



Thomas R. Guskey

Using five critical levels of evaluation, you can improve your school's professional development program. But be sure to start with the desired result—improved student outcomes.

Educators have long considered professional development to be their right—something they deserve as dedicated and hardworking individuals. But legislators and policymakers have recently begun to question that right. As education budgets grow tight, they look at what schools spend on professional development and want to know, Does the investment yield tangible payoffs or could that money be spent in better ways? Such questions make effective evaluation of professional development programs more important than ever.

Traditionally, educators haven't paid much attention to evaluating their professional development efforts. Many consider evaluation a costly, time-consuming process that diverts attention from more important activities such as planning, implementation, and follow-up. Others feel they lack the skill and expertise to become involved in rigorous evaluations; as a result, they either neglect evaluation issues completely or leave them to "evaluation experts."

Good evaluations don't have to be complicated. They simply require thoughtful planning, the ability to ask good questions, and a basic understanding of how to find valid answers. What's more, they can provide meaningful information that you can use to make thoughtful, responsible decisions about professional development processes and effects.

What Is Evaluation?

In simplest terms, evaluation is "the systematic investigation of merit or worth" (Joint Committee on Standards for Educational Evaluation, 1994, p. 3). *Systematic* implies a focused, thoughtful, and intentional process. We conduct evaluations for clear reasons and with explicit intent. *Investigation* refers to the collection and analysis of pertinent information through appropriate methods and techniques. *Merit or worth* denotes appraisal and judgment. We use evaluations to determine the value of something—to help answer such questions as, Is this program or activity achieving its intended results? Is it better than what was done in the past? Is it better than another, competing activity? Is it worth the costs?

Some educators understand the importance of evaluation for event-driven professional development activities, such as workshops and seminars, but forget the wide range of less formal, ongoing, job-embedded professional development activities—study groups, action research, collaborative planning, curriculum development, structured observations, peer coaching, mentoring, and so on. But regardless of its form, professional development should be a purposeful endeavor. Through evaluation, you can determine whether these activities are achieving their purposes.

Critical Levels of Professional Development Evaluation

Effective professional development evaluations require the collection and analysis of the five critical levels of information shown in Figure 1 (Guskey, 2000a). With each succeeding level, the process of gathering evaluation information gets a bit more complex. And because each level builds on those that come before, success at one level is usually necessary for success at higher levels.

Figure 1. Five Levels of Professional Development Evaluation

Evaluation Level	What Questions Are Addressed?	How Will Information Be Gathered?	What Is Measured or Assessed?	How Will Information Be Used?
1. Participants' Reactions	<p>Did they like it?</p> <p>Was their time well spent?</p> <p>Did the material make sense?</p> <p>Will it be useful?</p> <p>Was the leader knowledgeable and helpful?</p> <p>Were the refreshments fresh and tasty?</p> <p>Was the room the right temperature?</p> <p>Were the chairs comfortable?</p>	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?	<ul style="list-style-type: none"> • 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 & Change	<p>Was implementation advocated, facilitated, and supported?</p> <p>Was the support public and overt?</p> <p>Were problems addressed quickly and efficiently?</p> <p>Were sufficient resources made available?</p>	<ul style="list-style-type: none"> • District and school records • Minutes from follow-up meetings • Questionnaires • Structured interviews with participants and district or school administrators • Participant portfolios 	The organization's advocacy, support, accommodation, facilitation, and recognition	<p>To document and improve organization support</p> <p>To inform future change efforts</p>

	<p>Were successes recognized and shared?</p> <p>What was the impact on the organization?</p> <p>Did it affect the organization's climate and procedures?</p>			
4. Participants' Use of New Knowledge and Skills	<p>Did participants effectively apply the new knowledge and skills?</p>	<ul style="list-style-type: none"> • Questionnaires • Structured interviews with participants and their supervisors • Participant reflections (oral and/or written) • Participant portfolios • Direct observations • Video or audio tapes 	<p>Degree and quality of implementation</p>	<p>To document and improve the implementation of program content</p>
5. Student Learning Outcomes	<p>What was the impact on students?</p> <p>Did it affect student performance or achievement?</p> <p>Did it influence students' physical or emotional well-being?</p> <p>Are students more confident as learners?</p> <p>Is student attendance improving?</p> <p>Are dropouts decreasing?</p>	<ul style="list-style-type: none"> • Student records • School records • Questionnaires • Structured interviews with students, parents, teachers, and/or administrators • Participant portfolios 	<p>Student learning outcomes:</p> <ul style="list-style-type: none"> • Cognitive (Performance & Achievement) • Affective (Attitudes & Dispositions) • Psychomotor (Skills & Behaviors) 	<p>To focus and improve all aspects of program design, implementation, and follow-up</p> <p>To demonstrate the overall impact of professional development</p>

Level 1: Participants' Reactions

The first level of evaluation looks at participants' reactions to the professional development experience. This is the most common form of professional development evaluations, and the easiest type of information to gather and analyze.

At Level 1, you address questions focusing on whether or not participants liked the experience. Did they feel their time was well spent? Did the material make sense to them? Were the activities well planned and meaningful? Was the leader knowledgeable and helpful? Did the participants find the information useful?

Important questions for professional development workshops and seminars also include, Was the coffee hot and ready on time? Was the room at the right temperature? Were the chairs comfortable? To some, questions such as these may seem silly and inconsequential. But experienced professional developers know the importance of attending to these basic human needs.

Information on participants' reactions is generally gathered through questionnaires handed out at the end of a session or activity. These questionnaires typically include a combination of rating-scale items and open-ended response questions that allow participants to make personal comments. Because of the general nature of this information, many organizations use the same questionnaire for all their professional development activities.

Some educators refer to these measures of participants' reactions as "happiness quotients," insisting that they reveal only the entertainment value of an activity, not its quality or worth. But measuring participants' initial satisfaction with the experience can help you improve the design and delivery of programs or activities in valid ways.

Level 2: Participants' Learning

In addition to liking their professional development experience, we also hope that participants learn something from it. Level 2 focuses on measuring the knowledge and skills that participants gained. Depending on the goals of the program or activity, this can involve anything from a pencil-and-paper assessment (Can participants describe the crucial attributes of mastery learning and give examples of how these might be applied in typical classroom situations?) to a simulation or full-scale skill demonstration (Presented with a variety of classroom conflicts, can participants diagnose each situation and then prescribe and carry out a fair and workable solution?). You can also use oral personal reflections or portfolios that participants assemble to document their learning.

Although you can usually gather Level 2 evaluation information at the completion of a professional development activity, it requires more than a standardized form. Measures must show attainment of specific learning goals. This means that indicators of successful learning need to be outlined before activities begin. You can use this information as a basis for improving the content, format, and organization of the program or activities.

Level 3: Organization Support and Change

At Level 3, the focus shifts to the organization. Lack of organization support and change can sabotage any professional development effort, even when all the individual aspects of professional development are done right.

Suppose, for example, that several secondary school educators participate in a professional development program on cooperative learning. They gain a thorough understanding of the theory and develop a variety of classroom activities based on cooperative learning principles. Following their training, they try to implement these activities in schools where students are graded "on the curve"—according to their relative standing among classmates—and great importance is attached to selecting the class valedictorian. Organization policies and practices such as these make learning highly competitive and will thwart the most valiant efforts to have students cooperate and help one another learn (Guskey, 2000b).

The lack of positive results in this case doesn't reflect poor training or inadequate learning, but rather organization policies that undermine implementation efforts. Problems at Level 3 have essentially canceled the gains made at Levels 1 and 2 (Sparks & Hirsh, 1997). That's why professional development evaluations must include information on organization support and change.

At Level 3, you need to focus on questions about the organization characteristics and attributes necessary for success. Did the professional development activities promote changes that were aligned with the mission of the school and district? Were changes at the individual level encouraged and supported at all levels? Were sufficient resources made available, including time for sharing and reflection? Were successes recognized and shared? Issues such as these can play a large part in determining the success of any professional development effort.

Gathering information at Level 3 is generally more complicated than at previous levels. Procedures differ depending on the goals of the program or activity. They may involve analyzing district or school records, examining the minutes from follow-up meetings, administering questionnaires, and interviewing participants and school administrators. You can use this information not only to document and improve organization support but also to inform future change initiatives.

Level 4: Participants' Use of New Knowledge and Skills

At Level 4 we ask, Did the new knowledge and skills that participants learned make a difference in their professional practice? The key to gathering relevant information at this level rests in specifying clear indicators of both the degree and the quality of implementation. Unlike Levels 1 and 2, this information cannot be gathered at the end of a professional development session. Enough time must pass to allow participants to adapt the new ideas and practices to their settings. Because implementation is often a gradual and uneven process, you may also need to measure progress at several time intervals.

You may gather this information through questionnaires or structured interviews with participants and their supervisors, oral or written personal reflections, or examination of participants' journals or portfolios. The most accurate information typically comes from direct observations, either with trained observers or by reviewing video- or audiotapes. These observations, however, should be kept as unobtrusive as possible (for examples, see Hall & Hord, 1987).

You can analyze this information to help restructure future programs and activities to facilitate better and more consistent implementation.

Level 5: Student Learning Outcomes

Level 5 addresses "the bottom line": How did the professional development activity affect students? Did it benefit them in any way? The particular student learning outcomes of interest depend, of course, on the goals of that specific professional development effort.

In addition to the stated goals, the activity may result in important unintended outcomes. For this reason, evaluations should always include multiple measures of student learning (Joyce, 1993). Consider, for example, elementary school educators who participate in study groups dedicated to finding ways to improve the quality of students' writing and devise a series of strategies that they believe will work for their students. In gathering Level 5 information, they find that their students' scores on measures of writing ability over the school year increased significantly compared with those of comparable students whose teachers did not use these strategies.

On further analysis, however, they discover that their students' scores on mathematics achievement declined compared with those of the other students. This unintended outcome apparently occurred because the teachers inadvertently sacrificed instructional time in mathematics to provide more time for

writing. Had information at Level 5 been restricted to the single measure of students' writing, this important unintended result might have gone unnoticed.

Measures of student learning typically include cognitive indicators of student performance and achievement, such as portfolio evaluations, grades, and scores from standardized tests. In addition, you may want to measure affective outcomes (attitudes and dispositions) and psychomotor outcomes (skills and behaviors). Examples include students' self-concepts, study habits, school attendance, homework completion rates, and classroom behaviors. You can also consider such schoolwide indicators as enrollment in advanced classes, memberships in honor societies, participation in school-related activities, disciplinary actions, and retention or drop-out rates. Student and school records provide the majority of such information. You can also include results from questionnaires and structured interviews with students, parents, teachers, and administrators.

Level 5 information about a program's overall impact can guide improvements in all aspects of professional development, including program design, implementation, and follow-up. In some cases, information on student learning outcomes is used to estimate the cost effectiveness of professional development, sometimes referred to as "return on investment" or "ROI evaluation" (Parry, 1996; Todnem & Warner, 1993).

Look for Evidence, Not Proof

Using these five levels of information in professional development evaluations, are you ready to "prove" that professional development programs make a difference? Can you now demonstrate that a particular professional development program, and nothing else, is solely responsible for the school's 10 percent increase in student achievement scores or its 50 percent reduction in discipline referrals?

Of course not. Nearly all professional development takes place in real-world settings. The relationship between professional development and improvements in student learning in these real-world settings is far too complex and includes too many intervening variables to permit simple causal inferences (Guskey, 1997; Guskey & Sparks, 1996). What's more, most schools are engaged in systemic reform initiatives that involve the simultaneous implementation of multiple innovations (Fullan, 1992). Isolating the effects of a single program or activity under such conditions is usually impossible.

But in the absence of proof, you *can* collect good evidence about whether a professional development program has contributed to specific gains in student learning. Superintendents, board members, and parents rarely ask, "Can you prove it?" Instead, they ask for evidence. Above all, be sure to gather evidence on measures that are meaningful to stakeholders in the evaluation process.

Consider, for example, the use of anecdotes and testimonials. From a methodological perspective, they are a poor source of data. They are typically highly subjective, and they may be inconsistent and unreliable. Nevertheless, as any trial attorney will tell you, they offer the kind of personalized evidence that most people believe, and they should not be ignored as a source of information. Of course, anecdotes and testimonials should never form the basis of an entire evaluation. Setting up meaningful comparison groups and using appropriate pre- and post-measures provide valuable information. Time-series designs that include multiple measures collected before and after implementation are another useful alternative.

Keep in mind, too, that good evidence isn't hard to come by if you know what you're looking for before you begin. Many educators find evaluation at Levels 4 and 5 difficult, expensive, and time-consuming because they are coming in after the fact to search for results (Gordon, 1991). If you don't know where you are going, it's very difficult to tell whether you've arrived. But if you clarify your goals up front, most evaluation issues fall into place.

Working Backward Through the Five Levels

Three important implications stem from this model for evaluating professional development. First, each of these five levels is important. The information gathered at each level provides vital data for improving the quality of professional development programs.

Second, tracking effectiveness at one level tells you nothing about the impact at the next. Although success at an early level may be necessary for positive results at the next higher one, it's clearly not sufficient. Breakdowns can occur at any point along the way. It's important to be aware of the difficulties involved in moving from professional development experiences (Level 1) to improvements in student learning (Level 5) and to plan for the time and effort required to build this connection.

The third implication, and perhaps the most important, is this: In planning professional development to improve student learning, *the order of these levels must be reversed*. You must plan "backward" (Guskey, 2001), starting where you want to end and then working back.

In backward planning, you first consider the student learning outcomes that you want to achieve (Level 5). For example, do you want to improve students' reading comprehension, enhance their skills in problem solving, develop their sense of confidence in learning situations, or improve their collaboration with classmates? Critical analyses of relevant data from assessments of student learning, examples of student work, and school records are especially useful in identifying these student learning goals.

Then you determine, on the basis of pertinent research evidence, what instructional practices and policies will most effectively and efficiently produce those outcomes (Level 4). You need to ask, "What evidence verifies that these particular practices and policies will lead to the desired results? How good or reliable is that evidence? Was it gathered in a context similar to ours? Watch out for popular innovations that are more opinion-based than research-based, promoted by people more concerned with "what sells" than with "what works." You need to be cautious before jumping on any education bandwagon, always making sure that trustworthy evidence validates whatever approach you choose.

Next, consider what aspects of organization support need to be in place for those practices and policies to be implemented (Level 3). Sometimes, as I mentioned earlier, aspects of the organization actually pose barriers to implementation. "No tolerance" policies regarding student discipline and grading, for example, may limit teachers' options in dealing with students' behavioral or learning problems. A big part of planning involves ensuring that organization elements are in place to support the desired practices and policies.

Then, decide what knowledge and skills the participating professionals must have to implement the prescribed practices and policies (Level 2). What must they know and be able to do to successfully adapt the innovation to their specific situation and bring about the sought-after change?

Finally, consider what set of experiences will enable participants to acquire the needed knowledge and skills (Level 1). Workshops and seminars, especially when paired with collaborative planning and structured opportunities for practice with feedback, action research projects, organized study groups, and a wide range of other activities can all be effective, depending on the specified purpose of the professional development.

This backward planning process is so important because the decisions made at each level profoundly affect those at the next. For example, the particular student learning outcomes you want to achieve influence the kinds of practices and policies you implement. Likewise, the practices and policies you want to implement influence the kinds of organization support or change required, and so on.

The context-specific nature of this work complicates matters further. Even if we agree on the student learning outcomes that we want to achieve, what works best in one context with a particular community of educators

and a particular group of students might not work as well in another context with different educators and different students. This is what makes developing examples of truly universal "best practices" in professional development so difficult. What works always depends on where, when, and with whom.

Unfortunately, professional developers can fall into the same trap in planning that teachers sometimes do—making plans in terms of what they are going to do, instead of what they want their students to know and be able to do. Professional developers often plan in terms of what they will do (workshops, seminars, institutes) or how they will do it (study groups, action research, peer coaching). This diminishes the effectiveness of their efforts and makes evaluation much more difficult.

Instead, begin planning professional development with what you want to achieve in terms of learning and learners and then work backward from there. Planning will be much more efficient and the results will be much easier to evaluate.

Making Evaluation Central

A lot of good things are done in the name of professional development. But so are a lot of rotten things. What educators haven't done is provide evidence to document the difference between the two.

Evaluation provides the key to making that distinction. By including systematic information gathering and analysis as a central component of all professional development activities, we can enhance the success of professional development efforts everywhere.

References

- Fullan, M.G. (1992). Visions that blind. *Educational Leadership*, 49(5) 19–20.
- Gordon, J. (1991, August). Measuring the "goodness" of training. *Training*, 19–25.
- Guskey, T.R. (1997). Research needs to link professional development and student learning. *Journal of Staff Development*, 18(2), 36–40.
- Guskey, T.R. (2000a). *Evaluating professional development*. Thousand Oaks, CA: Corwin.
- Guskey, T.R. (2000b). Grading policies that work against standards and how to fix them. *NASSP Bulletin*, 84(620), 20–29.
- Guskey, T.R. (2001). The backward approach. *Journal of Staff Development*, 22(3), 60.
- Guskey, T.R., & Sparks, D. (1996). Exploring the relationship between staff development and improvements in student learning. *Journal of Staff Development*, 17(4), 34–38.
- Hall, G.E., & Hord, S.M. (1987). *Change in schools: Facilitating the process*. Albany, NY: SUNY Press.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.
- Joyce, B. (1993). The link is there, but where do we go from here? *Journal of Staff Development*, 14(3), 10–12.
- Parry, S.B. (1996). Measuring training's ROI. *Training & Development*, 50(5), 72–75.
- Sparks, D. (1996, February). Viewing reform from a systems perspective. *The Developer*, 2, 6.
- Sparks, D., & Hirsh, S. (1997). *A new vision for staff development*. Alexandria, VA: ASCD.
- Todnem, G., & Warner, M.P. (1993). Using ROI to assess staff development efforts. *Journal of Staff Development*, 14(3), 32–34.