

## Domain of Evaluation

Since the 1930's, the idea of using evaluation as a way to judge the value of educational materials has evolved into one of the most integral components within the field of instructional technology and education as a whole. In order to determine the adequacy of instruction and learning, an instructional designer may need to utilize a variety of models, processes and tools depending upon the purpose of the evaluation. There are two levels of evaluation, each with a different purpose. An evaluation could be conducted for an entire program at the macro-level, or for a learning module at the micro level. There are also three different types of evaluation including formative, summative and confirmative evaluation.

There are several instructional design models available to help guide the designer in conducting evaluation including the Kemp, Morrison and Ross Model (1994), Dick and Carey Model (2005), or Phillips Return on Investment (ROI) Model (1996). In addition, there are a number of program/course evaluation models to assist the designer through the process including CIPP (Context, Input, Process & Product) (Stufflebeam, 1998), ETEM (Education and Training Evaluation Model), (ASC, 2003), and Kirkpatrick's 4-level Model (1994).

Within the evaluation domain, there are four sub-domains: problem analysis, criterion reference measurement, formative and summative evaluation.

### Problem Analysis

At the beginning stages of an instructional design project, an evaluation is conducted in order to determine what problems currently exist within the system and what solutions would be appropriate for solving them. This process is generally referred to as a problem analysis or front-end analysis and involves gathering both qualitative and quantitative information from a variety of sources using tools such as surveys, questionnaires, interviews, observations and focus groups.

In order to accomplish this task, an instructional technologist should utilize several different techniques including: needs assessment, extant data analysis, task analysis and subject matter analysis (Rossett, 1987). This information can be triangulated by the designer in order to get a clear description of the nature and parameters of any identified problems or needs. Allison Rossett (1987) describes "needs" as being the difference between current performance or knowledge and optimal performance or knowledge. Once the problem analysis is complete, the results or identified needs are used to determine whether training is necessary to solve the performance issue or is there is another non-instructional solution.

### Criterion-Referenced Measurement

This type of measurement can be used to assess a learner's mastery of knowledge, attitudes or skills in relation to the identified instructional goals. An instructional

designer would use the information obtained from the task analysis to formulate performance objectives, which are then used to develop the criterion-referenced test. These learner-centered measurements can then be used to assess the learner's ability to meet those performance objectives. The learner's performance is compared to that standard or optimal performance as opposed to being compared relative to group test scores as in norm-referenced testing (Lynn & Miller, 2005).

## **Formative Evaluation**

During the development phases of an instructional design project, a series of evaluations are conducted with the purpose of measuring how adequate a product is to meet the instructional goals, which also enables the designer to identify specific areas needing further revision. In order to accomplish this task, an instructional designer may choose to use a model such as Dick & Carey's (2005) model which involves using three different levels of evaluation techniques including: one-to-one, small-groups and field trials (Dick & Carey, 2005). The qualitative and quantitative data collected during these phases is usually informal and collected using short tests, observation and interviews (Seels & Richey, 1994). The results of the evaluations are typically presented in a detailed report with specific recommendations for improving the materials.

As previously mentioned, there are many models from which an instructional designer may choose to aid them in planning and conducting formative evaluations. For example, Gooler (1980) outlines the following eight-step approach to planning formative evaluations, which guides the designer by posing a series of questions.

- **Purpose**
  - Why is the evaluation being conducted?
- **Audiences**
  - Who are the target recipients of the evaluation results?
- **Issues**
  - What are the major questions/objectives of the evaluation?
- **Resources**
  - What resources will be needed to conduct the evaluation?
- **Evidence**
  - What type of data or information will be needed to answer the evaluation questions?
- **Data-gathering Techniques**
  - What methods are needed to collect the evidence needed?
- **Analysis**
  - How will the evidence collected be analyzed?
- **Reporting**
  - How, to whom, and when will the results of the evaluation be reported?

This formative evaluation process is critical in terms of developing instructional materials that can be used effectively by learners to accomplish instructional objectives, and it is also cost effective. The instructional designer saves the client both time and

money by going through this iterative process, because making changes after a product is produced can be very costly and time consuming.

### Summative Evaluation

Once the formative evaluation of an instructional design project is complete, and the final materials have been produced, a summative evaluation is conducted usually during and after the implementation phase. The purpose of the summative evaluation is to gather data, which will serve to verify the effectiveness of the instruction for, and acceptance by, the target audience (Dick & Carey, 2001). To accomplish this, Kirkpatrick's Four Levels of Evaluation model (1994) is one that is commonly used throughout the field to guide the instructional designer through the summative evaluation process. The steps or levels within this model build successively upon each other beginning with learner reactions as illustrated in figure 1.

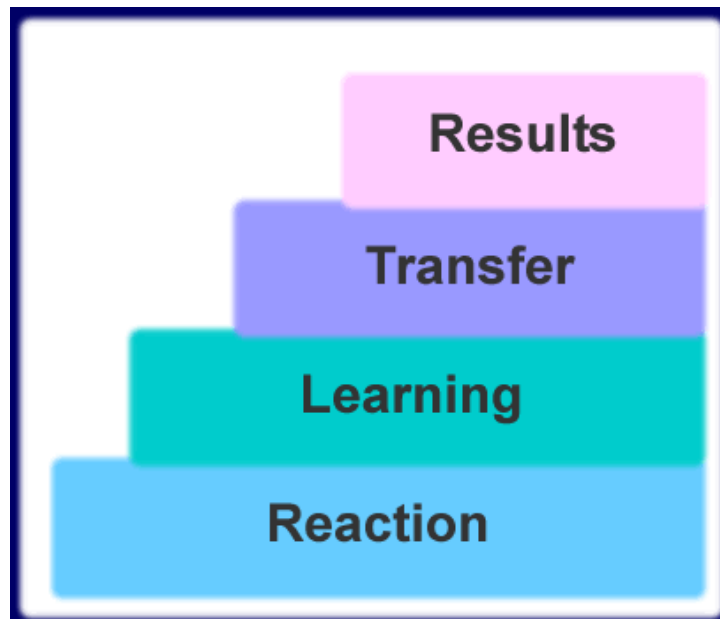


Figure 1. Kirkpatrick's Four-Level Summative Model Adapted from "Evaluating the Effectiveness of Training Programs," by P. Boverie, D.S. Mulcahy, and J.A. Zondolo, 1994, *Developing Human Resources*. Copyright 1994 by Pfeiffer & Company.

1. Reaction – How satisfied are the learners with the training and what is their level of acceptance regarding materials, environment and the process in general?
2. Learning – According to the level of knowledge, attitudes and skills measured before and directly after instruction, were the instructional objectives met by the learners?
3. Behavior – Were the learners able to retain the information after training and transfer the new knowledge and skills to their real-world job successfully?
4. Results – What impact did this training have on the organization as a whole and is the associated cost justifiable?

The methods used to gather the summative evaluation data are typically more formal than the formative evaluation in regards to collection methods which include: pre/post tests, reaction tests, embedded tests or performance tests (Chen, 2002).

Another significant method that differentiates summative from formative evaluation is that it is usually conducted by a third party as opposed to the instructional development team. The reason for this is because team members are usually personally invested in their project and may not be able to remain objective throughout the assessment process (Seels & Richey, 1994).