PDP Communiqué

A Newsletter from the
Professional Development Program, Rockefeller College

UNIVERSITY AT ALBANY
State University of New York

Measuring Success: Accountability and Professional Development

Dianna Newman, Ph.D.

Professional development is increasing efforts to demonstrate accountability of resources spent on the development process, the immediate outcomes associated with it, and changes in organizational performance by the end consumer. Over the past two decades, tremendous strides have been made in developing scientifically proven practices that will improve the quality of services, and ultimately, consumer outcomes, as researchers and practitioners generate improved models and methods of delivering and evaluating professional development.

Why is evaluation of professional development important?

There are several reasons why the evaluation of professional development has become increasingly important. The first and foremost reason is accountability of services to the consumer. This includes accountability of how we gain new knowledge and how that knowledge is used in practice; both of these involve processes, products, and systems. It is important that we provide credible evidence that we are offering consumers the newest, most informative knowledge in a way that increases not just short-term ability, but the application of that knowledge over time and setting. We must document the validity of knowledge, transfer and use of knowledge, and systemic integration of knowledge in a way that is credible to consumers.

A second reason supporting evaluation is the need for accountability to our funders — those taxpayers, agencies, organizations and businesses that support our professional development. As fiscal resources become more limited and the need for services grows, there is an increasing need to show that monies in support of professional development are not only well spent, but they are also irreplaceable and not duplicative of other opportunities. For the past decade, federal and flow-through state funding has required that professional development activities satisfy revised GPRA and PART mandates that address not only the numbers of persons trained but also the numbers of consumers served by those who are trained, cost benefit and cost effectiveness ratios reflecting savings in resources, and efficiency of training. Current mandates also are asking for the use of scientifically based evidence of outcomes as well as documentation of the maintainability and sustainability of the knowledge transfer process. Variables of interest include changes in the training process, improvement in the practice of those receiving the training, and support from the systems that will implement the new practices.

What do we need to change in our evaluation of professional development?

This broader more delineated definition of evaluation necessitates changes in our evaluation practices. One of the first things we need to change when

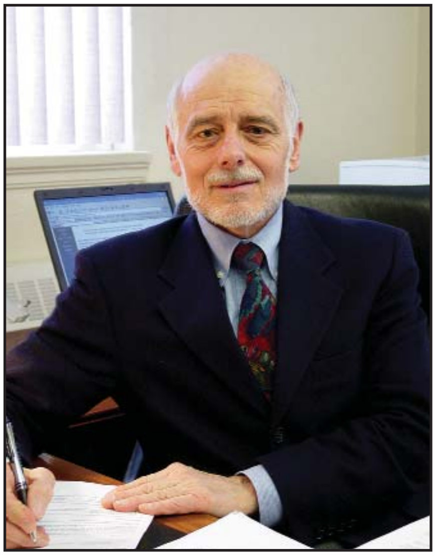
Continued on page 3

MARK SCHMIDT

Dr. Dianna Newman, UAlbany Associate Professor of Educational Psychology and Methodology

VOLUME 14, SPRING 2005

From the Director…



GEORGE DOWSE

Eugene J. Monaco, Director and Public Service Professor

Distinguished Extended Learning Leadership Award

MARK SCHMIDT

Dr. Sue Faerman, University at Albany Dean of Undergraduate Studies, Professor, and Collins Fellow receives the 2004 Distinguished Extended Learning Leadership Award at the November Rockefeller College Awards Program. The award is presented annually by PDP for exemplary leadership and significant contributions in the field of continuing professional education.

Kirkpatrick and Beyond: PDP Ups Measures for Accountability

Eugene J. Monaco,
Director, Professional Development Program

What is it we look for when evaluating our educational programs? Certainly, we need to determine the degree to which our objectives have been achieved. Yet, what else? To measure the gains in knowledge and skills or to change behaviors and attitudes? To enhance quality or to impact the bottom line? Whichever of these questions we may be seeking to answer, accountability is becoming the driving force. It is a complex accountability involving learners, employers, funding sources, professions, colleagues, and end customers. For more than a decade, educational institutions along with the public and private sectors have witnessed a dramatic shift toward making the training of individuals directly meet the organizational expectations of return on investment (ROI). During this period, program evaluation has changed to reflect the new conditions of business accountability metrics.

In this issue of the Communiqué, evaluation is looked at from two different perspectives. Dianna Newman from the UAlbany School of Education discusses the importance of evaluating professional development, and PDP's Andrew Sanchirico provides an overview of Donald Kirkpatrick's *Four Levels of Training Evaluation* first introduced in 1959, revised in 1975, 1979 and 1998. Both articles stress the need to conduct evaluation that reflects the changing expectations for training.

In evaluating contract training programs for government and not-for-profits, PDP has embraced a responsibility to our contract sponsors that takes us beyond merely determining learner satisfaction with our training programs. We instead have been generating data that measures how well learners apply the knowledge and skill they developed in the classroom environment to the jobs they are being asked to do in the workplace. In fact, we have led our contract agencies to expect this higher quality of evaluative information and find ourselves being measured by our effectiveness in delivering it. As our training has moved away from satisfying trainees to meeting the contracting organizations' expectations, more emphasis has been placed on outcome-based metrics such as aligning improved job performance with organizational objectives. This is the more sophisticated and demanding level of ROI evaluation that is discussed in the work of Jack Phillips (1996) who is well respected in the field of ROI analysis on training.

Although some suggest that the Kirkpatrick model has outlived its usefulness, and that other more sophisticated ROI models should be employed, our position at PDP is to be inclusive of both approaches rather than exclusive to only one perspective. PDP has looked well beyond Kirkpatrick's Level I evaluation of trainee satisfaction, and Level II continues to be important since it measures what was learned. However, we now find ourselves placing more emphasis on both the Level III measures of improvement in job performance and the Level IV metrics related to the business impact as today's workplace demands we quantify customer satisfaction, impact on business, and return on investment.

By taking the Kirkpatrick model up a notch, our program evaluations have responded to market demand. We've been successful at using this method of evaluation which provides the answer to the question —"Was it worth it?" For PDP to continue to be a leader in this crucial area, our success within our own market remains dependent upon our success with these efforts.

Continued from page 1

conducting evaluation of professional development is our definition of the outcome. The sole use of participant satisfaction surveys and seat/hour counts as a means of documenting impact must change. Professional development is a complex social learning process and as such is dynamic in nature. Past evaluations have clearly indicated the need for variations in documenting outcomes by discipline, complexity of content, background of learner, perceived status of the trainer, etc. In addition, we now know that the methods and tools of transfer will also impact outcomes. Professional development encompasses several models of transferring or improving knowledge, including: direct training; observation/assessment; involvement in a developmental/improvement process; study groups; involvement in inquiry/action research; individually guided activities; and mentoring. Each of these approaches support different intermediate outcomes, but acknowledge the same systems goals — improvement of services to consumers. Innovative use of alternative tools (e.g. video-conferencing and online discussions) also is in need of assessment. The path of evaluation for each will be different.

Paradoxically, as the above variations in professional development are offered and expanded, the need for more standardization of what is considered exemplary transfer of knowledge has consolidated. Just as models for what constitutes appropriate standards for math or science education (at both K-12 and higher ed settings) have evolved, so too have standards of what constitutes appropriate professional development. Reforms in professional development have resulted in standards for design and delivery that now include the use of research-based delivery of knowledge, context-based transfer of skills, embedded support for growth in leadership, sustainability of knowledge and its use, inclusion of all stakeholders, sound use of adult pedagogy, and results-driven goals, objectives, and assessment. Professional development standards provide us, as evaluators and decision makers, with benchmarks against which to judge the adequacy of educational offerings and a means by which programs can be compared to determine best practices.

Another urgent change needed in the design and implementation of evaluation of professional development is inclusion of longitudinal, cross-site, and meta-analytic studies. Short-term, site-specific evaluations offer good formative information for local decision makers but do not assist in documenting the long-term impact of a particular training, or the broader impact across multiple sites and offerings. Evaluation efforts need to be expanded to include multiple follow-ups that will document long-term changes in work practices, professional identity, and systems changes, as well as impact on fellow workers and consumers. If these data are collected in a replicable manner across sites, and across methods, program planners and funders will be able to improve the process of transferring knowledge to make it both more efficient and more effective.

These changes in evaluation call for a more extensive use of replicable, credible evaluation models. Kirkpatrick's model is frequently used to guide evaluation and documentation of professional development but there are other models that offer equally useful approaches. These include Tyler's objectives based approach, Stufflebeam's decision-making model, Stake's standards based reviews, and Scriven's consumer orientation. All of these models address the process of defining, developing, delivering, and assessing professional development as a programmatic effort that encompasses a series of interconnected goals, benchmarks, and assessment. All are considered replicable means of assessing the merit, worth and value of a program. Each, however, stresses a different role for the instructor, the learner, and the consumer in determining the process value. As evaluators of professional development we also should use standards for documenting the validity and reliability of the evaluation process. It is important that our users perceive both the process and evaluation product as credible and relevant.

Improving Practice and Looking Ahead

Obviously, there also are other things we should consider when conducting evaluation of professional development. These include the cost of evaluation, who evaluates, who gets the results and in what form, and how to determine when "enough is enough." All of these will impact the process of evaluation, the product, and the continued transfer of knowledge to those in the workplace. We should not forget, however, that evaluation is, in itself, a form of professional development. As we jointly and individually examine our training programs, workshops and educational efforts, we are also examining and restructuring our own knowledge base about professional development, the content being offered, the methods used, and what we consider to be acceptable outcomes. Reflection on the practice of transferring and gaining knowledge is a key characteristic of professional growth and should be encouraged as part of all our efforts. Improving what we do, whether it is assisting in transferring knowledge, delivering services, or evaluating how we do it, will assist us in developing better practices. Evaluation of professional development begins and ends with a desire for improvements that will enhance our ability to serve others.

Applying Kirkpatrick's Four-Level Model of Training Evaluation

Andrew Sanchirico, Ph.D., PDP Senior Research Scientist

The Kirkpatrick model of training evaluation is the established evaluation design used by the Professional Development Program (PDP). This model specifies four levels of training evaluation: (1) Reaction, (2) Learning, (3) Behavior, and (4) Results. The four levels are sequential; while the evaluation of a training program may include one or more of these levels, no level should be bypassed to reach a higher level (although exceptions may sometimes be made). This article describes how this model is applied to training evaluation and discusses some of the issues involved in its implementation.

Level 1: Reaction

This level of evaluation measures how the participants reacted to the training — i.e., what they thought about the quality of the training program. This is the most basic level of evaluation; it measures participant satisfaction generally by a questionnaire distributed to the participants immediately after the completion of training.

Level 2: Learning

The evaluation of learning measures the extent to which participants acquired knowledge or skills as a result of the training program. A pre-post test is the most common type of instrument used to assess learning. The information obtained at this level helps determine if the trainees achieved the expected level of learning and can help to identify training areas in need of revision.

Level 3: Behavior

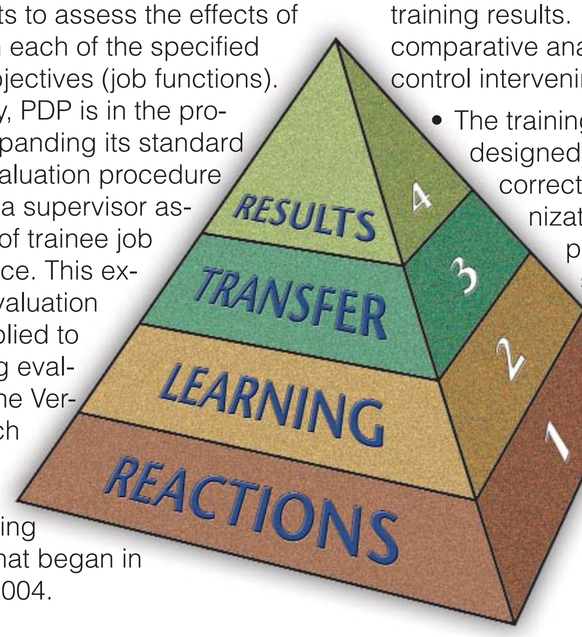
The evaluation of behavior examines the extent to which improvement in job performance occurred because the participant attended the training program. This level of evaluation is designed to measure transfer of learning to the workplace. Unlike the previous two levels, the assessment of behavior cannot usually be measured at the training site. In addition, there may be some training programs for which a Level 3 evaluation is not appropriate. If a Level 3 evaluation should be applied to a specific training course, the following criteria should be met:

- The training course should be specifically designed to improve trainee job performance on his or her *current job*. Skills that are learned at training, but are not regularly performed at the work site, cannot be objectively measured on the job. Conversely, skills that are required on the job, but that are not included in the training, cannot be used as legitimate measures of learning transfer.
- The training course should be evaluated at Levels 2 and 1.

For Level 3, several evaluation methods can be used: observation of job performance, pre-post measures of job performance, supervisor assessment of trainee job performance, and trainee self-assessment of job performance.

PDP typically uses a trainee self-assessment of job performance to measure transfer of learning. The survey is sent to the trainees at their work sites approximately 45 days after they complete training. It lists training objectives, which are keyed to the trainees' job functions, and asks the participants to assess the effects of training on each of the specified training objectives (job functions).

Currently, PDP is in the process of expanding its standard Level 3 evaluation procedure to include a supervisor assessment of trainee job performance. This expanded evaluation will be applied to the training evaluation of the Vermont Reach Up Case Management Training Program that began in October, 2004.



Level 4: Results

The fourth level of evaluation assesses the organizational outcomes that occur because of the training program. For example, if a training program was designed to train job counselors on how to help their clients find jobs, the outcome measure would be the rate of client job placement. As such, the evaluation of training results represents a major analytical leap in training evaluation. Unlike the previous three levels of evaluation, the fourth level shifts the unit of analysis to the trainees' place of employment. This shift involves the following criteria:

- Access to organizational data is required.
- As many factors besides training may influence organizational results, there is a need to identify and control intervening variables. For example, for a training program designed to improve client job acquisition, intervening variables would include the local economy and the labor supply.
- A comparative analysis is required if the training effects are to be accurately calculated. The comparison may be between groups (trainees vs. non-trainees) or it may be a comparison of pre-training vs. post-training results. If properly done, the comparative analysis can serve to control intervening variables.
- The training course should be designed to address and correct clearly stated organizational objectives or problems; and measures of organizational results should be specified prior to embarking upon a Level 4 evaluation.

Continued on page 6

PDP Creates New e-Learning Module for CPHP

Last summer the Center for Public Health Preparedness (CPHP) at UAlbany School of Public Health announced the public release of the online course "Terrorism, Preparedness, and Public Health: An Introduction." The Instructional Technologies Unit (ITU) of the Professional Development Program created the course with the Center's Director, Dr. Robert G. Westphal. The course provides an introduction to key concepts related to public health emergency preparedness and focuses on potential terrorist threats, their possible scope and impact, and possible responses to emergency events.

The School of Public Health used an e-mail distribution service to advertise the course to state and local health departments and public health professionals. Although expectations for this course ran high, no one anticipated that over 500 individuals would register in the first week, including: emergency medical technicians; police, fire, and emergency room personnel; and community partners in addition to the public health workforce. To date, over 6,000 people have taken the free course.

Based on this success, PDP is currently working on a new course that examines the risk of flu pandemic from a public health perspective.

For more information, visit the UAlbany Center for Public Health Preparedness web site at http://www.ualbanycphp.org/.

Did You Know?

- There are more than 50 million Chinese e-learning users online and some analysts claim that the number is doubling every 6 months.
  — Training, December 2004
- 77 percent of organizations use reaction measures; 36 percent use learning evaluations; 15 percent measure behavior change; and 8 percent measure results. Organizations that are maximizing the potential of money invested in learning processes are asking about measurable outcomes.
  — Learning Resources Network, January 2005
- The cost of an impact or ROI study is typically 5 to 10 percent of the total cost of the project — Make Training Evaluations Work, ASTD Press, November 2004
- Surveys show a strong movement toward more formal measurement of training investments. Because the greatest demand for analytics is above Level 3, organizations should strive to create the right processes and use the most efficient and effective technologies to make Level 3, 4, and ROI analyses practical. — The State of Learning Analytics, ASTD 2004
- As this generation of well-educated and creative minds prepares to take on more management positions, industry leaders are worried about decades of institutional knowledge leaving the workforce as 76 million "Baby Boomers" move into retirement. — Detroit Free Press, December 2004
- In a survey of 35 members of the International Federation of Training and Development Organizations, measuring return on investment was consistently rated the hottest topic among members of those organizations. — IFTDO, 2003
- Quantifying the positive economic value of training is necessary in today's economic climate. When managers are not able to measure the effectiveness of training, it's only to be expected that they will concentrate on those aspects of their job for which they are held accountable. This encourages managers to reduce or eliminate training expenses and sacrifice long-term profit gains — potentially exponential gains — in favor of short-term cost-cutting. — Human Resource Management Systems, 2005

- The training course should be evaluated at Levels 3, 2, and 1.

There are no standard instruments that are used for a fourth level evaluation. Instead, as explained above, what is needed is a well developed evaluation design and access to organizational data.

PDP is in the early stages of developing a Level 4 evaluation for a program that will train supervisors to use an online system for reviewing Temporary Assistance cases in New York City. One intended outcome of this training program is to reduce the error rates at HRA job centers.

In conclusion, the Kirkpatrick Model provides a logical and intuitive design for planning and implementing a training evaluation. If all four levels are fully implemented, virtually every aspect of the training program can be evaluated. Realistically, a full implementation of the model is difficult to achieve. However, the strength of the model is that it simplifies and systematizes a complicated set of procedures; and adherence to the Kirkpatrick Model increases the probability that an evaluation will produce meaningful information about a training program.

PDP Opens NYC Office to Enhance Services for HRA and NYS OTDA

PDP has established a New York City office to anchor our working partnership with the NYC Human Resources Administration and the NYS Office of Temporary and Disability Assistance, Bureau of Training and Management Analysis. Located at the corner of Martin Luther King Boulevard and 125th Street, the location houses five staff as well as meeting and training space.

PDP Communiqué is produced by the Professional Development Program, Nelson A. Rockefeller College of Public Affairs and Policy, University at Albany, State University of New York

Eugene J. Monaco
Director

Kary Jablonka
Editor

Christine Moore
Associate Editor

George Dowse
Graphic Designer

PDP

UNIVERSITY AT ALBANY
State University of New York

Professional Development Program
Rockefeller College
Richardson Hall 301
135 Western Avenue
Albany, New York 12222

(518) 442-5700; fax: 442-5768
www.pdp.albany.edu