

# Advances in Developing Human Resources

<http://adh.sagepub.com/>

---

## Training Evaluation: Knowing More Than Is Practiced

Greg G. Wang and Diane Wilcox

*Advances in Developing Human Resources* 2006 8: 528

DOI: 10.1177/1523422306293007

The online version of this article can be found at:

<http://adh.sagepub.com/content/8/4/528>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



[Academy of Human Resource Development](#)

**Additional services and information for *Advances in Developing Human Resources* can be found at:**

**Email Alerts:** <http://adh.sagepub.com/cgi/alerts>

**Subscriptions:** <http://adh.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://adh.sagepub.com/content/8/4/528.refs.html>

>> [Version of Record](#) - Oct 12, 2006

Downloaded from [adh.sagepub.com](http://adh.sagepub.com) at SAGE Publications on December 3, 2014

What is This?

# Training Evaluation: Knowing More Than Is Practiced

Greg G. Wang

Diane Wilcox

**The problem and the solution.** Training program evaluation is an important and culminating phase in the analysis, design, develop, implement, evaluate (ADDIE) process. However, evaluation has often been overlooked or not implemented to its full capacity. To assess and ensure the quality, effectiveness, and the impact of systematic training, this article emphasizes the importance of summative evaluation at the last phase of ADDIE and presents developments toward a summative evaluation framework of training program effectiveness. The focus is the connection of final summative evaluation to the direction provided by the analysis phase and the concerns of the host organization.

**Keywords:** *formative evaluation; summative evaluation; ISD; outcome evaluation; impact evaluation*

As a systematic process for developing needed workplace knowledge and expertise, instructional systems design requires an evaluation component to determine if the training program achieved its intended goal—if it did what it purported to do. However, evaluation, the last phase of the ADDIE (analysis, design, develop, implement, evaluate) model, is often overlooked when organizations create and implement training programs. Strictly speaking, the larger view of evaluation may not be treated as a separate phase during the process. It is indeed an ongoing effort throughout all phases of the ADDIE process (Hannum & Hansen, 1989) and culminating at the last phase.

A number of reasons have been noted for organizations failing to conduct systematic evaluations. First, many training professionals either do not believe in evaluation or do not possess the mind-set necessary to conduct evaluation (Swanson, 2005). Others do not wish to evaluate their training programs because of the lack of confidence in whether their programs add value to, or have impact on, organizations (Spitzer, 1999). Lack of evaluation in training was also attributed to the lack of resources and expertise, as well as lack of an organization culture that supports such efforts (Desimone, Werner, & Harris,

Advances in Developing Human Resources Vol. 8, No. 4 November 2006 528-539

DOI: 10.1177/1523422306293007

Copyright 2006 Sage Publications

2002; Moller, Benscoter, & Rohrer-Murphy, 2000). Even for limited efforts in training evaluation, most are retrospective in nature (Brown & Gerhardt, 2002; Wang & Wang, 2005). A study of a group of instructional design practitioners indicated that 89.5% of those conduct end-of-course evaluation, 71% evaluate learning; however, only 44% use acceptable techniques for measuring achievement. Yet merely 20% of those surveyed correctly identified methods for results evaluation (Moller & Mallin, 1996). Brown and Gerhardt (2002) concluded that companies expend even less effort in evaluating the instructional design process.

The purpose of this article is to examine program evaluation throughout the ADDIE process for systematic training design. In particular, we focus on the importance of summative or outcome evaluation at the evaluation phase. By presenting developments of summative evaluation, we argue that result-oriented evaluation is critical to ensure the quality and desired outcomes of systematic training.

## Formative and Summative Evaluation

In a broad sense, training program evaluation can be divided into two categories, formative evaluation and summative evaluation (Noe, 2002; Scriven, 1996). An evaluation intended to provide information on improving program design and development is called *formative evaluation* (Scriven, 1991). Specifically, the purpose of formative evaluation is to identify weakness in instructional material, methods, or learning objectives with the intention to develop prescriptive solutions during training program design and development (Brown & Gerhardt, 2002). A further purpose of formative evaluation is to help form and shape the program quality to perform better and should be built into each of the ADDIE phases. It should be an ongoing, integrated effort throughout all phases of the ADDIE process (Hannum & Hansen, 1989). Therefore, formative evaluation should be imbedded in the entire systematic training process.

In contrast, an evaluation conducted to determine whether intended training goals and outcomes are achieved is called a *summative evaluation* (Scriven, 1991). Summative evaluation is conducted after a training program has been delivered. It is the focus of this discussion of the fifth phase in the ADDIE model. A major purpose of summative evaluation is to render a summary judgment or conclusion through measurement and assessment of program outcomes.

## Summative Evaluation

Summative evaluation, the focus of the evaluation phase of systematic training, is centered on training outcomes (Alvarez, Salas, & Garofano, 2004). It seeks to identify the benefits of training to individuals and organizations in the form of learning and enhanced on-the-job performance that were specified

in the analysis phase. There are several major purposes for organizations to conduct summative evaluation after systematic training. First, summative evaluation connects all the ADDIE phases in systematic training with organizational goals and objectives. It will not only justify the training budget and human resource development (HRD) investment but also validate implemented interventions. More importantly, it demonstrates to the organization decision makers the value of training interventions. Second, systematic summative evaluation may discover the areas of training interventions that do not meet the stakeholders' expectations. Such evaluation will certainly provide opportunities for future improvement. Last, but not least, summative evaluation may assist and support future training and HRD investment. In today's competitive world, training and HRD are frequently competing with all other functions for organizational resources. Sound summative evaluation of systematic training demonstrates the accountability of training and HRD functions and supports decision making regarding future training investments.

Training and HRD professionals have been attempting to understand summative evaluation since late 1950s. Various classification systems have been proposed to specify the functions and purposes of evaluation in systematic training intervention. Kirkpatrick's (1998) four-level evaluation created in 1959 was the first classification schema or taxonomy specifically for outcome evaluation or summative evaluation as noted by a number of studies (Alliger & Janak, 1989; Holton, 1996; Wang, Dou, & Li, 2002; Wang & Wang, 2005). The function of the four-level evaluation was further identified as a communication tool, instead of being claimed as evaluation techniques or steps, for training evaluation practice (Wang et al., 2002). Based on the learning domain, training evaluation may be classified into three types: (a) cognitive: evaluating knowledge and cognitive strategies; (b) skill-based: evaluating constructs such as *automaticity* and *compilation*; and (c) affective: evaluating constructs such as *attitudes* and *motivation* (Kraiger, Ford, & Salas, 1993). Swanson and Holton (1999) also presented a comprehensive results assessment system focused on three domains: performance (system and financial), learning (knowledge and expertise), and perceptions (participant and stakeholder). Training evaluation can also be classified according to time frames involved, such as short-term or long-term impact evaluation.

Practically, the purpose of all classification systems in evaluation is to help conceptualize and understand the nature, functions, or purposes of evaluation from different perspectives for facilitating data collection and analysis aspects of program evaluation. It is a matter of preference, familiarity, or convention of practices to various classification regarding data collection and analysis. In other words, such classification schemes will not be able to provide analytical tools or techniques for evaluation, except for communicating purpose or focus of evaluations. It is the training and HRD professionals' task to determine what data collection techniques to use and what analytical tools to choose for the actual evaluation efforts.

## Challenge and Opportunity

It is expected that engaging in formative evaluation within the training phases of ADDIE and taking action on the findings will improve the quality of the training intervention. However, a sound formative evaluation result may not necessarily be able to guarantee positive summative evaluation results, particularly when summative evaluation conducted on learners returning to their performance setting. Almost always, the application of learned knowledge and expertise are intertwined with other organizational factors, such as organization support and the application environment (Holton, 2005; Wang et al., 2002).

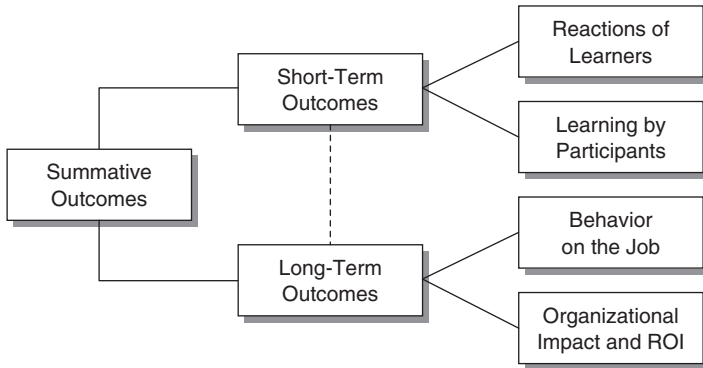
Summative evaluation has been experiencing difficulties and challenges in research and practice areas of training (Wang & Wang, 2005). Most recent data from American Society for Training and Development (ASTD) showed that only 12.9% of the largest organizations have conducted any kind of training impact evaluation (Sugrue & Rivera, 2005), the most important form of summative evaluation. This reveals an interesting and alarming question: If impact (summative) evaluation, including return-on-investment (ROI) measurement, is so important, why we still see so few organizations actually conduct such evaluation in training reality?

Wang and Spitzer (2005) divided the evaluation evolution in training and HRD into three stages. The first stage is practice-oriented atheoretical stage as represented by the Kirkpatrick four-level scheme, ranging from the late 1950s to the late 1980s. The second stage is a *process-driven operational stage*, represented by the ROI wave (e.g., Phillips, 2003), spanning from the late 1980s to the early 2000s. We are now at the beginning of the third stage, the research-oriented comprehensive stage. Recent years have witnessed a burgeoning of evaluation methods developed by scholars and practitioners (e.g., Wang & Spitzer, 2005), which offer training and HRD professionals more alternative approaches and opportunities to conduct outcome evaluation.

## Outcome Evaluation

Following the classification tradition initiated by Kirkpatrick (1998), this section discusses a new perspective of classification on outcome evaluation. In addition to demonstrating how current and previous evaluation taxonomy, theories, and approaches may fit in the framework, we argue that connecting outcome evaluation to the direction defined by the analysis phase and the concerns of the host organization should be the focus of the evaluation.

Summative evaluation at the systematic training evaluation phase can be further detailed and classified into short-term and long-term outcome evaluation. This classification is presented in Figure 1. In most situations, organizations need to identify immediate training outcomes after the program implementation. Based on the four-level evaluation scheme, short-term outcome may include participant reactions and reported or measured learning outcomes.



**FIGURE 1: Framework for Summative Evaluation Phase Within Systematic Training**

*Note:* ROI = return on investment.

### Short-Term Evaluation

As a short-term evaluation, measuring participants' perceptions and reactions on training occurs during and toward the end of the implement phase. An implicit assumption under perceptual evaluation is that the participants are clear about their learning needs, and the needs are consistent with those identified during analysis phase. Under this assumption, if a training program fails to satisfy the learning needs, a task for the evaluation is to identify whether it is the responsibility of the program design or delivery. However, results from a perceptual evaluation do not necessarily indicate what new skills and knowledge the learners have acquired because the above assumption is not often held true in training reality. Therefore, a more realistic way in reaction evaluation is to obtain learners' feedback on the interest, attention, and motivation to the learning subject. These areas are critical to the success of a training program. Intuitively, participants should not have negative feelings and reactions to training experiences.

Evaluation of learner's reactions as a short-term training outcome is often conducted with an attitudinal questionnaire survey with commonly used Likert-type rating scales. Normally, the survey should cover questions in the following areas:

- learning objectives, content and design
- instructional approaches
- learning environment and interactions.

For technology-based training, questionnaires should also include a component on learner's perceptions on technology usability and navigation. Regardless

of its rationale, evaluation on reaction is the first outcome feedback that training professionals receive regarding a specific program.

Evaluation of learning outcomes is the other form of short-term evaluation. The purpose of learning evaluation is to measure knowledge and skill enhancement by the learners due to the training program participation. Consequently, such evaluation is generally based on learning assessment and measurement and is conducted immediately after training implementation.

Traditionally, learning evaluation has been interpreted as posttest of learning (Phillips, 1997). However, testing should not be considered the only means of evaluating learning, especially in organizational settings where improving performance is the ultimate goal of learning. Depending on the domain of the subject covered by a training program, learning outcome evaluation in a training context frequently takes on formats other than knowledge testing. Assessing expertise requires the learner to do things such as demonstration, presentation, and hands-on projects. No matter what format one chooses to assess learning, it should be framed by the analysis phase embedded in the training program design process. Based on performance requirements identified during the analysis phase and learning objectives in the design phase, trainers should create measurement instruments that fit into the learning context. This fit is the core basis for evaluation of content validity.

It is critical to note that the two forms of short-term evaluation—reaction and learning assessment as Levels 1 and 2 defined by Kirkpatrick (1998) may not have any causal relationship in term of evaluation results. In other words, good results from a reaction evaluation do not necessarily warrant satisfactory learning outcomes (Alliger & Janak, 1989). However, results obtained from reactional and learning evaluation can be used to improve and enhance the training program in subsequent systematic training design.

### **Long-Term Evaluation**

In organization settings, the ultimate goal for design and conducting training is to improve individual and organizational performance. The nature of application-driven instructional design also determines the importance of long-term evaluation. In today's organizational reality, the "smile sheet" can no longer represent an acceptable evaluation of training effectiveness (Moller et al., 2000). Therefore, long-term evaluation of the impact of training programs has received increasing attention in organizations.

### **Evaluation of Learning Transfer**

It is important to evaluate learning transfer and behavior change because one of the primary purposes of systematic training is to improve individual on-the-job performance after the training. Such improvement can only be

identified when the participants apply newly learned skills and knowledge on the performance settings. This type of evaluation is also referred to as a Level 3 evaluation according to Kirkpatrick's classification. Because learners may not be confronted with opportunities to demonstrate their learned knowledge and skills immediately after a training program, evaluation of this aspect needs to allow sufficient time for such opportunities to occur in the performance settings. In practice, evaluation of knowledge and skills transfer requires a 3 to 6 months time period after a training program.

*Transfer of training* is defined as the degree to which learners apply the knowledge, skills, and attitudes gained in training to their jobs (Wexley & Latham, 1991). Transfer of training may be measured by the maintenance of the skills, knowledge, and attitudes during a certain period of time (Baldwin & Ford, 1988). Rouiller and Goldstein (1993) expanded the research on transfer of training to include the concept of a *transfer climate* that either inhibits or helps facilitate the transfer of learning into a job performance setting. Accurately measuring transfer climate is important because it can help HRD move beyond the question of whether training works, to analyzing why training works. Therefore, obtaining valid and reliable measures of transfer climate will help identify not only when an organization is ready for a training intervention but also when individuals, groups, and organizations are ready for such an intervention (Holton, Bates, Seyler, & Carvalho, 1997).

Holton, Bates, and Ruona (2000) further extended transfer climate to a learning transfer system. This system is defined as all factors in the person, training, and organization that influence transfer of learning to job performance. That study argued that transfer climate is only one subset of factors that influences learning transfer. Other influences on transfer may include training design, personal characteristics, opportunity to use training, and motivational influences. Transfer can only be completely understood and influenced by examining the entire system of influences (Holton, 2005). It is also a requirement for learning transfer to take place on the job. Thus, to effectively measure the learning transfer climate and potential subsequent behavioral change, Holton and colleagues (Holton et al., 1997; Holton et al., 2000) posited and validated the following variables in a proposed learning transfer system inventory (LTSI): learner readiness, motivation to transfer, positive and/or negative personal outcomes, personal capacity for transfer, peer and supervisor support, supervisor sanctions, transfer design, and opportunity to use.

LTSI as a latest development in measuring training outcomes takes into consideration of all influencing organization variables and constructs bearing on the transfer of learning to workplace behaviors. It clearly points to a future direction for enhancing learning transfer and behavior change after a training program at a longer term. Without considering the variables as defined in the above LTSI, it is difficult to effectively enhance the transfer process and close the performance gap identified in the analysis phase of systematic training.



Frequently, direct observations and assessment of actual workplace behaviors is also required to confirm actual transfer.

### **Evaluation of Organizational Impact**

As an important form of long-term evaluation, measuring the impact of training on organizational results poses unique problems. Traditional ADDIE models start with the assumption that training is needed and moves on to systematically create and deliver that requested training. In almost all cases, the core analysis as to what the organizational problem or goal was, and what it would truly take to make the desired gains, is outside the traditional ADDIE process. In this common scenario, Wang and Wang (2005) observed that the complexity and difficulty is caused by the fact that such evaluation is to analyze the impact of a training subsystem on an organizational overall system. Given the multiple subsystems coexisting and intertwining with each other in any organizational system, many view measurement of training programs' organizational results or ROI of training a thorny task, if not impossible. However, to demonstrate the contribution of systematic training to overall organizational performance, it has become a requirement to conduct impact evaluation in many organizations.

Occasionally, ROI formula may be applied to the evaluation settings as recommended by Phillips (1997). The formula is expressed as  $ROI = \text{Net Benefit} / \text{Total Cost} \times 100\%$ . The application of the ROI calculation, however, is subject to one restriction: A valid and reliable measure of the training program's net benefit ( $\text{Net Program Benefit} = \text{Total Program Benefit} - \text{Total Program Cost}$ ) is readily available. Yet in training reality, the net program benefit is often entangled with other organizational system variables and difficult to separate, although the term of total program cost may be easily obtainable. In fact, if one can calculate the net benefit for a training program, it may become unnecessary to determine the ROI because the net benefit is the organizational impact of training, or the contribution that a training program makes for the organization.

Control groups were recommended by Wang (2002) for identifying the net benefit of a training program. Wang (2002) defined four types of control groups based on experimental design methodology. While Type I control group serves as a benchmark to gauge the validity and reliability of the measurement, Type II control group can be used for two or more groups' comparison on training program benefit. Type III control group is a time-series measurement, with one physical group, that treats the training group as the control of its own. Meanwhile, Type IV control group setting is a combination of Type II and Type III which can be used for more complex measurement scenarios for obtaining the benefit information generated by training programs.

When more sophisticated ADDIE processes are applied over traditional ADDIE models, the up-front analyst takes off his or her training hat and puts on a performance improvement hat (Rummler & Brache, 1995; Swanson,

1996). When this happens, the analysis will focus on mission-related outputs of the organization. Performance goals related to outputs of the goods or services produced by the organization will almost always require an intervention that goes beyond training so as to consider all the elements required to assure performance improvement. After performance shortcomings are identified, and complete interventions are assessed to be at the root of improvement, the resulting gains in units of work productivity, and their conversion to financial benefits, are quite easy (Swanson, 2001).

### **Nonfinancial Alternatives**

At the beginning of the new measurement and evaluation stage, the research-oriented comprehensive stage as defined by Wang and Spitzer (2005), recent years have witnessed more summative evaluation approaches focusing on measuring organizational impact of training interventions. Several studies advocate systems thinking for conducting the impact measurement in organizations (e.g., Russ-Eft & Preskill, 2005; Wang & Wang, 2005). Such analyses consider the complexity of evaluation within an organizational systems context, and indicate that in training impact analysis one needs to consider what impact other organizational system variables may have on the outcome measure to effectively answer the organization's evaluative questions. Through case studies, Russ-Eft and Preskill (2005) demonstrated the application of a systems approach to measuring organizational impact of training programs qualitatively. That study showed that determining ROI is a multifaceted and complicated task within a complex system. Many of the measurement requests for ROI tend to be "knee-jerk" reactions, based on a lack of understanding and misconceptions about evaluation. In a similar systems framework, Wang and Wang (2005) derived a quantitative approach to measuring quantifiable organizational outcomes generated by the training function while considering all other factors' contributions to the organization.

In a similar vein, Brinkerhoff (2003) further argued that evaluation of training is a whole organization strategy. Training should not be the object of evaluation. Instead, what is needed is evaluation of how well organizations use training. This requires focusing evaluation inquiry on the larger process of training as it is integrated with performance management and includes those factors and actions that determine whether training can create performance results. By proposing a success case method, Brinkerhoff demonstrated that effective training impact should involve all relevant stakeholders of a training program. Likewise, Nickols (2005) suggested a stakeholder-based approach for measuring organizational impact of training. This approach requires training and HRD professionals to incorporate stakeholder requirements into the design, development, and delivery of training, increasing stakeholder interest in the outcomes and in evaluating those outcomes in ways that offer meaning

and value to all the stakeholders. Other latest methodological development of organizational impact for training programs also includes learning effectiveness measurement approach that focuses on the relationship between training and business causal chain (Spitzer, 2005), and critical outcome technique that concentrates on result-oriented program evaluation (Mattson, 2005).

No matter what methods are used in obtaining the information on training benefit regarding organizational impact, key metrics in three areas are almost always appealing to decision makers and may be effectively measured for training's impact. They are time, quality, and cost, as directly related to business principle of "faster, better, and cheaper." All the three areas of metrics are productivity driven. Although other more directive metrics, such as sales revenue or profitability, may be more attractive, given the multiple factors involved and the difficulties in separating training program effects, we do not recommend directly measuring such impact. In other words, if a training program, through a systematic ADDIE process, can generate results that save an organization time in delivering their product and services, or improve product or service quality, or reduce the cost, it will certainly receive more organizational support and investment. In measurement reality, the three areas of key metrics should be further specified at different levels according to particular measurement situations and requirements. For instance, time-related metrics may include project cycle time, facility downtime, time to market, and time to delivery; quality-related metrics may include defect rate, unscheduled maintenance, and customer and/or stakeholder satisfaction rate; and cost-related metrics may include cost reduction in any specific area. In short, measuring organizational impact of training is to identify specific metrics that an organization values and translate the training impact into quantifiable organizational outcomes with credible methods.

## Concluding Remark

When reviewing the theory and practice within the evaluation realm of systematic training, the training profession should know more than it is practicing. With increasingly available approaches and models for conducting training outcome and impact evaluation, a core challenge to the profession is to know the right approaches and techniques for specific evaluation needs and the concerns of the host organization. Furthermore, having measurement skills related to creating valid and reliable instruments is also critical to produce credible evaluation results.

## References

- Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of criteria: Thirty years later. *Personnel Psychology, 42*, 331-341.
- Alvarez, K., Salas, E., & Garofano, C. M. (2004). An integrated model of training evaluation and effectiveness. *Human Resource Development Review, 3*(4), 385-416.

- Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology, 41*, 63-105.
- Brinkerhoff, R. O. (2003). *The success case method*. San Francisco: Berrett-Koehler.
- Brown, K. G., & Gerhardt, M. W. (2002). Formative evaluation: An integrative practice model and case study. *Personnel Psychology, 55*, 951-983.
- Desimone, R. L., Werner, J. M., & Harris, D. M. (2002). *Human resource development*. Cincinnati, OH: South-Western.
- Hannum, W., & Hansen, C. (1989). *Instructional systems development in large organizations*. Englewood Cliffs, NJ: Educational Technology Publications.
- Holton, E. F., III. (1996). The flawed four level evaluation model. *Human Resource Development Quarterly, 7*, 5-21.
- Holton, E. F., III. (2005). Holton's evaluation model: New evidence and construct elaborations. *Advances in Developing Human Resources, 7*(1), 37-54.
- Holton, E. F., III, Bates, R. A., & Ruona, W. E. A. (2000). Development of a generalized learning transfer system inventory. *Human Resource Development Quarterly, 11*(4), 333-360.
- Holton, E. F., III, Bates, R., Seyler, D., & Carvalho, M. (1997). Toward a construct validation of a transfer climate instrument. *Human Resource Development Quarterly, 8*(2), 95-113.
- Kirkpatrick, D. (1998). *Evaluating training programs: The four levels* (2nd ed.). San Francisco: Berrett-Koehler.
- Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78*, 311-328.
- Mattson, B. W. (2005). Using the critical outcome techniques to demonstrate financial and organizational performance results. *Advances in Developing Human Resources, 7*(1), 102-120.
- Moller, L., Bencotter, B., & Rohrer-Murphy, L. (2000). Utilizing performance technology to improve evaluative practices of instructional designers. *Performance Improvement Quarterly, 13*(1), 84-95.
- Moller, L., & Mallin, P. (1996). Evaluation practices of instructional designers and organizational support or barriers. *Performance Improvement Quarterly, 9*(4), 82-92.
- Nickols, F. W. (2005). Why a stakeholder approach to evaluating training? *Advances in Developing Human Resources, 7*(1), 121-134.
- Noe, R. (2002). *Employee training and development* (2nd ed.). New York: McGraw-Hill.
- Phillips, J. (1997). *Handbook of training evaluation and measurement methods* (3rd ed.). Houston, TX: Gulf.
- Phillips, J. (2003). *Return on investment in training and performance improvement programs* (2nd ed.). Boston: Butterworth-Heinemann.
- Rouiller, J. Z., & Goldstein, I. L. (1993). The relationship between organizational transfer climate and positive transfer of training. *Human Resource Development Quarterly, 4*(4), 95-113.
- Rummler, G. A., & Brache, A. P. (1995). *Improving performance: How to manage the white space on the organization chart* (2nd ed.). San Francisco: Jossey-Bass.
- Russ-Eft, D., & Preskill, H. (2005). In search of the holy grail: Return on investment evaluation in human resource development. *Advances in Developing Human Resources, 7*(1), 71-85.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage.

- Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17(2), 151-161.
- Spitzer, D. R. (1999). Embracing evaluation. *Training*, 36(6), 42-47.
- Spitzer, D. R. (2005). Learning effectiveness: A new approach for measuring and managing learning to achieve business results. *Advances in Developing Human Resources*, 7(1), 55-70.
- Sugrue, B., & Rivera, R. J. (2005). *2005 state of the industry: ASTD annual review of trends in workplace learning and performance*. Alexandria, VA: American Society for Training and Development.
- Swanson, R. A. (1996). *Analysis for improving performance: Tools for diagnosing organizations and documenting workplace expertise*. San Francisco: Berrett-Koehler.
- Swanson, R. A. (2001). *Assessing the financial benefits of human resource development*. Cambridge, MA: Perseus.
- Swanson, R. A. (2005). Evaluation, a state of mind. *Advances in Developing Human Resources*, 7(1), 16-21.
- Swanson, R. A., & Holton, E. F. (1999). *Results: How to assess performance, learning, and perceptions in organizations*. San Francisco: Berrett-Koehler.
- Wang, G. (2002). Control groups for human performance technology (HPT) evaluation and measurement. *Performance Improvement Quarterly*, 15(2), 34-48.
- Wang, G., Dou, Z., & Li, N. (2002). A systems approach to measuring return on investment for HRD programs. *Human Resource Development Quarterly*, 13(2), 203-224.
- Wang, G. G., & Spitzer, D. R. (2005). HRD measurement & evaluation: Looking back and moving forward. *Advances in Developing Human Resources*, 7(1), 5-15.
- Wang, G. G., & Wang, J. (2005). HRD evaluation: Emerging market, barriers, and theory building. *Advances in Developing Human Resources*, 7(1), 22-36.
- Wexley, K. R., & Latham, G. P. (1991). *Developing and training human resources in organizations*. New York: HarperCollins.

**Greg G. Wang** is associate professor of human resource development (HRD) at Old Dominion University. His research focus has been on economic foundation of HRD, evaluation and measurement, e-learning interventions, and international HRD. Formerly with GE, Motorola, and Pennsylvania Power & Light, he is also actively involved in HRD practice. He has consulted with IBM, GE, Hewlett Packard, and other major corporations. He is the founder and moderator of the Internet forum, ROI<sub>net</sub>, with more than 2,300 researchers and practitioners from more than 40 countries.

**Diane Wilcox** is an assistant professor of human resource development (HRD) at James Madison University where she teaches HRD, instructional technology, and adult learning. Her research interests include the use of innovative technologies in training, performance assessment, and evaluation. She has 15 years of HRD practitioner experience and earned her PhD at the University of North Carolina at Chapel Hill in educational psychology.

- Wang, G. G., & Wilcox, D. (2006). Training evaluation: Knowing more than is practiced. *Advances in Developing Human Resources*, 8(4), 528-539.