Domain of Evaluation

The establishment of formalized education programs created the need for formalized evaluation programs. More systematic and scientific procedures were required to evaluate these programs (Seels & Richey, 1994). Seels and Richey define the domain of evaluation as "the process of determining the adequacy of instruction and learning" (Seels & Richey, 1994, p.54). The domain of evaluation examines the instruction; making changes where needed (formative evaluation) and determines the effectiveness of the instruction and the impact upon learners (summative evaluation).

The Joint Committee on Standards for Educational Evaluation (1981) describes three types of evaluations: program evaluation, project evaluation and materials evaluation (pp. 12, 13). Program evaluations assess educational activities which provide services on a continuing basis. Project evaluations assess activities that are funded for a defined period of time to perform a specific task. Materials evaluations assess the merit or worth of content-related physical items (Seels & Richey, 1994). The domain of evaluation is made up of five sub-domains: problem analysis, criterion-referenced measurement, formative evaluation, summative evaluation, and confirmative evaluation (Seels & Richey, 1994).

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 often the first step in the instructional design process. According to Rosset (1987), the purpose of the training needs assessment (TNA) is to gather information about the optimals, actuals, feelings, causes, and solutions. Optimals describe what should be occurring within the organization; whereas, actuals describe what is currently happening. Feelings describe the attitudes of those involved within the organization. Finally, causes and solutions focus on the opinions and data related to possible causes of the problem as well as possible solutions to resolve the problem.

Goals and priorities are set based on the defined needs, which are the differences between actual performance and optimal performance (Seels & Richey, 1994; Rossett, 1987). Instructional designers must collect data using appropriate data collection instruments to identify the problem and causes in order to provide solutions to the problem. Rossett (1987) recommended four data collection techniques: extant data analysis, needs assessment, subject matter analysis, and task analysis (Rossett, 1987).

Extant data analysis involves reviewing the information the company collects that represents the results of employee performance.

Needs assessment involves going out and seeking opinions on the optimals, actuals, feelings, causes and solutions from a variety of sources.

Subject matter analysis uses subject matter experts and various documents to uncover the optimals, the information that an informed employee possess which enables him or her to get the job done.

Task analysis relies heavily on observation to derive optimals attached to visible tasks.

The completion of a problem analysis determines if training or other solutions should be implemented.

Criterion-referenced Measurement

“Criterion-referenced measurement involves techniques for determining learner mastery of pre-specified content” (Seels & Richey, 1994, p.56). The purpose of criterion-referenced measurement is to determine if the student learned the intended information and to determine if the goal of the materials has been reached. Dick, Carey, and Carey (2005) stated that criterion-referenced assessments should be created in relation to the objectives of the materials.

Formative Evaluation

“Formative evaluation involves gathering information on adequacy and using this information as a basis for further development” (Seels & Richey, 1994, p.57). The purpose is to identify errors in the instructional materials, identify issues affecting learning outcomes, diagnose learning problems of users, and revise and improve the quality of materials and learning. Formative evaluation is conducted during the development or improvement of a program or product. It is conducted for the in-house staff of the program and normally remains in-house. It may be implemented by an internal or external evaluator or by both (Scriven, 1967). Formative evaluation consists of three phases: one-to-one evaluation, small-group evaluation, and field trial (Dick, Carey & Carey, 2005). Each phase is described below.

One-to-one evaluation- The instructional designer works directly with individual learners to obtain data for revision of the materials.

Small-group evaluation- The instructional designer will randomly select eight to twenty learners who represent the target population to review the materials. The learners will evaluate the changes made by the one-to-one evaluation and identify any remaining learning problems that may occur.

Field trial-The instructor will use a learning context that resembles the intended setting to determine whether the instruction is administratively possible to implement.

Summative Evaluation

"Summative evaluation involves gathering information on adequacy and using this information to make decisions about utilization" (Seels & Richey, 1994, p.57). It is conducted after the implementation of the program or product (Scriven, 1967). The principle behind summative evaluation is to determine whether to continue use of the current materials or to investigate something new that would better benefit the organization. Summative evaluation also provides information to justify the cost of implementation and offers ideas for future revisions. Dick, Carey, and Carey (2005) recommend someone other than the creator to perform the summative evaluation to ensure objectivity and validity of the findings.

There are four levels of summative evaluation according to Kirkpatrick: reaction, learning, transfer and results (Kirkpatrick, 1994). The lowest level of reaction gauges the learner's feelings about the training or implementation. The level of learning gauges the changes in learner's knowledge, attitude, and skill as a result of the instruction. The level of transfer measures the learner's behavior change or the ability to transfer the new information learned. The level of results measures the outcome of the training to the organization, such as increase in production, improvement in quality, decrease in cost, increase in profit, etc. See the model below.

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

Confirmative Evaluation

E.R. Misanchuk (1978) first introduced the concept of confirmative evaluation. Confirmative evaluation is a continuous type of evaluation used to determine the long-term organizational impact of the implementation (Morrison, Ross & Kemp, 2007; Reiser & Dempsey, 2007). The goal of the instructional designer, with regards to confirmative evaluation, is to provide the organization with "meaningful and valid evidence that the training has made a measurable difference" (Reiser & Dempsey, 2007, p.180).

Confirmative evaluation is conducted by a team of unbiased evaluators. The evaluators use tools such as interviews, surveys, performance assessments, and knowledge tests to gather information (Seels & Richey, 1994). Two questions guide confirmative evaluation: "Do materials still meet the original objectives?" and "Have learners maintained their level of competence?" (Seels & Richey, 1994, p.59) Confirmative evaluation should take place at least six months to a year after initial implementation (Morrison, Ross & Kemp, 2007).