. Stigler and Hiebert, p. 20.


19. Information on how to obtain the videotapes and supplementary information is available at http://nces.ed.gov/timss.

Members of the TIMSS Video Mathematics Research Group

The following individuals are members of the TIMSS Video Mathematics Research Group and joint authors of this article:

James Hiebert, University of Delaware
Ronald Gallimore, UCLA and LessonLab
Helen Gamier, UCLA and LessonLab
Karen Bogard Givvin, LessonLab
Hilary Hollingsworth, LessonLab
Jennifer Jacobs, LessonLab
Angel Miu-Ying Chui, LessonLab
Diana Wearne, University of Delaware
Margaret Smith, Iona College
Nicole Kersting, UCLA
Alfred Manaster, University of California, San Diego
Ellen Tseng, Open University of Hong Kong
Wallace Etterbeek, California State University, Sacramento
Carl Manaster, Christian-Albrechts University, Kiel, Germany
Patrick Gonzales, National Center for Education Statistics
James W. Stigler, UCLA and LessonLab

Copyright Notice
Phi Delta Kappa International, Inc., holds copyright to this article, which may be reproduced or otherwise used only in accordance with U.S. law governing fair use. MULTIPLE copies, in print and electronic formats, may not be made or distributed without express permission from Phi Delta Kappa International, Inc. All rights reserved.

Please fax permission requests to the attention of KAPPAN Permissions Editor at 812/339-0018 or e-mail permission requests to kappan@pdkintl.org.