

# Evaluating Professional Development: A Framework for Adult Education

A Publication of

Building Professional Development Partnerships  
for Adult Educators Project

PRO-NET

**May 1997**

---

Mark Kutner  
Renee Sherman  
John Tibbetts  
Larry Condelli  
Pelavin Research Institute

Sponsored by:  
U.S. Department of Education  
Division of Adult Education and Literacy

Ronald Pugsley, Division Director  
Jim Parker, Project Officer



# TABLE OF CONTENTS

|  | <u>PAGE</u> |
|--|-------------|
| <b>CHAPTER 1: EVALUATING: PROFESSIONAL DEVELOPMENT:<br/>AN OVERVIEW</b>                    |             |
| Introduction   | 1           |
| Purpose and Audience   | 2           |
| Organization of Monograph  | 3           |
| Evaluation and the Professional Development Process  | 4           |
| Issues in Evaluating Professional Development  | 5           |
| Summary  | 7           |
| <br>   |             |
| <b>CHAPTER 2: A FRAMEWORK FOR EVALUATING PROFESSIONAL<br/>DEVELOPMENT</b>                  |             |
| Introduction   | 9           |
| Comprehensive Professional Development: The Essential Triangle<br>For Change               | 10          |
| A Model for Evaluating Professional Development: Instructors<br>Program Services, Learners | 12          |
| Impacts on Instructors   | 13          |
| Evaluating Reactions   | 14          |
| Evaluating Acquisition of Skills and Knowledge   | 14          |
| Evaluating Changes in Instructional Behavior   | 15          |
| Impact on Programs   | 15          |
| Instructional Arrangements   | 16          |
| Program Processes  | 17          |
| Student Assessment   | 17          |
| Learner Supports   | 18          |
| Impact on Learners   | 18          |
| Evaluate Student Reactions   | 19          |
| Evaluate Acquisition of Knowledge and Skills   | 20          |
| Evaluate Changes in Student Behavior (Uses of Learning)                                    | 21          |
| Summary  | 22          |
| <br>   |             |
| <b>CHAPTER 3: STRATEGIES FOR EVALUATING PROFESSIONAL<br/>DEVELOPMENT</b>                   |             |
| Introduction   | 23          |
| Professional Development Evaluation Strategies   | 24          |
| Multi-purpose Evaluation Strategies  | 24          |
| Questionnaires   | 24          |
| Interviews   | 26          |
| Focus Groups   | 28          |
| Strategies for Assessing Instructors   | 29          |

## TABLE OF CONTENTS (Continued)

|                                   | <b><u>PAGE</u></b> |
|-----------------------------------|--------------------|
| Observation of Practice           | 29                 |
| Instructor Portfolios             | 31                 |
| Pracitioner Journals              | 32                 |
| Collaborative Teams               | 33                 |
| Strategies for Assessing Students | 34                 |
| Tests                             | 34                 |
| Alternative Assessment            | 35                 |
| Summary                           | 36                 |

### **CHAPTER 4: DESIGNING AND IMPLEMENTING PROFESSIONAL DEVELOPMENT EVALUATIONS: ISSUES AND RECOMMENDATIONS**

|  |    |
|--|----|
| Introduction   | 37 |
| The Context of Professional Development Evaluations: Issues Related to the Adult Education Service Delivery System | 38 |
| Part-time and Voluntary Nature of Service Delivery   | 38 |
| Part-time Instructors and High Rate of Instructor Turnover   | 39 |
| Lack of State Certification Requirements or Required Instructor Skills and Competencies                            | 39 |
| Components of a Professional Development Evaluation Plan   | 39 |
| Determine Evaluation Objectives, Audiences of Evaluation and Questions   | 40 |
| Audiences for Evaluation Data  | 41 |
| Evaluation Questions   | 42 |
| Identify Evaluation Strategies to be Used  | 42 |
| Consider Implementation Issues of Different Strategies   | 43 |
| Provide both Quantitative and Qualitative Data   | 43 |
| Determine Evaluation Methodology   | 44 |
| Sampling   | 44 |
| Data Reliability and Validity  | 45 |
| Data Confidentiality   | 45 |
| Determine Required Person and Fiscal Resources   | 45 |
| Establish a Time Frame for Collecting and Reporting Data   | 46 |
| Sufficient Time Must be Allowed Before Results Can Be Expected   | 46 |
| Measurements Should be Repeated at Appropriate Times   | 46 |
| Summary  | 47 |



---

---

## CHAPTER 1

### EVALUATING PROFESSIONAL DEVELOPMENT: AN OVERVIEW

#### Introduction

Ideally, professional development for adult educators is a process in which instructors gradually acquire a body of knowledge and skills to improve the quality of teaching for learners and, ultimately, to enhance learner outcomes. In fact, many adult education programs require instructors to participate in a minimum number of professional development activities each year. Some states (e.g., Alabama, Arkansas, California, Connecticut), as a way of improving program quality, have adopted teacher certification standards or competencies that often must be met within a specific time frame, and involve a specific number of hours of instruction.

For all of these requirements, little is known about the impact of professional development activities and certification standards on quality of teaching and student outcomes. Evaluations of professional development continue to emphasize what is characterized as the “happiness quotient” —involving an assessment of instructor reactions to a specific professional development activity immediately upon its completion. With growing concerns about the effectiveness of adult education and literacy services, and increasing competition for resources, evaluations of professional development are needed to assess changes in instructors’ knowledge, skill levels, and behavior, as well as to document changes in program practices and student outcomes.

**W**hat is the relationship among professional development activities, instructional behavior, program processes and learning outcomes in adult education?

Although the research literature consistently identifies evaluation as a critical component in the delivery of professional development to adult education instructors and other staff, evaluation is typically overlooked. This situation must change. Evaluations must be incorporated within all aspects of professional development activities, and must begin with an initial state-level planning process for professional development—if useful data are to be provided for state administrators, professional development providers, and local adult education program administrators, as well as instructors. The field must move to more substantive evaluations that consider professional development’s impact on instructors, program services, and learners themselves.

---

---

Keeping such substantive goals in mind, evaluation plays at least two roles in the professional development process: (1) to promote continuous program improvement; and (2) to ensure program accountability. Furthermore, if adult education practices are to respond to rapidly changing technological and social structures, professional development can be the primary vehicle for meeting that challenge. But sound information is needed to make informed decisions on *how* to change directions. Therefore, state adult education offices that provide a great deal of financial support for professional development activities, professional development agencies, and local adult education programs, must all concentrate on evaluating professional development.

### **Purpose and Audience**

This monograph presents a suggested framework, and specific practical strategies and procedures, for evaluating the impact of professional development activities. An underlying principle of the framework is that while professional development's most immediate impact must be on instructors, changes in programs (e.g., instructional practices, processes, student assessment, and learner supports) must also occur for there to be an effect on learner outcomes. While assessing professional development's impact on instructors can be accomplished in a fairly straightforward manner, efforts to relate professional development, first on programs and then on learner outcomes, are more complicated tasks that have generally not been addressed even in the K-12 arena. Complicating professional development evaluations in the adult education field is the relatively unsophisticated and underfunded program infrastructure supporting the adult education delivery system and the many intervening variables that influence student outcomes (e.g., class size, teacher characteristics, student learning styles, inconsistent student attendance). Yet, the adult education field no longer has the luxury of supporting unexamined professional development and must begin to incorporate evaluation into all components of professional development services.

Most states, professional development agencies, and local adult education programs will find many aspects of the framework—especially professional development's impact on instructors and programs—relatively easy to adopt. Other aspects of the framework, especially relating professional development with learner outcomes is likely to currently be beyond the capacity of most programs. Yet, because the ultimate objective of professional development is to enhance learner outcomes, states, professional development agencies, and local adult education programs must begin addressing the issues of how to assess such impacts.

State agencies must actively promote evaluations and should consider requiring evaluation data from professional development agencies and local adult education programs receiving funds, as well as providing additional resources, including money and technical assistance. Professional development

agencies should incorporate evaluation within all services provided and allocate resources necessary for successful evaluation efforts. Local programs should be prepared to participate in data collection activities as part of evaluation efforts, including providing instructors with the necessary supports to participate in evaluations.

Specifically, the monograph is designed to promote professional development evaluations by assisting individuals interested in sponsoring, conducting, and participating in evaluations, including:

- **State staff** responsible for distributing funds to support professional development activities and for ensuring the overall quality of adult education services;
- **Professional development agency staff** responsible for designing and providing professional development services and documenting their success and impact to state adult education and other agencies that provide their funding;
- **Local Administrators** that need data to determine whether their program is investing in appropriate professional development services and is having an impact on instructors, program services, and student learning. These data feed program change and are useful in reporting and marketing to stakeholders and policy makers; and
- **Instructors** who need to know if the strategies they are using make a difference in student learning.

### **Organization of Monograph**

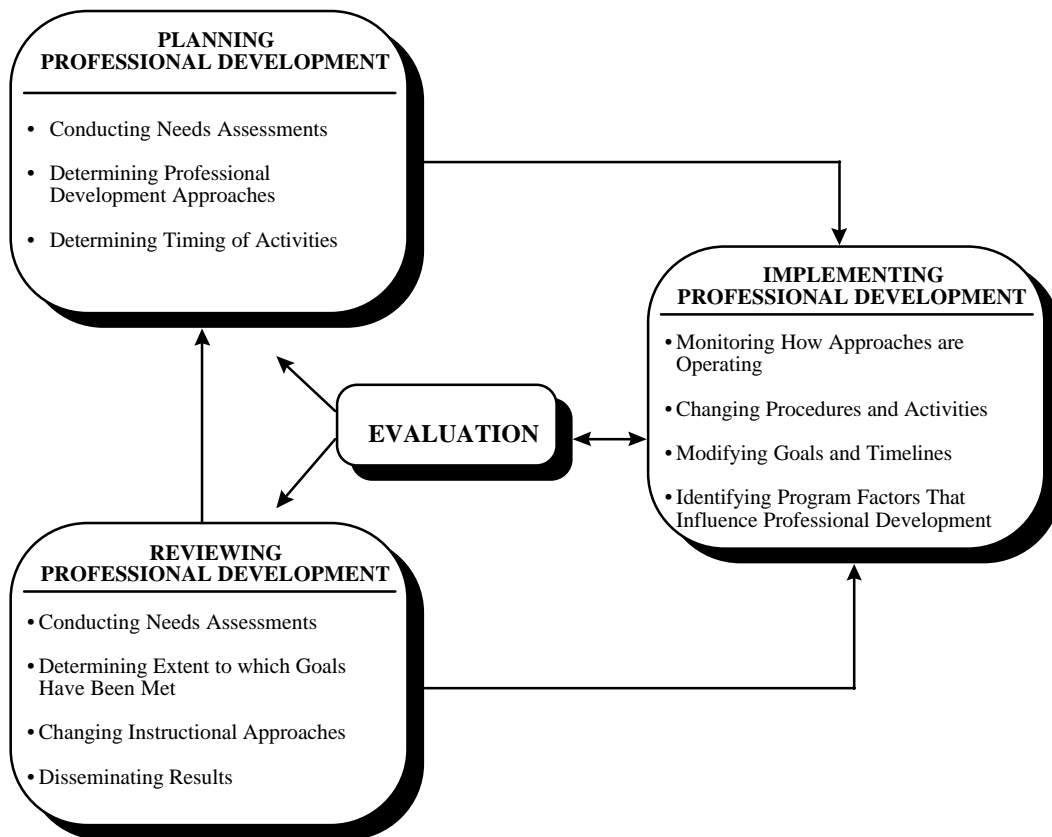
The remainder of this chapter discusses evaluation and the professional development process, including issues associated with evaluating the professional development process. Chapter 2 presents a framework for evaluating professional development that considers the three dimensions of the professional development change process: instructors, program services, and learners. In Chapter 3, specific strategies for evaluating professional development activities are discussed, as well as issues related to their implementation regarding instructors, program services, and student outcomes. Chapter 4 discusses issues related to the adult education delivery system that must be addressed when designing evaluations, and presents the necessary components of a professional development evaluation plan, to ensure all needed data are obtained.

## **Evaluation and the Professional Development Process**

Evaluation relates to all aspects of the professional development process and can inform continuous program improvement and accountability efforts by staff from state and local adult education programs and professional development agencies. As shown in Exhibit 1, professional development evaluations are integrated into planning, implementing, and reviewing professional development activities: Under this model *evaluation is a continuous activity rather than a single event that occurs at the end of professional development activities*. Hence, evaluation data must be collected on an ongoing basis and procedures to ensure this happens must be incorporated within the professional development process. When planning, providing, and reviewing professional development activities, strategies for collecting and analyzing evaluation data always must be considered and implemented by professional development agencies and adult education programs.

### **EXHIBIT 1**

#### **Evaluation as Part of an Ongoing Professional Development Process**



Evaluation data, by providing information on goal attainment and obstacles to anticipated outcomes, are used to refine and adjust professional development planning and implementation. In addition, data from professional development evaluation feeds back into the needs identification process and becomes part of the planning process for professional development. Evaluation data should be used by many different stakeholders, including state staff responsible for distributing funds that support professional development activities and agencies, state and local agencies that provide professional development, and local adult education programs whose instructors and staff participate in professional development activities. Also, legislators and policy makers from the national and state levels, as well as staff from other human service delivery areas who could benefit from professional development services, should be interested in data from professional development evaluations.

Given the nature of the professional development process, there are a number of issues that must be considered and addressed when developing and implementing the type of ongoing evaluation efforts necessary to assess the impact of adult education professional development.

### **Issues in Evaluating Professional Development**

Professional development is a *change process* whose effects on instructors, program services, and learners are not easy to document and must be measured over time. However, obtaining information about the effectiveness of professional development activities is complicated by the differences in evaluating: (1) single professional development activities and multiple or sequential activities occurring over a period of time; and (2) short-term and long-term changes. Therefore, when evaluating professional development activities, it is important to consider the following:

- ***Sufficient time is required to pass before professional development activities can be expected to show success.*** It is normally not useful to evaluate the effect of a single professional development activity on individual behaviors. For example, “one-shot” workshops are not likely to change behaviors, and are designed primarily to provide information and to raise awareness about a specific issue or topic. Evaluations focusing on a series of activities that have a more cumulative effect and are woven into a teacher’s long-term professional growth are more likely to provide useful data. However, evaluative data of single professional development activities often are requested by administrators who fund professional development and who are concerned about accountability.
- ***Impacts on instructors, program services, and learners must be measured.*** The ultimate reason for offering professional development services is to improve learner outcomes. First, however, professional development must be positively perceived by instructors who participate in such services; and it must enhance their knowledge and skills as well as have an impact upon their instructional behavior. If changes in instructor behavior are to influence learners, it is also necessary to alter program services, especially instructional practices. Evaluations must, therefore, assess the impact of professional development not only on instructors, but also on program services and ultimately on learners.

- ***Data must be collected concerning the context in which instruction takes place, the extent administrative and other support structures reinforce practices promoted through professional development, and the nature of the professional development activities, themselves.*** It is, after all, unrealistic to expect that professional development can have an impact on instruction and learners unless such activities reflect effective professional development practices, continue over time to reinforce changes in instructional behavior, are supported by program administrators, and are used in developing and delivering instruction. Data related to program context that should be collected include information about the receptivity of local program staff and administrators to professional development, the extent to which instructors have the opportunity and support to implement what was learned during professional development services, the number of instructors, the amount of time instructors engage in professional development, and the availability of multiple professional development approaches.
- ***Professional development activities can only result in improved instruction and better instructional practices if adult education programs encourage and support instructors, allow instructors access to special knowledge, provide instructors the time to focus on the requirements of a new task, and provide time to experiment and to observe others.*** In other words, there needs to be time to practice new instructional strategies, as information alone is not sufficient to produce change. Furthermore, organizations need to adapt to new structures and norms, if change is to come about (Cafarella, 1994). Some professional development activities (e.g., a curriculum development project) may require specific types of administrative support. To support the adaptation of a new curriculum, for example, administrators may need to help sustain study groups for teachers using the curriculum to obtain feedback and to discuss issues of implementation and ways to ease into the new curriculum and thus change teaching practices. Administrative support, in the form of study groups in this case, integrates professional development into the culture of the organization and helps foster change. Other support structures such as professional networks are important, as well, to fostering change. These structures may be formal networks (e.g., membership in a professional organization or involvement in a teacher network focusing on specific subject matter), or informal groups (e.g., peer groups, study groups, collaboration of teachers, administrators, non-instructional staff). Such structures support dialogue and enhance knowledge around pedagogy and subject-matter, and, in a collegial atmosphere, help instructors seek solutions to problems related to their practices. In short, such structures foster a commitment to improvement.

### **Summary**

Professional development is a change process whose effects on instructors, programs, and learners are usually not documented. Although evaluations have typically been an overlooked component of the professional development process, the adult education field no longer has the luxury of supporting unexamined professional development activities. To provide stakeholders with the data necessary to inform the continuous improvement of professional development services, and to ensure program accountability, evaluations must be incorporated within all aspects of the professional development process, including

planning, implementing, and reviewing activities. In planning and conducting professional development evaluations, it is, therefore, important to allow sufficient time to pass before expecting professional development to show success, and to include procedures for assessing impacts on instructors, program services, and learners, as well as to collect data regarding the context within which instruction occurs—including the extent of administrative support for professional development, and the very nature of professional development activities.

## **CHAPTER 2**

### **A FRAMEWORK FOR EVALUATING PROFESSIONAL**

#### **Introduction**

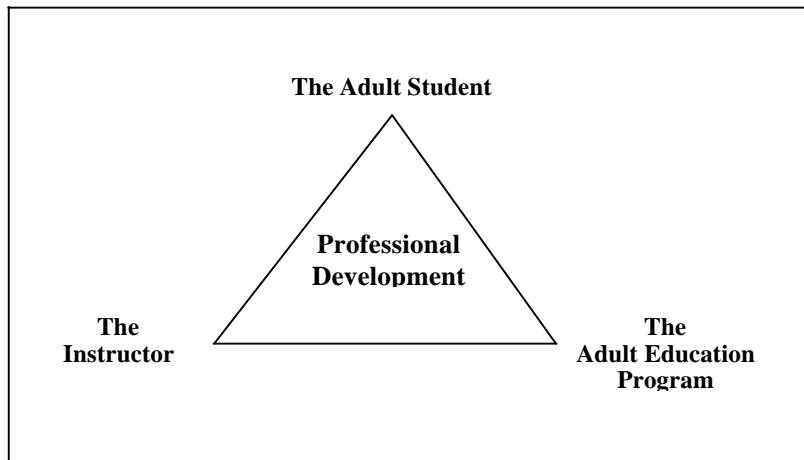
Successful change in education usually does not take place without some form of professional development. The sad history of educational reform is replete with examples of program changes that have been made without adequate professional development (New Math, Systems 90, The Open Classroom, to name a few). All failed, at least in part, because instructors were not adequately involved in the planning and implementation of such programs. Professional development conducted in isolation from the student and the system in which it occurs is likewise doomed.

Almost any adult educator would acknowledge that the "bottom line" of professional development is improving student learning. But often there has been a huge *assumption* that *whatever* is done in the name of professional development does, in fact, improve student learning. This assumption is usually not tested and is certainly not always true. A common perception of past practices of pre-service and in-service education and professional development is, in fact, that these efforts have most often targeted instructors as an end in themselves, without considering the effects of professional development on instructional programs and learners. Such a narrow focus has failed to produce the most sought after result—namely, improved learning.

The reason for this neglect is clear. Professional development evaluations of program services and learners are difficult and time consuming. Although determining whether or not *instruction* has changed as a result of professional development activity is not a daunting task, the many possible intervening variables between professional development, instruction, and any resulting *student achievement* make identifying the relationship among them considerably more difficult to identify. Yet, to succeed in affecting the necessary changes that will have an impact upon student outcomes, the two major targets of professional development must be instructors and program services.

**Comprehensive Professional Development: The Essential Triangle for Change**

Because student growth and achievement is the reason for adult education's existence, and, therefore, for all professional development activities, such activities should be always based upon the systematically determined needs of adult students. Likewise, successful professional development should change instructional practices to produce desired student growth. For that to happen, the organization (the local program and often the state system) must be willing to make whatever changes necessary to support and facilitate instructional practice. Exhibit 2 illustrates this essential triangle for change – the dynamic process between adult students, instructors, and the organizations that guide successful professional development to improve student learning.

**EXHIBIT 2****The Essential Triangle for Change**

As simple as this process appears to be, it is fraught with opportunities for failure at one or more junctures. Even a cursory review of professional development activities will reveal, for example, often traumatic events that were brought about by the failure to give *simultaneous* attention to all three targets of the essential triangle. To illustrate this point, the following anecdote illustrates what happens when a program's administrative structure does not support an instructional practice presented during professional development activities:

*Teachers in one state were provided professional development in the use of cooperative learning for adult basic education. One teacher was so successful and the students so enthusiastic that she decided to do a "jigsaw" cooperative learning activity for her annual*

*evaluation by the program director. As a result, she received the lowest rating she had ever received, because the director observed that students kept moving about the room and were animatedly talking during most of the session, causing the director to think the instructor was "not in control" of her class. The instructor also was criticized for not making a detailed "presentation" of material, which, the director noted, she had been hired to do. The instructor was in tears, as she felt she had been "set-up" for failure. She called the professional development specialist who made an appointment with the director to explain how cooperative learning can result in better understanding and retention, precisely because students construct or build their own concepts and then reinforce them with one another. The teacher's evaluation report was finally adjusted, with considerable embarrassment and resentment by all parties involved.*

As shown by this example, a key factor that often impedes the effectiveness of professional development activities is the organizational setting within which instructional practices supported by professional development activities are implemented. As Elmore (1996) cautions: "Schools may change all the time -- new structures, schedules, textbooks or teaching systems—yet never change in any fundamental way what teachers and students actually do when they are together in the classroom." Professional development systems are deeply institutionalized in patterns of organization, management and resource allocation within schools and school districts, and are structured by Federal, state, and district policies. Many interests are at stake in any proposals for the reform of professional development, which has a powerful resistance to change and is well-adapted to the ecology of "schooling" as we have come to know it (Sykes, 1996).

Despite these difficulties, a growing body of research demonstrates that reform-oriented teacher learning can, indeed, affect student learning. A mere 15 years ago, research identified practices such as cooperative learning as linked to positive increases in student learning and achievement (Sparks, 1983). Joyce and Showers (1995, p. 9) affirm that ". . . research on teaching and learning has resulted in a considerable array of curricular and instructional alternatives that have great promise for increasing student learning." Likewise, Hirsh and Ponder (1991) report that ". . . the body of research on effective teaching practices and a growing body of literature on adult learning now provide bridges to a more results-oriented view of schooling and staff development." Increases in student achievement as a result of staff development programs are documented by research studies in Richmond County, Georgia, and West Orange, New Jersey (Joyce, et. al 1989; Sousa, 1990) in order to construct just such a bridge.

It is difficult to conceive a solution to the problem of student learning without a solution to the problem of engaging teaching (Elmore, 1996). To link professional development activities and student learning there must be potent content and a good design, not only of professional development, but of the organization in which that content is to be implemented (Joyce and Showers, 1995). The *evaluation*

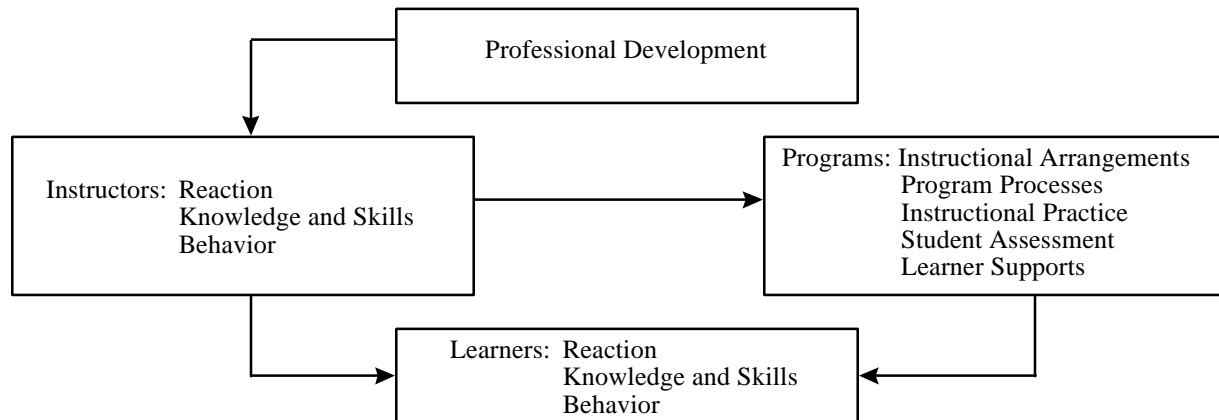
*process of linking professional development with student achievement* now becomes clearer: evaluation must be related to the professional development provided, and must result in changes in program services in order to impact student outcomes.

### **A Model for Evaluating Professional Development: Instructors, Program Services, Learners**

The essential triangle for change forms the basis of the following model for understanding and evaluating the impact of professional development activities. Evaluations of professional development, if they are to present an accurate assessment of such a change process, must document changes in instructors, program services, and, ultimately, impact on student outcomes. Exhibit 3 presents the framework for evaluating professional development that incorporates these three dimensions.

#### **EXHIBIT 3**

#### **Components of a Professional Development Evaluation Framework**



In the model, professional development is shown to have its most immediate impact on instructors, by their reactions to the training, the skills and knowledge they obtain and the subsequent effect this knowledge has on their behavior. Instructors in turn can change their program in different ways depending on the content of the professional development. The program can be changed by the new instructional practices learned by teachers or by a new approach to assessment or other program processes. Supports for learners can also be changed by the new knowledge. Both instructors, by their changed instructional behavior, and programs through new processes, supports and assessment, can affect learners reaction,

knowledge, skills and behavior. The evaluation model embodies two critical assumptions about professional development:

- The hallmark of successful professional development is change – in instructors, programs and learners; and
- Learners can be changed only through professional development, *after* teachers and programs change.

Below, we further demonstrate the model by describing the topics and research questions that can be addressed to understand the impact of professional development activities.

### **Impacts on Instructors**

Professional development has its most immediate and direct impact on instructors, since they are the direct target of the training. Evaluation can address three areas of such impact: (1) instructors' reactions to the professional development experiences; (2) their acquisition of knowledge and skills gained from the experience; and (3) changes in instructional behavior resulting from the experience. These aspects build on one another, and reactions are a necessary condition for the acquisition of skills and knowledge; and it is through the acquisition of skills and knowledge that change in instructional behavior can be expected (Kirkpatrick, 1994). Evaluation of instructors can provide answers to planners, administrators, and to instructors, themselves, to the following questions that help improve the professional development process:

- *What types of professional development activities are being provided?*
- *How are these activities meeting the needs of instructors?*
- *How many hours do instructors participate in different types of professional development activities (i.e., workshops/presentations, inquiry/research projects, observation/feedback, program/product development)?*
- *To what extent are instructors satisfied with the professional development opportunities they have experienced?*
- *What changes are there in the knowledge and skill level of instructors, as a result of professional development opportunities?*
- *Are instructors consistently using newly learned strategies in the ways they were intended to be used?*
- *What supports (i.e., administrative, informal and formal teacher networks) are in place that help foster change in instructional behavior?*

### **Evaluating Reactions**

The most basic level of evaluation is the instructor's own reaction to his or her professional development experience. Evaluators want to know if the experience was positive, as instructors are more likely to learn from positive experiences, and are more likely to want to incorporate the new learning into their teaching practices. Positive experiences also work as incentives for continuing to engage in professional development activities, including sharing information with others. While this type of evaluation, which generally occurs at the close of workshops or training institutes, is often, itself, considered to be assessing the "happiness quotient," it does serve an important purpose. Such an evaluation provides information for planning and for improving future professional development experiences (e.g., format, materials, presentations, activities). In addition, the data help in developing a standard by which to assess the quality of professional development activities.

### **Evaluating Acquisition of Skills and Knowledge**

At a more sophisticated level, evaluations can be used to gauge the knowledge and skills gained by the instructor from the professional development experience. It is important to measure changes in knowledge, as it is through increased knowledge that changes in behavior are expected to occur. "Knowledge," in this context, refers to such areas as *awareness of educational theories and practices, curricula, and academic content*. Skills may refer to such behaviors as (1) a demonstration of questioning strategies that encourage students to move from lower-level to higher-level thinking or (2) the ability to monitor student performance.

Measuring the acquisition of knowledge and skills is significant for two reasons. First, the data from the evaluation provide insight into how effective the professional development experience is in establishing a new knowledge base or in developing a higher skill level. Second, it allows the instructor, and the individuals responsible for staff development in the program, to assess how well the goals set out in the professional development plan have been met. For example, if the goal is to understand the theory behind communicative ESL and to be able to use this approach appropriately in the classroom, the assessment will provide information on whether more professional development in this area is required.

### **Evaluating Changes in Instructional Behavior**

Assessing reactions to professional development experiences or measuring whether knowledge and skills are acquired are only the first levels of the evaluation process. A more important question to ask is *How has instructional behavior changed as a result of the professional development experience?* In other words, what levels of use and what degrees of transfer were achieved as a result of the experience? This *transfer* of learning represents the consistent and appropriate use of new skills and strategies for instruction

(Joyce and Showers, 1995). Transfer of learning into actual practice is dependent upon several factors, including:

- The level and complexity of new knowledge and skills to be acquired;
- Teachers' perceptions of how new learning fits into existing instructional practices;
- The support structures within the program that allow teachers to solve implementation problems with other staff, including peers and administrators; and
- Opportunities for the essential practice to develop new skills.

Increases in the transfer of training to a classroom setting is more likely to occur where in-class coaching occurs, along with providing theory, demonstration, practice, and feedback (Joyce and Showers, 1995).

While data reported in the K-12 literature demonstrate that professional development is effective in bringing about significant instructional change for the teachers who participated, less is known about these effects in adult education. Key issues in assessing change, therefore, revolve around identifying how much of the new learning finds its way into instructor's practice, and whether the learning persists. What is already surmised is that more complex learning theories, knowledge and skills, or learning that is new to the instructor, will be less likely to find its way into practice. Some learning may translate into practice, but changes may be short lived if instructors do not perceive the changes as having positive effects on students, or they do not perceive support from administrators. Short-term changes, may occur right after the learning experience, as instructors "experiment" with newly acquired skills. Integrating new knowledge and skills into daily practice, therefore, may be difficult unless support structures, teacher incentives, and opportunities for practice (e.g., mentoring, coaching), as well as feedback are in place.

### **Impact on Programs**

To have a meaningful impact upon learners, effective professional development activities must not only influence instructor knowledge and behavior, but must bring about actual changes in programs. Which is to say, one instructor learning and implementing better techniques for cooperative learning will not sufficiently change the program to influence new learners. New instructional behaviors require institutional support and broad implementation of changed instructional arrangements to make a difference; and program administrators must be supportive of instructors' efforts and be willing to make changes to the program to accommodate innovations.

The specific program change to evaluate depends upon the topic of the professional development activity. While much professional development focuses on changing instruction, other program processes,

student assessment methods, and student support services also may be targeted for change, as illustrated by the questions below:

- *How have instructional arrangements changed throughout the program?*
- *What new program processes, such as recruitment, intake or student orientation changed?*
- *Has professional development improved student assessment methods – both for placement and progress?*
- *What types of learner supports are being provided as a result of professional development?*
- *What changes in program processes, instruction, assessment, and learner supports have occurred because of the professional development activities in which instructors participated?*

### **Instructional Arrangements**

Teachers often return from a professional development activity excited about a new instructional method they have learned. Hence, they may try to implement this method in their classrooms, and hope to see both student interaction and learning improve in their class. However, a few teachers changing what they do is insufficient to produce lasting change in students. A lasting impact requires that program administration and other teachers support the new instructional practices and establish them within the larger program. Without institutional support, teachers will not implement the new practices consistently over time and the use of the practice will end when the teachers leave.

Successful evaluation of the impact of professional development demands a careful assessment of how instructional arrangements within the program *overall* have changed. Such evaluation may require a review of materials and activities used by teachers, attitudes of teachers and program administrators, and the use of instructional practices by all teachers in the program. If change, in fact, has occurred, there should be a relationship between the professional development activities, the materials used, and the curriculum. If the program adopts a whole language approach to reading instruction, for instance, classroom activities and materials should reflect this overall approach. Whatever the instructional arrangement adopted, though, it should follow established principles of adult learning as well as allow for flexibility to meet the needs of adult learning styles.

### **Program Processes**

Professional development activities often target change to program processes other than instruction. For example, student recruitment, intake, and orientation procedures are subject to improvement, as are other, general program-management activities. An evaluation of program processes

thus needs to assess the implementation of changed procedures in these areas, when they are addressed by professional development.

Improved recruitment and intake procedures should be oriented toward the needs of the adult learner. The program's instructional goals, for example, should match learner needs; and outreach activities should be directed toward appropriate student populations. Improved intake and orientation methods must include a comprehensive assessment of student needs and abilities, as well as a strong orientation to the program, in order to identify instructional and other activities through which the student can benefit. Improved program management activities may include more involvement of instructors, other staff, and even students in determining program operations and policies, as such involvement can make a program more responsive to the needs of learners in the community.

### **Student Assessment**

Improvement of methods for assessing students is a common topic of professional development activities. An effective program has appropriate methods for assessing student needs at intake to place them, as well as methods for assessing learners progress throughout the program. Implementation of effective assessment includes using the appropriate assessment instruments; training of staff in the correct use, administration and interpretation of the instrument; and incorporating assessment findings into instruction. Due to shortcomings in the various assessment methods (discussed in Chapter 3), most practitioners and researchers recommend using both standardized and alternative assessment procedures to obtain an accurate picture of students' abilities and progress.

Evaluation of a program's assessment methods involves reviewing the actual instruments used, determining the training given to staff, and identifying how the information has been applied. Assessment procedures should be followed consistently and used widely in the program, if they are to affect student learning.

### **Learner Supports**

Professional development activities may inform teachers and staff on the types of needs learners have and how to identify and address what are called support needs. While transportation and child care are the most commonly considered support services, academic counseling, instructional assistance (such as providing tutors), and assistance to transition to further education are equally important services that learners need in order to succeed.

Most programs cannot afford to provide extensive support services to learners, but can do so through coordination arrangements with other agencies. Programs often provide child care and transportation in this way. Depending upon learner needs, programs also may provide social-service and

job-placement assistance through coordination with welfare and employment agencies. Many programs often provide instructional support through the use of volunteers; and counseling as well as transitional services often are provided directly in larger- or better-funded programs.

Considering the above, any evaluation of the implementation of learner support services in a program should examine the following issues:

- the type of services offered by the program,
- the method used to identify learners in need of service,
- whether or not these services meet the needs of learners,
- how these services are provided,
- the proportion of learners in need who actually receive the service, and
- unmet support-service needs.

Professional development aimed at improving program support services should have an impact in these areas, if student outcomes are expected from the activity.

### **Impact on Learners**

Assessing the impact of professional development on student learning is an almost completely ignored component of professional development evaluations. This neglect is almost entirely due to the inherent difficulties of measuring student impact, especially because of the relatively unsophisticated program structure supporting the adult education delivery system, and the difficulties of determining whether any observed change was, in fact, due to professional development. At present, the resources and technical expertise to conduct such evaluations put them beyond the reach of most programs. Nonetheless, broad evaluations can be conducted (e.g., sponsored through state or local funds), as such relationships to student learning development *must* be identified to truly determine the impact of professional development. Again, even though most programs cannot conduct this type of evaluation, the basic issues need to be considered as professional development activities are planned and evaluated because improving learner outcomes is the ultimate objective of professional development. The key evaluation questions, therefore, are:

- *Are students more satisfied with the instructional services as a result of professional development?*
- *Are students making learning gains?*

- *Has student behavior changed as a result of learning (i.e., are they able to use what they have learned in their work, family, or community)?*

As with the evaluation of professional development of instructors, answering these questions follows the same three basic steps: (1) evaluate reactions to new content and approaches (student feelings, emotional reactions, and values); (2) evaluate acquisition of knowledge and skills; and (3) evaluate changes in student behavior (doing things they could not do before), and the application of new learning (transfer to new situations).

### **Evaluate Student Reactions**

As instructors use new strategies or content material, it is important to explain to adult students why these strategies are being used, and what is expected from them as a result of the new strategies. Likewise, instructors should evaluate student reactions to these new strategies or content. Both research and common sense tell us that students are unlikely to learn what is distasteful to them, or to learn content that they fear they will not be able to master—two emotions that are probably related. Students who dislike and are afraid of math, for instance, will find it nearly impossible to succeed in that field; and can probably recall some school subject which we disliked, because we did not do well and, therefore, avoided it as much as possible.

Sylwester (1997, p. 17) agrees that the best teachers know that students learn more readily when they are emotionally involved in learning ". . . because emotion drives attention, which drives learning and memory. It's biologically impossible to learn anything [when] you're not paying attention." A side effect of professional development, then, is the raising of instructor morale in order to affect student learning in a positive way. "Nothing can make teachers feel better than learning to help students succeed who would have failed!" (Joyce and Showers, 1995, p. 58). Instructors' excitement also can infect students—thereby further raising student reactions and, hopefully, outcomes.

Once instructors determine how students feel about the content and process of instruction, they can to help students adjust their own attitudes. Instructors can devise instructional strategies to help students overcome dislike and fear, for example. Thus, the importance of evaluating for student reactions is a crucial step in the process, just as it is important to instructors in their own professional development.

### **Evaluate Acquisition of Knowledge and Skills**

As we have noted, the underlying rationale for professional development activities is that they result in greater student learning. Evaluating student knowledge and skills to assess instructional impact is an activity instructors understand well, and one they do most (although that may not be what they do best). In contrast to instructor learning, where reactions are most often sought for evaluation, adult student

learning must be concerned with *content*. In addition to the necessary skills of listening, speaking, reading, writing, computing, and learning new technology skills, controversy continues to abound concerning what subject-matter content *should* be taught. However, evaluation of content as currently performed, focuses on low-levels of cognition. The National Center for Educational Statistics reports (1993, p. 4) that ". . . students do reasonably well on lower level tests of achievement and cognitive skills but relatively poorly on tests that require complex reasoning, inference, judgment and transfer of knowledge from one type of problem to another." The reason is obvious: Recall of information is easiest to teach and to test.

A strong case could be made that the most important knowledge and skills students should learn are "concepts" rather than "facts." Knowledge about, and skills for, accessing rapidly changing, and growing information also is essential, as are thinking and organizational skills—all of which enable the complex processing and use of information. Such a shift in knowledge and skill emphasis requires a shift in the kinds of evaluation instruments and procedures used to determine mastery. Research supports the notion that staff development, using carefully selected strategies, can produce dramatic student gains. For example, Baveja compared the effectiveness of an inductive, concept-based approach with an intensive tutorial treatment. Pre- and post-tests were administered. "The scores by students from the inductive, concept-based treatment group were *eight times* higher than the scores for the tutorial group" (reported in Joyce and Showers, 1995, pp. 49-50).

### **Evaluate Changes in Student Behavior (*Uses of Learning*)**

The final and most complex aspect of evaluating professional development is to determine whether students use the knowledge and skills they learn to enhance their lives. How, though, can instructors determine *if* students can use what they have learned beyond repeating it back to instructors in oral or written form? One dimension of how students use learning involves observing that learning process in situations as close to real life as possible. For example, students can write real letters or do creative writing; and many ABE and ESL classes produce literary publications or present theater performances to demonstrate acquired literacy skills.

Another dimension for the use of learning is the ability to transfer that learning to new situations. There is general agreement that "transfer" is one of the least successful products of the teaching-learning process (Gardner, 1992; Joyce and Showers, 1995). As programs broaden the scope of their educational activities from knowledge and skill development into other life roles, however, it has become more important to evaluate whether adult education can prepare students for the different roles they need or want to adopt in their lives. To address that issue, several state and national efforts are currently underway to define the uses of adult education for learners. The *Equipped for the Future* Project (National Institute for

Literacy, 1995) is developing consensus within the field about what adults need to know in relation to their roles as parents, citizens, and workers. Several other states, including California (see Exhibit 4), Florida, Oregon, Rhode Island, and Wisconsin are conducting major efforts to develop curriculum standards for adult education instruction (Kutner, Webb, and Matheson, 1996).

These are difficult issues to assess, as follow-up of mobile students is both difficult and costly. In any case, the importance of collecting data only on material related to actual instruction cannot be overstated. Doing so would have avoided the problems encountered when one correctional program focused its instruction on basic skills, but did a follow-up study on decreased recidivism and increased employment. The results were not encouraging and the program was strongly criticized for failing in its mission. Since the only skills taught were basic skills, only those should have been the focus of any follow-up activity. The follow up did not, in effect, reflect reality.

**EXHIBIT 4****California Curriculum Standards for Students**

| <b>Productive Worker</b>   | <b>Involved Community Member</b>  |
|--|---|
| <ul style="list-style-type: none"><li>- improve basic skills for employment</li><li>- identify and demonstrate maturity skills</li><li>- get a job, better job, or promotion</li><li>- improve communication skills</li></ul>                                | <ul style="list-style-type: none"><li>- pass a citizenship examination</li><li>- become a citizen</li><li>- register to vote</li><li>- increase and demonstrate knowledge of history, government, civics</li><li>- become involved in civic/community affairs</li></ul> |
| <b>Lifelong Learner</b>  | <b>Effective Family Member</b>  |
| <ul style="list-style-type: none"><li>- improve basic skills</li><li>- use library and other community resources</li><li>- obtain high school diploma or GED</li><li>- enroll in other adult education classes</li><li>- enter a community college</li></ul> | <ul style="list-style-type: none"><li>- serve as role model for children</li><li>- become involved in children's schools</li><li>- help children with homework</li><li>- read to/with their children</li><li>- demonstrate parent skills</li></ul>                      |

**Summary**

An appropriate framework for evaluating professional development in one which regards professional development as a *change process*. Using that rationale, changes related to instructors, program services, and learners must be documented in order to provide an accurate assessment of the impact of professional development, which means that professional development's most immediate impact is on instructors—their reactions to professional development opportunities, the skills and knowledge they obtain, and the resulting changes on their instructional behavior. If, however, professional development has any hope of affecting learners, themselves, changes in program services (i.e., program adoption of a new instructional strategy or approach to assessment) also must occur; and although the resources and technical expertise necessary to measure professional development's impact on learners is unfortunately beyond the capacity of most programs, the basic issues related to assessing the impact on learners should still be considered as professional development activities are planned, provided, and evaluated.

## **CHAPTER 3**

### **STRATEGIES FOR EVALUATING PROFESSIONAL DEVELOPMENT**

#### **Introduction**

Several evaluation strategies have been used in both the K-12 and the adult education arena to assess the impact of professional development on instructors, programs, and learners. Some of these strategies rely on self-reports (e.g., questionnaires, interviews, focus groups, instructor portfolios, and practitioner journals); others involve feedback from administrators, peers or students (e.g., observations, written questionnaires, or interviews); and still others are competency tests and simulation exercises.

Within the context of day-to-day teaching and learning, selecting and implementing appropriate evaluation strategies for professional development is not an exact science. Such decisions, therefore, should be based upon a number of considerations, including: the specific questions to be answered; the focus of the evaluation effort (i.e., instructors, programs, adult learners); and the resources (both staff time and financial) available to allocate for the evaluation. Whenever possible, it also is important to use a diverse group of strategies for evaluating professional development.

Specific strategies that are appropriate for evaluating the effect of professional development on instructors, program services, and adult learners are discussed in this chapter, which includes a description of each approach, its advantages and disadvantages, and then gives examples of how they can be used for evaluation purposes.

#### **Professional Development Evaluation Strategies**

Exhibit 5 summarizes evaluation strategies appropriate for assessing the different dimensions of professional development, and the areas they are best used for assessing. What can be understood is that three strategies can be used in a wide variety of situations to address many different types of research questions. These multi-purpose strategies – *questionnaires*, *interview guides*, and *focus groups* – are best for collecting information about students, instructors, and program staff reactions (e.g., attitudes and opinions), knowledge and skills, and behavior.

#### **EXHIBIT 5**

#### **Professional Development Evaluation Strategies**

| <b>Type</b>                          | <b>Strategy</b>   | <b>Assessment Area</b>   |
|--------------------------------------|---|--|
| Multi-purpose Strategies             | Questionnaires<br>Interviews<br>Focus Groups  | Reactions<br>Knowledge and Skills<br>Behavior                          |
| Strategies for Assessing Instructors | Observations<br>Instructor Portfolios<br>Practitioner Journals<br>Collaborative Teams | Instructional Activities and Behaviors<br>Teacher Knowledge and Skills |
| Strategies for Assessing Students    | Tests<br>Alternative Assessment   | Student Knowledge and Skills<br>Student Behaviors                      |

Strategies best suited for evaluating an instructor’s classroom activities and behavior, as well as knowledge, include *observations*; *instructor portfolios* (with samples of a teacher’s work); and *practitioner journals*, wherein teachers record their activities and reflect upon what they have done in the classroom. *Collaborative teams*, in which groups of teachers work to implement and evaluate their own practices, are another way teachers can evaluate their performance.

Two strategies well-suited to evaluating student performance include *tests* and *alternative assessment* of student work.

### **Multi-purpose Evaluation Strategies**

Multi-purpose strategies are commonly used measurement methods for evaluation and research. These methods—questionnaires, interviews and focus groups—can be used to obtain information on a wide variety of topics, and from several types of respondents. For evaluating professional development, these strategies can collect information on reactions, knowledge, skills, and behaviors from teachers and students.

#### **Questionnaires**

Questionnaires, either administered by mail or by phone, make up one of the most frequently used type of self-report evaluation, typically used to collect information about perceptions. Questions may be either close-ended or open-ended, or a combination of both. Close-ended questions require the selection of one or more of several possible responses. Responses, in other words, may be selected from several choices (as in a multiple choice or true and false question), or from a checklist. Responses also may be designated on a scale, such as a Likert, which asks respondents to indicate their opinions on a numerical

scale that focuses on a single dimension of an issue.<sup>1</sup> The responses, usually moving from low to high, or agreement to disagreement, show how strongly an individual feels about a particular issue. Open-ended questions, on the other hand, ask respondents to answer questions in their own words. (Sometimes the questionnaire will include several sub-questions to facilitate the respondent's answer.)

Questionnaires are useful instruments for evaluating the impact of professional development on instructors, program services, and learners. They can be used with all types of professional development approaches for gathering information on reactions to professional development experiences, acquisition of knowledge and skills, and changes in behavior. Data about the impact of professional development on program services, including general organizational processes, instruction, assessment, and learner supports, also can be obtained from administrators, instructors, and learners. Questionnaires also are a useful mechanism for collecting information about the impact of professional development on instructor and learner reactions and behaviors.

***Implications for implementation.*** There are several recommendations that should be considered in the construction of a questionnaire to evaluate professional development. First, it is important to tailor the instrument to a specific event, to ensure that all the questions are appropriate. Second, it is better to ask a few questions than it is to ask a series of redundant questions that do not provide any new information. Third, it is better to ask respondents about specific questions regarding the content, process, and provider characteristics, rather than ask an array of questions about peripherals, such as room arrangements, temperature, audio-visual aids, refreshments, and timing. A respondent's negative opinions about refreshments may have a "spill over" effect on his or her responses about content or about the presenter, thus providing inaccurate results. Furthermore, focusing on the presenter, content, and process provides a more accurate picture of the event, as there tends to be a high correlation among responses to questions about the presenter, the content, and the process (Joyce and Showers, 1995).

Close-ended questions are easy to administer and analyze, although the range of possible responses is restricted; and respondents may be influenced by the structure of the questionnaire. For example, responses based on a checklist or Likert scale may be influenced by the items on the checklist or on how the Likert scale is presented. Also, on a Likert scale, there is a tendency to pick the middle choice when five choices are offered. For that reason, alone, many questionnaire developers offer only four choices, which are thought to force a more positive or negative response.

---

<sup>1</sup> Examples of dimensions used in scales are: strongly agree to strongly disagree, very useful to not useful, very favorable to unfavorable. Often scales are odd-numbered (e.g., 1-5) with 3 being a neutral or undecided value. Even-numbered scales (e.g., 1-4) force the respondent to lean toward one or the other end of the scale.

Responses on open-ended questionnaires are less likely to be influenced by the structure of the questionnaire. However, open-ended responses are more difficult and time consuming to analyze. To facilitate analysis, a code, containing key words, needs to be developed, to ensure that responses can be compared across completed questionnaires.

An advantage of questionnaires, especially when administered by mail, is that they can be administered to large audiences with relative ease, and they tend to be easy to analyze. The value of questionnaires, however, depends upon the quality of the instrument (well-constructed and pilot-tested) and on the sincerity and knowledge of the respondents. Procedures also must be established for conducting mail surveys, including tracking respondents, following-up with non-respondents, and reviewing and coding completed questionnaires, including telephone calls to clarify responses. It must be added that one drawback to using questionnaires is that response rates are typically low—unless there is a built-in incentive, or the evaluator makes a special effort to get people to complete them. Questionnaires also do not provide much depth of information.

### **Interviews**

Interviews can be valuable evaluation instruments, as they provide, with the interviewer's ability to probe for answers, more in-depth information than do questionnaires. Interviews can be administered by administrators, peers, or outside evaluators (who are often used to maintain objectivity and to solicit honest responses). Interviews generally are administered using protocols, which can be structured or open-ended. Structured protocols are essentially similar to questionnaires that are administered by mail or by telephone, and also may include both close-ended and open-ended questions. Open-ended questions are likely to include a series of "probes," which serve as prompts to help ensure that the respondent fully understand the questions and considers all possible options when responding.

With interviews, respondents can summarize in their own words answers to questions such as, *Why was this a valuable professional development experience? How do you plan to implement the new learning into your teaching environment? Why are you unable to adapt the new curriculum to your classroom setting?* Confidentiality is an important part of the interview, as the respondent must feel free to speak openly and honestly. (Interviews can be used to validate information recorded by teachers in their own logs or journals.)

Interview protocols can be used to assess all types of professional development experiences. They also can be used with all levels of evaluation, but seem to be particularly valuable in obtaining reports of changes in behavior. Interviews with administrators and supervisors validates information gathered through other evaluative mechanisms. In their supervisory roles, these persons can provide information

gathered through their own interviews, observations, and interactions with instructors; and they can provide data on the organizational climate and support provided for professional development. Such information places change in instructional behavior within the context of the organization.

Student responses to interviews can provide their perceptions on what occurs in a classroom or learning lab. Students also can report on how what they have learned has affected their lives. Data gathered from students, when used in conjunction with other evaluation strategies, could be used to validate findings.

***Implications for implementation.*** There are several drawbacks to using interviews as an evaluation tool. First, they are time-consuming for the interviewer. One interview may take 30 minutes or more to conduct, thus reducing the number of interviews that might be conducted. Second, interviews require trained interviewers who can avoid recording only answers they agree with and who can process the responses provided and ask follow-up questions to gather more in-depth information. Finally, interviews require well-designed protocols (sets of questions) to obtain appropriate information. Hence, interview protocols are often field-tested and revised before they are actually implemented.

There also is a danger for bias, if the interviewers are not trained and monitored on an ongoing basis. Results of interviews depend upon the skill and sensitivity of the interviewer and the care taken in the constructing of questions and follow-up probes. Consequently, when using interviews as an evaluation strategy, it is essential for interviewers to receive appropriate training.

There are several drawbacks to relying upon student responses to interview questions. Students may be reluctant to be critical about their instructors and may provide responses they think are expected, rather than report on actual behavior. Adult learners also may have preconceived notions about appropriate instructional strategies that may bias their responses. In addition, adult learners often have erratic attendance, which may make it difficult to accurately evaluate implementation of instructional strategies. Finally, gathering data from ESL students may be difficult. Language limitations and cultural inhibitions may cause them to be unwilling or unable to communicate as freely as students whose first language is English (Pennington and Young, 1989).

### **Focus Groups**

Focus groups are a common market research tool, used initially by the business sector to help position new products in the marketplace. Now, they have become popular with social science researchers. Focus groups generally involve a homogeneous group of 8 to 12 participants, a trained facilitator, and a prepared protocol containing open-ended questions. Discussions generally last between one and two hours. As typically conducted in educational research, focus groups are actually “group interviews,” and thus can

capture more in-depth information than can questionnaires. Focus group protocols typically include a series of general questions and thorough probes, through which facilitators can obtain responses to “Why” and “How” questions that provide more specific information. Since these activities are conducted with groups, only a limited number of questions can be asked, due to time constraints.

Focus groups can be used to assess all professional development experiences, and they are particularly useful for gathering reactions of instructors to professional development activities, and learners to instructor behavior and program services. As an evaluation strategy, focus groups also are useful for identifying what works well and what does not work well, as well as for generating ideas for improving the design and delivery of subsequent professional development activities.

*Issues for implementation.* Focus group interviews can be conducted with groups of practitioners or learners in the same program, or with practitioners who have had similar types of learning experiences in different adult education programs. A key advantage of a focus group is the potential to obtain information from more people than can be collected through one-on-one interviews.

As with other evaluation strategies, there are several disadvantages to using focus groups. Hearing other respondents, for example, may influence an individual’s responses, thus coloring results. Facilitators, like interviewers, need to be well-trained in leading discussions. In addition, a second person, who serves as a recorder, is often required to capture the information accurately. Alternatively, responses can be taped, but are costly to transcribe. Further, arrangements must be made for translators to participate in focus group interviews with learners whose first language is not English.

Another drawback is the difficulty in scheduling the discussion, given the part-time nature of many adult education instructors and the time commitments of adult students. One strategy for working around these kinds of scheduling difficulties is to conduct focus groups when instructors come together for a training conference or around a class session. Scheduling difficulties also affect focus groups for adult learners who may need to miss a class session in order to participate.

### **Strategies for Assessing Instructors**

While interviews, surveys, and focus groups are useful for assessing the reactions of instructors to professional development opportunities, the strategies presented in this section are particularly useful for assessing professional development’s impact upon instructor knowledge and skills, as well as upon actual behavior. Especially well suited for studying instructional practice, these methods are:

- Observation of Practice
- Instructor Portfolios
- Practitioner Journals

- Collaborative Teams

### **Observation of Practice**

Observations of individual teachers, over a period of weeks, months, or on shorter terms, provide useful information on teaching practices. There are two types of observations: structured and unstructured. Structured observations, using systematic observation and a coding instrument, narrow the focus of the observation and provide data on specific instructional behavior. Items on a checklist, for example, can be related to the specific skills and knowledge addressed in the instructor's professional development activities. Unstructured observations, in which the observer views the overall classroom and assesses a wide variety of teaching behaviors, provide general information about instructional practice and student responses, and are informative when used in conjunction with teacher interviews.

Observations require detailed notes for accuracy, rather than summary statements compiled after the observation is completed. In addition, a pre-observation conference may help to establish the focus of the observation, as it allows the instructor to relate to the observer the skills and knowledge on which to focus the observation; and it provides an opportunity to select the observation methods (e.g., use of a tape recorder or video tape), as well as allow the observer to gather information prior to the actual observation, thus enhancing the validity and reliability of the observation.

There are various observation techniques that can be used to gather data on classroom activities. Some of these techniques provide a detailed picture of what is occurring in the classroom setting; others focus on the specific aspects of the classroom environment. Included among these techniques are running transcripts, grids, tally sheets, lesson plans, time notation, dialogue recording, and videotapes. Once the data have been analyzed, post-observation conferences provide an opportunity for feedback to the instructor.

Observations are a useful evaluation mechanism, as they allow for an authentic assessment of instructor's behavior in their own working environment. Observations also validate information collected through teacher interviews, and may become part of an instructor's portfolio. In practice, observations may be conducted by peers, coaches, mentors, providers of professional development services, and administrators. Reliability can be increased through the use of structured observations, and by establishing, in advance, the focus of the observation and the agreed-upon format.

***Implications for implementation.*** The observation/feedback process is not only an evaluation strategy, but a professional development approach that can foster changes in behavior. Instructors can use the information gained through the process to help them assess how well they are achieving goals set out in

their professional development plan; and it helps them to identify strengths, weaknesses, and areas in which further practice and professional development is required. Observations can be used to assess the impact of all professional development approaches; and are a useful strategy for identifying changes in instructional behavior—particularly when used over a period of time.

Like other evaluation strategies, there are drawbacks to using this approach. Training is required in observation methods, data analysis, and feedback. It also can be a time-consuming process, particularly if there are pre- and post-observation conferences. Observations also may require adequate sampling of instructional behavior, as one observation may not be sufficient in demonstrating an instructor’s usual performance. Rather, a number of observations under comparable conditions may be required over a period of time. Scheduling and logistics also may present problems, especially if observations are conducted by peers who are part-time, and classrooms are at sites that are considerable distances from one another. Despite these drawbacks, however, observations are considered one of the best methods for bringing about change in instructional behavior (Joyce and Showers, 1983).

There are several ways that observations can be scheduled in adult education in order to overcome some of the obstacles mentioned above. Administrators can serve as substitutes while peers or mentors conduct observations, or, perhaps, where back-to-back classes are scheduled, these can be combined periodically to allow instructors to take turns trying out strategies and observing one another in the process.

### **Instructor Portfolios**

Instructor portfolios are a purposeful collection of instructional materials selected by the instructor, and they do not contain everything a teacher does. In this regard, they are tailored to the individual instructor. Portfolios foster instructor reflection, as items are selected, and as the contents are assessed over time; and they provide an ongoing record of instructor growth—and may be used by a mentor, trainer, or administrator to assess change over time and to plan further professional development activities. From K-12, they are used for “low-stake” purposes, such as pre-service training, self-evaluation, and professional development; in some states, they are used for “high-stake” purposes, such as promotions and salary increases.

Items selected for a portfolio can demonstrate acquisition of knowledge and skills and how this newly acquired learning is put into practice; and they can be used to assess changes resulting from any professional development experience. They are particularly well suited for practitioner inquiry, as these experiences usually occur over a period of time.

***Implications for implementation.*** Materials included in the portfolio should reflect the professional development of the instructor over time, so that items may include:

- the instructor's short and long term professional goals,
- statements indicating the extent to which the instructor believes the stated goals have been achieved, and
- tangible evidence and artifacts related to the instructor's achievements that reflect professional development experiences, including materials produced by the instructor (e.g., reflective essays, instructional activities), products of teaching (e.g., student work, videotapes of the learning environment), and information from others (e.g., results of peer observations, student evaluations).

As practitioners select materials for inclusion into their portfolios, they should consider the following questions: *Why did you include this entry and what is the most interesting thing about it for you? How does the entry relate to your work and the other entries you have selected? Why does this entry show most clearly what you have learned? How does this entry show your strengths and weaknesses in what you have learned?* (Feldman, 1992). Responses to these types of questions help practitioners to critically analyze their entries, and gain a better understanding of their own professional development.

There are several advantages to the use of portfolios as an evaluation tool. Compared to other evaluation approaches, such as observations, interviews, or tests, portfolios contain a variety of materials, can reflect many different tasks, and can cover an array of teaching behaviors. Furthermore, they lend themselves to use by a range of instructors, and can be adapted to any type of teaching situation (e.g., classroom instruction, labs). In addition, portfolios can include evidence of performance, such as letters and publications, or products not readily available through other evaluation strategies; and they also can include evidence from multiple sources, including other teachers, mentors, coaches and students.

The very process of developing portfolios may encourage instructors to work together, thus creating an atmosphere of collaboration among practitioners. Overall, the process of compiling a portfolio encourages self-evaluation, reflection, self-improvement, and professional development.

The successful implementation of portfolios as an evaluation strategy requires support from administrators, as well as "a buy-in" from instructors. Doolittle (1994) suggests that administrators and professional development coordinators should consider the following steps in implementing portfolio assessments: start slowly, gain acceptance, instill ownership, communicate implementation, use models of other instructors, provide training, and provide time for instructors.

Portfolios, however, take a significant amount of time to compile, and, to be developmental, they must be compiled over a period of time. Part-time instructors may not have adequate time to devote to organizing such portfolios. In addition, instructors need guidance in developing portfolios and in selecting

items for inclusion. Oftentimes, there is a tendency for instructors to focus on their best efforts, rather than on compiling a developmental portfolio.

Finally, there may be difficulties in analyzing and interpreting data collected in the portfolio. Since each portfolio is distinct and tailored to the individual, it may be difficult to compare information from different portfolios. As is the case with most performance assessments, some rubric (rule or explanation) for evaluating the data must be developed. Furthermore, given the distinct nature of each portfolio, it is difficult to aggregate the data and report it to state and Federal agencies.

### **Practitioner Journals**

Practitioner journals are written logs that contain an instructor's description of, and reflection upon, teaching practices over time. Journals can be used as a means of documenting progress in using new skills or in applying new learning. Journal entries, therefore, may include the number of times a new strategy is used, student responses to practices, ways the instructor plans to modify practices, or the kinds of professional development or support necessary to enhance the implementation of the practice. Sections of the journals may be included as part of the instructor's portfolio, as a means of documenting professional growth.

*Implications for implementation.* The primary advantage of practitioner journals is that they encourage instructors to take time to think about their practices, and to reflect upon what works and what does not work in their classrooms and how they can improve their teaching. Thus, it is a valuable self-evaluation tool. There also is evidence that structured journals, in which teachers record their daily activities, can provide data consistent with classroom observations (Porter et., al, 1993).

The drawbacks of maintaining a journal are similar to those of developing instructor portfolios. Journals may be too time-consuming, especially for part-time instructors. If too much time elapses between entries, it is difficult to get a clear picture of continuous growth. In addition, training is required, if instructors are to derive the most value out of this tool. Finally, it is difficult to analyze and interpret the data, since each journal records the individual's own experiences and thoughts.

### **Collaborative Teams**

In this approach, teams of teachers, working collaboratively, can plan how they will monitor their implementation of new instructional strategies resulting from their professional development experiences—and how they will determine the effect of such strategies on students. Once they agree upon a design, instructors monitor their own efforts for implementing the strategy; and they discuss at team meetings what works and what does not work, brainstorming methods for improvement. Team meetings also serve as forums to discuss how students are responding to new instructional strategy. Working through ways to

improve implementation of new instructional approaches becomes part of the professional development cycle, as instructors plan how they will foster change. This approach, which can be used to assess the impact of all professional development experiences, combines self-report with working collaboratively with peers.

*Implications for implementation.* The advantage of collaborative teams is that they provide the opportunity for instructors to obtain feedback on their own efforts to implement new instructional strategies, and to learn from one another. It also is taken for granted that there is a more concerted effort for planning change when instructors work together in harmony.

One drawback, however, is providing enough time for instructors to meet collaboratively. The part-time nature of the adult education staff may, itself, make it difficult to schedule meetings. Support from the administrative staff is, therefore, required; and may include hiring substitutes to cover classes to allow teachers time to collaborate, or providing stipends for teachers to meet on the weekend to discuss how the implementation of new strategies is proceeding.

### **Strategies for Assessing Students**

Assessing learning and the use of information learned is one of the biggest issues confronting adult education. For that reason, two common methods of assessing students are presented here: tests and performance-based assessments. Standardized, curriculum-based, and program- and teacher-developed tests are the most common method for evaluating students, although many practitioners believe that tests are less than satisfactory in capturing what adult students learn. “Alternative assessment of student performance” is a broad term that encompasses several non-standardized ways of evaluating students, such as observing them or developing a portfolio of student work.

#### **Tests**

As previously stated, assessing the acquisition of skills and knowledge through tests is the most commonly used mechanism to evaluate the impact of professional development experiences on learners. Such tests include both teacher- or program-developed assessments and standardized assessments. Generally, teacher- or program-developed assessments consist of paper and pencil tests of specific content and skills covered during instruction.

Standardized tests can be useful to assess the extent to which participants have acquired new knowledge and skills. They are most appropriately used following a workshop/presentation approach, wherein contents and technique are the focus. Commercially available standardized tests, such as the Test of Adult basic Education (TABE) and the Adult Basic Learning Examination (ABLE), predominate in adult basic education. For ESL students, the most commonly used tests are the Basic English Skills Test

(BEST) and the CASAS ESL Appraisal (Burt and Keenan, 1995). Important advantages of using tests as an evaluation strategy are that they eliminate individual bias, if objectively scored, and they guarantee a minimum standard of knowledge (as well as being legally defensible in a court of law) (Darling-Hammond, Wise and Pease, 1983). Standardized tests, however, often do not assess individual learner strengths and weaknesses, and often do not measure what has been taught (nor address student goals). (Burt and Keenan, 1995; Lytle and Wolf, 1989; Wrigley, 1992).

***Implications for implementation.*** Tests can produce a picture of student achievement at any given point, although it must also be recognized that there is a movement away from relying on tests toward more authentic assessments, such as portfolios and observations. To determine whether change or growth has taken place, however, requires some sort of pre- and post-measures. After a particular instructional sequence, for example, instructors often test for the content or processes taught. The *assumption* is that the results of the test reflect what students learned during that instruction. This is a risky assumption, as the level of the student's prior knowledge is not known without a pre-test.

To determine accurately the success of an instructional sequence, instructors must pre-test students at the outset, and, using the same or a similar instrument, test them at specified intervals throughout the instructional process. Those measures will reveal what students have actually learned from instruction. (In adult education, intervals must be shorter than sometimes desired, because of the high turnover or absenteeism of adult students.)

Four ways to maximize the benefits of testing are to: (1) choose tests that match the demographic and educational backgrounds of the learners; (2) interpret scores carefully; (3) ensure that test objectives match the program objectives and curricular content; and (4) use additional instruments to measure learner achievement (Burt and Keenan, 1995).

### **Alternative Assessment**

The limitations of traditional testing have made alternative assessment methods attractive to many adult educators. Alternative assessment, which has come to mean any type of student assessment that does not rely on standardized tests, include student portfolios, and performance-based assessment. Much like the instructor portfolios discussed above, student portfolios are collections of student works—such as writing samples, student reflections on their progress, book reports, classroom work, and test scores. The teacher and student maintain the portfolio over time, and student learning and progress may be observed by comparing later work to earlier work.

Performance-based assessment involves testing students or observing them as they perform tasks, using authentic materials in the actual context in which they are likely to be encountered. An employee, for

instance, may be observed reading a machine-operating manual on the job. The assessment may be made using a test, or through a structured observation.

*Implications for implementation.* Alternative assessment methods for students have the same advantages and drawbacks as the instructor portfolio assessment described above. Many types of alternative assessments are well suited to adult education's individualized approach; and an assessment can be tailored to the student's specific goals, as well as include instruction received. Portfolios can contain a variety of materials, can lend themselves to use by a range of instructors, can be adapted to any type of teaching, and can include evidence of student performance, such as letters or products not readily available through other evaluation strategies. The information in student portfolios is especially useful to instructors, who can then tailor lessons to areas of an individual student's interests, strengths, and weaknesses.

For the purposes of evaluation, however, alternative assessments, since they do not produce quantitative data, can be difficult to analyze. Rating or scoring student work also is very difficult to standardize—and different raters may derive different assessments of learner progress and needs. Maintaining a student portfolio is very time-consuming for both students and teachers, and also requires that the student remain in the program for some time. The student must stay long enough for a body of work to develop and be collected. It also is time-consuming to review and evaluate student work, making portfolio assessment impractical for large-scale evaluations.

### **Summary**

Many strategies are available for evaluating professional development and decisions about which strategies to use should be based on the specific evaluation objectives, evaluation questions, and audiences for the information. Whenever possible, it is preferable to use a variety of strategies. Many strategies can be used for evaluating impact on instructors, programs, and learners, although some are better suited for specific areas. Three strategies—questionnaires, interview guides, and focus groups—are best for collecting information about instructor, program staff, and student reactions, knowledge and skills gained, and behavior. Strategies best suited for evaluating impact on instructors are observations, portfolios, journals, and collaborative teams. Two strategies especially well-suited for assessing the impact of professional development on students are tests and alternative assessments of students.

## **CHAPTER 4**

### **DESIGNING A PROFESSIONAL DEVELOPMENT EVALUATION PLAN**

#### **Introduction**

This monograph has emphasized how efforts to evaluate the effectiveness of professional development activities must be incorporated within different levels of the professional development process, if, indeed, information from evaluations are to inform program accountability and program improvement efforts. Changes in instructors, programs, and students must be documented and evaluated, to assess the impact of professional development activities; and conducting effective evaluations requires sufficient attention be paid to the context within which professional development services are provided (i.e., the adult education delivery system).

Assessing professional development's impacts on instructors and even on programs are fairly easy to accomplish. Determining the relationship between professional development and learner outcomes is a more complicated undertaking. Broad evaluations sponsored by states or local funds can examine the relationships between professional development and student outcomes, although most individual programs may not currently possess the necessary resources or technical expertise to conduct such evaluations.

Because enhancing learner outcomes is the ultimate objective of professional development, states, professional development agencies, and local adult education programs must develop plans to assess the relationship between professional development and learner outcomes. Successful professional development evaluations require careful planning and sufficient resources, including time, to conduct the evaluation, before the beginning of which, the specific questions that will be answered are developed, and appropriate evaluation strategies, and plans for analysis and reporting data need to be made. In this concluding chapter, specific issues related to the adult education program infrastructure that affect professional development evaluations are presented, followed by a discussion about the components of a professional development evaluation plan.

#### **The Context of Professional Development Evaluations: Issues Related to the Adult Education Service Delivery System**

Even under the best of circumstances, conducting evaluations is a complex process and is by no means an exact science. Program providers naturally prefer to concentrate their energies on providing services, rather than on collecting data, especially if they do not understand how information can be used to

enhance program services and improve learner outcomes. Other service delivery areas that receive substantially more financial support than adult education (e.g., JTPA, Title I, special education) have both more experience with evaluation and a program infrastructure that is able to provide data in response to fairly stringent reporting requirements. This is certainly not the situation in the adult education field, which has for so long been under funded and under appreciated.

When designing professional development evaluations, it is essential to consider and address difficulties associated with a relatively unsophisticated program structure, which include a part-time delivery system, a mostly part-time instructional force, a high rate of instructor turnover, and a lack of certification requirements or commonly accepted instructor competencies and skills.

### **Part-time and Voluntary Nature of Service Delivery**

Efforts to evaluate professional development activities, particularly student outcomes, must consider that adults generally attend classes on a voluntary, part-time basis, and can stop attending classes whenever they wish. In addition, to accommodate complicated schedules, instruction is often organized in an “open-entry and open-exit” manner, with great flexibility as to when students can begin and end a course of study. Many programs do not operate on a strict semester basis, as in elementary and secondary education. Frequently, instruction is offered in short cycles (e.g., nine weeks) over the course of a year.

This part-time and volunteer nature of the adult education service delivery system complicates efforts to determine student learning gains. Not only do many students exit the program before gains can be registered, but they often stop attending at any time—usually without formally withdrawing or notifying the instructor. This is a problem that plagues most efforts to evaluate adult education services and also, of course, professional development evaluations seeking to assess professional development’s impact on learners.

### **Part-Time Instructors and High Rate of Instructor Turnover**

With instructors typically working at adult education programs on a part-time basis and usually employed full- or part-time at other jobs, efforts to schedule professional development opportunities are complicated, and the types of available professional development are limited. Furthermore, many instructors are not compensated for any planning time or professional development time.

The added high rate of instructor turnover means that programs must offer introductory professional development opportunities almost continuously. It is certainly more difficult to demonstrate professional development’s impact, if instructors do not continue teaching long enough to effect instructional change, and if professional development agencies must continuously repeat activities because there are always so many new instructors.

These characteristics of the adult education workforce can make evaluation of professional development impossible when the instructors who receive the training leave before they can implement changes to the program, or before they can be interviewed or observed.

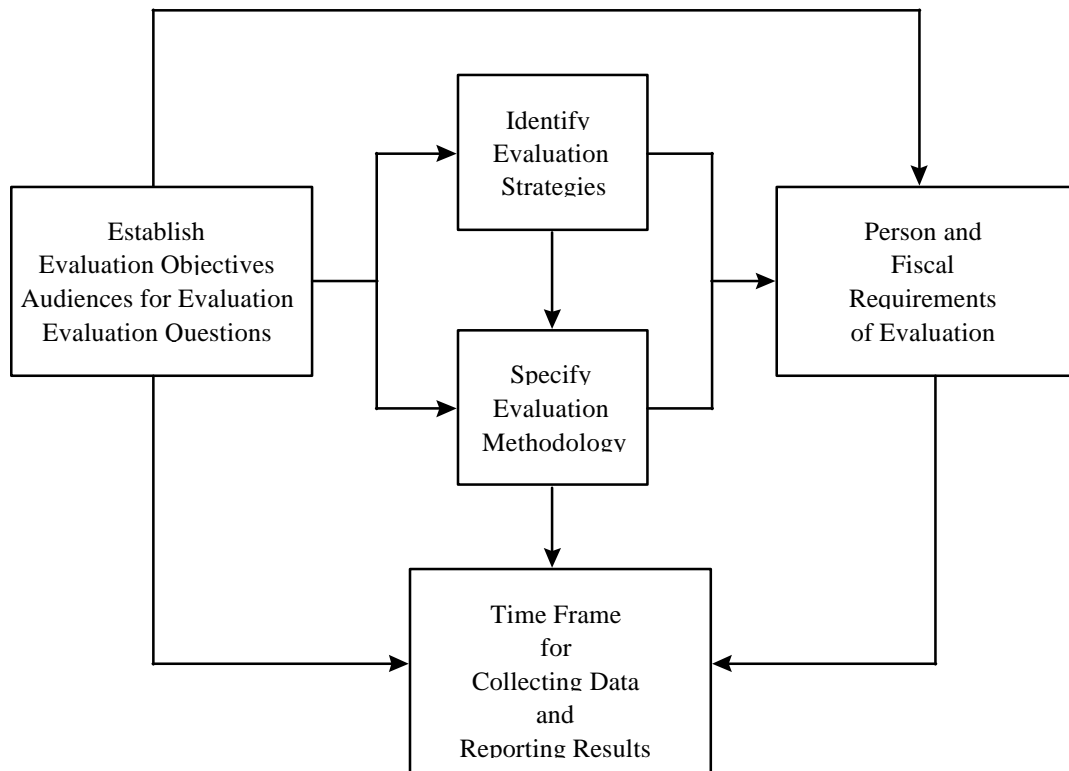
### **Lack of State Certification Requirements or Required Instructor Skills and Competencies**

Unlike the elementary and secondary education arena, where instructors must meet state certification requirements before they are allowed to teach, few states have specific certification requirements for adult education instructors. When states require instructors to be certified it generally consists of elementary and secondary education certification. Several states have identified competencies and skills expected of adult education instructors, and the 10 instructional packets prepared through the *Study of ABE/ESL Instructor Training Approaches* identifies a series of broad competencies that instructors should possess. Yet, the absence of a national consensus about instructor skills and competencies complicates efforts to evaluate professional development services.

## **Components of a Professional Development Evaluation Plan**

When designing professional development activities, having a detailed plan for collecting evaluation data that will inform both efforts to ensure continuous program improvement and program accountability must be developed. This plan serves as a “road map” for professional development evaluations and will help ensure that appropriate data are collected for both program improvement and accountability purposes. Exhibit 6 presents the components of a professional development evaluation plan and their interrelationship, and begins with a first component that is the determination of the evaluation’s objectives, audience, and evaluation questions. This component,

**EXHIBIT 6**  
**Components of a Professional Development Evaluation Plan**



in short, establishes the parameters for all other plan components. The specific evaluation strategies and evaluation methodologies chosen are, then, largely based upon the evaluation’s objectives, audience, and evaluation questions, as well as upon the level of person and fiscal resources required to complete the evaluation and the time frame for conducting the evaluation.

**Determine Evaluation Objectives, Audiences for Evaluation, and Questions**

As discussed in the first chapter, data reporting the impact of professional development activities on instructors, programs, and learners may be used in many different ways that relate to the rather broad objectives of program accountability and program improvement purposes. Thus, the first step in developing a plan for evaluating professional development is to determine specific evaluation objectives and audiences for the evaluation effort. The objectives and questions should be as specific as possible, and should be related to the target of the evaluation (i.e., teacher, program, or student).

### **Audiences for Evaluation Data**

Specific data and information needs from professional development evaluations will depend on the specific audiences for the evaluation data. Potential audiences for data from evaluations of professional development activities include:

- legislators and policy makers at the Federal, state, and local levels,
- state adult education staff who are responsible for distributing funds that support professional development activities,
- state and local staff from other agencies who are interested in benefiting from professional development activities provided by professional development agencies,
- local program administrators who participate in professional development activities and whose staff participate in professional development activities, and
- instructors who participate in professional development opportunities.

The reason it is important to determine your audience is that individuals representing many different roles want to know “something different” from an evaluation. For example, legislators are often focused on the “bottom line”: the cost effectiveness of professional development. Instructors on the other hand, are more concerned about whether professional development makes them better teachers. Furthermore, information from professional development evaluations can inform local programs about the mix of professional development services that are most effective in situations similar to their own program’s environment. Evaluation data also can offer programs insights into what administrative supports and changes in program practices are necessary to facilitate changes in instruction.

Professional development agencies may be interested in knowing participants’ reactions to training. This and other information from the evaluation also can help them in determining the mix of activities that is most likely to result in increased instructor knowledge and skills and changed instructional behavior.

State staff most typically conduct evaluations of professional development to help them make informed decisions about whether to fund professional development agencies, refocus the orientation and

focus of professional development services, or to promote different instructional practices in local adult education programs.

### **Evaluation Questions**

Specific evaluation questions should be developed once evaluation objectives have been determined and the specific audiences for the evaluation have been identified because, as shown above, different audiences will want different types of information from professional development evaluations. These questions will drive the decisions that must be made regarding other components of the evaluation plan, including the specific evaluation strategies and methodologies selected, necessary person and fiscal requirements, and needed time period for completing data collection and reporting activities.

Developing evaluation questions is especially important when a central objective of the evaluation effort is to provide information for policy makers about the effectiveness of professional development, and the specific questions must be directly related to the evaluation's intended audience and their interest in professional development. When the principal objective of a professional development evaluation is to obtain information that can be used to improve program services, identifying specific evaluation questions is a more straightforward process.

Whatever the purposes of the evaluation, when developing specific evaluation questions it is important to ensure that they:

- Have a direct and clear relationship with the purposes and objectives of the professional development activities, program, or system being evaluated;
- Will provide information that can be used by the intended audiences of the evaluation in order to improve program services or promote accountability;
- Can be answered given, the context of the adult education service delivery system and its professional development system, and with the person and financial resources available to conduct the evaluation; and
- Will result in data and information that is meaningful and easy to understand.

Evaluation questions also provide guidance for analyzing the data collected from the evaluation. The analysis can revolve, in fact, around answering those questions.

### **Identify Evaluation Strategies to be Used**

As discussed in Chapter 3, an *evaluation strategy* is a method of collecting information about the impact of a specific component of professional development. Decisions on what strategy to use typically involve considering the specific objectives of the evaluation effort, including potential audiences for the information and the resources available for the evaluation. It also is of particular importance when deciding

on evaluation strategies to consider the implementation issues associated with different strategies, and select those that provide both quantitative and qualitative data, as well as reflect the different mix of approaches being used to provide professional development.

### **Consider Implementation Issues of Different Strategies**

As discussed earlier, the implementation issues to consider when selecting an evaluation strategy typically revolve around instrument development, data collection training, and a willingness of respondents to participate in evaluation activities. Therefore, when developing data collection protocols, care must be given to ensuring that the questions are asked in an unbiased manner, and that respondents will understand the information that questions are intended to elicit. Such protocols, including surveys, interviews, and focus group instruments, should be field-tested in “real-world” situations.

Training data collectors is another issue that must be addressed when selecting specific strategies. When using phone or in-person interviewers or focus group moderators, the individuals conducting the interviews must receive sufficient training to ensure that they are comfortable with the instrument, are knowledgeable enough about the program and professional development, so that they can understand the range of possible responses, and will ask questions in a neutral manner, so as not to lead the respondent to an answer. When using inquiry research or journal evaluation strategies, care must be given to ensure that respondents have received sufficient training and technical assistance with these strategies and are willing to commit a sufficient amount of time to complete their tasks.

### **Provide both Quantitative and Qualitative Data**

The evaluation strategies discussed in Chapter 3 provide both quantitative and qualitative data that can be used to assess impact of professional development strategies on instructors, programs, and learners. Quantitative data, such as test scores, are easily aggregated. The advantages of these strategies are obvious: data can be quickly tabulated, manipulated and graphed. Comparisons are easy to see, as well, although they can be misleading if not properly presented. The general public and most funders and fond of these data because, aside from having an intuitive appeal, they make it easy to compare successes and failures. However, unless the data are valid and reliable, they cannot provide a complete or accurate picture of outcomes. If not valid, numeric

estimates may be based upon unsupported statistical assumptions and modeling. In addition, the methodology used to gather the data may be suspect.

Qualitative data, often obtained through self-report devices (surveys, interviews, questionnaires, and the like), allow for richer, although more subjective, information, as making certain that self-report data are accurate is a difficult task. On the other hand, the information may be more valuable than any test score in determining whether student goals have been met. Likewise, “soft data” in the form of anecdotes often can be more persuasive than hard data precisely because people tend to recall the anecdote's emotional, and perhaps even intellectual impact. New brain research also is beginning to demonstrate how vital a role emotions play in learning. Furthermore, criteria or “rubrics” can be established to make soft data more valid.

### **Determine Evaluation Methodology**

Selecting the evaluation methodology generally involves considering a range of technical and statistical issues, as well as accounting for the resources available for data collection and analysis. Hence, decisions about the specific evaluation methodologies used should be related to the objectives and evaluation strategies chosen to collect data. Other technical issues that must be addressed at the outset of the evaluation effort include whether to collect data from all participants (i.e., instructors, programs, or learners), or to select a sample, the reliability and validity of the data being collected, and the importance of maintaining data confidentiality.

#### **Sampling**

Evaluations, whether seeking to obtain data that will inform either program improvement or program accountability purposes, can collect data from all instructors, programs, and learners, or from a scientifically drawn sample that can provide equally valid data at less cost. A state interested in evaluating the professional development services it funds can collect data from all of its professional development grantees and a sample of local adult education programs receiving professional development from each agency. Sampling methods often are used with methodologies that involve follow-up data collection activities with instructors or learners. It is also appropriate to employ sampling methods when using alternative learner assessments whose scoring costs can be substantial.

When considering using a sample, it is important to remember that although highly valid sampling plans usually can be devised, successful implementation of a sampling methodology often is beyond the capacity of most program providers. Sampling methods require extensive knowledge of the characteristics of participants, rigorous procedures for drawing a sample. Numerous potential selection biases also can plague a sample, thus invalidating data..

### **Data Reliability and Validity**

All evaluations should be concerned with ensuring the reliability and validity of the data collected. *Reliability* refers to whether measurement will be the same when conducted over time and by different people. A test that is perfectly reliable and administered properly, will provide the same result no matter who administers it and no matter where it is given. This is an especially important issue when many different individuals, such as instructors at many adult education programs, will be collecting data over time (e.g., through the use of interviews and surveys). Inter-rater reliability, for example, is an issue that must be addressed when using alternative assessments as an evaluation strategy.

*Validity* refers to whether the data collected measures what it is intended to measure. When there is a direct relationship between a data collection strategy and the intended outcome there is typically high validity. Ideally, a test to measure instructor knowledge in a content area or a standardized assessment to measure learning gains should have high validity. In general, threats to validity include poorly designed questionnaires or interview protocols, inadequately trained data collectors, flawed sampling plans, and insufficient follow-up of program participants,.

### **Data Confidentiality**

Data collection methodologies that rely upon data sharing, where multiple agencies may collect and share information (i.e., professional development agencies, local adult education programs, and for learners other service providers) must deal with issues of data confidentiality. There may be a need, therefore, for a *common identifier* to facilitate collecting data. *Control* of data is a related issue, especially what data are reported and by whom.

### **Determine Required Person and Fiscal Resources**

Evaluation strategy, methodology, and time frame determine the number of persons and fiscal resources needed to complete the evaluation. These combined resources always are finite, and are typically not large, which often influences what can actually be accomplished. Person resources are likely to extend beyond the staff conducting the evaluation, such as the person time required by professional development providers, local administrators, instructors, and learners.

When deciding about which mix of evaluation strategies to use and the types of data analysis techniques to employ, it is essential to consider the cost implications of all decisions. With limited resources, what must be weighed includes the relative benefits of using different evaluation strategies and the resulting effects on the type of data collected, as well as the potential uses of that information. More resources typically produce better data validity and reliability, but, depending upon the evaluation's

specific objectives, the audience for the evaluation effort, and the potential ways that the data are expected to be used, it may be reasonable to have lower quality data, if they can be collected at a lower cost.

### **Establish a Time Frame for Collecting and Reporting Data**

It may take considerable time before the impact of professional development activities can appear. Teachers must have that time to develop new practices and lessons; programs need time to implement new procedures; and students need instructional time, before any change can be expected. This time lapse must be considered when establishing the time frame for the evaluation effort, including the schedule for collecting and reporting. Since changes often occur in stages, or slowly over time, it often will be necessary to include in the evaluation methodology data collection procedures that allow for repeated measurements over time.

### **Sufficient Time Must be Allowed Before Results Can Be Expected**

During the first year of a professional development initiative, data collection efforts should focus on changes in instructor knowledge and behavior, with perhaps a preliminary assessment of program services and learners. It is only after instructors have the opportunity to practice the skills that professional development activities are intended to promote, and to participate in repeated professional development activities that reinforce and enhance what is being learned, that an evaluation will be able to reveal effects on program services. Finally, improvements in learner outcomes can be realized only after program services have changed in response to instructor changes that result from professional development activities.

### **Measurements Should Be Repeated at Appropriate Times**

Since professional development is a “change process,” it is necessary to repeat data collection activities. In fact, to determine the extent of professional development’s effect on instructors and learners, it is necessary to assess their knowledge and skill levels before professional development activities have started. Such measurements must be repeated on a regular basis to determine the overall effects of professional development, and to identify the specific and effective professional development activities. Since it is unlikely for professional development to affect learner outcomes without first changing program services, it is essential to document the program’s efforts to make that change, and to be able to relate such changes to specific professional development activities. It is equally important to measure student learning over time, and to be able to tie learning to instructor and programmatic changes.

## **Summary**

Even under the best of circumstances, conducting evaluations is a complex process, and is by no means an exact science. When designing professional development evaluations, it is, therefore, essential to consider and address difficulties associated with a relatively unsophisticated program structure, that includes a part-time delivery system, a mostly part-time instructional force, a high rate of instructor turnover, and a lack of certification requirements or commonly accepted competencies and skills. Preparing an evaluation plan will further help to ensure the success of a professional development evaluation, with components of such a plan including: (1) determining the evaluation's objectives, audience, and specific evaluation questions; (2) evaluation strategies to be used; (3) evaluation methodology, including data collection procedures; (4) person and fiscal resource requirements; and (5) time frame for collecting data and reporting results.