

Thomas R. Guskey is an international expert in evaluation design, analysis, and educational reform. His essay explains why the Data standard for professional learning is an essential



foundation for all of the other standards. "Because of its indispensable and fundamental nature,"

Guskey writes, "no other standard is more important or more vital to the purpose of the Standards for Professional Learning."

The full essay explores the meaning of data in the context of professional learning. In this excerpt, Guskey examines the use of data in the systemic evaluation of professional learning.

# GAUGE IMPACT WITH 5 LEVELS OF DATA

By Thomas R. Guskey

Effective professional learning evaluation requires consideration of five critical stages or levels of information (Guskey, 2000a, 2002a, 2005). These five levels represent an adaptation of an evaluation model developed by Kirkpatrick (1959, 1998) for judging the value of supervisory training programs in business and industry. Kirkpatrick's model, although widely applied, has seen limited use in education because of inadequate explanatory power. While helpful in addressing a broad range of "what" questions, many find it lacking when it comes to explaining "why" (Alliger & Janak, 1989; Holton, 1996).

The five levels in this model are hierarchically arranged, from simple to more complex. With each succeeding level, the process of gathering evaluation data requires more time and resources. And because each level builds on those that come before, success at one level is usually necessary for success at higher levels.

## LEVEL 1: PARTICIPANTS' REACTIONS

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

At this level, questions focus on whether participants liked the experience. Did they feel their time was well spent? Did the content and material make sense to them? Were the activities well-planned and meaningful? Was the leader knowledgeable, credible, and helpful? Did they find the information useful?

Also important for some professional learning experiences are questions related to the context: Was the room the right temperature? Were the chairs comfortable? Were the refreshments fresh and tasty? To some, questions such

Professional learning that increases educator effectiveness and results for all students **uses a variety of sources and types of student, educator, and system data to plan, assess, and evaluate professional learning.**



**ABOUT THE BOOK**  
**Guskey, T.R., Roy, P., & von Frank, V. (2014).** *Reach the highest standard in professional learning: Data.* Thousand Oaks, CA: Corwin. Excerpted with permission.

as these may seem silly and inconsequential. But experienced professional development leaders know the importance of attending to these basic human needs.

Data on participants' reactions are usually gathered through questionnaires handed out at the end of a program or activity, or by online surveys distributed later through email. These questionnaires and surveys typically include a combination of rating-scale items and open-ended response questions that allow participants to provide more personalized comments. Because of the general nature of this information, many organizations

use the same questionnaire or survey for all professional learning, regardless of the format.

Some educators refer to these measures of participants' reactions as "happiness quotients," insisting that they reveal only the entertainment value of an experience or activity, not its quality or worth. But measuring participants' initial satisfaction provides data that can help improve the design and facilitation of professional learning in valid ways. In addition, positive reactions from participants are usually a necessary prerequisite to higher-level evaluation results.

## LEVEL 2: PARTICIPANTS' LEARNING

In addition to liking their professional learning experiences, participants should learn something from them. Level 2 focuses on measuring the new knowledge, skills, and perhaps attitudes or dispositions that participants gain (Guskey, 2002b).

Depending on the goals of the professional learning program or activity, this can involve anything from a pencil-and-paper assessment (Can participants describe the critical attributes of effective questioning techniques and give examples of how these might be applied in common classroom situations?) to a simulation or full-scale skill demonstration (Presented with a variety of classroom conflicts, can participants diagnose each situation, then prescribe and carry out a fair and workable solution?). Oral or written personal reflections or examinations of the portfolios that participants assemble can also document their learning.

Although Level 2 evaluation data often can be gathered at the completion of a professional learning program or activity, it usually requires more than a standardized form. And because measures must show attainment of specific learning goals, professional learning leaders need to outline indicators of successful learning before activities begin.

Careful evaluators also consider possible unintended learning outcomes, both positive and negative. Professional learning that engages teachers and school leaders in collaboration, for example, can foster a positive sense of community and shared purpose among participants (Supovitz, 2002). But in some instances, individuals collaborate to block change or inhibit advancement (Corcoran, Fuhrman, & Belcher, 2001; Little, 1990). Investigations further show that collaborative efforts sometimes run headlong into conflicts over professional beliefs and practices that can impede progress (Achinstein, 2002). Thus even the best-planned professional learning occasionally yields unanticipated negative consequences.

If there is concern that participants may already possess the requisite knowledge and skills, evaluators may require some form of pre- and post-assessment. Analyzing this data provides a basis for improving the professional learning's content, format, and organization.

## LEVEL 3: ORGANIZATIONAL SUPPORT AND CHANGE

At Level 3, the focus shifts from participants to organizational dimensions that may be vital to the success of the professional learning experience. Organizational elements also can sometimes hinder or prevent success, even when the individual aspects of professional development are done right (Sparks, 1996).

Suppose, for example, that a group of secondary educators participates in professional learning on aspects of cooperative learning. As part of their experience, they gain an in-depth understanding of cooperative learning theory and organize a variety of classroom activities based on cooperative learning principles.

THOMAS R. GUSKEY is professor of educational psychology in the College of Education at the University of Kentucky. He served on the policy research team of the National Commission on Teaching and America's Future and on the task force to develop the Standards for Staff Development. He was named a fellow in the American Educational Research Association, which also honored him in 2006 for his outstanding contribution relating research to practice.

Following their learning experience, they implement these activities in classes where students are graded or marked on the curve — according to their relative standing among classmates — and great importance is attached to each student's individual class rank.

Organizational grading policies and practices such as these, however, make learning highly competitive and thwart the most valiant efforts to have students cooperate and help each other learn. When graded on the curve, students must compete against each other for the few scarce rewards (high grades) dispensed by the teacher. Cooperation is discouraged since helping other students succeed lessens the helper's chance of success (Guskey, 2000b).

The lack of positive results in this case does not reflect poor training or inadequate learning on the part of the participating teachers, but rather organizational policies that are incompatible with implementation efforts. Problems at Level 3 have essentially canceled the gains made at Levels 1 and 2 (Sparks & Hirsh, 1997). That is why professional learning evaluations must include data on organizational support and change.

Level 3 questions focus on the organizational characteristics and attributes necessary for success. Did the professional learning promote changes that were aligned with the mission of the school? Were changes at the individual level encouraged and supported at the building and district levels (Corcoran et al., 2001)? Were sufficient resources made available, including time for sharing and reflection (Colton & Langer, 2005; Langer & Colton, 1994)? Were successes recognized and shared? Issues such as these often play a large part in determining the success of any professional learning.

Procedures for gathering data at Level 3 differ depending on the goals of the professional learning. They may involve analyzing school records, examining the minutes from follow-up meetings, and administering questionnaires that tap issues related to the organization's advocacy, support, accommodation, facilitation, and recognition of change efforts.

Structured interviews with participants and school administrators also can be helpful. These data are used not only to

document and improve organizational support for professional learning, but also to inform future change initiatives.

#### **LEVEL 4: PARTICIPANTS' USE OF NEW KNOWLEDGE AND SKILLS**

At Level 4, the primary question is: Did the new knowledge and skills that participants learned make a difference in their professional practice? The key to gathering relevant data at this level of evaluation rests in specifying clear indicators of both the degree and quality of implementation.

Unlike Levels 1 and 2, these data cannot be gathered at the end of a professional learning program or activity. 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, evaluators may need to gather measures of progress at several time intervals.

Depending on the goals of the professional learning, these data may involve questionnaires or structured interviews with participants and their school leaders. Evaluators might consider oral or written personal reflections or examinations of participants' journals or portfolios. The most accurate data typically come from direct observations, either by trained observers or using digital recordings. These observations, however, should be kept as unobtrusive as possible (for examples, see Hall & Hord, 1987).

Analyzing these data provides evidence on current levels of use. It also helps professional development leaders restructure future programs and activities to facilitate better, more consistent implementation.

#### **LEVEL 5: STUDENT LEARNING OUTCOMES**

Level 5 addresses the bottom line in education: What was the impact on students? Did the professional learning benefit them in any way? The particular student learning outcomes of interest will depend, of course, on the goals of that specific professional learning endeavor.

In addition to the stated goals, the program or activity may result in important unintended outcomes. Suppose, for example, that students' average scores on large-scale assessments went up, but so did the school dropout rate. Mixed results such as this are typical in education improvement efforts and reiterate the importance of including multiple measures of student learning in all evaluations (Chester, 2005; Guskey, 2007).

Furthermore, since stakeholders vary in their trust of different sources of evidence, it is unlikely that any single indicator of success will prove adequate or sufficient to all. When providing acceptable data for judging the effects of professional learning, evaluators should always include multiple sources of evidence. In addition, evaluators must carefully match these sources of data to the needs and perceptions of different stakeholder groups (Guskey, 2012).

Results from large-scale assessments and nationally normed

standardized exams may be important for accountability purposes and will need to be included. In addition, school leaders often consider these measures to be valid indicators of success. Teachers, however, generally see limitations in large-scale assessment results.

These types of assessments are typically administered once a year, and results may not be available until several months later. By that time, the school year may have ended and students promoted to another teacher's class. So, although important, many teachers do not find such results particularly useful (Guskey, 2007).

Teachers put more trust in results from their own assessments of student learning — classroom assessments, common formative assessments, and portfolios of student work. They turn to these sources of evidence for feedback to determine if the new strategies or practices they are implementing make a difference.

Classroom assessments provide timely, targeted, and instructionally relevant information that can be used to plan revisions when needed. Since teachers comprise a major stakeholder group in any professional learning, sources of evidence that they trust and believe are particularly important.

Measures of student learning typically include cognitive indicators of student performance and achievement, such as assessment results, portfolio evaluations, marks or grades, and scores from standardized tests. Affective and psychomotor or behavioral indicators of student performance can be relevant as well.

Student surveys designed to measure how much students like school; their perceptions of teachers, fellow students, and themselves; their sense of self-efficacy; and their confidence in new learning situations can be especially informative. Evidence on school attendance, enrollment patterns, dropout rates, class disruptions, and disciplinary actions are also important outcomes.

In some areas, parents' or families' perceptions may be a vital consideration. This is especially true in initiatives that involve changes in grading practices, report cards, or other aspects of school-to-home and home-to-school communication (Epstein & Associates, 2009; Guskey, 2002c).

#### **MEANINGFUL COMPARISONS**

Evaluations of professional learning that extend to Level 5 should be made as methodologically rigorous as possible. Rigor, however, does not imply that only one evaluation method or design can produce credible evidence. Although randomized designs (i.e. true experimental studies) represent the gold standard in scientific research, especially in studies of causal effects, a wide range of quasi-experimental designs can produce valid results.

When evaluations are replicated with similar findings, that validity is further enhanced. One of the best ways to enhance an evaluation's methodological rigor is to plan for meaningful comparisons.

In many cases, data on outcomes at Level 5 are gathered from a single school or school district in a single setting for a

restricted time period. From a design perspective, such data lack reliability and validity. Regardless of whether results are positive or not, so many alternative explanations may account for the results that most authorities would consider such outcomes dubious at best and meaningless at worst (Guskey & Yoon, 2009).

It may be, for example, that the professional learning did lead to noted improvements. But maybe the improvements were the result of a change in leadership or personnel instead. Maybe the community or student population changed. Maybe changes in government policies or assessments made a difference. Maybe other simultaneously implemented interventions were responsible. The possibility that these or other extraneous factors influenced results makes it impossible to draw definitive conclusions.

The best way to counter these threats to the validity of results is to include a comparison group — another similar group of educators or schools not involved in the current activity or perhaps engaged in a different activity.

Ideal comparisons involve the random assignment of students, teachers, or schools to different groups. But because that is rarely possible in most education settings, finding similar classrooms, schools, or school districts provides the next best option.

In some cases, involvement in a professional learning activity can be staggered so that half of the group of teachers or schools that volunteer can be selected randomly to take part initially while the others delay involvement and serve as the comparison group. In other cases, comparisons can be made to matched classrooms, schools, or school districts that share similar characteristics related to motivation, size, and demographics.

Using comparison groups does not eliminate the effects of extraneous factors that might influence results. It simply allows planners greater confidence in attributing the results attained to the particular program or activity being considered. In addition, other investigative methods may be used to formulate important questions and develop new measures relating to professional growth (Raudenbush, 2005).

Student and school records provide the majority of data at Level 5. Results from questionnaires and structured interviews with students, parents, teachers, and administrators could be included as well. Level 5 data are used summatively to document a program or activity's overall impact.

Formatively, Level 5 can help guide improvements in all aspects of professional learning, including design, implementation, and follow-up. In some cases, data on student learning outcomes are used to estimate the cost-effectiveness of professional learning, sometimes referred to as return on investment (Parry, 1996; Phillips, 1997; Todnem & Warner, 1993).

### IMPLICATIONS FOR IMPROVEMENT

Three important implications stem from this model for evaluating professional learning.

**1. Each of the five evaluation levels is important.** Although evaluation at any level can be done well or poorly, the data

gathered at each level provide vital information for improving the quality of professional learning. And while each level relies on different types of information that may be gathered at different times, no level can be neglected.

- 2. Tracking effectiveness at one level tells little about impact at the next level.** Although success at an early level may be necessary for positive results at the next higher one, it is clearly not sufficient (Cody & Guskey, 1997). Breakdowns can occur at any point along the way. Sadly, most government officials and policymakers fail to recognize the difficulties involved in moving from professional learning experiences (Level 1) to improvements in student learning (Level 5). They also tend to be unaware of the complexity of this process, as well as the time and effort required to build this connection (Guskey, 1997; Guskey & Sparks, 2004).
- 3. Perhaps most important is this: When planning professional learning to impact student learning, the order of these levels must be reversed.** In other words, education leaders must plan backward (Guskey, 2001a, 2001b, 2003, 2014), starting where they want to end up and then working back (Hirsh, 2012).

### THE IMPORTANCE OF BACKWARD PLANNING

In backward planning, educators first decide what student learning outcomes they want to achieve and what data best reflect those outcomes (Level 5). Next they must determine, on the basis of pertinent research, what instructional practices and policies will most effectively and efficiently produce those outcomes (Level 4).

After that, leaders need to consider what aspects of organizational support need to be in place for those practices and policies to be implemented (Level 3). Then leaders must decide what knowledge and skills the participating professionals must have in order to implement the prescribed practices and policies (Level 2).

Finally, consideration turns to what set of experiences will enable participants to acquire the needed knowledge and skills (Level 1). What makes this backward planning process so important is that the decisions made at each level profoundly affect those made at the next.

The most effective professional learning planning begins with clear specification of the student learning outcomes to be achieved and the sources of data that best reflect those outcomes. With those goals articulated, school leaders and teachers then work backward.

Not only will this make planning much more efficient, but it also provides a format for addressing the issues most crucial to evaluation. As a result, it makes evaluation a natural part of the planning process and offers a basis for accountability.

### REFERENCES

- Achinstein, B. (2002).** Conflict amid community: The micropolitics of teacher collaboration. *Teachers College Record, 104*(3), 421-455.

- Alliger, G.M. & Janak, E.A. (1989).** Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology*, 42(2), 331-342.
- Chester, M.D. (2005).** Making valid and consistent inferences about school effectiveness from multiple measures. *Educational Measurement: Issues and Practice*, 24(4), 40-52.
- Cody, C.B. & Guskey, T.R. (1997).** Professional development. In J.C. Lindle, J.M. Petrosko, & R.S. Pankratz (Eds.), *1996 Review of research on the Kentucky Education Reform Act* (pp. 191-209). Frankfort, KY: The Kentucky Institute for Education Research.
- Colton, A.B. & Langer, G.M. (2005).** Looking at student work. *Educational Leadership*, 62(5), 22.
- Corcoran, T., Fuhrman, S.H., & Belcher, C.L. (2001).** The district role in instructional improvement. *Phi Delta Kappan*, 83(1), 78-84.
- Epstein, J.L. & Associates. (2009).** *School, family, and community partnerships: Your handbook for action* (3rd ed.). Thousand Oaks, CA: Corwin.
- 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. (2001a).** Backward planning: An outcomes-based strategy for professional development. *Curriculum in Context*, 28(2), 18-20.
- Guskey, T.R. (2001b).** The backward approach. *Journal of Staff Development*, 22(3), 60.
- Guskey, T.R. (2002a).** Does it make a difference? Evaluating professional development. *Educational Leadership*, 59(6), 45-51.
- Guskey, T.R. (2002b).** Professional development and teacher change. *Teachers and Teaching: Theory and Practice*, 8(3), 381-391.
- Guskey, T.R. (2002c).** *How's my kid doing? A parents' guide to grades, marks, and report cards*. San Francisco, CA: Jossey-Bass.
- Guskey, T.R. (2003).** Scooping up meaningful evidence. *JSD*, 24(4), 27-30.
- Guskey, T.R. (2005).** Taking a second look at accountability. *JSD*, 26(1), 10-18.
- Guskey, T.R. (2007).** Multiple sources of evidence: An analysis of stakeholders' perceptions of various indicators of student learning. *Educational Measurement: Issues and Practice*, 26(1), 19-27.
- Guskey, T.R. (2012).** The rules of evidence. *JSD*, 33(4), 40-43.
- Guskey, T.R. (2014).** Planning professional learning. *Educational Leadership*, 71(8), 10-16.
- Guskey, T.R. & Sparks, D. (2004).** Linking professional development to improvements in student learning. In E.M. Guyton & J.R. Dangel (Eds.), *Teacher education yearbook XII: Research linking teacher preparation and student performance*. Dubuque, IA: Kendall/Hunt.
- Guskey, T.R. & Yoon, K.S. (2009).** What works in professional development? *Phi Delta Kappan*, 90(7), 495-500.
- Hall, G.E. & Hord, S.M. (1987).** *Change in schools: Facilitating the process*. Albany, NY: SUNY Press.
- Hirsh, S. (2012).** Student outcomes are the driving force behind professional learning decisions. *JSD*, 33(5), 72.
- Holton, E.F. (1996).** The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7(1), 5-21.
- Kirkpatrick, D.L. (1959).** Techniques for evaluating training programs. A four-part series beginning in the November issue (Vol. 13, No. 11) of *Training and Development Journal* (then titled *Journal for the American Society of Training Directors*).
- Kirkpatrick, D.L. (1998).** *Evaluating training programs: The four levels* (2nd ed.). San Francisco, CA: Berrett-Koehler Publishers.
- Langer, G.M. & Colton, A.B. (1994).** Reflective decision making: The cornerstone of school reform. *Journal of Staff Development*, 15(1), 2-7.
- Little, J.W. (1990).** The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, 91(4), 509-536.
- Parry, S.B. (1996).** Measuring training's ROI. *Training & Development*, 50(5), 72-75.
- Phillips, J.J. (1997).** *Return on investment in training and performance improvement programs*. Houston, TX: Gulf Publishing Co.
- Raudenbush, S.W. (2005).** Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, 34(5), 25-31.
- 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.
- Supovitz, J.A. (2002).** Developing communities of instructional practice. *Teachers College Record*, 104(8), 1591-1626.
- Todnem, G. & Warner, M.P. (1993).** Using ROI to assess staff development efforts. *Journal of Staff Development*, 14(3), 32-34. ■

**The most effective professional learning planning begins with clear specification of the student learning outcomes to be achieved and the sources of data that best reflect those outcomes.**