

# **Evaluation of Training**

## **By Harold D. Stolovitch and Erica J. Keeps**

True learning and performance professionals focus on ensuring that any training they produce is developed according to principles that science and best practice have demonstrated to be effective within the workplace setting. As importantly, they also apply rigorous evaluation processes and standards, taking into account what is feasible to verify that their training is sound and can demonstrate quality and effectiveness. What exactly does "*rigorous evaluation processes*" mean?

### **Some Basic Evaluation Concepts**

Evaluation is the verification of the value of something. It is a process that produces a judgment – a determination – of value based on fixed or generally accepted standards. Within the workplace learning and performance context, evaluation includes the application of a valid and reliable methodology along with suitable procedures and means of measurement that permit an organization to determine whether or not individuals or groups have met standards of learning and/or performance required by the job. This means that prior to evaluation taking place, job related learning and/or performance requirements have been validly derived from state-of-the-art content and practices, requisites of the job, the conditions of the work context, regulations, standards and all other sources relevant to establishing what is required of trainees as appropriate for their work and grade level.

Within the training context, depending on desired outcomes, teams of experts and instructional designers establish specific and verifiable objectives. These are validated by other independent and qualified experts, and are matched against work, regulatory, safety and other requirements.

For the process of evaluation to produce credible, usable results, it must meet three criteria. It must be: systematic (planned and applied in a disciplined manner by qualified professionals); objective and honest (free from bias, favoritism or political influence); conducted in a manner that produces credible data which can be readily treated, analyzed and reported in a form that facilitates management decisions and actions.

### **The Purpose of Evaluation**

Proper evaluation of training programs requires money, resources and time. There must be solid reasons for conducting such an intense, resource-consuming activity. Evaluation in the learning and performance context has two major purposes:

*To prove (summative evaluation)* – In other words, to demonstrate in a believable, data-based manner how well (or not) a training program is performing in terms of both learning and on-job performance. The recipients of this proof are: management; those responsible for training; departmental supervisors; governmental and regulatory

agencies; the trainees themselves; all others who have a stake in the success of the training.

*To improve (formative evaluation)* – To gather information about how well training programs are performing, based on demonstrated trainee learning and performance (from collected data), what areas are working well – producing desired results – and which ones require improvements/revisions. The recipients of the evaluation results are those responsible for the management, design, development, delivery and implementation of the training.

The overall purpose of evaluation efforts and investment, therefore, is to produce and deliver the best training and performance support resources possible, resulting in professionally prepared individuals and teams who perform in the most effective, efficient and safe manner possible. It is also to provide information to all stakeholders on effectiveness or needed changes.

## **Levels of Evaluation**

Throughout the world of work, from business and industry, the military, transportation, medicine and others, there are numerous, serious attempts to conduct meaningful evaluations of training and other performance improvement interventions. An accepted vocabulary has emerged based on the writings and practices of Donald Kirkpatrick, Jack Phillips as well as other evaluation specialists. Of particular importance has been the acceptance of what are called, The Levels of Evaluation. While there is much academic discussion about these levels, there is widespread consensus that attention to each of them is extremely worthwhile within the workplace learning and performance setting. What follows draws from various writings about Kirkpatrick's four levels of evaluation and an additional level from Phillips.

## **Kirkpatrick's Levels of Evaluation**

Level 1: **Reaction** - What participants thought and felt about as well as how they judged the value of the training (satisfaction; sometimes derogatorily called "smile sheets"). Most work organizations conduct Level 1 evaluations for almost all of their courses. Level 1 includes items that are not only concerned with trainee thoughts and feelings, but also about content relevance/applicability to the job, use of examples, instructional methods and other types of information concerning the quality and utility of the training. The purpose is to verify immediate reactions to the training as well as gain information on how to improve it.

Level 2: **Learning** - The resulting increase in knowledge and/or skills, and changes in attitudes from a training program. This evaluation may occur during and/or at the end of a training session, course or program in the form of either a knowledge/skill demonstration or a verbal test, if appropriate. The purpose is to verify how much immediate learning actually occurred as a

result of the training and to identify knowledge/skill portions that require revision to improve trainee results. Many organizations employ pre- and post-tests to verify knowledge/skill gain from the training. Level 2 testing requires that test items, including practical trainee demonstrations, **must** be tightly linked to the training objectives.

Level 3: **Transfer** – Demonstrated application of knowledge, skills and/or attitudes from the training setting to the job (change in job behavior due to training). This evaluation generally occurs three to six months post training while the trainee is performing the job. Evaluation usually occurs through observation. When the context makes this too difficult to operationalize due to unpredictability of events and risks to participating parties, alternate methods are employed. These generally take the form of surveys on the application of the training that was provided. Trained experts may also conduct on-site interviews with both former trainees and supervisors. In the process, the interviewers may collect examples of application. Sometimes, as part of the Level 3 process, organizations administer delayed training course post-tests with trainees, three to six months after the training event has been completed. The purpose is to verify *retention* (storage in long-term memory) of what was acquired during the training. Data from the delayed post tests are compared with pre and post training scores to verify the staying power of the training. The ideal, however, is to collect hard data, first-hand, at locations where the trainees work through direct on-job observation and examination of work results. Feasibility and cost are the critical factors for accomplishing this.

Level 4: **Results** - The final pay-off that occurred because of attendance and participation in a training program. Too few organizations verify links between its training and changes in bottom-line results (e.g. decrease in incidents, accidents, injuries and fatalities; more rapid response time; prevention/avoidance of potential problems; increased revenues; decreased time to market). Performing Level 4 evaluation has significant value for making decisions about future learning and performance investments.

Several authors, especially Phillips, have suggested the addition of a fifth level of evaluation. They call for a Return-on-Investment (ROI) level, which is essentially about comparing the fourth level of the standard Kirkpatrick model to the overall costs of training. Once again, although not as difficult to conduct as many believe, it is nevertheless infrequently done. Models and examples appear in many publications. (See Jack and Patty Phillips' books and articles or our *Beyond Training Ain't Performance Fieldbook*.)

## Conclusion

Training still remains the most frequently employed means for achieving workplace performance success. Is the training your organization delivers necessary or even

appropriate to attain the ends sought? Is it sufficient? Is the quality adequate? Is it complete, current or efficient? Are the methods for training the best ones? So many questions – all relevant given the time, cost and resources that training requires. Without well-conducted, comprehensive evaluations, we can only guess at the answers. Is this acceptable to your organization or to you as a learning and performance professional?

Harold Stolovitch and Erica Keeps are the principals of [HSA LEARNING & PERFORMANCE SOLUTIONS LLC](#), a leader in workplace learning and performance improvement. Their proven learning and performance solutions have helped maximize employee performance at hundreds of organizations throughout the world. Together they have devoted a combined total of over 80 years to make workplace learning and performance both enjoyable and effective.