A Proposal for Improving Classroom Teaching: Lessons from the TIMSS Video Study

James Hiebert
University of Delaware

James W. Stigler
University of California, Los Angeles

Abstract

Results from the Third International Mathematics and Science Study (TIMSS) Video Study of Teaching show that many teachers in the United States believe they are changing the way they teach while they retain the core of traditional practice. Results from the video study, which also included teachers from Germany and Japan, along with comparisons of teacher development systems, help to explain the persistence of traditional teaching methods and point to a new model for improving teaching. After presenting results from the video study regarding teachers’ perceptions of the effects of recent reforms on their practice, we propose a research and development system for improving teaching that builds on the Japanese process of lesson study. We describe the process as it works in Japan and outline the conditions that would enable such a process to function effectively in the United States.

Results from the Video Study of Teaching, conducted as part of the Third International Mathematics and Science Study (TIMSS), confirm what has been reported in many observation and case studies: efforts to reform education in the United States often influence classroom teaching at the margins, if at all (Cuban, 1993; Education Evaluation and Policy Analysis [EEPA], 1990; Ferrini-Mundy & Schram, 1997; Griffin, 1995; McLeod, Stake, Schappelle, Mellissinos, & Gierl, 1996; Spillane & Jennings, 1997; Stake & Easley, 1978; Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999; Stigler & Hiebert, 1997, 1999; Tyack & Cuban, 1995). In this article we present data from the TIMSS Video Study on teachers’ views of how the reform movement in mathematics education has affected their teaching and then abstract from the study an explanation for the
apparent mismatch between what teachers say and what they do. Based on these observations and on a comparison of teaching improvement systems in Japan and the United States, we propose a model for improving teaching that requires a fundamental shift away from traditional American methods of reform.

The TIMSS Video Study

The video study of classroom mathematics teaching was included as part of TIMSS because the study planners were interested in collecting information on school factors that might be related to the anticipated achievement differences. Instructional practices were assumed to be a central feature of schooling. The goals of the video study were (1) to learn how eighth-grade mathematics is taught in the United States, (2) to compare American teaching with teaching in Germany and Japan, and (3) to examine the influence of the reform recommendations on classroom practice.

Overview of Methods

Sample. The video sample was constructed to be nationally representative by drawing a random subsample from the full TIMSS sample. Not only were specific teachers selected but specific class periods as well. No substitutions were allowed, either by another teacher within the same school or by another class period with the same teacher. The final video sample included 231 classrooms (one lesson per classroom): 100 in Germany, 50 in Japan, and 81 in the United States.

Data collection. Data were collected during the 1994–1995 school year and included the videotaped lesson and the teacher’s responses on a questionnaire. To get useful, comparable video in all classrooms, numerous videography issues, such as where to point the camera at any given moment, had to be anticipated and resolved in advance. Standardized procedures for camera use were developed, tested, and revised, and videographers were trained.

Only one camera was used in each classroom, and it focused on what an ideal student would be focusing on—usually the teacher.

Teachers filled out a questionnaire about the taped lesson describing the goal of the lesson, its place within the current sequence of lessons, and how typical the lesson was. The questionnaire also asked how aware the teacher was of “current ideas” on teaching and how the taped lesson matched those ideas.

Lesson coding. Tapes were encoded and stored digitally on CD-ROM and were accessed and analyzed using multimedia database software developed for this project. All lessons were transcribed and then analyzed on a number of dimensions by teams of coders who were native speakers of the three languages.

The coding scheme was designed to capture those aspects that might make a real difference in the mathematics the students were learning and to yield valid and informative descriptions of instruction across the three cultures. The first goal was addressed by consulting the research on teaching and learning mathematics as well as reform documents such as Professional Standards for Teaching Mathematics (National Council of Teachers of Mathematics [NCTM], 1991). To address the second goal, a team of six code developers, two from each country, spent the summer of 1994 watching and discussing 27 field test tapes. Out of these intensive discussions emerged the initial coding system. The system was refined regularly as the primary coding team began applying it to the actual study tapes and as intercoder reliability checks revealed categories needing further definition. (For more information on background, sampling, and procedures of the video study, see Stigler et al. [1999].)

Results: Believing in Reform but Retaining Practice

Beliefs about implementing reform recommendations. Based on teachers’ re-
sponses on the questionnaire, most U.S. teachers are aware of "current ideas about teaching and learning mathematics" and believe that these ideas are reflected in their teaching. Interestingly, more U.S. teachers responded positively to these questions than did their German and Japanese colleagues.

Figure 1 shows that 95% of U.S. teachers said they were somewhat or very aware of current ideas about teaching mathematics. Responding to open-ended questions, they said that outside-of-school workshops and seminars, and school-based programs were the most likely sources of information about current ideas. Most U.S. teachers (59%) mentioned the NCTM Standards (NCTM, 1989, 1991) or another NCTM publication when asked what they had read to stay informed about current ideas.

Not only did U.S. teachers profess awareness of current ideas, most believed they were implementing such ideas in their teaching, including in the lesson that was videotaped. As seen in Figure 2, 70% of U.S. teachers said that the videotaped lesson was at least a "fair amount" in accord with these ideas. Again, this percentage was higher than for their German and Japanese colleagues.

**Classroom practice.** Despite the strength of U.S. teachers' beliefs that they know about current ideas, as expressed in the NCTM Standards, for example, and are implementing them in their classrooms, the evidence from their actual practice suggests otherwise. Full reports of the video data, as well as summaries, are available elsewhere (Manaster, 1998; Stigler et al., 1999; Stigler & Hiebert, 1997, 1998b, 1999), but a few indices will convey the nature of the results.

One indicator showed that, in the U.S. lessons, teachers usually just stated mathematical concepts rather than developing them (see Fig. 3). For example, U.S. teachers were more inclined simply to state the rule for adding exponents when multiplying algebraic expressions than to show why the rule makes sense. The reverse tendency was found in other countries.

A second indicator revealed no instances of working through proofs or reasoning deductively in U.S. lessons. In contrast, 10% of
German lessons and 53% of Japanese lessons included proofs. A third indicator showed that almost all of U.S. students' time during seat work (96%) was devoted to practicing procedures. Japanese students, in contrast, split their time evenly between practicing procedures and doing original work—inventing procedures or analyzing mathematical situations in new ways.

Changes at the margins. How is it that classroom practice in the United States seems to be at odds with the core of current ideas of teaching—in how mathematical ideas are developed and in the kind of thinking in which students engage (NCTM, 1989, 1991)—when many teachers believe they are faithful to these ideas? Additional findings from the video study provide some initial clues. When teachers who believed their taped lesson was consistent with current ideas were asked to describe specific aspects of the lesson that exemplified these ideas, they identified a wide range of features, the vast majority of which could be classified into three categories: 38% of responses focused on real-world and/or hands-on activities; 31% described cooperative learning; and 19% talked about mathematical thinking and problem solving. This means that more than two-thirds of the teachers focused on activities or organizational features of the classrooms, features that can be implemented easily without altering the way in which students and teachers do mathematics.

To investigate further, the lessons of two groups of U.S. teachers were reanalyzed: those who responded "a lot" when asked the degree to which the taped lesson was in accord with current ideas (N = 22) and those who answered either "a little" or "not at all" (N = 19). No significant differences were found between these groups on any of the indicators described earlier. Across all indicators, only the following statistically significant differences appeared.

The first group, the "reformers," allocated a higher percentage of lesson time to seat work (43%) than did the second group, the "nonreformers" (29%). Furthermore, 66% of the reformers arranged students in small groups at some point during this seat work time, whereas only 24% of the nonreformers did so. Finally, the reformers used the textbook less often (18% of lessons) than the nonreformers (63% of the lessons). Although the absence of other between-group differences might have been due, in part, to the lack of statistical power because of the small sample size, we do not believe this is the primary reason. In our judgment, there were few distinctions between the lessons of these two groups of teachers. The distinctions that did exist, like those above, were at the margins.

An example helps to clarify the point. In one lesson we videotaped, the teacher asked students to work in groups to "try to figure out the name of a three-dimensional object with 12 faces." The students had been studying properties of polyhedra. The students queried each other within their groups about whether anyone knew the name. One student did and informed her group mates; the rest of the students did not. In either case, the discussions in most groups turned quickly to nonmathematical topics. The teacher then asked the groups to volunteer their "findings." One group suggested dodecahedron; the other groups reported that they had not been able to think of the name. The task, of course, simply asked students to recall a label, so no reasoning was required, or even possible. Although students were working in groups, the nature of the mathematics they were doing had not changed. Arranging students in small groups was a change at the margins.

The fact that teachers' practice can be aligned with their descriptions only in superficial ways is not a new finding. Many researchers have found similar occurrences, especially in this era of educational reform (EEPA, 1990; Ferrini-Mundy & Schram, 1997; Griffin, 1995; Spillane & Jennings, 1997). Why is this so common?

SEPTEMBER 2000
Lessons about the Nature of Teaching

Two truths about teaching, identified by a number of educators in a variety of contexts, find new meaning in the video data and help to explain the resiliency of traditional instruction. One is that teaching is a system of interacting elements, not just a collection of features (Jackson, 1968; Sarason, 1997). The second is that this system is embedded in a cultural context that overdetermines its nature (Cazden, 1988; Cazden, John, & Hymes, 1972; Gallimore, 1996; Tyack & Cuban, 1995).

Teaching Is a System

The act of teaching is a complex interaction of numerous elements: the behaviors of the teacher, the expectations and behaviors of the students, the physical setting, the participation structures that guide social interactions and discourse, the lesson activities, the curriculum, the materials, and so on. When the system runs smoothly, these all reinforce each other and move the system toward shared goals. This makes it difficult to change just one element. If the changed element does not function in the same way as its predecessor, either the system will be impaired or it will swallow up the new element and run as before.

An example of changes impairing the system comes from a group of teachers who were meeting regularly to analyze videotapes of alternative mathematics instruction. One sixth-grade teacher decided to change his traditional approach to a more problem-solving approach shown in the tapes. He carefully planned a new lesson to try out the approach. Instead of asking short-answer questions, he began his lesson by presenting a challenging problem and asking students to spend 10 minutes working on a solution. Although the teacher changed his behavior to correspond with the teacher on the videotape, the students, not having seen the videotape and discussed its implications for them, did not change their behavior. They waited for the teacher to show them how to solve the problem. The lesson did not succeed.

The kinds of changes envisioned by reformers require changes not only in the features of instruction but in the very goals of the teaching system. Although some teachers in the video study did change features of their instruction, very few appeared to shift their goals toward deeper mathematical understanding. One item on the questionnaire asked, “What was the main thing you wanted students to learn from today’s lesson?” As seen in Figure 4, 61% of U.S. teachers focused on the development and execution of skills. It is no wonder that the changes they wrote about in their questionnaire responses ended up looking superficial in practice; many teachers were working toward the same goals.

The videotaped lesson summarized earlier is a good example. Students were working together in small groups and were asked to share their thinking, lesson features that could be vehicles for engaging mathematics in a deeper way. But the teacher’s goal still was rather narrow—learning terms and definitions—so these classroom changes became changes about form rather than substance.

Teaching Is a Cultural Activity

Cultural activities are common everyday routines that have evolved, in ways that are consistent with underlying assumptions and beliefs, to deal with recur-
ring tasks. Grocery shopping, driving cars, and family dinners are all cultural activities for many Americans. A significant feature of cultural activities is that they are so widely shared that they are nearly invisible. People do not stop to think about them unless they differ significantly from their expectations, as might occur when traveling to a different culture.

In our view, teaching is a cultural activity in just this sense (Stigler & Hiebert, 1998a, 1999). Teaching has evolved, in ways consistent with society's most common beliefs about how and what students should learn, to deal with the task of one teacher educating 30 students during 50-minute lessons. The way in which teaching is conducted within a culture is so widely shared that anyone who has grown up in the culture probably could enter a classroom tomorrow and act like a teacher. (Indeed, that is the explicit presumption of many alternative teaching certification programs that require little or no pedagogical training.)

This "script" for teaching, that everyone shares, is so common that it is invisible, at least in part, even to those who teach. This yields a pregnant irony in American education: although teachers presumably are free to create their own teaching methods, they are, in fact, bound by the U.S. cultural script for teaching.

Evidence for cultural scripts of teaching is contained in the patterns of teaching that are visible across the lessons within countries. When we began the video study, we were not aware of the strong pull of cultural scripts and we expected to see great variability in teaching methods in the United States. And there was variation. Some teachers began the lesson with warm-up activities, others by checking homework, still others by working the examples for the day. Some teachers demonstrated procedures to solve problems in a lecture form, others by asking students to fill in the next step, still others by passing out a reference sheet and walking students through the examples. Some teachers asked students to work on the assigned problems in small groups; others asked students to work independently. What did not strike us, until we watched lessons from other countries, is that U.S. teachers almost always showed students what to do, with relatively little mathematical development, and then assigned problems similar to the worked examples. This pattern took on real significance when we detected very different patterns in the other two countries (Stigler & Hiebert, 1999). In the end, the differences within countries paled in comparison to the differences among countries.

The Need for a New Model of Improvement

The Tradition of Reform, American Style

One way of explaining America's continuing dilemma of reforming again and again with relatively little change in classroom practice (Cuban, 1990, 1993; Elmore & McLaughlin, 1988; Fullan, 1991; House, 1996; Knapp, 1997) is that many reformers in the United States have been using methods of change suited to noncultural activities. Noncultural activities can be characterized, in part, as those that are learned only through deliberate, systematic study, such as filing taxes for a new small business or, for older persons, using a computer. The methods traditionally used to improve teaching are more appropriate for these kinds of activities, methods such as composing and distributing written documents from experts; analyzing the activity into separate parts and replacing individual parts (e.g., the curriculum, classroom organization, such as cooperative groups, and so on); dividing labor so that some people create the new recommendations and others implement them; and assuming that methods that work well in one context will work well in another.

Improving Systemic, Cultural Practices

Improving teaching at its core—in the way students and teachers interact about...
subject matter—will require methods attuned to complex cultural activities. These will be methods that take a long-term view, that recognize the gradual and incremental nature of change (Gallimore, 1996; Tyack & Cuban, 1995). In spite of the waves of reform, this is, in fact, how teaching has changed. Dramatic and fundamental improvements can occur, but these will result from accumulating small changes in core classroom processes over time.

Methods that are most effective will place a premium on working out shared goals for the system (Joyce, Wolf, & Calhoun, 1993; Simon, 1996). Educators generally agree that improving students’ learning is the ultimate goal, but this goal is lost as particular features of instruction become ends in themselves. Success is judged mistakenly on whether teachers are using, say, cooperative groups or real-life problems. When these changes become ends in themselves, they can be implemented without touching the core processes in the classroom (Cuban, 1993; EEPA, 1990). They become changes at the margins that have little effect on students’ learning. Joyce and colleagues (1993) noted, after reviewing school reform efforts, that they did not find a single case in the literature where students’ learning increased without such increase having been a central goal of the reform.

In addition, effective methods will afford opportunities for teachers to learn in contest (Borko & Putnam, 1998; Darling-Hammond, 1998; Darling-Hammond & Sykes, 1999; Lave, 1988). It is widely agreed that students learn best when taught in the context in which their knowledge is to be applied. But this principle has not been applied when teachers are the learners. Teachers perform their work in classrooms but rarely does their training occur in classrooms (Sarason, 1983; Schaefer, 1967).

Finally, methods that work will place decision-making power in the hands of those who have the best information. In analyzing how complex systems function most effectively for achieving their goals, Simon (1996) argues that a major challenge is managing the information-processing limitations. No one person in the system can have access to, or process, all of the information. Systems improve their functioning by locating decisions where the most relevant information is available. This principle is rarely employed in education. Rather than locating substantive decisions for improving teaching within the schools and classrooms, where teaching occurs, the analysis and recommendations often are pushed up the bureaucratic ladder (Sarason, 1997).

Methods that instantiate these principles for improving classroom teaching have been demonstrated in a variety of local settings, usually initiated by special research and development efforts (Borko & Putnam, 1998; Brown, Smith, & Stein, 1996; Cobb et al., 1991; Cognition and Technology Group at Vanderbilt [CTGV], 1997; Darling-Hammond, 1998; Elmore, Peterson, & McCarthey, 1996; Fennema et al., 1996; Little, 1982, 1993; Schifter & Fosnot, 1993; Swafford, Jones, & Thornton, 1997; Tharp & Gallimore, 1988). Unfortunately, the average teacher in the United States still does not have access to these essential learning opportunities (Ball, 1996; Cohen & Hill, 1998; Darling-Hammond, 1997; Elmore, 1996; Lubeck, 1996). How can a system for improving teaching be built that would extend the opportunity for all teachers to improve classroom teaching gradually?

A Research and Development System for Improving Teaching

A successful research and development system for improving teaching must not only be effective for individual teachers and schools who receive intensive assistance, it must also be capable of scaling up. There must be a way to start small and gradually expand. This requires two things. First, the system must be able to learn from its own experiences. Unfortunately, the American educational system is better known for its
pendulum swings than for its ability to show cumulative progress (Wilson & Daviss, 1994). As Sarason (1997) noted, "There are no means, procedures, forums through which the system 'learns' " (p. 37). But improving teaching on a large scale requires precisely such means. There must be a way of accumulating the knowledge gained over time and sharing it across space and time.

The second requirement for scaling up is that the system must be modular. It must be capable of working well, as a system, on a small scale and then growing cumulatively, working equally well on a larger scale, until it can serve the entire state or nation.

The challenge, then, is to design a research and development (R & D) system for improving teaching consistent with the general principles outlined earlier, with local instantiations containing the features already found to be effective for professional development but with an activity structure for teachers that affords shareable knowledge and scaling up.

Lesson Study: Japan's Countrywide System of Improvement

One way to imagine alternatives to U.S. models for teaching improvement is to step outside the American culture and look at how others handle similar issues. Japan provides an eye-opening contrast for several reasons. One is that the video study found many high-quality lessons in Japan, lessons that were, in some ways, more consistent with current reform recommendations than lessons in the United States. How do teachers learn to teach in this way? A second reason for considering Japan's response to improving teaching is that they have institutionalized a system of improvement built squarely on the idea that teaching is a complex, cultural activity. Their countrywide system is consistent with the principles proposed earlier for changing cultural activities and manifests the features of effective teacher development found in this country. This makes the contrast especially intriguing.

The process of lesson study. Many Japanese teachers participate, throughout their careers, in a continuing in-service program built around the lesson-study group (Lewis & Tsuchida, 1997, 1998; Shimahara, 1998; Shimahara & Sakai, 1995; Takemura & Shimizu, 1993; Yoshida, 1999). Small groups of teachers meet regularly, once a week for several hours, to collaboratively plan, implement, evaluate, and revise lessons. Many groups focus on only a few lessons over the year with the aim of perfecting these. A group of fourth-grade teachers, for example, might be dissatisfied with their current lessons on adding fractions with unlike denominators. They begin the process of improving them by reading about what other teachers have done, what ideas are recommended by various educational groups, what has been reported on students' learning of this topic, and so on. They design several lessons, one group member tries them out while the others observe and evaluate what works and what does not, and they revise the lessons. They often base their changes on specific misunderstandings students evidence as the lesson progresses. They might change the wording of the opening problem, or the kinds of follow-up questions they ask, or they might use the information about the methods the students are likely to invent to change the order in which methods are presented during the whole-class discussion. Then, they try out the lessons again, perhaps with other teachers watching. This process may go on for several months, or several years. When the replacement lessons are ready, complete with development and test information, including expected student responses for each question and problem, they are shared with other teachers and other schools.

Lesson as the unit of analysis and improvement. All efforts to improve teaching must start somewhere in their attempts to change the system. For example, Cognitively Guided Instruction helps teachers un-
nderstand students’ thinking in particular mathematical domains (Fennema et al., 1996). The Jasper Project emphasizes the nature of the curriculum and the environmental supports needed to engage students in solving genuine problems (CTGV, 1997). Japan’s approach is somewhat different. Japanese teachers see the lesson as the place where all of the relevant factors are woven together—goals for students’ learning, attention to students’ thinking, analyses of curriculum and pedagogy, and so on. Lesson study requires focusing on the interactions among these elements. These different approaches to dealing with the systemic nature of teaching raise an important question: What is the most appropriate unit for analyzing and improving teaching?

We believe the approach of the Japanese teachers is worth pursuing because, given the way teaching is practiced in most countries, the classroom lesson becomes the smallest unit that preserves the system of teaching. The lesson is where the interactions occur among all the individual features of teaching. Students’ thinking, curriculum projects, pedagogical moves all must come together in a lesson. Multilesson topical units or other larger entities also would contain these interactions but they can overwhelm efforts to analyze classroom processes in detail. The classroom lesson seems just at the right level—large enough to capture the system, small enough to afford analysis and improvement.

Of course, teaching is more than individual lessons. Lessons must be sequenced and connected in ways that allow large ideas to be developed over time. Lesson study attends to relations between lessons as well as the makeup of individual lessons. But the individual lesson is the primary unit.

A corollary of the fact that lessons capture the system of teaching is that lessons afford teachers the opportunity of working toward improving teaching through careful planning rather than relying solely on becoming more skillful on-the-fly decision makers. Planning that includes anticipated student responses and productive interactions is impossible if individual features of teaching, for example, learning to ask higher-order questions, or using a new material, are singled out. Detailed planning requires anticipating how the individual features will work together. Lessons contain these interactions but are manageable for designing, testing, revising, testing again, and refining. Through this process, teachers learn what kinds of interactions they are likely to have and they can plan how to weave these together to address the points they wish to make.

It is useful to notice that the process of lesson study would have yielded something different from the failed lesson on problem solving described earlier. Because much attention is given to anticipating students’ responses, the collective knowledge of teachers would have predicted, almost certainly, that the students were not yet prepared to respond in the new ways the lesson demanded.

Lesson study generates shareable knowledge. Sharing knowledge about teaching is not a simple task. Codifying discoveries as written recommendations distributed in research and policy reports has produced disappointing results. Teachers interpret the recommendations in widely different ways and have trouble adapting them to the realities of classroom life. Pulling out specific tips and techniques and favorite activities for sharing during workshops can lead to the changes at the margins described earlier.

The lesson study process, as practiced in Japan, involves collecting and sharing information in a very different form. The report of a group’s 1–2-year effort contains descriptions of the learning goals and activities, the rationale for the lesson design, anticipated responses of students, and suggested responses by the teacher. These reports are theories linked with examples. Hypotheses about how to help students reach particular learning goals are linked to
actual lessons and students; practical suggestions are linked to the teachers' theoretical analysis of the learning goals and ways in which students might achieve them. This blend of theory and example gives other teachers the information they need to relate the work to their own classrooms and to think through the changes that might be required to achieve the goals in different contexts.

The knowledge being shared through lesson study is not just collections of lesson plans that teachers can pull off the shelves and use. The goal of lesson study is not just to produce lessons that can be copied but to produce knowledge about teaching upon which colleagues can build (see Ball & Cohen, 1996). Such a knowledge base grows as teachers reflect on and improve what others have done, working to understand the basis for the improvements.

**Lesson study as teacher development.**

As practiced in Japan, lesson study reverses the relation prevalent in the United States between improving teaching and improving teachers. Working on improving teaching yields teacher development, rather than vice versa. Designing and testing lessons provides a rich context in which teachers can improve their own knowledge and skills. While teachers are producing shareable work, they are engaged in exactly the kind of learning that they need to become more effective teachers. They must learn more about the subject, about their students' thinking, about alternative pedagogies. The lesson study process assumes this kind of learning and offers opportunities to learn through collaboration with and observation of colleagues, as well as time for study and reflection.

Not surprisingly, lesson study provides the backbone for Japan's system of teacher development (Lewis & Tsuchida, 1997, 1998; Shimahara, 1998; Shimahara & Sakai, 1995; Takemura & Shimizu, 1993; Yoshida, 1999). The full system of school-based professional development is referred to as **kounaiikenshuu**. This includes informal or teacher-initiated activities as well as formal, government-mandated programs. Perhaps the most common and valued aspect of **kounaiikenshuu**, at least at the elementary and middle school level, is teacher-initiated lesson study (**jugyou kenkyuu**). It is through this process that many Japanese teachers believe they learn to teach and, indeed, become teachers. As one Japanese teacher said, when asked why she invests so much time in improving lessons, "Why do we do research lessons? I don’t think there are any laws. But if we didn’t do research lessons, we wouldn’t be teachers" (Lewis, 1997, p. 3).

**Lesson Study: The Engine for an R & D System in the United States**

Can the process of lesson study, or something like it, work in the United States? There is reason for skepticism, because it would require significant changes in the culture of U.S. schools. More than that, it would require changes in the way teachers think about teaching and planning for teaching. Many teachers in the United States do not plan instruction by constructing detailed lesson plans linked to learning goals (Jackson, 1968; Richardson, 1994; Shavelson & Stern, 1981).

At the same time, there are reasons for optimism. Lesson study fits well with two features of the current American educational landscape. One is the spreading culture of teacher as researcher (Burnaford, Fischer, & Hobson, 1996; Cochran-Smith & Lytle, 1990, 1993; Franke, Carpenter, Fennema, Ansell, & Behrend, 1998; Hollingsworth & Sackett, 1994; Richardson, 1990, 1994; Wagner, 1997). The lesson study process is consistent with building a disposition of inquiry among teachers, a disposition toward investigating one’s own practice.

A second feature of the current American educational scene is that islands of change are springing up around the country. Individual schools and districts are setting priorities on improving teaching and

---

**SEPTEMBER 2000**
on enhancing teachers' opportunities to learn. Although some are using traditional workshops (McDiarmid & Kelly, 1997), others are finding that alternative forms for learning are more effective. The local forms differ from each other in some respects, but, as cited earlier, many are consistent with the principles for improving complex human practices. Northeastern Elementary School (Elmore et al., 1996) and Portsmouth Middle School (Stein, Silver, & Smith, 1998), for example, independently developed school-based, collaborative, curriculum-linked, ongoing professional development programs. Although neither is identical to the lesson study process we described, the process would fit well within the emerging traditions at such schools.

There is another reason for optimism about the feasibility of lesson study in the United States. The lesson study process would appeal to teachers caught in a persistent dilemma, now heightened by calls for reform-minded instruction (NCTM, 1991). On the one hand, teachers understandably choose methods that they feel they can master and implement effectively (Doyle & Ponder, 1977; Smith, 1996). On the other hand, the new methods are said to involve, necessarily, a great deal of uncertainty (Ball, 1993, 1996; Cohen, 1988; Lampert, 1985; McDonald, 1992; Schiffrer, 1996; Schifter & Fosnot, 1993). Students are encouraged to become active participants in the classroom and this, presumably, means that teachers must expect the unexpected. This more ambitious teaching places a premium on the individual teacher's skill in orchestrating the unpredictable.

Lesson study provides a way of dealing with this dilemma by shifting a large part of the burden from extemporaneous decision making during the lesson to careful investigation and planning before the lesson. Possessing information about how students are likely to respond at critical points allows the teacher to implement adventurous lessons (for the students) while retaining control. Lesson study allows teachers to reduce uncertainty through collaborative preparation rather than through compromising their learning goals for students.

There is likely to be resistance to this resolution, stemming largely from an unfortunate notion in the United States that reducing uncertainty in teaching is synonymous with teacher-proofing the classroom and deskilling teachers (Eisner, 1979; McDonald, 1992). The coupling of these notions may be due, in part, to the U.S. tradition of reform, which often has attempted to impose teaching practices from outside the classroom. It is an unfortunate coupling because it reinforces the belief that improvement depends on the extemporaneous, natural artistry of individual teachers. Once the tradition of imposing teaching practices is changed and teachers are allowed to improve teaching through extensive, collaborative preparation, testing, and refinement of lessons, then it is possible to see that reducing uncertainty empowers teachers rather than deskills them.

What Else Does the R & D System Need?

It would be foolish to expect a lesson study process to work effectively if dropped into the current U.S. culture of educational improvement, even with the trends noted above. Several initiatives need to be launched concomitantly to enable lesson study to take hold.

Set clear learning goals for students. Lessons are improved only with respect to specific learning goals. The standards movement in the United States is, in part, an effort to set learning goals for students. The contribution of this activity to improving teaching depends on whether the standards are expressed clearly enough that teachers can use them for planning instruction.

Not only must the learning goals be clear, they also must be widely accepted. The success of the lesson study process depends on teachers working together, with the same goals in mind. Otherwise, the
products of one group are of little value to others. The United States does not have national learning goals and, in some cases, learning goals are not even shared within a district (Spillane, 1998). Policy makers tend to rationalize this situation by referring to local and individual needs of different groups of students in different parts of the country and even in different schools and classrooms. Although local contexts may differ in some respects, shared learning goals are a minimum requirement for teachers to collaborate effectively. This does not mean that a set of national goals needs to be enforced, but it does mean that shared goals must be developed at least at the district level.

It should be noted that lesson study, as we propose it, is driven by achieving clear learning goals for students, not by implementing particular pedagogies or curricula, nor even by meeting specific sets of teaching standards. Due to the nature of the lesson study process, pedagogies and curricula that most effectively help students achieve the learning goals will become, over time, the ones that are most widely shared and implemented. If the learning goals accepted within a school district are narrow sets of skills, for example, it is likely that lesson study would yield more effective versions of some traditional lessons. If, however, the district adopts more ambitious learning goals that include conceptual understanding and critical reasoning, it is likely that lesson study would yield, over time, lessons that employ more ambitious curricula and pedagogies.

Given the fact that in the United States curricula and assessments usually are chosen at the district or state level, not by individual teachers, it follows that lesson study can work well only if the district’s curricula and assessments are consistent with the learning goals. If teachers are to work on lessons that more effectively help students achieve the goals, they must have access to curriculum materials designed for this purpose. Otherwise, teachers must design lessons to achieve one set of learning goals using curriculum suited for a different set.

Similarly, teachers must be able to measure students’ progress toward the goals using assessments aligned with these goals. The critical function of appropriate assessments may not be immediately apparent but cannot be overstated because assessment information is the primary means by which teachers can evaluate their instruction. If teachers are to improve their practice, they must have access to a steady flow of information about the effectiveness of their lessons. They must be able to sort out improvements from mere changes. Clear goals, linked to assessments, and both tied to the curriculum, provide the context teachers need to plan and evaluate their teaching.

Restructure schools as places where teachers can learn. An underappreciated fact in educational change is that every recommendation for improving teaching requires teachers to learn (Cohen & Barnes, 1993). From the principles for improving complex human practices, such learning makes most sense in the context of practice. This points to schools as the ideal sites for learning (Eisner, 1979; Sarason, 1983; Schaefer, 1967). But the concept of schools as places for teachers to learn is nearly foreign in American education. “Schools have no well-defined structures for helping teachers learn from the everyday experience of teaching” (Nemser, 1983, p. 163). This may be explained, in part, by a set of cultural beliefs about teachers and teaching that minimize the importance of teacher learning. Good teachers are “naturals”; they are born, not made. “We regard inspired and demanding teaching as an individual trait, much like hair color or shoe size, rather than as a professional norm” (Elmore, 1996, p. 5). This attitude undermines support for investing in ongoing learning opportunities for teachers.

In order for schools to become sites for teacher learning, numerous changes are
needed, changes that must be supported by key personnel—parents, principals, super-
intendents, and school board members. We
mention here only issues of time and the
leadership role of the district in restructur-
ing. (For a more complete treatment of these
issues with respect to developing a culture
of lesson study, see Stigler & Hiebert [1999].)

Schools will not become learning places
for teachers unless teachers have time—
time for collaborative planning and inves-
tigation, time to work together to analyze
and design and test new lessons. This time
must be built into their weekly schedules.
But finding the time presents a challenge to
the typical school district. The cost seems
prohibitive. However, there are low-cost
options that can work and that are working
in a number of school districts across the
country. Darling-Hammond (1997) details
the way in which a variety of school dis-
tricts have created time for teacher learning.
For example, New York City's Community
District No. 2 has targeted teachers' devel-
opment as its highest priority and has re-
structured schools to provide time for
teacher collaboration and even cross-school
observations.

It is worth noting that significant
changes can be achieved without additional
resources. Now imagine what would hap-
pen if Americans took the millions of dol-
ars spent every year to reform education
and used them to provide the time and re-
sources teachers need to improve teaching.
The United States could support a vigorous,
highly focused research and development
system.

Only if districts take the lead is it likely
that schools will be able to implement the
changes needed to enable teachers to collab-
orate productively. For several reasons, the
district probably is the smallest unit that
can restructure successfully (Joyce et al.,
1993). First, the district office is where most
of the budgetary control resides, so deci-
sions about funding and staff allocations
usually are made here. Second, district-
wide support provides teachers with a criti-
cal but often overlooked opportunity for
professional growth. The R & D system we
propose asks teachers to take on a challeng-
ing task—to improve the quality of teach-
ing. This means more than improving
teaching in their own classrooms; it means
raising the standard of good teaching
within the profession. This demands that
teachers work together, sharing what they
learn in their classrooms to leverage im-
provements on a wider scale. Improving the
professional norm requires that teachers see
themselves as part of this larger effort, as
responsible for building the knowledge
base for teaching. Teachers must have the
opportunity to enlarge their horizons, to see
beyond their own classrooms and even
their own schools. They must experience
the power of sharing knowledge across
schools.

Foster a steady infusion of new ideas.
Systems, and individuals, have a difficult
time learning without a steady diet of vari-
ability (Siegler, 1996). Innovations, alterna-
tive images, different ways of doing things,
new information, all are needed to create
new experiences from which the system can
learn. Yet, because teaching is a cultural
activity, based on methods widely practiced
within a culture, teachers will not fre-
cently encounter unique and different
ways of teaching. New ideas must be
sought out.

Fortunately, new ideas for teaching in
the United States emerge from a variety of
sources. Individual teachers who have bro-
en from convention and experimented
with alternatives are prime sources. Most
research centers on teaching and learning
have candidates for improved instruction
ready and waiting. New, more ambitious
curricula are appearing at a rapid rate. Rad-
ically different approaches to teaching also
can be found in other cultures. Looking
across cultures can be especially useful be-
cause it helps to break outside of the fami-
lar cultural model of teaching.

The problem, of course, is that new ideas
for teaching, from whatever source, usually
have had little effect on classroom practice. Teachers in the United States have had no mechanism by which they could process the ideas, try them out, refine them, and share the new information. The R & D system we are proposing, and the lesson study process, provides one solution to this problem. New ideas are fed into the process, lessons are designed that turn the ideas into practice, and feedback is collected.

In addition to the intellectual benefits of this process, it also resolves a persistent political problem. In the old tradition of reform, high status was ascribed to those who developed the ideas and low status to those who implemented them (Johnson, 1987; Lagemann, 1996). This, among other things, created a reaction against importing ideas for teaching from the outside and bolstered the view that teachers must generate their own ideas (McDonald, 1992). We agree that teachers have good ideas for improving teaching, but we believe that all ideas with promise should be considered, regardless of their source. Lesson study invites all good ideas into the process and gives teachers the responsibility and freedom to digest and test them. The politics surrounding the source of ideas disappears.

Scalability Revisited

Elmore (1996) identifies scalability as the most serious problem facing educational reformers. Pockets of exemplary practice already exist around the country, but they have not touched the average classroom. How do small, intensive efforts scale up? We argued earlier that such efforts require two things: (1) they must learn from their own experience and what they learn must be shareable; and (2) they must be modular in the sense that they can be combined and accumulated to yield large-scale systems that work as well as small ones. We proposed an R & D system, built on lesson study, with these requirements in mind. Does such a system scale up?

What makes the lesson study process especially appealing is that it satisfies these requirements in multiple ways. First, the fact that the daily classroom lesson serves as the unit of analysis and improvement places the process at the intersection of these two requirements. With respect to preserving teachers' learning and sharing it with others, the full reports of lessons serve as ideal repositories because they contain information on students' thinking, the curriculum, pedagogy, and how these interact, all in a form that is meaningful to other teachers.

Sharing knowledge of the kind constructed through lesson study is relatively straightforward because all teachers teach lessons. Demand for this knowledge will increase to the extent that teachers share learning goals and a curriculum and thus expect to teach similar lessons. New technologically enhanced ways of sharing the knowledge have recently become reality. Video examples of classroom lessons, linked with evolving theoretical understandings of teaching, could be stored in large digital libraries. Video, audio, images of student work, and commentary by teachers/researchers could be merged into a single, integrated database accessible over the Internet. Teacher groups could post the results of their work and actually collaborate with other groups working on similar lessons. It is not unreasonable to think that these technologies could play an exciting role in the scaling up process.

Lessons also are modular. Because the lesson preserves the interactions between the features of teaching and because almost all teachers teach through daily lessons, lessons can be accumulated, from many lesson study groups, to gradually extend the effects of lesson study across the school year. This modular characteristic does not hold for some of the "units" that have been used to study and improve teaching, such as the nature of teachers' questions. Results cannot be combined with results from studies of other features, such as instructional tasks, in a straightforward way because these features interact. They slice the teaching sys-
tem into nonmodular forms. But lessons can be connected, with appropriate transitions, in ways that allow work on individual lessons to fit together; connecting lessons does not require the work on individual lessons to be completely redone.

A second way in which lesson study meets the scaling up requirements is through its role as professional development. The Japanese experience suggests that teachers can learn more general knowledge and skills, of exactly the kind they need, by working deeply with a few lessons. Improving teaching does not depend on eventually perfecting 182 lessons but rather on engaging intensively with the issues involved in teaching any lesson. According to Japanese teachers, the yearlong process of perfecting a few lessons helps them teach other lessons more effectively (Lewis & Tsuchida, 1998; Takemura & Shimizu, 1993).

Finally, the R & D system we propose meets the scalability requirement of modularity in an institutional sense. The system can start in a single district and gradually spread as districts link with each other. This institutional modularity is crucial because change inevitably will begin within individual schools and districts. Then, as new districts take on the challenge, there will be more and more experienced districts to collaborate and form expanding networks. This affords increasing opportunities for teachers to share and receive the work of colleagues, which means that the lesson study process gains even more strength.

An Idea Worth Pursuing, Not a Template for Sure Success

Applying ideas gained through cross-cultural comparisons is always tricky. On the one hand, it is foolish to think that programs simply can be imported and show the same results in their new home. On the other hand, it is equally foolish to dismiss promising ideas that lie behind these programs simply because they are working well in a different culture. Whether the notion of lesson study can be the catalyst for developing an effective R & D system in the United States for improving teaching is an open question. But, because lesson study is consistent with the cultural, systemic nature of teaching and because it shares many of the principles found in a variety of recent, effective teacher development efforts in the United States, we believe it is an idea worth examining.

Chances of success, however, are closely tied to society’s willingness to recalibrate its expectations for change. Reforms in the United States come with imperatives to show quick and dramatic change. But the R & D system proposed here will show gradual, incremental change. Such changes can add up to something, over time, that completely reshapes teaching, but the changes will not come quickly. If the United States is in a hurry, it should take the first step soon to establish a long-term improvement strategy that guarantees teaching will be more effective in the future, not 1 year from now but 20 years from now.

Note

The TIMSS Video Study was funded by the U.S. Department of Education’s National Center for Education Statistics. The study was conducted in collaboration with Jürgen Baumert and Rainer Lehmann in Germany, and Toshio Sawada in Japan. In addition, Patrick Gonzales, Takako Kawanaka, Steffen Knoll, and Ana Serrano functioned as primary researchers on the study, and Eric Derghazarian, Gundula Huber, Fumiko Ichikawa, and Nicole Kersting served as primary coders.

References


and what we need to learn. *Phi Delta Kappan*, 77, 500–508.


Griffin, G. A. (1995). Influences of shared decision making on school and classroom activ-
Improving Teaching

19


professional development: Teaching as craft. Teaching and Teacher Education, 14, 451–462.
Contributors

☐ James Hiebert is H. Rodney Sharp Professor of Education at the University of Delaware in Newark.

☐ James W. Stigler is professor in the Department of Psychology at the University of California, Los Angeles.

☐ William H. Jeynes is assistant professor of education at Hillsdale College in Hillsdale, MI.

☐ Stephen W. Littell teaches at Carleton Washburne School in Winnetka, IL.

☐ James Griffith is senior research analyst with the Montgomery County Public Schools in Rockville, MD.

☐ Karen M. La Paro is a research faculty member in the Department of Human Services, and Robert Pianta is professor in the Departments of Human Services and School and Clinical Psychology, at the University of Virginia in Charlottesville.

☐ Martha Cox is professor at the Frank Porter Graham Child Development Center at the University of North Carolina, Chapel Hill.

☐ Shelley Peterson is assistant professor in the Department of Curriculum, Teaching, and Learning at the Ontario Institute for Studies in Education/University of Toronto, in Toronto, Ontario, Canada.

☐ Susan L. Kessler-Sklar is a management consultant at William M. Mercer, Inc., in New York, NY.

☐ Amy J. L. Baker is director of research and program evaluation at The Children's Village in Dobbs Ferry, NY.