

Domain of Evaluation

Evaluation is “the process of determining the adequacy of instruction and learning” (Seels & Richey, 1994, p. 54). The instructional designer engages in several forms of evaluation during the course of a project, including problem analysis, criterion-referenced measurement, formative evaluation, and summative evaluation. Each of these different types of evaluation is performed at a different level. The problem analysis uncovers knowledge and/or performance gap(s); criterion-referenced measurement tests learners’ mastery of the performance objectives; formative evaluation identifies and makes necessary changes to the product while it is still in development, before implementation; summative evaluation determines the success of a project after it has been in use for some time; and confirmative evaluation determines whether a course is still effective after the summative evaluation has been administered.

Problem Analysis

“Problem analysis involves determining the nature and parameters of the problem by using information-gathering and decision-making strategies” (Seels & Richey, 1994, p. 56). Problem analysis is also known as front end analysis or needs assessment, and it is often the first step in the instructional design process. Instructional designers often use a set of guidelines and principles to conduct needs analysis. For example, in her book *Training Needs Assessment*, Allison Rossett (1987) outlines the five items that instructional designers should seek information about during the needs assessment:

- Optimal performance or knowledge
- Actual or current performance or knowledge
- Feelings of trainees and significant others
- Causes of the problem from many perspectives
- Solutions to the problem from many perspectives

Upon collection of the information, instructional designers conduct a gap analysis. The gap between what should be occurring (optimal) and what is really occurring (actual) is identified as a discrepancy or need (Kaufman, 1982). A mathematical equation illustrates this concept:

$$\text{OPTIMAL} - \text{ACTUAL} = \text{NEED}$$

In order to collect proper data, instructional designers use four main techniques: extant data analysis, needs assessment, subject matter analysis, and task analysis. The extant data analysis consists of delving into the information that the organization has collected which represents the results of employee performance (Rossett, 1987). The extant data analysis is performed in order to reveal actuals. The needs assessment discovers opinions on optimal, actuals, feelings, causes, and solutions, and it is performed on several different sources in order to garner a wide range of opinions. The subject matter analysis uses subject matter experts and various documents to uncover the optimal which tell the instructional designer exactly what it is an informed employee possesses in order to get

the job done right. And finally, the task analysis, which relies mainly on observation, unearth optimal regarding visible tasks. The various analyses are carried out with tools such as interviews, observations, group facilitation, and surveys or questionnaires.

When the problem analysis is complete, the gap between the optimal and the actual is identified as the problem, which is used to determine if training or another solution should be implemented.

Criterion-referenced Measurement

Criterion-reference measurement “involves techniques for determining learner mastery of pre-specified content” (Seels & Richey, 1994, p. 56). Instructional designers use instructional goals and performance objectives to identify assessment measures. The process of using goals and performance objectives to identify and construct assessment measures results in criterion-referenced measures. Criterion-referenced measurements are learner-centered assessments because they are aligned with expected learning outcomes rather than the content or subject matter (Dick, Carey & Carey, 2005). As opposed to norm-referenced testing which is based on experts’ views of the content of performance objectives (the expert decides what measures should be used to assess the learners), criterion-referenced measurement provides information about the learner’s mastery of knowledge, attitudes or skills as it pertains to the objectives (Seels & Glasgow, 1990).

Formative Evaluation

“Formative evaluation is the process designers use to obtain data that can be used to revise their instruction to make it more efficient and effective” (Dick, Carey, & Carey, 2005, p. 278). The instructional designer conducts three basic phases during formative evaluation (Dick, Carey, & Carey, 2005). Prior to the three phases, however, subject matter experts who are not directly involved in the project review the instruction for errors prior to testing it out in front of the learners. The first phase of formative evaluation is one-to-one evaluation. In this phase, the designer works directly with individual learners to obtain data for revision of the materials. Next is the small-group evaluation in which eight to twenty learners representative of the target population go through the materials individually followed by being tested in order for the designer to collect the required data. Third, and last, is the field trial. Roughly thirty learners participate in a “real world” simulation of the procedures required for the installation of the instruction.

The formative evaluation is usually conducted internally by a member of the project team. This person most likely has a personal investment in the project and therefore is well suited to gather information in order to make the materials the best they can possibly be (Dick, Carey, & Carey, 2005). At the conclusion of the formative evaluation, the instructional designer and his or her team revise the materials based on the feedback gained through all three formative evaluation phases. It is essential to fix all problems with the instructional materials prior to implementation. Otherwise, it is costly with both time and funds to make changes after the materials have been produced.

As a result of formative evaluation, the instructional designer is able to make any necessary corrections to the material while it is still in development and before it is implemented, thus saving valuable time and cost.

Summative Evaluation

Dick, Carey, and Carey (2005, p. 340) define summative evaluation as “the design of evaluation studies and the collection of data to verify the effectiveness of instructional materials with target learners.” The purpose of summative evaluation is to make decisions regarding whether to maintain instructional materials which are currently in use or whether to go with something new which might better meet the instructional needs of the organization.

The instructional designer conducts two stages within summative evaluation: expert judgment and field trial. In the expert judgment phase, the instructional designer determines whether the instruction currently in use or whether new instruction will better meet the organization’s instructional needs. In the field trial phase, the instructional designer samples the possible new instruction with the target audience in the actual instructional setting.

A third party usually conducts the summative evaluation. Since members of the project team are personally invested in the project, it likely is difficult for them to remain objective during the summative evaluation process.

A summative evaluation which is in wide use today is Kirkpatrick’s Four Levels of Evaluation (Kirkpatrick, 1994). Level one, which serves as the base of the model, moves to level two, until each successive layer has built upon the previous one. Kirkpatrick’s model is illustrated in Figure 3 below:

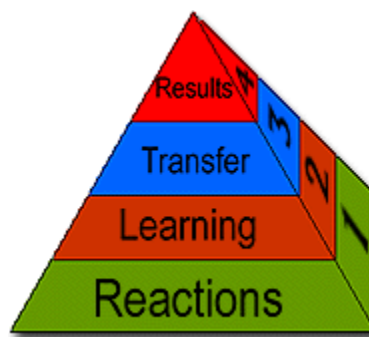


Figure 3. Kirkpatrick’s Four Levels of Evaluation (Kirkpatrick, 1994)
Adopted from <http://coe.sdsu.edu/eet/Articles/k4levels/index.htm>

Level 1 – Reactions – measures the initial reaction of the participants

Level 2 – Learning – moves beyond learner satisfaction to gauge the changes in learners’ knowledge, attitude, and skill as a result of the instruction

Level 3 – Transfer – measures how much the newly-acquired knowledge, attitude, and skill is being used in the learners' everyday environment

Level 4 – Results – measures the organization's "bottom line" – increase in production, improvement in quality, decrease in cost, increase in profit, etc.