Selected Publications on Professional Development and Instructional Improvement

**Evaluating Professional Development.**
* NSDC Book of the Year – 2000

**Professional Development in Education: New Paradigms and Practices.** (with M. Huberman).
* NSDC Book of the Year – 1996

“Closing the Knowledge Gap on Effective Professional Development.” *Educational Horizons,* 88(3).


* NSDC Best Non-Dissertation Research Award – 2003


* NSDC Article of the Year – 2002


“What Does It Mean To Be ‘Research-Based’?” *The Developer* (November 1992, p. 5).


“The Effects of Staff Development on Teachers’ Perceptions About Effective Teaching.” *Journal of Educational Research,* 1985 78(6), 378-381.

“Staff Development and Teacher Change.” *Educational Leadership,* 1985, 42(7), 57-60.

Dr. Guskey is Professor of Educational Psychology in the College of Education at the University of Kentucky, and well known for his work in professional development and educational change. He is included in the National Staff Development Council’s *Leaders in Staff Development,* and is the only person to have twice won the Council’s Book of the Year Award and three times the Article of the Year Award. Dr. Guskey served on the Policy Research Team of the *National Commission on Teaching & America’s Future,* on the Task Force to develop the *National Standards for Staff Development,* and recently was named a Fellow in the American Educational Research Association, one of the Association’s highest honors. He is the author/editor of 18 books and over 200 articles and papers, and has been featured on National Public Radio’s “Morning Edition” and “Talk of the Nation.” As a consultant to schools and education agencies around the world, he brings insights to education’s most complex problems.

Thomas R. Guskey
College of Education, University of Kentucky, Lexington, KY 40506
(Phone: 859-272-5841 E-mail: Guskey@uky.edu)
How do we know if professional development improves student learning?

Evaluating Professional Development

Thomas R. Guskey
Foreword by Dennis Sparks, Executive Director of NSDC

How do we determine the effects and effectiveness of activities designed to enhance the professional knowledge and skills of educators, so that they might improve the learning of students?

The purpose of this book is to offer educators practical guidance in asking good questions and gathering valid information, both to document the effects of professional development and to describe more precisely what contributes to its effectiveness.

Thomas R. Guskey explores the processes and procedures involved in evaluating professional development, from the very simple to the very complex, at five increasing levels of sophistication:

- Participants’ reactions to professional development
- How much participants learn
- Evaluating organizational support and change
- How participants use their new knowledge and skills
- Improvement in student learning

...complete with sample evaluation forms, checklists, and helpful hints and tips.

This is a comprehensive volume for use by district and site administrators, program evaluators, staff developers, and university faculty.

$34.95 paperback, $74.95 library edition

Use the convenient order form on the other side!
PROFESSIONAL DEVELOPMENT
IN EDUCATION
New Paradigms and Practices

Thomas R. Guskey and Michael Huberman
Foreword by Matthew B. Miles

The knowledge base in education is constantly expanding. Practitioners in education, like those in other professional fields, must keep abreast of this emerging knowledge base and must use it to upgrade their craft skills regularly. How this is to be accomplished can be viewed from a variety of perspectives, each with its own conceptual premises, each informed by different bodies of research, and each offering different prescriptions for improvement.

This book contains a collection of these varied perspectives, with individual chapters written by researchers and theorists renowned for their work in the area of professional and career development. Their contributions not only detail their particular perspective and the conceptual ground from which it derives, they also offer specific prescriptions for practice based on that perspective. Each chapter’s discussion is framed by a common set of issue questions provided by the editors; as a result, the reader can draw upon these unifying themes to compare the various perspectives.

In addition to the volume editors, the distinguished and internationally known contributors to this book include Hilda Borko, Michael Erath, Ralph Fessier, Michael Fulton, Andy Hargreaves, Zemira Mevarech, Ralph T. Putnam, Mark A. Smylie, John Smyth, Harm Tillemans, and Jeroen Imants.

Audience: Policy makers, researchers, and practitioners; graduate courses in educational administration, staff development, inservice education, and educational foundations.

Thomas R. Guskey, Professor of Educational Policy Studies and Evaluation, University of Kentucky, and Michael Huberman, visiting professor, Harvard Graduate School of Education and Senior Research Associate, The Network, Inc.

Questionnaires
and
Activities
**Professional Development Questionnaire**  
© Thomas R. Guskey

**Directions:** In the blanks below first describe the characteristics of the *most* effective (best) professional development experiences in which you have been involved. Then describe the characteristics of the *least* effective (worst) experiences.

<table>
<thead>
<tr>
<th>Question</th>
<th>Most Effective</th>
<th>Least Effective</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What was the topic or title?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Who planned the experience?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Who participated in the experience?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Who lead the experience? (Consultants or Staff)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. How large was the group that participated?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. When was the experience held?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. How long was the training portion of the experience?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. What types of activities were involved?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Did the experience involve changes in practice?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. How extensive were the suggested changes?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. How complex or difficult were the suggested changes?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Were there follow-up activities involved?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Who led the follow-up activities?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. What improvements did the experience bring?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Copies of Slides
Enhancing and Evaluating the Effectiveness of Professional Development

Thomas R. Guskey

Systemic Change

- Change is a Highly Complex Process
- Professional Development is Essential

Change is a Prerequisite for Improvement!

What Makes Professional Development Effective?

Consider your BEST and WORST Professional Development Experiences:

1. What was the Topic or Content?
2. Who Planned it?
3. Who Led it?
4. What were the Results?

Five Levels of Professional Development Evaluation:

1. Participants’ Reactions to the Experience
2. Participants’ Learning from the Experience
3. Organization Support & Change
4. Participants’ Use of New Knowledge & Skill
5. Results: Student Learning Outcomes
Five Levels of Professional Development Planning:

5. Results: Student Learning Outcomes
4. Participants' Use of New Knowledge & Skill
3. Organization Support & Change
2. Participants' Learning from the Experience
1. Participants' Reactions to the Experience

1. Program Topic?
2. Program Planning?
3. Program Participants?
4. Program Leadership?

5. Group Size?
6. Program Timing?
7. Length of Training?
8. Types of Activities?

9. Extent & Complexity of Change?
10. Follow-up & Improvements?

What Does Make Professional Development Effective?

Effective Professional Development

1. Begin with a clear focus on Learning and Learners!
Five Levels of Professional Development Planning:

5. Results: Student Learning Outcomes
4. Participants' Use of New Knowledge & Skill
3. Organization Support & Change
2. Participants' Learning from the Experience
1. Participants' Reactions to the Experience

Effective Professional Development

✓ There must be **Clear Goals**
   along with **Assessment Procedures**
   to document progress.

Solution: 3-Step Planning:

1. How does this activity relate to the school mission?
2. What are the intended student learning outcomes?
3. What evidence best reflects those outcomes?

Effective Professional Development

✓ Clear goals help mobilize **Everyone**
   and keep efforts **On Task**.

Effective Professional Development

✓ Goals prevent distraction by peripheral issues
   that **Waste Crucial Time**
   and **Divert Energy**.

Effective Professional Development

2. Engage in **Rigorous Self-Analysis**.
Effective Professional Development

**Self-Analysis requires:**
1. The **courage** to ask tough questions.
2. The **skill** to find honest answers.

Effective Professional Development

**Reaching your **Goals** requires a clear sense of Where You Are.**

Effective Professional Development

**We must continually ask:**
1. Who is not learning?
2. Why?
3. What can we do about it?

Effective Professional Development

**3. Recognize change is an Individual and Organizational process.**

Remember the **Stages of Concern**
1. Personal
2. Management
3. Impact

Consider the **Order of Change**
- Teacher Attitudes and Beliefs
- Teaching Practices
- Student Learning

---


Effective Professional Development

4. Think **Big**, but Start **Small**!

5. Ensure procedures are **Ongoing** and **Professionally Embedded**.

6. Work in **Teams** to maintain support!

---

- Change is **Dynamic** and **Large Scale**, but implemented through a series of **Smaller Steps**.

- We must balance the **Urgency for Improvement** with the **Need to Validate Effectiveness**.

- Change is a **Natural** and **Recurring Process** that is essential to **All Learning Environments**.
Effective Professional Development

7. Use **Available Resources!**

Effective Professional Development

8. Include procedures for **Feedback** on results!

Effective Professional Development

9. Provide continued **Follow-up, Support, and Pressure!**

Effective Professional Development

10. **Integrate Programs!**

**Major Models of Professional Development**

1. Training
2. Observation / Assessment
3. Involvement in a Development / Improvement Process
4. Study Groups
5. Inquiry / Action Research
6. Individually-Guided Activities
7. Mentoring

**Integrating Innovations**

- Performance-Based Education
- Critical Thinking Skills
- Learning Styles / Modalities
- Feedback, Correctives & Enrichment
- Summative Evaluation
- Effective Schools
- Mastery Teaching
- Mastery Learning
- Cooperative Learning
- TESA
- Instruction
- Formative Assessment
Professional Development and Improved Student Learning

Evaluating Professional Development

Thomas R. Guskey

Five Levels of Professional Development Planning:
1. Participants’ Reactions to the Experience
2. Participants’ Learning from the Experience
3. Organization Support & Change
4. Participants’ Use of New Knowledge & Skill
5. Results: Student Learning Outcomes

At Each Level We Must Consider:
A. What questions are addressed?
B. How will information be gathered?
C. What is measured or assessed?
D. How will the information be used?

Five Levels of Professional Development Evaluation:
1. Participants’ Reactions to the Experience
2. Participants’ Learning from the Experience
3. Organization Support & Change
4. Participants’ Use of New Knowledge & Skill
5. Results: Student Learning Outcomes
1. Participants’ Reactions
A. What questions are addressed?
   - Did they like it?
   - Was their time well spent?
   - Did the material make sense?
   - Will it be useful?
   - Was the leader knowledgeable & helpful?
   - Were the refreshments fresh & tasty?
   - Was the room the right temperature?
   - Were the chairs comfortable?

B. How will information be gathered?
   - Questionnaires and surveys administered at the end of session(s).

C. What is measured or assessed?
   - Initial satisfaction with the experience.

D. How will the information be used?
   - To improve program planning, design, and delivery.

2. Participants’ Learning
A. What questions are addressed?
   - Did participants acquire the intended knowledge and skills?

B. How will information be gathered?
   - Paper-and-pencil instruments
   - Performance assessments / tasks
   - Simulations
   - Demonstrations
   - Participant reflections
   - Participant portfolios
2. Participants’ Learning

C. What is measured or assessed?
   - New knowledge and skills of participants.

D. How will the information be used?
   - To improve the content, format, and organization of the program.

3. Organization Support & Change

A. What questions are addressed?
   - Was implementation advocated, supported, and facilitated?
   - Was the support public and overt?
   - Were problems addressed quickly and efficiently?
   - Were sufficient resources made available?
   - Were successes recognized and shared?
   - What was the impact on the organization?
   - Did it affect organizational climate and procedures?

B. How will information be gathered?
   - District and school records
   - Minutes from follow-up meetings
   - Questionnaires / Surveys
   - Interviews with participants or administrators
   - Participant reflections/portfolios

C. What is measured or assessed?
   - The organization’s advocacy, support, accommodation, facilitation, & recognition.

D. How will the information be used?
   - To document and improve organizational support.
   - To inform future change efforts.
4. Participants’ Use of New Knowledge & Skills

A. What questions are addressed?
   - Did participants effectively apply or implement the new knowledge & skills?

C. What is measured or assessed?
   - Degree and quality of implementation.

B. How will information be gathered?
   - Questionnaires / Surveys
   - Interviews with participants and their supervisors
   - Participant reflections
   - Direct observations
   - Video or audio tapes

D. How will the information be used?
   - To document and improve the implementation of program content.

5. Results: Student Learning Outcomes

A. What questions are addressed?
   - What was the impact on students?
   - Did it affect student performance or achievement?
   - Did it influence students’ attitudes, dispositions, or behaviors?
   - Are students more confident as learners?
   - Is student attendance improving?
   - Are dropouts decreasing?

B. How will information be gathered?
   - Student records
   - School records
   - Questionnaires / Surveys
   - Interviews with students, parents, teachers, and administrators
   - Participant portfolios
5. Results: Student Learning Outcomes

C. What is measured or assessed?
   - Student learning outcomes:
     a. Cognitive (Performance & Achievement)
     b. Affective (Attitudes & Dispositions)
     c. Psychomotor (Skills & Behaviors)

D. How will the information be used?
   - To focus and improve all aspects of program design, implementation, and follow-up.
   - To demonstrate the "ultimate impact" of professional development.

An Important Distinction:

Managers know how to do things right.

Leaders know the right things to do.

For Help or Additional Information:

Thomas R. Guskey
College of Education
University of Kentucky
Lexington, KY 40506

Phone: 859-272-5841
E-mail: Guskey@uky.edu
Readings
Results-oriented Professional Development: In Search of an Optimal Mix of Effective Practices

Procedural guidelines provide a framework for developing the optimal mix of professional development processes and technologies that will work best in a specific context at a particular point in time.

Thomas R. Guskey

Never before in the history of education has there been greater recognition of the importance of professional development. Every proposal to reform, restructure, or transform schools emphasizes professional development as a primary vehicle in efforts to bring about needed change. With this increased recognition, however, has come increased scrutiny. Questions are being raised about the effectiveness of all forms of professional development. And with these questions have come increased demands for demonstrable results. Legislators, policy makers, funding agencies, and the general public all want to know if professional development programs really make a difference. If they do, what evidence is there to show they are effective?

To address these questions professional developers are considering more seriously the issues of program evaluation. They are beginning to gather information more regularly on the outcomes of professional development activities. And this information is no longer limited to surveys of teachers' attitudes and practices. Increasingly, information on crucial measures of student learning is also being considered (Guskey & Sparks, 1991).

But perhaps more importantly, professional developers are looking more seriously at the research on professional development in education. They are examining what is known about the various forms of professional development, not only for teachers but for all those involved in the educational process. They also are considering what is known about various organizational characteristics and structures, especially those that facilitate ongoing professional growth.

This article will examine what that research says about the effectiveness of professional development. In particular, it will consider the mixed messages reporters are getting from this research and how we might

Thomas R. Guskey is a professor, University of Kentucky, Dept. of Educational Policy Studies and Evaluation, 131 Taylor Education Building, Lexington, KY 40506, 606/257-8600.

Journal of Staff Development

42 Fall 1994, Vol. 15, No. 4
make sense of diverse messages. It will then turn to a series of guidelines (or professional development goals) drawn principally from the research on individual and organizational change. Finally, the paper will turn our attention to the potential impact of implementing these guidelines.

Research on Professional Development

The research base on professional development in education is quite extensive. For the most part, however, it is research that has documented the inadequacies of professional development and, occasionally, proposed solutions (Epstein, Lockard, & Docher, 1988; Griffin, 1983; Guiseley, 1986; Joyner & Showers, 1988; Lichtenberg & Miller, 1976; Ornich, 1989; Wood & Thompson, 1980; 1993).

Still, reformers attempting to make sense of these various solutions quickly find themselves faced with seemingly incompatible dichotomies. For instance:

- Some researchers suggest that professional development efforts designed to facilitate change must be teacher specific and focus on the day-to-day activities of the classroom teacher (McLaughlin & Linsky, 1980; Witt, 1991). Others indicate that an emphasis on individual goals is detrimental to progress and more systemic or organizational approaches are necessary (Tye & Tye, 1984; Waugh & Punct, 1987).
- Some experts stress that reforms in professional development must be initiated and carried out by individual teachers and school-based personnel (Joyce, McNair, Ditz, & McKibbin, 1976; Lamberti, 1988; Lawrence, 1974; Massarella, 1980). Others emphasize that the most successful programs are guided by a clear vision that transcends the walls of individual classrooms and schools, since individual teachers and school-based individuals generally lack the opportunity to conceive and implement worthwhile improvements (Bhath, 1991; Clune, 1991; Mann, 1986; Wade, 1984).
- Some researchers argue that the most effective professional development programs are those that approach change in a gradual and incremental fashion, not expecting too much at one time (Dwyer & Dwyer, 1977; Fullan, 1985; Mann, 1978; Sparks, 1983). Others insist that the broader the scope of a professional development program, the more effort required of teachers, and the greater the overall change in teaching style attempted, the more likely the program is to elicit the enthusiasm of teachers and to be implemented well (Berman, 1978; McLaughlin & Marsh, 1978).

These and other similar dichotomies in the professional development literature leave reformers feeling confused. Many question how one can be expected to design, implement successful professional development programs when even researchers and experts in the field cannot agree on what should be done. While the critical issues of clear, solutions remain illusive. As a result, reformers are likely, desirably, in their attempts to address educators' many and highly diverse professional development needs.

The Search For An Optimal Mix

A major problem in these efforts to identify elements of successful professional development programs is that they are generally looking for "one right answer." Most begin by gathering evidence from a variety of studies, investigations, and program evaluations. This evidence is then combined and synthesized to identify those characteristics that are consistently associated with some measure of effectiveness. The modern technique, many researchers use to conduct such a synthesis is called "meta-analysis" (Hedges & Olkin, 1985). In most cases, program effectiveness is judged by an index of participants' satisfaction with the program or some indication of change in participants' professional knowledge. Only rarely is change in professional practice considered and rarer still is the assessment of impact on student learning (Guiseley & Sparks, 1991). The result of such an effort is usually a prescription comprised of general practices described in broad and nebulous terms. Unfortunately, such pre-scriptions offer little guidance to practically minded reformers who want to know precisely what to do and how to do it.

What is neglected in nearly all of these efforts is the powerful impact of context. In fact, synthesizing the evidence across studies is done specifically to eliminate the effects of context, or to decontextualize the data. Yet as Clark, Lotto, and Asato (1984), Fiske, and Corslett (1987), Fullan (1985), Huberman and Miles (1984), and others suggest, the uniqueness of the individual setting will always be a critical factor in education. While there may be some general principles that apply throughout, most will need to be adapted, at least in part, to the unique characteristics of that setting.

Businesses and industries operating in different sizes, in different countries or in different regions around the world may successfully utilize different processes to produce the same quality product. But reforms based upon assumptions of uniformity in the educational system repeatedly fail (Elmore & McLaughlin, 1988). The teaching and learning process is a complex endeavor that is embedded in contexts that are highly diverse. This combination of complexity and diversity makes it difficult, if not impossible, for researchers to come up with universal solutions (Guiseley, 1993; Huberman, 1983, 1985). We know with certainty that reforms in education today succeed to the degree that they adapt to and capitalize on this variability. In other words, they must be shaped and integrated in ways that best suit local values, norms, policies, structures, resources, and processes (Griffin & Barnes, 1984; McLaughlin, 1990; Tabber, McLaughlin, & Rowan, 1993).

Recognizing the importance of contextual differences brings clarity to the dichotomies described earlier. That is, successful change efforts in some contexts require professional development that focuses on teacher specific activities (Porter, 1986; Wise, 1991); in other contexts demand a more systemic or organizational approach (Sarasohn, 1990). In some contexts teacher-initiated efforts work best (Wentworth & Lipsky, 1977), while in others a more administratively directed approach may be needed (Mann, 1986). And while some contexts demand that professional development take a gradual approach to change (Sparks, 1983), others require immediate and drastic alterations at all levels of the organization (McLaughlin, 1990).

Fall 1994, Vol. 15, No. 4 43

Reforms based upon assumptions of uniformity in the educational system repeatedly fail.
Acknowledging the powerful influence of context also shows the futility of any search for “one right answer.” Rather than “one right answer,” there will be a collection of answers, each specific to a context. Our search must focus, therefore, on finding the optimal mix—that assortment of professional development processes and technologies that will work best in a specific context at a particular point in time.

In reviewing these guidelines, it is important to keep in mind that we know far more about professional development processes that fail than we do about those that succeed (Gall & Renczl, 1985; Gosselin, 1989; Joyce & Showers, 1987). There is no guarantee, therefore, that following these guidelines will result in successful professional development programs. Nevertheless, substantial evidence indicates that neglecting the issues described in these guidelines will most likely prevent success and, worst, will result in programs that fail to bring about significant or enduring change.

1. Recognize that change is both an individual and organizational process. An important lesson learned from the past is that we cannot improve schools without improving the skills and abilities of the teachers and principals within them. In other words, we must see change as an individual process and be willing to invest in the intellectual capital of those individuals who staff our schools (Wine, 1991).

Success in any improvement effort always hinges on the smallest unit of the organization and, in education, that is the classroom (McLaughlin, 1990). Teachers are the ones chiefly responsible for implementing change. Therefore professional development processes, regardless of their form (Sparks & Loosac-Horvitz, 1989), must not only relate to teachers, but must directly address their needs and concerns (Hall & Loosac, 1978; Weatherly & Lipsky, 1977).

Yet to see change as only an individual process can limit the effectiveness of professional development. Even changes that are empowering bring a certain amount of anxiety. And teachers, like professionals in many fields, are reluctant to adopt new practices or procedures unless they feel sure they can make them work (Lortie, 1975). To change or to try something new means to risk failure, and that is both highly embarrassing and threatening to one’s sense of professional pride (Pejoulx, 1990).

2. Form and maintain a level of comfort and skill that are sufficient to keep in mind that organizations, like individuals, must change (Sarason, 1973; Shreiner, 1990; Waugh & Punch, 1987). To focus exclusively on individuals in a professional development efforts and to neglect factors such as organizational features and system politics severely limits the likelihood of success (Berman, 1978; Clot, Holland, & Veal, 1990; Deal, 1987; Fullan & Pomfret, 1977; Parker, 1980). A debilitating environment can squash any change effort, no matter how much we expend individuals to persist (Beane, 1991).

To focus on change as only an organizational matter, however, is equally ineffectual. Fiddling with the organizational structure is a favorite device of education policy makers and administrators because it commutates to the public symbolically that they are concerned with the performance of the system. But as Elmore (1992) argues, evidence is scant that such structural change leads in any reliable way to changes in how teachers teach, what they teach, or how students learn. McLaughlin (1990) describes this as the difference between macro-level confusion and micro-level readiness. To facilitate change, we must look beyond policy and consider the embedded structure that most directly affects the actions and choices of the individuals involved.

The key is to find the optimal mix of individual and organizational processes that will contribute to success in a particular context. In some situations, individual initiative and motivation might be quite high, but organizational structures stand in the way of significant improvement. In other, progressive and supportive organizational structures may be in place, but the lack of personal incentives for collaboration and experimentation inhibit any meaningful change in classroom practice. Viewing change as both an individual and organizational process that must be adapted to contextual characteristics will help clarify the steps necessary for success in professional development.

3. Think big but start small. There is no easier way to sabotage change efforts than to take on too much at one time. In fact, if there is one truism in the vast research literature on change it is that the magnitude of change people are asked to make is inversely related to their likelihood of making it (Guskey, 1991). Professionals at all levels generally
Because of the powerful and dynamic influence of context, it is impossible to make precise statements about the elements of an effective professional development program. Even programs that share a common vision and seek to attain comparable goals may need to follow very different pathways to succeed.

oppose radical alternatives to their present procedures. Hence, the probability of their implementing a new program or innovation depends largely on their judgment of the magnitude of change required for implementation (Dyke & Ponder, 1977; Fullan, 1982; Mann, 1978).

Successful professional development programs are those that approach change in a gradual and incremental fashion. Efforts are made to illustrate how the new practices can be implemented without causing too disruptive or require a great deal of extra work (Sparks, 1983). If a new program does require major changes, it is best to ease into its use rather than expect comprehensive implementation at once (Fullan, 1985).

But while the changes advocated in a professional development program may not be so dramatic that they require too much too soon, they need to be sufficient in scope to challenge professionals and kindle interest (McLaughlin, 1990; Crandall, Eisemann, and Louis, 1986) argue that the greatest success is likely when the size of the change is not so massive that typical users find it necessary to adopt a coping strategy that seriously dissonant the change, but large enough to require noticeable, sustained effort. Modest, narrowly conceived projects seldom bring the enthusiasm necessary for implementation. This is what is meant by "think big."

The key again is to find the optimal mix. Professional development efforts should be designed with long-term goals based on the assumption that they are not too small, and those larger goals should be accompanied by a strategic plan that includes specific incremental goals for three to five years into the future, gradually expanding on what is successful in that context and offering support to those engaged in the change (Fullan, 1992; Louis & Miles, 1990).

3. Work in teams to maintain support

The discomfort that accompanies change is greatly compounded if the individuals involved perceive they have no say in the process or if they feel isolated and detached from the implementation efforts. For this reason it is imperative that all aspects of a professional development program be fashioned to involve teams of individuals working together. This means that planning, implementation, and follow-up activities should be seen as joint efforts, providing opportunities for those with diverse interests and responsibilities to offer their input and advice (Massarelli, 1980).

To ensure that the teams function well and gain broad-based support for professional development efforts, it is important they involve individuals from all levels of the organization. In school improvement programs, for example, the key professional development teams include teachers, non-instructional staff members, and building and central office administrators (Caldwell & Wood, 1988).

In some contexts the involvement of parents and community members also can be helpful (Lerom, 1989). Although the roles and responsibilities of these individuals in the professional development process will be different, all have valuable insights and expertise to offer.

Still, the notion of teamwork must be balanced. There is evidence to show, for instance, that large-scale participation during the early stages of a change effort is sometimes counterproductive (Huberman & Miles, 1984). Elaborate needs assessments, endless committee and taskforce meetings, and long and tedious planning sessions can create confusion and alienation in the absence of any action. Extensive planning can also exhaust the energy needed for implementation so that by the time change is to be enacted, people are burned out (Fullan, 1991).

Furthermore, broad-based participation in many decisions is not always essential or possible on a large scale (Dawson, 1981; Hoop & Blackwell, 1980). As Little (1989) argues, there is nothing particularly virtuous about teamwork or collaboration per se. It can serve to block change or inhibit progress just as easily as it can serve to enhance the process.

To facilitate change, teamwork must be linked to established norms of continuous improvement and experimentation. In other words, teamwork and collaboration must be balanced with the expectation that all involved in the process—teachers, administrators, and non-instructional staff members—are constantly seeking and assessing potentially better practices (Little, 1989). Such a balance promotes collegial interac-

Acknowledgments

The authors would like to thank the members of the National Staff Development Council for their assistance in the development of this article.
A debilitating environment can squash any change effort, no matter how much we exhort individuals to persist.

Role-playing is essential to the success of any professional development effort. Personal feedback on results can be provided in a variety of ways, depending on the context. In professional development programs involving the implementation of mastery learning (Bloom, 1968, 1971), for example, teachers receive this feedback from their students through popular formative assessments (Bloom, Madaus & Haris, 1981).

In mastery learning classrooms, formative assessments are used to guide student with feedback on their learning progress and to diagnose learning problems. As such, they can take many forms, including writing samples, skill demonstrations, projects, reports, performance tasks, or other, more objective assessment devices such as quizzes or tests. These assessments are then paired with corrective activities designed to help students remedy any learning errors identified through the assessment.

Yet despite its importance, procedures for gathering feedback on results must be balanced with other concerns. The methods used to obtain feedback, for example, must not be disruptive to instructional procedures. Furthermore, they should not require immediate amounts of time or extra work from those engaged in the difficult process of implementation.

Tuning issues are also critical, for it is unfair to expect too much too soon from those involved in implementation. As Locksley-Horsley et al. (1987) point out, this is analogous to pulling a plow out of the ground each day to check its root of growth. In other words, the need for feedback must be adapted to the characteristics of the program and the setting. Feedback procedures must focus on outcomes that are meaningful to the professionals involved, but also aimed to be at best an object of the program and the setting. Feedback procedures must focus on outcomes that are meaningful to the professionals involved, but also aimed to be at best an object of the program and the setting.

5. Provide continued follow-up, support, and pressure. Few persons can move from a professional development experience directly into successful implementation. In fact, few will even venture into the uncertainty of implementation unless there is an appreciation of the difficulties and potential problems involved (Fullane & Miles, 1992).

Fitting new practices and techniques to match the job that is to be done is an open-ended process that requires time and extra effort, especially with beginning teachers (Berman, 1978; Joyce and Showers, 1980). Guidance, direction, and support with pressure are crucial when these adaptations are being made (Baldridge & Deal, 1975; Fullane; 1991; Parry, 1985; Wragg and Park, 1987). What makes the early stages of implementation so complicated is that the problems encountered at this time are often multiple, pervasive, and unanticipated.

Miles and Louis (1990) point out that developing the capacity to deal with these problems promptly, actively, and in some depth may be "the single biggest determinant of program success" (p. 60). And regardless of how much advanced planning or preparation takes place, it is when professionals actually implement the new ideas or practices that they have the most specific problems and doubts (Berman, 1978; Fullane & Miles, 1992).

Support coupled with pressure at this time are vital for continuation. Support allows those engaged in the difficult process of implementation to tolerate the anxiety of occasional failures. Pressure is often necessary to initiate change among those with little self-confidence for change (Aratian, 1987; Huberman & Crandall, 1983). In addition, it provides the encouragement, motivation, and occasional nudging that many practitioners require to persist in the challenging tasks that are intrinsic to all change efforts.

Of all aspects of professional development, this aspect of support with pressure is perhaps the most neglected. It makes clear that if we are to be successful, professional development must be set in a process, not an event (Locksley-Horsley et al., 1987). Learning to be proficient at something new or finding meaning in a new way of doing things is difficult and sometimes painful. Furthermore, any change that holds great promise for increasing individuals' competence or enhancing an organization's effectiveness is likely to be slow and require extra work (Huberman & Miles, 1984). It is impera-

The key is to find the optimal mix of individual and organizational processes that will contribute to success in a particular context.

46 Journal of Staff Development
If a new program or innovation is to be implemented well, it must become a natural part of practitioners' repertoire of professional skills and built into the normal structures and practices of the organization (Fullan & Miles, 1992; Miles & Louis, 1987). For advances to be made and professional improvements to continue, the new practices and techniques that were the focus of the professional development effort must be used automatically. And for this to occur, continued support and encouragement, pared with subtle pressure to persist, are essential.

The crucial support with pressure can be offered in a variety of ways. McLaughlin and Marsh (1978) recommend that local resource personnel or consultants be available to provide on-line assistance when difficulties arise. They emphasize, however, that the quality of the assistance is critical and that it is better to offer no assistance than poor or inappropriate assistance.

Joyce and Showers (1988), suggest that support for change take the form of coaching. They advise guiding teacher in adapting and applying the new practice to their unique contextual conditions, helping them to analyze the effects of their efforts, and urging them to continue despite minor setbacks. In other words, coaching is personal, practical, on-the-job assistance that can be provided by consultants, supervisors, directors, or professional colleagues. Simply offering opportunities for practitioners to interact and share ideas with each other also can be valuable (Massarella, 1980; McLaughlin & Marsh, 1978).

Here again, the notion of balance is critical. Resource contexts a substantial amount of pressure from leaders may be necessary to overcome inertia, recalcitrance, or outright resistance (Munn, 1986). It is possible, for example, when making decisions about instructional practices to overemphasize teachers' personal preferences and underemphasize concerns about student learning (Buchmann, 1986). Yet in contexts in which there is considerable individual initiative, such pressure may be seen as a strong-armed metric and unprofessional (Leiter & Cooper, 1978). The key is to find the optimal mix for that context, understanding well the interpersonal dynamics of the individuals involved and the culture of the organization in which they work.

6. Integrate Programs: More so than any other professional, education seems fraught with innovation. In fact, innovations seem to come and go in education about as regularly as the seasons change. Each year new programs are introduced in schools without any effort to show how they relate to the ones that came before or those that may come afterward. Furthermore, there is little mention of how these various innovations contribute to a growing professional knowledge base. The result is learning the peculiar load of fragmented, uncoordinated, and ephemeral attempts at change (Fullan & Miles, 1992).

The steady stream of innovations in education causes many practitioners to view all new programs as isolated fads that will soon be gone, only to be replaced by yet another bandwagon (Latham, 1988). This pattern of constant yet unrelated, short-term innovations not only obscures improvement and provides cynicism, it also imposes a sense of affliction. Hearing a multitude of innovations come into and go out of fashion, veteran teachers frequently calm the fears of their less experienced colleagues who express concern about implementing a new program with the advice, "Don't worry, this too shall pass."

If professional development efforts that focus on the implementation of innovations are to succeed, they must include precise descriptions of how these innovations can be integrated. That is, each new innovation must be presented as part of a coherent framework for improvement. It is difficult enough for practitioners to learn the particular features of one innovation, let alone to figure out how it can be combined with others. And because no single innovation is totally comprehensive, implementing only one will leave many problems unresolved. It is only when several strategies are carefully and systematically integrated that substantial improvements become possible. Doyle (1992), Sarason (1990), and others also emphasize that coordinating programs and combining ideas releases great energy in the improvement process.

In recent years, insightful researchers have described how different combinations of innovations can yield impressive results (e.g., Arredondo & Block, 1990; Davidson & O'Leary, 1990; Guskey, 1988, 1990a; Mesvarech, 1985; Welter, 1990). In addition, several frameworks for integrating a collec-
tion of programs or innovations have been developed that practitioners are finding especially useful.

One example is a framework developed by Marzano, Pickering, and Brandt (1999) based on various dimensions of development. Another developed by Guskey (1990b) is built around five major components in the teaching and learning process. These frameworks allow skilled practitioners to see more clearly the linkages between various innovations. They also offer guidance to the efforts of seriously minded reformers seeking to pull together programs that collectively address the problems they are most pressing in a particular context.

A crucial point here is that the particular collection of programs or innovations that is best understood will vary from setting to setting. As a result, the way linkages are established and applications integrated will need to vary as well.

Fullan (1992) stresses that, "schools are not in the business of managing single innovations; they are in the business of composing with multiple innovations simultaneously." (p. 19). By recognizing the dimensions of learning that a particular innovation stresses, the components of the teaching and learning process it emphasizes, savvy educators can pull together innovations that collectively address what is most needed in that context at a particular point in time.

Conclusion

The ideas presented in these procedural guidelines are not really new and certainly cannot be considered revolutionary. They may, in fact, appear obvious to those with extensive experience in the professional development process. Yet as self-evident as they may seem, it is rare to find a professional development program today that is designed and implemented with thorough attention to these guidelines or the factors that underlie them. It is rare to find professional development programs that evaluate the implementation of these guidelines in terms of their effects on student learning.

What is evident from these guidelines is that the key to greater success in professional development, which translates to improvements in student learning, rests not so much in the discovery of new knowledge, but in our capacity to deliberately and wisely use the knowledge we have. This is true regardless of whether professional development is viewed as an integral part of one's career cycle, as a self-directed journey of finding meaning and appreciation in one's work, or as a path to growth and advancement in one's field. To develop this capacity requires a clear vision of our goals and a thorough understanding of the process by which those goals can be achieved.

In the minds of many today, there is a clear vision of what would be ideal in professional development, but little sense of how educators at all levels constantly in search of new and better ways to address the diverse learning needs of their students. Yet it's schools as a whole that are continually engaged in inquiry and stimulating discourse. They see practitioners in education respected for their professional knowledge and pedagogical skill.

The exact process by which that vision can be accomplished, however, is blurred and confused. The reason, as we have argued here, is that the process is so highly contextualized. There is no "one right answer" or "one best way." Rather, there are a multitude of ways, all adapted to the complex and dynamic characteristics of specific contexts. Success, therefore, rests in finding the optimal mix of process elements and strategies that keep professionals' interest sated, and thoughtfully applied in a particular setting.

While it is true that the ideas presented here offer an optimistic perspective on the potential of professional development in education, these ideas are not far-fetched. They illustrate that although the process of change is difficult and complex, we are beginning to understand how to facilitate that process through pragmatic adaptations to specific contexts so that ongoing professional growth and improved professional practice are ensured. Doing so is essential to the improved learning of all students.

Reference Note

This article is adapted from a chapter in J.R. Guskey and M. Huberman (Eds.), New Paradigms and Practices in Professional Development, soon to be published by Teachers College Press. Copyright T. Guskey.

References


ding the connections between thinking skills and mastery learning. Educational Leadership, 47(5), 4-10.


Fall 1994, Vol. 12, No. 4
Interviews With Educational Leaders

by Dennis Sparks

Interviews with a number of educational leaders have been printed in the Journal of Staff Development and The Developer on topics such as staff development and school improvement, change, cognitive coaching, cooperative schools, leadership, and research into practice.

Interviews with some of these educational leaders are now collected into one volume. Those interviewed include:

- Ernest Bover
- Michael Fullan
- Robert Garrison
- David and Robert Joens
- Judy-Ann Krupp
- Susan Lookis-Horsley
- Phil Schlechter
- J. Martin Shanker
- Leo Shulman
- Inge Stallings
- Richard Wallace
- Pat Wolfe

Copies may be ordered for $12.50, or $10 for NSDC members.

Send orders to:
NSDC
P.O. Box 240
Oxford, OH 45056
Telephone (513) 523-6029

Available from the National Staff Development Council

Fall 1994, Vol. 10, No. 4
Educational Leadership

March 2002 | Volume 59 | Number 6
Redesigning Professional Development | Pages 45-51

Does It Make a Difference?
Evaluating Professional Development

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 skills 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**

|------------------|--------------------------------|----------------------------------|-------------------------------|--------------------------------|
| **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 | Paper-and-pencil instruments  
New knowledge and skills of participants | | To improve program content, format, |
<table>
<thead>
<tr>
<th>Intended Knowledge and Skills?</th>
<th>Simulations, Demonstrations, Participant Reflections (oral and/or written), Participant Portfolios</th>
<th>and Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3. Organization Support &amp; Change</strong></td>
<td><strong>Was implementation advocated, facilitated, and supported?</strong></td>
<td><strong>District and school records, Minutes from follow-up meetings, Questionnaires, Structured interviews with participants and district or school administrators, Participant portfolios</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Was the support public and overt?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Were problems addressed quickly and efficiently?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Were sufficient resources made available?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Were successes recognized and shared?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>What was the impact on the organization?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Did it affect the organization's climate and procedures?</strong></td>
<td></td>
</tr>
<tr>
<td><strong>4. Participants' Use of New Knowledge and Skills</strong></td>
<td><strong>Did participants effectively apply the new knowledge and skills?</strong></td>
<td><strong>Questionnaires, Structured interviews with participants and their</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 5. Student Learning Outcomes

| 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? | Student records | School records | Questionnaires | Structured interviews with students, parents, teachers, and/or administrators | Participant portfolios | Student learning outcomes: |
|---------------------------------|-----------------------------------------------|-------------------------------------------------|---------------------------------|----------------------------------|--------------------------------|-------------------------|-----------------|----------------|--------------------------------|------------------|-----------------------------------------------|-----------------|-----------------------------------------------|-----------------|
|                                 |                                               |                                                 |                                 |                                  |                                |                         |                 |                | *Cognitive (Performance & Achievement)* |                                | *Affective (Attitudes & Dispositions)* |                   |                                | *Psychomotor (Skills & Behaviors)* |
|                                 |                                               |                                                 |                                 |                                  |                                |                         |                 |                |                                 |                                |                                 |                   |                                |                                 |

**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, 2002).

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 & Hirsch, 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 example, 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, membership 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, 2002), 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


Thomas R. Guskey (guskey@kent.edu) is a professor of education policy studies and evaluation, College of Education, University of Kentucky, Taylor Education Building, Lexington, KY 40506, and author of Evaluating Professional Development (Corwin Press, 2000).

Copyright © 2002 by Thomas R. Guskey

© Copyright ASCD. All rights reserved.
Analyzing Lists of the Characteristics of Effective Professional Development to Promote Visionary Leadership

Thomas A. Guskey

In recent years, different researchers and research agencies, teacher associations, national education organizations, and the U.S. Department of Education have published lists of the characteristics of effective professional development to guide school leaders in their improvement efforts. This study analyzed 13 of the better known of these lists to determine whether they were derived through similar procedures, based on similar frames of reference, and included the same elements or characteristics. Results show that individual characteristics vary widely in their frequency of inclusion in the lists and that no characteristic is consistently named in all lists. In addition, research evidence supporting most of the identified characteristics is insufficient and often contradictory. Implications for leaders interested in improving professional development activities are discussed, as well as ways to enhance efforts to identify the characteristics of effective professional development.

The education reforms of the past decade brought new prominence to the role of professional development. Recognizing that schools can be no better than the teachers and administrators who work in them, policymakers emphasized professional development as a key component in nearly every education improvement plan. The recently enacted No Child Left Behind Act of 2001 (NCLB, 2002), for example, stresses the importance of high-quality professional development to guarantee that all teachers are "highly qualified" and that all students reach high levels of achievement. This emphasis compels visionary leaders to reexamine professional development efforts with a keen eye for issues of value and worth.

With this increased prominence, however, has come increased scrutiny. Professional development's tainted histories of effectiveness and the lack of strong evidence showing its direct link to improvement as student learning outcomes (see, e.g., Corcoran, 1995; Frechtling, Staff, Carew, & Baden-Krieman, 1995; Newman, King, & Youngs, 2000; Wang, Frechtling, & Sanders, 1995) led those same policymakers to demand assurances of quality in these endeavors. This, in turn, prompted the publication and dissemination of an assortment of "lists" describing the characteristics of high-quality, effective professional development.

Those involved in professional development research or practice, as well as school leaders at all levels, are undoubtedly familiar with several of these lists. They have been produced in recent years by researchers and research agencies (e.g., Educational Research Service and Educational Testing Service), teacher associations (e.g., American Federation of Teachers [AFT]), national education organizations (e.g., National Partnership for Excellence and Accountability in Teaching), and the U.S. Department of Education.

A quick inspection shows, however, that although some overlap exists among the identified characteristics, these lists are not identical. Some characteristics are frequently mentioned, whereas others appear unique to a particular list. Such a lack of consensus undoubtedly frustrates and confuses those responsible for designing and implementing high-quality professional development programs for educators. How can they ensure that their efforts will be effective when there appears to be little agreement about the factors that contribute to effectiveness?

To clarify these efforts and to determine the degree of consensus that exists among the individuals and organizations responsible for developing these lists, 13 of the better known lists of the characteristics of effective professional development were selected for analysis in this investigation. The analysis had three goals: The first goal was to determine whether the lists were derived in comparable ways. In other words, were they drawn from the same sources of evidence and based on similar frames of reference? The second goal was to ascertain whether specific characteristics appear on all of the lists. This would indicate strong consensus regarding the contribution of these characteristics to the effectiveness of professional development. The final goal was to verify how well the identified characteristics correspond to the recently revised Standards for Staff Development, (National Staff Development Council [NSDC], 2001), published to help educators at all levels improve the quality of professional development endeavors.

The Lists

The lists included in the analysis are described below in chronological order, based on their publication date. They were collected from a search of the Educational Resources Information Center (ERIC) system, from a targeted Internet search, and from related professional development literature (i.e., publications from NSDC). Although all of the lists were developed within the last decade, no claim of completeness can be made for this list of lists, and other equally valid lists may exist. Nevertheless, it is believed these 13 lists comprise a fairly representative sample of those developed in recent years, and clearly rank among those best known to researchers and school leaders.
The first and earliest list considered was assembled by Thomas B. Corcoran (1995a, 1995b) as part of a project sponsored by the National Governors' Association and coordinated by the Consortium for Policy Research in Education. Corcoran reviewed papers and reports published by individuals well known in professional development circles (e.g., Griffin-Hodges, Joyce & Showers, Loucks-Horsley, Little, Price, and Zimpher & Hower) and the Standards for Staff Development, Middle Level Edition, prepared by the NSDC (1994). From these sources, he identified eight guiding principles that "a number of experts and organizations have suggested contribute to the more promising professional development programs or policies" (1995a, p. 3).

The second was a list of nine guidelines published by AFT (1996) in Principles for Professional Development. These guidelines were derived from teachers' views of effective professional development.

The third list came from a report by Willis D. Hlavay and Linda Valli (1995, 1999), who reviewed the literature on professional development and identified eight characteristics associated with effective professional development programs. According to Hlavay and Valli, these eight characteristics represent "an almost unprecedented consensus... among researchers, professional development specialists, and key policymakers on a vision of ways to substantially increase the knowledge and skills of educators" (1996, p. 1).

The fourth list was based on the work of Susan Loucks-Horsley, Katherine Stiles, and Peter Hewson (1996), who examined a variety of standards and related materials as part of the Professional Development Project of the National Institute for Science Education. Their goal was to determine whether the science, mathematics, and professional development communities share a common understanding of effective professional development. This effort led to the identification of seven principles that consistently marked "the best professional development experiences for science and mathematics educators" (p. 1).

The fifth list came from a 1997 report published by the U.S. Department of Education titled National Study of Achievement Emmy Clark. This report included 10 principles for high-quality professional development on which "experts agree" (p. 3); however, the report provided no information as to who are these "experts."

The sixth list was part of a research summary conducted by the Educational Research Service (1998) titled Professional Development for Teachers: Challenges and Trends. This summary drew on reports by NSDC (1994), the National Commission on Teaching America's Future (1996), the National Foundation for the Improvement of Education (1996), and the U.S. Department of Education (1998). It extended a review by Sparks and Loucks-Horsley (1998) of the characteristics of effective professional development by articulating 16 additional, "research-based principles and standards" (p. 5).

The seventh list was developed from a research project sponsored by the National Institute for Science Education and conducted by Mary M. Kennedy (1998). For this project, Kennedy identified a pool of 95 studies that examined the effectiveness of various approaches to continuing teacher education in either mathematics or science. Kennedy limited her review, however, to the 10 investigations that included evidence of benefits to students and reached six overall conclusions about factors that contribute to this "ultimate benefit."

The eighth list came from the work of Beaucage F. Burns, Laura Desimone, Andrew C. Porter, and Michael S. Garet (2000; see also Garet, Porter, Desimone, Burns, & Yoon, 2001). These researchers surveyed a nationally representative sample of more than 1,000 teachers who participated in professional development sponsored in part by the federal government's Eisenhower Professional Development Program. As part of their national evaluation, they also conducted six exploratory case studies and 10 indepth case studies in five states (Garet, Burns, Porter, Desimone, Herman, & Yoon, 1999). By studying the literature and analyzing the survey data, they identified three structural features and three core features that characterize effective professional development experiences.

The ninth list was developed as part of the California Professional Development Reform Initiative. In this project, Karen Rent and Carol Lingusen (2006) led a working committee drawn from three partner organizations: the California Staff Development Council, the Center for the Future of Teaching and Learning, a private nonprofit organization, and the California Department of Education. After reading summaries of recent professional development reports and articles (e.g., Sparks, 1997; Sparks & Hirsh, 1997), committee members drafted "organizing themes around the nature of quality professional development for teachers" (p. 34). The group then tested these themes by gathering interview data from teachers, and finally developed a list of 10 design elements for high-quality professional development.

The tenth list stemmed from the work the National Partnership for Excellence and Accountability in Teaching (2000). Based on presentations and interviews with participants from the 1999 Conference of Teachers Professional Development and stories from several schools and districts that have been using student-centered strategies for professional development, the Partnership developed a list of nine "research-based principles" for improving professional development.

The eleventh list came from a report prepared for the Urban Special Education Leadership Collaborative by Mary Terris (2000) of the Education Development Center. Terris synthesized the current literature on effective professional development for special school reform and outlined eight specific design features that contribute to effectiveness.

The twelfth list was derived from a study by Harold Wenglinsky (2002) of the Educational Testing Service. Wenglinsky applied multilevel structural equation modeling to data drawn from the 1996 National Assessment of
The results from the analysis are in some ways validating and in other ways surprising. They show, for example, that there is no relationship between the presence of certain characteristics and the performance of professionals. In some cases, the characteristics were not even associated with any performance measure. This suggests that there may be other factors at play that influence performance, such as individual skills or experience.

The analysis also revealed some interesting patterns. For instance, it was found that professionals with higher levels of experience were more likely to exhibit certain characteristics, such as a strong sense of mission and a commitment to continuous improvement. This is important because it suggests that training and development programs may need to focus on these areas to effectively improve performance.

The findings also highlight the importance of a comprehensive approach to professional development. Instead of focusing solely on individual characteristics, it is important to consider the role of the organization and the broader work environment. This can help to create a culture that supports the development of professional competencies and enhances overall performance.

In conclusion, the results of this analysis provide valuable insights into the characteristics of effective professionals and the factors that influence their performance. By understanding these factors, organizations can develop more effective training and development programs that will help professionals to achieve their full potential.
<table>
<thead>
<tr>
<th>Characteristic/principle</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Enhances teachers' content and pedagogic knowledge</td>
<td>3, 1, 2, 5, 7, 15</td>
</tr>
<tr>
<td>2. Provides sufficient time and other resources</td>
<td>7</td>
</tr>
<tr>
<td>3. Promotes collegiality and collaboration</td>
<td>5</td>
</tr>
<tr>
<td>4. Includes procedures for evaluation</td>
<td>5</td>
</tr>
<tr>
<td>5. Aligns with other reform initiatives</td>
<td>6</td>
</tr>
<tr>
<td>6. Models high-quality instruction</td>
<td>4</td>
</tr>
<tr>
<td>7. Is school or site based</td>
<td>1</td>
</tr>
<tr>
<td>8. Builds leadership capacity</td>
<td>5</td>
</tr>
<tr>
<td>9. Based on teachers' identified needs</td>
<td>8</td>
</tr>
<tr>
<td>10. Driven by analyses of student learning data</td>
<td>1</td>
</tr>
<tr>
<td>11. Focuses on individual and organizational improvement</td>
<td>2</td>
</tr>
<tr>
<td>12. Includes follow-up and support</td>
<td>7</td>
</tr>
<tr>
<td>13. Is ongoing and/or embedded</td>
<td>8</td>
</tr>
<tr>
<td>14. Helps accommodate diversity and promote equity</td>
<td>10</td>
</tr>
<tr>
<td>15. Based on best available research evidence</td>
<td>4</td>
</tr>
<tr>
<td>16. Takes a variety of forms</td>
<td>9</td>
</tr>
<tr>
<td>17. Provides opportunities for theoretical understanding</td>
<td>7</td>
</tr>
<tr>
<td>18. Driven by an image of effective teaching and learning</td>
<td>8</td>
</tr>
<tr>
<td>19. Provides for different phases of change</td>
<td>8</td>
</tr>
<tr>
<td>20. Promotes continuous inquiry and reflection</td>
<td>6</td>
</tr>
<tr>
<td>21. Involves families and other stakeholders</td>
<td>12</td>
</tr>
</tbody>
</table>

Hill, 1998, 2000. Kennedy, 1998, 1999. Wenglinsky, 2002. A few smaller studies have considered the importance of teachers’ literacy context knowledge. (McGinley, 1993; Moser, 1994), but these have focused narrowly on special education contexts. For the most part, whether this relationship is true for achievement in language arts, social studies, foreign languages, or other subject areas has not been thoroughly investigated. Ten of the lists include the provision of sufficient time and other resources as essential to effective professional development. Obviously, educators need time to deepen their understanding, analyze students’ work, and develop new (continues)
approaches to instruction. Data gathered by Birman et al. (2000) further showed that "activities of longer duration have more subject-area content, focus on more practices, and involve more coherence with teachers' other experiences than do shorter activities" (p. 30). But significant contrary evidence exists. Kenyon's (1998) analysis of differences in the time spent in professional development activities were unrelated to improvements in student outcomes. How time was distributed also yielded mixed results: with no differences found in mathematics studies and only modest effects found in science investigations (Wenger 2002). Study of mathematics achievement found that among the various aspects of professional development investigated, "the amount of time is not significantly related to achievement" (p. 10). Therefore, although effective professional development clearly requires time, it also seems clear that such time must be well organized, carefully structured, and purposefully directed (Guskey, 1999).

Another consistently noted characteristic is the promotion of collegiality and collaborative exchange. Educators at all levels value opportunities to work together, reflect on their practices, exchange ideas, and share strategies and expertise. Collaboration also helps build a sense of community (Supovitz, 2002). But as Little (1998) pointed out, there is nothing particularly virtuous about collaboration: people come to block change or inhibit progress just as easily as they can to enhance the process. Recent investigations showed that many collaborative efforts run headlong into enormous conflicts over professional beliefs and practices for which practitioners are generally ill-prepared (Achinstein, 2002). For collaboration to bring its intended benefits it too, needs to be structured and purposeful, with efforts guided by clear goals for improving student learning.

Most of the lists and the NSDC Standards stress the inclusion of specific evaluation procedures. This emphasis may stem from growing awareness among educators at all levels of the need to gather regular, informative information to guide their improvement efforts (Guskey, 2000, 2002). Of course, policymakers' new demands for accountability in professional development undoubtedly contribute to an emphasis on evaluation as well (Fuhrman, 1999; Guskey, 1998; Killion, 2002).

The majority of lists recognize the need for professional development activities to be aligned with other reform initiatives and to model high-quality instruction. These characteristics likely come from appreciation of the largescale, comprehensive, and systemic nature of most modern education reform initiatives and from increased awareness of similarities in the learning patterns of adults and children (Consortium for Policy Research in Education, 1996; Darling-Hammond & McLaughlin, 1995; Fullan, 1995).

Most lists also stress that professional development should be school or site based, even though significant research evidence suggests this may not always be effective (see Hollower, 1990; Latanem, 1998). A recent review by Corcoran, Fuhrman, and Belcher (2001), for example, found when decisions about professional development were primarily school based, "school staff invested paid lip service to the use of research and were more interested in designing that they already felt were good 'in that design that were producing results'" (p. 81). According to these researchers, "the decentralization of decision making appears to be undermining the use of knowledge rather than promoting it" (p. 81). A well-planned, carefully organized collaboration between district-level leaders who have broader perspective on problems and site-based educators who are keenly aware of critical contextual characteristics seems essential to optimize the effectiveness of professional development (Guskey, 1996).

Building leadership capacity is included in many lists and in the NSDC Standards, reflecting current emphasis on teacher leadership and the need for principals to be strong and consistent instructional leaders (Phillips & Bredeson, 2002). Several lists also note that professional development activities should be based on teachers' identified needs, even though noteworthy evidence shows that teachers rarely are able to articulate their needs. Although they have no difficulty identifying problems, dilemmas, concerns, and fantasies, teachers tend to describe symptoms of needs that must be diagnosed more thoroughly and interpreted more broadly (Jones & Hayes, 1988).

Despite the current emphasis in reform initiatives on improving student performance (see Linn, 2000), less than half the lists mention the importance of using careful analyses of student learning data to drive professional development activities. Likewise, fewer than half of the lists cite the importance of including follow up, ensuring that professional development activities are ongoing and job embedded, or focusing on individual and organizational improvement.

Only three lists along with the NSDC Standards stress that professional development should be based on the best available research evidence. This is particularly striking given the longstanding criticism of professional development activities that focus on fads and bandwagon movements rather than solid evidence of what works with students (Coban, 1998; Ravitch, 2006). It may also explain the stipulation in the recently enacted NCLB legislation that only those strategies and methods "proven effective by the standard of scientifically based research should be included in school reform programs" (U.S. Department of Education, 2002, p. 2).

The remaining characteristics appear on three or fewer lists. These include the following: (a) helps to accommodate diversity and promote equity; (b) takes a variety of forms, (c) provides opportunities for theoretical understanding, (d) is driven by an image of effective teaching and learning, (e) provides for different phases of change, and (f) promotes continuous inquiry and
reflection. Inconsistencies exist even in these characteristics. Wraggoski's (1992) study showed, for example, that not all indicators of dealing with diversity and teaching special populations of students contribute substantially to improvements in student learning. Indicators related to teaching students from culturally diverse backgrounds and students with limited English proficiency were strongly related to student achievement measures, but teaching special needs students was not.

Surprisingly, none of the NSDC Standards associated with involving families and other stakeholders. This may be because the NSDC Standards describe a broad and comprehensive list of characteristics that experts believe professional development could or should be, rather than what strong evidence supports.

Summary

This analysis of lists of the characteristics of effective professional development yielded three conclusions. First, these appear to be no agreement among professional development researchers or practitioners regarding the criteria for effectiveness in professional development. Some define it in terms of teachers' self-reports of professional development features that influence increases in their knowledge and changes in their instructional practices. Others look to consensus in the opinions of professional development writers and researchers. Although lists of the characteristics of effective professional development based on such diverse criteria are valuable, they provide only a starting point in efforts to improve the quality of professional development programs and activities.

It could be argued, of course, that these lists were prepared for different purposes and audiences, and that the differences likely affect how the lists were configured: Some of the lists were prepared as policy documents (e.g., the U.S. Department of Education list), whereas others were prepared for audiences of professional development practitioners and school leaders (e.g., the AFT list or the NSDC list). As such, the lists review some portion of the existing literature and packages it for a particular audience. The needs and characteristics of each audience may have as much to do with the content of the lists as the evidence on which each is based. The differences that exist, therefore, are merely a function of how each list was prepared, not the underlying body of evidence.

To suggest that the criteria for determining the effectiveness of professional development vary depending on the intended audience implies, however, that effectiveness, like beauty, is in the eye of the beholder. How can we advance as a profession without at least a nominal level of consensus among policymakers, researchers, practitioners, and school leaders about the criteria for judging effectiveness in professional development? Such consensus brings essential focus to the goals of professional development and provides a foundation for the kinds of systematic inquiry needed to build a professional knowledge base. Without a clear sense of professional development's goals and learning outcomes, evaluation and improvement in the quality of educators' professional development experiences is unlikely.

In order to gain authentic evidence and make serious improvements, visionary leaders must push beyond the starting point set by these myriad efforts and move toward what Kennedy (1992) referred to as professional development's "ultimate benefit"—demonstrable improvements in student learning outcomes. All school leaders, and especially school principals, must insist that professional development planning focus on two critical questions: How will this help our students? What evidence will we use to verify that it does? That evidence might include a variety of indicators of student achievement, such as assessment or test results, portfolio evaluations, marks or grades, or scores from standardized examinations. It might also include affective and behavioral indicators, such as students' attitudes, study habits, homework completion rates, or classroom behaviors. Schoolwide indicators such as attendance rates, dropout statistics, reductions in discipline problems, enrollment in advanced classes, membership in honor societies, and participation in school-related activities might be considered as well. Significant advances in professional development will come only when researchers, practitioners, and visionary school leaders insist on this fundamental goal of improvement in student learning outcomes as the principal criteria of effectiveness.

Second, perhaps because of the lack of agreement on the criteria for effectiveness, many of the currently identified characteristics of effective professional development can be best described as "Yes, but..." statements. For example, yes, enhancing teachers' content and pedagogic knowledge is important, but existing research is limited mainly to investigations of mathematics and science instruction. Yes, professional development should provide sufficient time and other resources, but that time and those resources must be used wisely, focusing on activities that positively affect learning and learners. Yes, professional development activities should promote collegiality and collaboration, but without clear direction and purpose, individuals can collaborate to resist reform and thwart improvement efforts. Yes, professional development should include procedures for evaluation, but evaluations that focus narrowly on educators' self-reported satisfaction with professional development activities offer inadequate guidance and direction to improvement efforts. And so on.

The problem with such "Yes, but..." statements is that they frustrate policymakers and school leaders who want simple answers to their questions about effective professional development. They also tend to diminish the
References


The backward approach

Thomas R. Guskey is professor of educational policy studies and evaluation at the University of Kentucky. You can contact him at the College of Education, Taylor Education Building, University of Kentucky, Lexington, KY 40506, (606) 257-8646, fax (859) 357-4245, e-mail: guskey@pop.uky.edu.

In backward planning, you first consider the student learning outcomes you want to achieve (Level 5). Do you want to improve students' reading comprehension, their skills in problem solving, their sense of confidence in learning situations, their persistence in school? Then you determine what instructional practices and policies will yield those outcomes (Level 4). It is vital to consider relevant research at this stage. Next, you consider what organizational support you need for those practices and policies to be implemented (Level 3). Some aspects of the organization may need to be altered, especially those that block implementation. Then you decide what knowledge and skills the participants providing needs to implement the prescribed practices and policies (Level 2). And finally, you consider how you will give participants the opportunity to acquire knowledge and those skills (Level 1).

What makes this process so critical is that the decisions made at each level profoundly affect those to be made 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 organizational support or change required, and so on. This is why staff development planning that focuses on "events" is so ineffective. There are many vital decisions to be made before we consider the "events."

Complicating matters further is the context-specific nature of this work. Even if we agree on the student learning outcomes we want to achieve, the best practices or policies to attain those outcomes might differ depending on the context. What works best in one context with a particular community of educators and a particular group of students may not work equally well with different educators and different students. That's what makes developing examples of truly generalizable "best practices" so difficult. Still, collecting meaningful "formative" evaluation evidence along the way can help steer us in the right direction.

If we begin our planning using what we want to achieve in terms of learning and learning work backward from there, not only will planning be a lot more efficient, but the results will be much more effective for the staff.

REFERENCES
What do we really know about the relationship between professional development and improvements in student learning? What evidence validates that relationship, and how trustworthy is that evidence? What does that evidence tell us about the characteristics of truly effective professional development activities?

These questions guided one of the largest and most inclusive syntheses of research on effective professional development conducted to date. Scholars from the American Institutes for Research analyzed findings from over 1,300 studies that potentially address the effect of professional development on student learning outcomes. The project was sponsored by the Regional Education Laboratory-Southwest (REL SW) and funded by the Institute of Education Sciences of the U.S. Department of Education.

The findings from this comprehensive analysis, titled *Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement* (Yoon et al. 2007), shed new light on the complex relationship between professional development and improvements in student learning. It is hoped that they also will lead to new and better research on this vitally important dimension of the educational improvement
process. (Details of the research are available in the online version of this article.)

One of the most discouraging findings in the project was the discovery that only nine of the original list of 1,343 studies met the standards of credible evidence set by the What Works Clearinghouse, the arm of the U.S. Department of Education that is charged with providing educators, policy makers, researchers, and the public with scientific evidence about “what works” in education. All nine studies focused on elementary schools and were conducted between 1986 and 2003. No studies of professional development at the middle school or high school levels met the WWC standards, nor did any of the studies published between 2004 and 2006.

**Workshops are not the poster child of ineffective practice that they are often made out to be.**

Four of the investigations included measures of student learning in reading and language arts. Two studies focused on mathematics, one on science, and two on language arts, mathematics, and science. Among the achievement measures considered, seven studies used standardized assessments of achievement, one involved researcher-developed measures of students’ knowledge of fractions, and one used Piagetian conservation tasks. The number of teachers involved in these studies ranged from five to 44; the number of students from 98 to 779.

**What Was Learned**

Researchers reviewed these nine well-designed investigations to determine whether the professional development efforts on which they focused shared common elements or characteristics. They noted that information about the professional development described in the studies was far from perfect and varied in its quality and effect. In addition, given an initial pool of more than 1,300 citations, the nine studies represent a relatively modest research base. Nevertheless, several common elements emerged. These shared characteristics were not what many would have guessed, and several differ from those factors frequently noted as contributing to the effectiveness of professional development endeavors (Guskey 2003).

**Workshops.** Of all professional development activities, none has been more disparaged in recent years than workshops, particularly those of short duration. Criticized as the epitome of ineffective practice, many education leaders regard workshops as a waste of both time and money. And, indeed, a lot of workshops are wasteful, especially the one-shot variety that offers no genuine follow-up or sustained support. But ironically, all of the studies that showed a positive relationship between professional development and improvements in student learning involved workshops or summer institutes. These workshops focused on the implementation of research-based instructional practices, involved active-learning experiences for participants, and provided teachers with opportunities to adapt the practices to their unique classroom situations. So while undoubtedly many workshops are poorly organized and focus on unproven ideas and strategies, as a form of professional development, they are not the poster child of ineffective practice that they are often made out to be.

**Outside Experts.** Many writers in education today stress that professional development should be strictly site-based and should build on the combined expertise of in-house staff members. They believe that the most effective way to bring improvement is to have educators in each school meet regularly to explore common problems and seek solutions based on shared experiences and collective wisdom. But while this may be an appropriate starting point, it is seldom, if ever, sufficient (Holloway 2000; Latham 1998). A review by Thomas Corcoran and his colleagues found, for example, that when decisions about professional development were primarily school-based, “school staff members paid lip service to the use of research” and “were more interested in designs that drew on research about practices that they already felt were ‘good’ than in designs that were producing results.” According to these researchers, “the decentralization of decision making appear[s] to be undermining the use of knowledge rather than promoting it” (2001, p. 81).

In the current analysis, the professional development efforts that brought improvements in student learning focused principally on ideas gained through the involvement of outside experts. These individuals were either program authors or researchers who presented ideas directly to teachers and then helped facilitate implementation. None of the successful efforts used a train-the-trainer approach, peer coaching, collaborative problem solving, or other forms of school-based professional learning. This does not imply that these practices are ineffective. Rather, it simply points out that at the present time, we have no strong, valid, and scientifically defensible evidence
demonstrating that they are effective.

Time. Professional development advocates have long lamented the lack of sufficient time for staff members to engage in high-quality professional learning. Obviously, educators need time to deepen their understanding, analyze students’ work, and develop new approaches to instruction. But simply providing more time for professional development yields no benefit if that time is not used wisely. An analysis by Mary Kennedy (1998) showed, in fact, that differences in the time spent in professional development activities were unrelated to improvements in student outcomes. Why? Presumably because doing ineffective things longer does not make them any better.

In this analysis, time was found to be a crucial factor to success. While the number of contact hours ranged widely, from five to over 100 hours depending on the study, those initiatives that showed positive effects included 30 or more contact hours. It thus seems clear that effective professional development requires considerable time, and that time must be well organized, carefully structured, purposefully directed, and focused on content or pedagogy or both (Birman et al. 2000; Garet et al. 2001; Guskey 1999).

Follow-up. For decades professional development experts have stressed the importance of follow-up activities. Educators at all levels need just-in-time, job-embedded assistance as they struggle to adapt new curricula and new instructional practices to their unique classroom contexts. This analysis confirmed the vital importance of follow-up. Virtually all of the studies that showed positive improvements in student learning included significant amounts of structured and sustained follow-up after the main professional development activities.

Activities. Discussions about “best practices” have dominated professional development circles in recent years. Debates frequently arise from these discussions about what particular professional development activities or designs are most effective and work best (Easton 2004). Yet this analysis of well-designed studies identified no set of common activities or designs linked to effect on student learning outcomes. In each case, the structural features of the professional development activity were determined by the specific content involved, the nature of the work, and the context in which that work took place. This corroborates the position taken by the National Staff Development Council (2001), which argues that the most effective professional development comes not from the implementation of a particular set of “best practices,” but from the careful adaptation of varied practices to specific content, process, and context elements.

Content. Equally debated in recent years is what professional development content is most likely to lead to improvements in student learning. The analysis noted considerable consistency regarding this aspect. The nine studies focused on specific subject-related content or pedagogic practices. In addition, most also emphasized teacher discretion in implementing that content and those pedagogic practices, justified by how students learn. In other words, the professional development efforts in every one of these investigations centered directly on enhancing teachers’ content knowledge and their pedagogic content knowledge (Shulman 1986). The activities were designed to help teachers better understand both what they teach and how students acquire specific content knowledge and skill.

Interpreting the Findings

Many professional developers are likely to be surprised by these results, and some may be disappointed. Many will be stunned, just as we were, to learn that only nine investigations from a pool of over 1,300 potentially useful citations met the WWC standards for inclusion in the analysis. Obviously, these findings paint a dismal picture of our knowledge about the relationship between professional development and improvements in student learning. Such a paucity of rigorous studies of the impact of professional development on student learning outcomes was corroborated by the recent National Mathematics Advisory Panel’s report (2008), which concluded that most studies of professional development in mathematics were descriptive in nature and lacking in the methodological rigor needed to warrant sound causal inferences (e.g., “one-group pretest/posttest designs” without a comparison group).

Nevertheless, these results should not be taken as an indictment of professional development advocates or their work. In the history of education, no improvement effort has ever succeeded in the absence of thoughtfully planned and well-implemented pro-
fessional development. This analysis shows simply that sound, trustworthy, and scientifically valid evidence on the specific aspects of professional development that contribute to such improvement is in dreadfully short supply and that dedicated efforts to enhance that body of evidence are sorely needed. Furthermore, this research synthesis confirms the difficulty of linking professional development to specific student achievement gains despite the intuitive and logical connection. It is hoped that a better understanding of what the current evidence reveals will help guide those efforts.

Educators at all levels need just-in-time, job-embedded assistance as they struggle to adapt new curricula and new instructional practices to their unique classroom contexts.

We also want to emphasize that the results from this analysis should not be taken to mean that alternative professional development activities and designs — such as coaching, the use of collective internal expertise, different allocations of time, or other types of professional development content — do not work. Rather, the results illustrate that at this time, we simply have no reliable, valid, and scientifically defensible data to show that these strategies do work. The best that can be said is that their true value has yet to be determined.

Some might argue that the “What Works Clearinghouse Evidence Standards” used to select the studies included in this analysis are unduly rigorous and that their use eliminated many good studies that other adequate but less restrictive criteria would not. Including these other studies might substantially change the complexion of the analysis and yield quite different results. Mary Kennedy’s review (1998), for example, included a different set of investigations, mostly due to different selection criteria. Using less stringent criteria could have yielded a broader range of effective professional development models, activities, and designs.

In defense of these criteria, however, we would counter that when educators ask what professional development approaches are most likely to lead to improvements in student learning, answers should be based on the most valid and scientifically defensible evidence available. The results from carefully designed, experimental or quasi-experimental studies provide such evidence. Furthermore, if the advocates of alternative professional development models, practices, and designs want their approaches to gain professional credibility and acceptance, then they should take responsibility for demonstrating effectiveness through rigorous and scientifically valid means. In other words, rather than simply appealing to practitioners’ intuition and making claims of common sense, take the time to conduct thorough and systematic investigations of the true effects. Doing so will not only establish credibility, it will go far in enhancing the professionalism of our field.

Implications

The implications of this analysis for professional developers are fourfold. First, at all levels of education, those responsible for planning and implementing professional development must learn how to critically assess and evaluate the effectiveness of what they do. This means that discussions about the specific goals of professional development, what evidence best reflects the achievement of those goals, and how that evidence can be gathered in meaningful and scientifically defensible ways must become the starting point for all planning activities (Guskey 2000; Guskey 2001). Only when gathering data on the effectiveness of professional development becomes a central focus in the planning process will the pool of valid and trustworthy evidence expand.

Second, practitioners at all levels must demand better evidence from consultants and purveyors of new strategies and practices. Stories about what happened at one time in a single school or district may be interesting, but they do not justify broader implementation. What we need is trustworthy, verifiable, replicable, and comparative data. In addition, those promoting particular ideas or techniques often preface their comments with the phrase, “Research says . . .” in order to enhance presumed credibility. School-based educators must be prepared to dispute such claims, asking such questions as: “What research?” “When was it conducted?” “Was it done in contexts similar to ours?” “Are the results applicable to our setting?” and “How trustworthy are those results?” Consultants have the responsibility to know that research in sufficient depth to answer these questions. And if they do not, then at least they should have the courage and integrity to say, “I don’t know.”

Third, implementation of any new professional development strategy should always begin with small-scale, carefully controlled, pilot studies designed to...
test its effectiveness. Before embracing any new strategy or committing large amounts of time, money, and other resources to any new approach, that new strategy should be carefully examined in that context to determine if the promised effects in terms of student learning gains can be realized. Comparing the progress of one group of educators engaged in the new approach with that of another, matched group of educators in similar teaching situations can yield important evidence on the likelihood of success. Positive results will enhance the credibility of the new approach and will provide a foundation on which larger scale implementation and evaluation can build. In the absence of positive results, either needed adaptations can be considered or resources can be redirected to other, more promising approaches.

Finally, researchers as well as practitioners must pursue greater rigor in the study of professional development. If public schools are spending about $20 billion annually on professional development activities (NCES 2008), then it merits serious study. The research community must dramatically improve the precision of studies of the relationship between professional development, changes in teaching practices, and improvements in student learning. Practitioners likewise should insist on better evidence when making decisions about how to spend their limited professional development resources.

Rigor, however, does not imply that only one method of inquiry is required to 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 such studies are replicated with similar findings, that validity is further enhanced. Comparing the progress of one group to a similar group that has been “matched” on relevant measures, for example, can be especially useful if data are available on pertinent background characteristics of the participating teachers and their students. Randomly selecting half of those who volunteer to take part in a new approach and then comparing their results with those from the other half who were not included but will be next year also can offer valuable information. In addition, other investigative methods may be used to formulate important research questions and develop new measures relating to professional growth (Raudenbush 2005).

The amount of valid and scientifically defensible evidence we currently have on the relationship between professional development and improvements in student learning is exceptionally modest. Nine studies from an initial group of 1,343 potentially relevant citations represent a very small percentage. Given this limited number of studies, we also have to be cautious about making a definitive conclusion about the effectiveness of specific elements of professional development. This conservative stance is echoed by the National Mathematics Advisory Panel. It concluded that “Although the Panel did find some positive effects of PD on students’ achievement gains, research does not yield sufficient evidence on the features of any particular approach to permit detailed conclusions about the forms of or approaches to effective PD” (2008, p. 40).

Still, we are now in a better position than ever before to organize and conduct professional development so that valid evidence can be gathered, both to determine the effectiveness of
Those responsible for planning and implementing professional development must learn how to critically assess and evaluate the effectiveness of what they do.

could not be answered in this analysis. Efforts are also being made to improve the rigor of studies specifically designed to examine this important relationship (Wayne et al. 2008). Moving in this direction will improve the likelihood of success and also elevate professional development to an inquiry-based profession, rather than a haphazard set of activities based on intuition, hearsay, tradition, and folklore.

REFERENCES


**DESIGN OF THE SYNTHESIS**

This broad research synthesis involved a series of carefully planned steps. It began with keyword searches of seven electronic databases: ERIC, PsycINFO, ProQuest, EBSCO’s Professional Development Collection, Dissertation Abstracts, Sociological Collection, and Campbell Collaboration. A deliberately wide net of keywords was used to capture literature on professional development and student learning in three core content areas: language arts, mathematics, and science. The search identified 1,343 citations as potentially addressing the impact of professional development on student learning outcomes.

Next, prescreening was performed by scanning the abstracts or full texts of the 1,343 studies to determine if they met broad relevance and methodology criteria (e.g., an empirical study involving professional development and some measure of student achievement). The prescreening process reduced the list to 132 studies that were considered relevant for systematic review. These studies were then subjected to three stages of coding.

Stage 1 coding examined the relevance of the studies using the following criteria:

- **Topic.** The study had to deal with the effects of professional development on student learning in at least one of three core content areas (language arts, mathematics, and science).
- **Population.** The sample had to include teachers of language arts, mathematics, or science and their students in grades K-12.
- **Outcome.** The study had to measure student learning outcomes.
- **Study design.** The study had to be empirically based and use randomized controlled trials or some form of quasi-experimental design.
- **Time.** The study had to be published between 1986 and 2006.
- **Country.** The study had to take place in Australia, Canada, the United Kingdom, or the United States, due to concerns about the external validity of the findings.

The results of this stage of coding yielded 27 relevant studies that were eligible for review in terms of study quality ratings.

Stage 2 coding focused on quality ratings of the 27 eligible studies using the U.S. Department of Education’s What Works Clearinghouse (WWC) Evidence Standards (see http://ies.ed.gov/ncee/wwc/overview/). At this stage, each study was given one of three possible ratings in accordance with the WWC technical guidelines:

- “Meets Evidence Standards” (e.g., randomized controlled trials that provided the strongest evidence of causal validity).
- “Meets Evidence Standards with Reservations” (e.g., quasi-experimental studies or controlled trials that had problems with randomization, attrition, teacher-intervention confound, or disruption).
- “Does Not Meet Evidence Standards” (e.g., studies that did not provide strong evidence of causal validity).

Only nine of the 27 studies were rated at the first or second level as having met the WWC Evidence Standards. The other 18 studies were rated at the third level: “Does Not Meet Evidence Standards.”

**Descriptive Results**

The next step in the analysis was to review the selected studies for shared descriptive characteristics. Among the nine studies that met the What Works Clearinghouse Evidence Standards for causal validity, six were published in peer-reviewed journals, while three were unpublished doctoral dissertations. All of the studies focused on elementary schools and were conducted between 1986 and 2003. No studies of professional development conducted at the middle school or high school levels met the standards, nor did any of the studies published more recently, between 2004 and 2006.

Four of the investigations included measures of student learning in reading and language arts. Two studies focused on mathematics, one on science, and two on language arts, mathematics, and science. Among the achievement measures considered, seven studies used standardized assessments of achievement, one involved researcher-developed measures of students’ knowledge of fractions, and one used Piagetian conservation tasks. The number of teachers involved in these studies ranged from five to 44, the number of students from 98 to 779.

Twenty different effect sizes were computed across the nine studies, ranging from -0.53 to +2.4. Eighteen of these effect sizes were positive, one was zero, and another was negative but not statistically significant. Eight of the 20 effect sizes proved statistically significant, and 12 were not. But among those 12, nine would be considered substantively important accord-
Analytic Results

Following the descriptive analysis, the researchers reviewed these well-designed investigations to determine whether or not the professional development efforts on which they focused shared common elements or characteristics. They noted that information about the professional development activities described in the studies was far from perfect and varied in its quality and effect. In addition, given an initial pool of more than 1,300 citations that were found in electronic literature searches to be linked to the keywords of professional development and improvements in student learning, the nine studies that met the guidelines of causal validity represent a relatively modest research base. Nevertheless, several common elements emerged from the research synthesis. Surprisingly, these shared characteristics were not what many would have guessed, and several differ from the factors frequently noted as contributing to the effectiveness of professional development endeavors.