On Research on Teaching: 
A Conversation with Lee Shulman

Stanford Professor Lee Shulman reflects on the limitations of generalizations about effective teaching and describes how case histories — specific stories about classroom experience — can enrich our collective "wisdom of practice."

Ron Brandt

In the early 1980s, educators were excited because we seemed to have the beginnings of a firm knowledge base for our profession. Barak Rosenshine and Jere Brophy, Dave Berliner, and you were among those who summarized this "effective teaching" research. But in the last five or six years we haven't heard much about it. What happened?

Well, I think we began to realize the limitations of that kind of research. We realized that it provided answers to a particular kind of question, but we were interested in asking other questions, too. If you ask what in general characterizes good teaching; if you're looking for a general set of descriptions and practices that apply no matter what the grade level, subject matter, or context, you're likely to get a set of principles much like those of effective teaching. I think they are reasonable and warranted — although I'd offer several caveats.

For example, those principles were developed from studies in which teachers were observed 5 or 10 times during a school year, and then the data from all 10 times were aggregated. A teacher of mathematics might be introducing the topic of estimation on day one, having the kids do guided practice on factoring on day two, preparing the kids for a quiz on the topic of exponents on day three, and so on. So, when you collapse all that together, you're putting into the same pot a lot of very different teaching situations, even for teaching mathematics to 2nd graders. Then the researcher would correlate the frequencies with which the teachers did a variety of things in all of those sessions against the performance of the students two to nine months later on a standardized achievement test. That's why the research is called "process-product" research.

Would you explain that term a little more completely?

Sure. The term was originated by Donald Medley and Harold Mitzel and elaborated by my colleague Nathaniel...
Gage. The notion is that you meticulously describe all the processes you possibly can that teachers and kids engage in and then relate those processes to products: to outcomes — attitudinal, cognitive, affective — that you can measure at some later time.

Now, you were beginning to explain some problems with that kind of research, and I thought you were going to say that a big problem is that it's all measured by standardized tests — because many educators feel that those tests don't fully measure the outcomes we're looking for.

Yes, I was about to get to that. The other flaw is the criterion measure — but let me remind you that in the '60s and early '70s when this kind of research was pioneered, it was really creative, a brilliant piece of research design. It permitted us to get a handle on what people had claimed was ineffable: the nature of good teaching.

Remember that in the late '60s we had the Coleman Report. We had Jencks; we had Jensen. We had a number of authorities arguing that schools didn't make a bit of difference — and if they didn't, neither did teachers or curriculum. That was the policy environment in which the process-product research was done.

We had a group of very creative researchers who said, "That's got to be nonsense. We've got to refute that claim. Let's measure variations in the ways teachers teach and meticulously document those variations." So they used variations in types of questioning, wait time, in the frequency with which teachers used praise, and so on. And they demonstrated that with the criterion measures mostly widely accepted in the policy arena — standardized tests — there were differences that could be attributed to teaching.

That was exciting — but why aren't researchers doing more of that research now? What happened?

Well, it may be that we learned as much as we were going to learn from asking that kind of question. And of course we entered an era in which we became painfully aware of the limitations of standardized tests.

Are you saying we weren't aware of those limitations at the time?

We weren't as sensitive to them then. No, we were at a point in the late '60s and early '70s when we had to demonstrate that something worked — that we made some sort of difference — and standardized tests were all we had. But anyway, in the last 10 years, we've learned to ask a new set of questions.

Let's talk about that in just a minute. You began by saying that not only are these process-product results dependent on standardized test scores, but they are also very generic. You were saying that when researchers brought these data collected at various points together, they created a kind of idealized model of the teacher, but that there really wasn't any single teacher who necessarily did everything just that way.

Yes. And that's one of the reasons why there is now as a counterweight a burgeoning literature of case studies of individual teachers. Once we began to establish these very general principles of teaching, lots of us began to see that what was true perhaps in general and for the most part wasn't necessarily true in particular and in special circumstances.

I'd like to talk more about that, but first let me ask you a very important question. Several states and a lot of school systems offer staff development courses built around the effective teaching research that you say the researchers have moved beyond. We said most of that research was designed to establish the kinds of teaching behaviors that produce higher test scores. If we now believe that those tests do not measure very well the outcomes we value most, then is any of that research valid? How much of it should we be teaching to teachers?

I get concerned when districts not only teach effective teaching principles in staff development programs but also translate them into the instruments through which they evaluate teachers.

Well, I don't think it's doing too much harm — although we may be boring some very good teachers occasionally. But I suspect the problem is not with the principles of effective teaching. The problem is in the implicit claim that those who practice them have achieved sufficiency. I get especially concerned when districts not only teach the principles of effective teaching in their staff development programs but also translate them into the instruments through which they
evaluate teachers, as if to imply that if teachers are enacting these general principles they are effective and even excellent teachers, and that we need not ask any more questions about their teaching.

Let me be clear on this point. I have no doubt that teachers who employ anticipatory set are better than teachers who never do. I have no doubt that teachers who get feedback and knowledge of results regularly from students are better than teachers who do not. I don't want to give the impression of denying the usefulness of those generic principles entirely.

Perhaps many of them are useful, but Lauren Resnick, for example, says that some aspects of the effective teaching research are in fact counter to what we really want. For example, asking many short, focused questions does not necessarily lead to higher-order thinking.

But that's clear. It's quite consistent with what we've already said: standardized tests don't measure higher-order thinking very well. If when the effective teaching research was first developed, it had been communicated with a little more modesty, we wouldn't have this problem.

You've been talking about effective teaching. David Berliner has been doing research on what he calls "expert" teachers. How do your terms "good" and "effective" relate to "expert"?

What I worry about is the technical use of the term "expert" and the technical use of the term "effective," both of which come out of research literatures that in one way or another label certain teachers as being in those categories, and then ask how their behavior is different from the masses. I think it's a mistake to label some teachers as "expert" and then treat everything they do as "expert teaching." Most teachers are better at some things than at others; they teach some topics better than others, or some kids better than others. They do a better job on some days than others.

Can you give an example of "wisdom of practice"?

As you know, I've done a lot of research on medical education, so let me use an analogy from medicine. In some ways the generic teaching research is like whatever research there might have been that produced the general model for how you do a medical work-up. There are certain well-established principles basic to taking a patient's history, to physical diagnosis, to using the laboratory, and I can teach you those principles in a two-day workshop. I can teach you how to listen to the heart, for example— but will you know what to listen for? Will you have enough content knowledge to distinguish a sound that means "a" from a sound that means "b"? I can tell you that you should gather information in such a way that within the first two or three minutes you've got two firm alternative diagnostic hypotheses to work with. That's great. But in comes a patient, and how do you know which hypotheses to generate?

It's like telling a teacher to begin with an anticipatory set. All right, I'm teaching a unit on the Civil War. What's the anticipatory set that makes the most sense? Which kinds of set are likely to lead to the broadest and most useful kinds of learning? And which will be ineffective, because even though they grab the students' interest, they lead to distorted notions of what the Civil War was about?

You come in with your generic evaluation form to evaluate my teaching on the first day of my unit on the Civil War. But you've been trained as a math teacher. I give an anticipatory set that's clearly identifiable as anticipatory set, but from the perspective of good history teaching, it will loop the student into deep left field. Now if there had been no anticipatory set at all, odds are the teaching would have been even worse, so the point is that it's not a bad idea; it's just that it takes you only so far.

Let's turn, then, to some of the research on teaching that researchers are doing now.

Okay. Two approaches are being used increasingly. They're similar except for the relationship between who's doing the research and who is being studied. First, we're seeing more detailed case studies of teaching. This is very different from the earlier research where the researcher would drop in at 10 different times during the
year randomly selected anywhere from two days to three weeks apart and document what the teacher was doing on those occasions. In a case study, the researcher wants to examine a natural segment of instruction, including its beginning, middle, and end — because there’s something dramaturgical about teaching; there’s a narrative or plot. For most of us as teachers, the unit of instruction is not the lesson; the lesson is just a piece of something larger. The better we get as teachers, the more we think about how our lessons hang together. In fact, when most teachers think about our instruction, we think in large chunks first and only then begin to decompose them into individual lessons.

Parenthetically, that’s one of the problems with the earlier effective teaching research. By sampling teaching a lesson at a time, the researchers implicitly defined the teaching as something that occurs primarily within individual lessons.

So one big difference in today’s research is more case studies.

Yes. And the people doing the case studies are no longer doing “one lesson stands.” In the case of a high school English class, you get there in time for the beginning of *Moby Dick*, and you don’t leave until it’s over. Now this not only changes the way you do the research; it changes the relationship between you and the colleague teacher. It changes how you report the work.

The second thing that is happening — and I find this very exciting — is that more and more researchers are beginning to treat their own teaching as a proper object of research. They themselves try to teach something over an extended period of time in a real school setting, and systematically inquire into their own teaching. The names that come to mind here are a number of colleagues at my former home at Michigan State: Maggie Lampert and Deborah Ball in elementary mathematics, Kathy Ross in elementary science, and Suzanne Wilson in elementary social studies are just a few examples. And there are secondary school examples, as well.

These university faculty members negotiate a relationship with a public school and public school teacher, and they take responsibility for a chunk of the curriculum. They go in; they teach that chunk of the curriculum. The teaching is documented in a variety of ways: videotape, audiotape, journals, observations, and so on, and it becomes an object of inquiry. And some of that work is astoundingly good.

I’m sure it’s fascinating — but what are we learning from it? What do we now know about teaching that we didn’t know a couple of years ago as a result of this kind of research?

One of the things we’ve learned is to stop asking that question. When you ask the question that way, you beg for general principles, but what we’re learning from this kind of research doesn’t necessarily assume that propositional form. Let me give you an example.

Here are a couple of ways of looking at the role of prior knowledge in teaching. We’ve got a lot of general research telling us that students bring a great deal of prior knowledge to the classroom and that that prior knowledge interacts with whatever you teach in ways that can often distort the message you think you are transmitting. Therefore, says the proposition, pay more attention to prior knowledge. Now that’s a good general principle. The problem is that it ignores how very difficult it often is to establish an anticipatory set unless you know of the content and something about the way students are probably thinking. Now do the case studies teach us new propositions? No. What they often do is tell us stories; they provide us with telling examples.

For example, some researchers were studying the teaching of “light and vision” in 5th grade, helping kids understand the relationship between the presence or absence of light and the ability to see. They taught the principle that light comes from a light source, bounces off the thing we’re looking at, and then bounces into our eye, which then receives this bounce of light, making vision possible. The kids all passed the test. Then they did clinical interviews with the kids and found that most of them still believed that the way we see is that our eyes shoot out rays like Superman, the rays bounce off the object we’re looking at and bounce back to the eye — so it’s like sonar or radar or something. Hey, it makes a lot more sense, doesn’t it?

Well, now that the researchers recognized the prior knowledge that students brought to the situation, they worked with the teachers to develop some preparatory lessons in which the

---

I find it very exciting that more and more researchers are beginning to treat their own teaching as a proper object of research.
kids first offered their own preexisting theories. The teachers explored those theories with them and then began to introduce the new ideas in relation to the prior knowledge of the students. That way, the kids incorporated the new ideas much more effectively.

Well, that's the story. It's a narrative. The general principle still concerns prior knowledge, but is that what you learn from the story? No. I would argue that this story and others like it are part of an emerging case literature that will stay in our memories, become part of our pedagogical intelligence in ways that are much more productive than the propositions standing up there kind of naked and alone. One of the things, then, we have to be prepared to say is, "If you've got a few hours, let me tell you some stories."

But practicing educators are very busy people. They don't have a lot of hours to spare.

Let's not assume that everybody has to read all of the great books. If I'm a 3rd grade teacher having a devil of a time teaching introduction to fractions, I don't want to read all the case literature of teaching. I want some good cases of fraction teaching. And I'll tell you something. If I'm having trouble teaching fractions and you give me a set of cases of fraction teaching, you won't have to motivate me to look at them, because my self-respect is on the line tomorrow.

That brings us again to the importance of content-specific pedagogy. For years you've been saying that the generic research is useful, but that teachers need more specific professional knowledge.

Well, that's obvious, isn't it? You learn about it from working closely with practitioners and seeing what they cope with every day. Example: a young woman in our teacher training program has been writing some cases of her own teaching — as are all the students in our program. One of the cases she wrote is about teaching the concept of the atom to her class. Now in 8th grade general science, you've got to teach what an atom is. Okay, how do you do that? It's interesting that, even if she solves all the generic problems of teaching — getting the classroom running smoothly, doing general lesson planning — she's still got to figure out how to teach the atom. So she writes about trying to teach what an atom is, and then teaching about the periodic table of elements, which helps students understand the way in which elements are made up of different sorts of atoms, you know.

Well, she gets to a point where she describes telling the kids that in this column of the periodic table all of the elements are what we call "noble gasses," and a kid asks, "Why?" Now, what kind of knowledge does she need? Well, you could say she needs knowledge of the history of chemistry, because there must be a reason why those are called noble gasses, but is that the question the kid was really asking? Did he want the history of chemistry at that point? The teacher thought for a moment and asked, "What do you associate with people who are nobles?" The kids played around with it and one of them said, "They're rich." Then she asked them what rich people need, and they said, "Not much, they have everything they need." And she said, "Well, noble gasses are that way too. They have all the atoms they need. They don't have to borrow from any other elements. That's why they're called noble gasses." End of story.

Now, we need lots more stories like that. We need a literature that focuses on the intersection of content and pedagogy. Literature that brings together the wisdom of practice on a topic-by-topic, idea-by-idea basis. I don't want to lock us into a 19th century organization of the disciplines, but it has to be particular.

This intersection of content and pedagogy is what you call pedagogical content knowledge, isn't it? A concept that finally puts to rest the arguments we've had through the years about whether process or content is more important. It helps explain why not everybody with a degree in physics is a good teacher of physics. Teachers have to know a lot about how to explain physics.

Yes. Here's another story. One of the enduring problems of math teaching is trying to teach kids about dividing by zero. Of course you can get kids to say that when you divide something by zero the result is — not zero — but undefined. But what does that mean? Well, one of the case studies we looked at last week was on that topic. Suzanne Wilson, one of my former
students who is now on the faculty at Michigan State, turned this problem into a story. She said that if you divide by smaller and smaller numbers — first by 1, then by .9, then by .8 and so on — as you keep on asking how many times the shrinking denominator can go into the number on top, the kids realize that as you get closer and closer to zero, it just sort of explodes. How big does it explode when it gets to zero? Well, it’s — undefined.

I thought I knew about dividing by zero before, but when she told that little narrative, I understood it at another level. A story had enriched my understanding. Now if that story could be added to the case knowledge of teachers of mathematics — along with lots of other stories — it would be immensely helpful to them. Sure, it would be pretty imposing after a while: we’d have to learn to arrange it — but people in law have learned how to take lots and lots of stories they call cases and organize them. We can, too.

But wait a minute. Suppose I’m a principal or director of staff development. I don’t know French, and I don’t know much about biology or art. I’m a generalist. What am I supposed to do with all of this content-specific research? How are generalists supposed to supervise teachers, all of whom have specialized interests?

Well, I guess I’d quote Whitehead, who advised us to “seek generalizations and distrust them.” I think you have to engage in some very reflective soul-searching to examine what it is you are competent to do as a generalist, and what you can responsibly do only in collaboration with colleagues who know more about particular things than you do.

We of the education community have gotten into a very bad habit; we have become consumers and not critics.

I teach a course for secondary education majors in every one of the content areas where we prepare teachers, but I try to teach generically through particulars. The first day of class we read a case of a teacher trying to teach Romeo and Juliet to a recalcitrant group of 10th graders. So that the class doesn’t lose sight of how very difficult it is to teach Romeo and Juliet, I pass out a copy of the first act, and I purposely ask two of the math majors to read aloud, beginning with the prologue, which talks about “star-crossed lovers.”

After we’ve worked on that particular case together, I ask how typical the teacher’s problem is. Almost every time I can expect a young woman in social studies to raise her hand and say, “It reminds me of my class yesterday.” And the discussion takes off from there with my asking, “What were the similarities? What were some very particular differences?” I think there is room for generalists — at some level we’re all generalists — but we have to be sensitive to our limitations. We have to begin collaborating much more actively with more specialized colleagues. And quite frankly, I think many of us have to take some time to study and read more. I’m very serious. Most of us who are generalists ought to decide that one of the ways we might understand more about teaching would be to pick an area that we find interesting and that is within our purview and get smarter about it: read some of this literature, read some of these cases, and so on. But the solution is not to try to rely solely on general pedagogical knowledge.

For years we educators have talked about “putting research into practice,” but from what you’re saying, that phrase is taking on a different meaning.

Yes, I hope so. First, we’re coming to recognize that as I said earlier, research answers the questions you ask, and that we of the education community have gotten into a very bad habit; we have become consumers and not critics.

My late, great teacher Joseph Schwab, in his work on science teaching, asked us to keep in mind the distinction between a rhetoric of conclusions and a narrative of inquiry. We in the educational practice community accepted the results of research on teaching as a rhetoric of conclusions. We permitted the answers to be divorced from the questions. We have to become a much more critical and aggressive group of consumers who regularly ask, “What were the questions to which these are the answers?” If we do that, we can see where the research on effective teaching is useful and ought to be incorporated into our work, and where we need to look for other kinds of research that may provide answers to different questions.


Lee S. Shulman is Charles E. Ducommun Professor of Education, Stanford University, School of Education, 330 Bonair Siding, Stanford, CA 94305-3084. Ron Brandt is ASCD’s Executive Editor.