



A Conversation with Thomas R. Guskey



Thomas Guskey

*Thomas R. Guskey, Ph.D., is a professor in the College of Education at the University of Kentucky and an expert in research and evaluation who has authored or edited 12 books, including *Evaluating Professional Development* (Corwin, 2000). He has twice won the National Staff Development Council's prestigious *Book of the Year Award* and three times won the *Article of the Year Award*. Below, he discusses his five-step process for evaluating professional development in education and its connection to professional development planning.*

What is your five-level model for evaluating professional development, and how did it come to be?

My thinking was influenced by the work of Donald Kirkpatrick, who developed a model for evaluating training programs in business and industry. Kirkpatrick described four levels of evaluation that he found necessary in determining the value and worth of training programs. The first was participants' reactions to the training—whether they liked it or not. A second level was what new knowledge and/or skills participants gained from the training. A third level was how it influenced what they did on the job. And a fourth level considered how the training affected their productivity.

I thought this model could be useful for what we do in professional development in education. As we applied the model, however, we found that professional development efforts still were not yielding positive results—but nothing in the model explained why. Examining programs more closely, I found that things were done right from a training perspective, but educators were then sent back to organizations that did not support them in what we asked them to do. Things broke down at the organization level. So I added a new level in the middle of the model, labeled “organizational support and change,” to consider those aspects of the organization that have critical influence on the implementation of new policies and practices. (See Figure 1 for the model.)

What do you hope people take away from your model?

There are three major aspects of the model that I hope people will consider. First, each of these five levels is important in its own right. Each level provides different types of information that can be used in both formative and summative ways. Formatively, we need to find out at each level what's been done well and, if not done well, how it can be improved. Summatively, we need to know the effectiveness of elements at each level to judge the true value and worth of any professional development endeavor.

Second, each level builds on those that come before. For example, people must have a positive reaction to a professional development experience before we can expect them to learn anything from it. They need to gain specific knowledge and skills before we look to the organization for critical aspects of support or change. Organizational support is necessary to gain high quality implementation of new policies and practices. And appropriate implementation is a prerequisite to seeing improvements in student learning. Things can break down at any point along the way, and once they break down, the improvement process comes to a screeching halt.

Third, many educators are now finding how useful it can be to reverse these five levels in professional development planning. In other words, the first thing people need to do when they plan professional development is to specify what impact they want to have on student learning. They begin planning by asking, “What improvements in student learning do we want to attain and what evidence best reflects those improvements?” Then they step back and ask, “If

that's the impact we want, what new policies or practices must be implemented to gain that impact?" Next, they consider what types of organizational support or change are needed to facilitate that implementation, and so forth. This planning process compels educators to plan not in terms of what they are going to do but in terms of what they want to accomplish with their students. All other decisions are then based on that fundamental premise.

I argue that most of the critical evaluation questions that need to be addressed in determining a professional development program's effectiveness should be asked in the planning stage. Planning more carefully and more intentionally not only makes evaluation easier, it also leads to much more effective professional development. Increasingly, educators at all levels are coming to view professional development as a purposeful and intentional endeavor that should be designed with specific goals in mind.

Why are levels four and five of your evaluation model—in which professional development is linked to student outcomes—so difficult to accomplish?

The primary reason is that getting information at those levels must be delayed. Immediately following any professional development activity, I can gather information about levels one and two—finding out if people liked it and what they gained from that experience in terms of new knowledge and skills. But information on levels three, four, and five cannot be gathered at that time. Again, planning backward makes this clearer. If I know what I want to accomplish and what evidence best reflects those goals, it's easier for me to decide how and when I'm going to gather that evidence and what I will do with it once I have it.

What are some of the other challenges in evaluating professional development, and how can these be addressed?

Many professional development leaders avoid systematic evaluations for fear that the evaluation won't yield "proof" that what they're doing leads to improvements in student learning. And if this is the case, funding may be withdrawn. Recognizing the distinction between "evidence" and "proof," however, can help resolve this dilemma.

To obtain proof—by which I mean to show that professional development uniquely and alone leads to improvements in student learning—is very difficult. It requires a level of experimental rigor that is hard and often impossible to attain in practical school settings. But most policymakers, legislators, and school leaders are not asking for ironclad proof. What they want is evidence that things are getting better. They want to see improvements in assessment results or test scores, increased attendance, fewer discipline problems, or decreased dropout rates. Historically, professional development leaders haven't done a very good job of providing any such evidence.

A related challenge concerns the nature of that evidence, especially its credibility and its timing. I recently discovered, for example, that not all stakeholders in professional development trust the same evidence. I conducted a study in which groups of educators were asked to rank order 15 different indicators of student learning in terms of which they believed provided the most valid evidence. When I compared administrators' and teachers' rankings, I found they were almost exactly reversed! Administrators rated national and state tests highly, while teachers trusted their own, more immediate sources of evidence. From a policy perspective, that indicates to me that no single source of evidence is going to be adequate. Instead, we need to consider multiple indicators. We also need to involve multiple stakeholders in the planning process to identify the sources of evidence that they believe provide the best and most valid representation of success.

Some experts suggest that when educators engage in professional development endeavors, results might not be evident for two or three years. But when teachers are experimenting with new approaches to instruction or a new curriculum, they need to gain evidence rapidly to show that it's making a difference. If they don't see such evidence, they quite naturally revert back to the tried and true things they've done in the past. This isn't because they are afraid of change. Rather, it's because they are so committed their students and fear that the new approach might lead to less positive results. So, in planning professional development, we must include some mechanism whereby those responsible for implementation can gain evidence of success from their students rather quickly—within the first month of implementation.

Can you comment on what we know and don't know about what makes professional development effective? How can we go about reaching some consensus about what is important?

A couple of years ago, I identified thirteen lists of characteristics of effective professional development that had been assembled by different professional organizations and research groups. In analyzing these lists, I found very little consensus. There wasn't even agreement on the criteria for effectiveness. Some lists were based on the concurrence opinions among researchers, others used teacher self-reports, and only a few looked at impact on student learning. My conclusion was that we may not have a true consensus on what makes professional development effective, and that moving toward one may be more complicated than most people think.

I helped to develop the Standards for Staff Development published by the National Staff Development Council. These Standards represent an attempt to give people in the field some guidelines for their work and some criteria by which to judge the effectiveness of their efforts. Because of their general nature, however, these Standards leave a lot of room for interpretation. For example, they describe the importance of extended time for professional development

and the need to ensure that activities are ongoing and job-embedded. Researchers have shown, however, that simply adding more time for job-embedded activities is insufficient. Doing ineffective things longer doesn't make them any better. Instead, we must ensure that the extended time provided for professional development is structured carefully and used wisely, engaging educators in activities shown to yield improved results.

How do you think the federal No Child Left Behind Act (NCLB) is impacting professional development and its evaluation?

I believe that certain aspects of the No Child Left Behind Act are motivated by frustration on the part of the federal government in allocating funds to education and not seeing much come from it. Too often in the past, educators have planned professional development based on what's new and what's hot, rather than on what is known to work with students. In NCLB, the federal government imposes specific requirements that compel educators to consider only programs and innovations that are "scientifically based research." Educators must now verify the research behind different programs and innovations. They must ensure that research comes from reliable sources, specifically peer-reviewed journals. They must show that the program has been applied in a wide variety of contexts and that its effects evaluated by third parties. They must demonstrate that the evidence of effects has been gathered over a significant period of time so that the program can be shown to be sustainable.

I agree with those who suggest that insistence on this definition of "scientifically based research" may be too restricting. A lot of valuable research does not meet the criteria of randomized designs, but can provide us with good, important evidence. Still, NCLB and other national efforts are moving us in the right direction.

This past year, I've met with leaders in the U.S. Department of Education and various philanthropic organizations, who are considering changing the request for proposal process to be more specific with regard to evaluation. In particular, they want people, within proposals, to outline specifically how they will gather evidence at each of the five levels in the evaluation model. Their hope is that this will lead to improved results from various funded programs. I share their hope.

Figure I. Five Levels of Professional Development Evaluation¹

Evaluation Level	What Questions Are Addressed?	How Will Information Be Gathered?	What Is Measured or Assessed?	How Will Information Be Used?
1. Participants' Reactions	Did they like it? Was their time well spent? Did the material make sense? Will it be useful? Was the leader knowledgeable and helpful? Were the refreshments fresh and tasty? Was the room the right temperature? Were the chairs comfortable?	Questionnaires administered at the end of the session	Initial satisfaction with the experience	To improve program design and delivery
2. Participants' Learning	Did participants acquire the intended knowledge and skills?	Paper-and-pencil instruments Simulations Demonstrations Participant reflections (oral and/or written) Participant portfolios	New knowledge and skills of participants	To improve program content, format, and organization
3. Organization Support and Change	What was the impact on the organization? Did it affect	District and school records Minutes from follow-up meetings	The organization's advocacy, support, accommodation, facilitation, and	To document and improve organizational support

	<p>organizational climate and procedures? Was implementation advocated, facilitated, and supported? Was the support public and overt? Were problems addressed quickly and efficiently? Were sufficient resources made available? Were successes recognized and shared?</p>	<p>Questionnaires Structured interviews with participants and district or school administrators Participant portfolios</p>	<p>recognition</p>	<p>To inform future change efforts</p>
<p>4. Participants' Use of New Knowledge and Skills</p>	<p>Did participants effectively apply the new knowledge and skills?</p>	<p>Questionnaires Structured interviews with participants and their supervisors Participant reflections (oral and/or written) Participant portfolios Direct observations Video or audio tapes</p>	<p>Degree and quality of implementation</p>	<p>To document and improve the implementation of program content</p>
<p>5. Student Learning Outcomes</p>	<p>What was the impact on students? Did it affect student performance or achievement? Did it influence students' physical or emotional well-being? Are students more confident as learners? Is student attendance improving? Are dropouts decreasing?</p>	<p>Student records School records Questionnaires Structured interviews with students, parents, teachers, and/ or administrators Participant portfolios</p>	<p>Student learning outcomes: • Cognitive (Performance & Achievement) • Affective (Attitudes & Dispositions) • Psychomotor (Skills & Behaviors)</p>	<p>To focus and improve all aspects of program design, implementation, and follow-up To demonstrate the overall impact of professional development</p>

¹ Guskey, T. R. (2000). Evaluating Professional Development. Thousand Oaks, CA: Corwin Press.

Holly Kreider, Project Manager, HFRP
Suzanne Bouffard, Research Analyst, HFRP