Selecting Criteria for Evaluating Training

02/02/2009

Training Evaluation Team

training_eval@adlnet.gov

Introduction

Effectively evaluating training requires the systematic collection of information from a variety of sources.As organizations use training to achieve a variety of organizational goals, there is no universal approach to evaluating training—each organization must select the criteria that are most relevant to their organizational objectives.This whitepaper will assist organizations in identifying appropriate criteria for assessing their training programs.

What Should Be Evaluated?

When choosing evaluation criteria, it is critical to identify what questions need adressing in the evaluation. Within the training community, the dominant approach to training evaluation categorizes criteria into four levels. As shown in Table 1, each type of outcome addresses a different evaluation question.

Table 1. Training Evaluation Outcomes(Kirkpatrick, 1976)

Reactions	What did the trainees think of the training program?
Learning	Did the trainees learn the principles, techniques, and attitudes presented in training?
Transfer	Did the trainees transfer the principles, techniques, and attitudes presented in training to the workplace?
Results	Did the training program address the organization's objectives?

The first two levels (reactions and learning) tend to require assessing immediately after training, while the second two levels (behavior and results) require assessing after the learners have completed training and have returned to the job (generally one month to one year after training). Each of the four levels is described in more detail on the following pages.

Reactions

The first criterion for training evaluation is reactions or trainees' perceptions of a course. This level of evaluation is the most widely used type of training assessment. A survey by the American Society of Training and Development revealed that 91% of training courses use a reaction measure at the conclusion of training to evaluate the course (Sugrue & Rivera, 2005).

Assessing reactions allows trainers to measure if trainees are satisfied with the course and if they feel that they are learning from the training. Reaction data can provide trainers with valuable diagnostic feedback that they can use to modify the courses to meet the needs of trainees and their organizations.

Types of Reactions: As there are multiple aspects of a course that can influence trainee satisfaction, trainers should assess the dimensions that are relevant to their courses. The five main categories of training reaction measures are below. Appendix A contains a list of reaction items for each dimension.

- *Affective reactions*—assess whether or not the trainees liked or enjoyed the training.
- *Utility reactions*—assess the trainees' perceptions that the skills taught in training were useful and relevant to their jobs.
- *Instructor reactions*—assess the learners' perceptions of the instructor's contributions to learning.
- *Delivery reactions*—assess the students' perceptions that the material was presented in an organized and coherent manner.
- *Technology reactions*—assess the trainees' satisfaction with the technology used, and their perceptions that the technology was easy to use and facilitated learning.

Collecting Reactions Data: After deciding which types of training reaction measures are relevant, instructors should pull the questions together in a questionnaire (paper or online) and administer it to learners at the conclusion of training.

Tips for Collecting Reaction Data

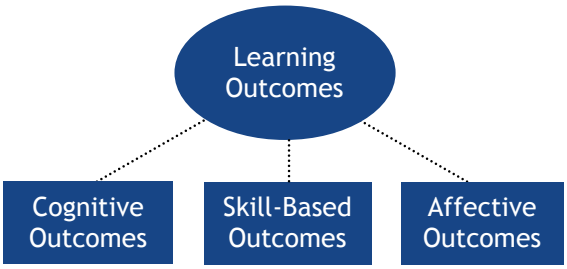
- Design the instrument so that results can be tabulated and quantified.
- To obtain more honest opinions, allow surveys to be completed anonymously.
- Provide a space for trainees to write-in about topics not covered in the survey.

Although reaction measures are useful for providing feedback on course characteristics, they do not provide evidence as to whether the training influenced learners' knowledge (Sitzmann, Brown, Casper, Ely, & Zimmerman, 2008).The following section provides recommendations for evaluating training with learning measures.

Learning

The second level of a training evaluation involves assessing what the students learned in the training. In measuring learning, three types of outcomes are generally measured: cognitive, skill-based, and affective. Cognitive outcomes include facts and information presented in training, while skill-based outcomes include knowledge of how to perform the tasks or skills presented in training. Occasionally, the important outcomes of training are not declarative or procedural knowledge, but affective changes in learners' attitudes or motivation. Figure 1 outlines the three categories of learning outcomes.

Figure 1. Training Evaluation Outcomes (Kraiger, Ford, & Salas, 1993).



Depending on whether the training objectives focus on cognitive, skill-based, or affective outcomes, there are different formats that are more appropriate for assessing knowledge gains. If the training objective is for learners to recognize and recall training content, a multiple-choice test would be appropriate, but if the training objective is executing a specific skill, scoring trainees' performance while performing the skill would be more appropriate.

One example of an affective outcome is self-efficacy, or trainees' confidence in their understanding of the training material and their belief in their ability to apply the material they learned in the workplace. Research has shown that self-efficacy is a strong predictor of training transfer (Sitzmann et al., 2008). Sample self-efficacy items are included in Appendix B.

It is important to remember that just because trainees do well on a post-training exam does not mean that they learned the material during training—the students could have pre-existing knowledge or could have learned the material somewhere else (e.g., on the job). In order to conclude if learning is due to training it is important to have a comparison point. When possible, it is recommended to compare trainees' post-training test scores with pre-training test scores, or with a control group who has not yet attended the training.

Transfer

Showing that trainees learned the material presented in training does not necessarily mean that the trainees will transfer the learning outcomes back to the workplace. In order to assess changes in behaviors on the job, it is important to have a comparison point of behaviors before the training in order to quantify improvements. As shown in the box below, there are multiple information sources that can be used to assess on-the-job behavior.

Information Sources to Assess Transfer

- Objective measurements of actual job behavior (e.g., number of errors made)
- Trained observers' assessments of job performance
- Performance appraisals conducted by the trainee, trainee's coworkers, supervisors, and subordinates

Evaluating training programs using on-the-job behavior is more difficult than using reaction or learning data as it requires a more systematic approach to collect pre-training and post-training data. Assessing post-training performance should be delayed at least three months after training to allow the trainees the opportunity to implement the changes in their performance (Kirkpatrick, 1976).

Results

Results refer to the degree to which the training met the organization's objectives. In assessing results, it is important to identify the organization's objectives and how the training influenced these objectives. For example, if an organization implements a safety training

program they could compare organizational records of on-the-job accidents before and after the training. Other results-level indicators that can be examined include costs, turnover, absenteeism, grievances, and morale (Goldstein & Ford, 2002). Results-level outcomes are the most challenging evaluation criteria to assess, although it is generally the outcome that organizations find the most valuable.

References

Goldstein, I. L., & Ford, J. K. (2002). Training in organizations (4th edition). Belmont, CA: Wadsworth.

Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig (Ed.),Training and development handbook: A guide to human resource development. New York: McGraw Hill.

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.

Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. Journal of Applied Psychology, 93, 280-295.

Sugrue, B., & Rivera, R. J. (2005). State of the industry: ASTD's annual review of trends in workplace learning and performance. Alexandria, VA: ASTD.

Appendix A: Sample Reactions Items

Reaction measures are used to assess students' satisfaction with a course. While reaction items do not assess actual learning, they provide instructors with valuable diagnostic feedback that can be used to modify courses to meet the needs of students and their organizations. The following pages include training reaction items compiled by ADL that instructors can use to evaluate students' reactions to training.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Affective Reactions					
1. I enjoyed the training.	1	2	3	4	5
2. This course was fun to complete.	1	2	3	4	5
3. Overall I am satisfied with this course.	1	2	3	4	5
4. I am enthusiastic about what I learned in this course.	1	2	3	4	5
5. Duringthis course I thought about how much I enjoyed it.	1	2	3	4	5
6. This course was boring.*	1	2	3	4	5
7. During this course I became frustrated about some of the material.*	1	2	3	4	5
Utility Reactions					
1. The information presented in this course is relevant to my job.	1	2	3	4	5
2. The training will help me perform my job.	1	2	3	4	5
3. This training will have a positive impact on my job performance.	1	2	3	4	5
4. I do not think I will use what I learned in this class.*	1	2	3	4	5
5. The training was relevant to my job.	1	2	3	4	5
Instructor Reactions (only for instructor-led courses)					
1. The instructor explained things clearly.	1	2	3	4	5
2. The instructor was prepared for every class.	1	2	3	4	5
3. The instructor was competent.	1	2	3	4	5
4. The instructor was knowledgeable about the training content.	1	2	3	4	5
5. Overall, this instructor was effective at teaching this course.	1	2	3	4	5
Training Delivery Reactions					
1. The course content was well organized.	1	2	3	4	5
2. The material presented was appropriate for students at my level of experience.	1	2	3	4	5
3. The structure of the course made it easy to learn the material.	1	2	3	4	5
4. The pace of the course was appropriate.	1	2	3	4	5
5. The training was coherent.	1	2	3	4	5
Technology Reactions (only for technology-based courses)					
1. The technology enhanced my learning experience.	1	2	3	4	5
2. The technology interface was difficult to use.*	1	2	3	4	5
3. Overall, I am satisfied with the technology used in this course.	1	2	3	4	5
4. The technology helped me learn the training content.	1	2	3	4	5

Note on Using Reactions Items: An asterisk (*) indicates that the item needs to be reverse-scored. When analyzing responses, high scores should indicate that trainees are more satisfied than low scores. Item for which a low score indicates greater satisfaction (items marked with an *) must be reverse scored so that higher scores indicate that the trainees were more satisfied with the course.

Appendix B: Sample Self-Efficacy Items

Self-efficacy is trainees' confidence in their ability to reach their training goals and transfer the information to their jobs. It is the best survey tool available for predicting skill-based knowledge and training transfer. After controlling for pretraining knowledge, post-training self-efficacy predicts 14% of the variance in post-training skill-based knowledge and 24% of the variance in training transfer.

Below are a few examples of self-efficacy measures. The first measure is a general scale while the second and third measures were designed for specific courses. We do not recommend using these exact items. Rather, the items should be tailored to be specific to a training course.

General Self-Efficacy

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1. I will be able to achieve most of the goals that I have set for myself.	1	2	3	4	5
2. When facing difficult tasks, I am certain that I will accomplish them.	1	2	3	4	5
3. In general, I think that I can obtain outcomes that are important to me.	1	2	3	4	5
4. I believe I can succeed at most any endeavor to which I set my mind.	1	2	3	4	5
5. I will be able to successfully overcome many challenges.	1	2	3	4	5
6. I am confident that I can perform effectively on many different tasks.	1	2	3	4	5
7. Compared to other people, I can do most tasks very well.	1	2	3	4	5
8. Even when things are tough, I can perform quite well.	1	2	3	4	5

Self-Efficacy for a Specific Simulation

1. I can meet the challenges of this simulation.	1	2	3	4	5
2. I am confident in my understanding of how information cues are related to decisions.	1	2	3	4	5
3. I can deal with decisions under ambiguous conditions.	1	2	3	4	5
4. I am certain that I can manage the requirements of this task.	1	2	3	4	5
5. I believe that I will fare well in this task if the workload is increased.	1	2	3	4	5
6. I am confident that I can cope with this simulation if it becomes more complex.	1	2	3	4	5
7. I believe I can develop methods to handle changing aspects of this task.	1	2	3	4	5
8. I am certain I can cope with task components competing for my time.	1	2	3	4	5

Self-Efficacy For Transfer

1. I am confident that I can apply the material that I learned in the course to my job.	1	2	3	4	5
2. I believe that I can transfer what I have learned to my job.	1	2	3	4	5
3. I am certain that I can use the skills I learned in training to improve my job performance.	1	2	3	4	5
4. I am confident that I have learned the material presented in training.	1	2	3	4	5
5. I believe that I have improved my work-related skills during the training course.	1	2	3	4	5