The Domain of Evaluation

The domain of evaluation is the "process of determining the adequacy of instruction and learning" (Seels & Richey, 1994, p.54). Evaluation is used by the instructional technologists for a number of purposes. The first is feedback, to evaluate the materials and the delivery methods, ranging from the identification of basic errors in instruction to determining whether a transfer of learning in the instructional and workplace environment occurred. The second purpose of evaluation is control. In other words, did the program meet the needs of the organization and was it worth the cost? The third purpose of evaluation is research, used to improve and justify future instructional processes and needs. Intervention, the fourth purpose of evaluation, examines an instructional program and identifiable outcomes affecting management, the work environment and the training department. And finally, the fifth purpose of evaluation is power, giving instructional technologists the data to support the adoption, expansion or abortion of a program (Indiana University, 2006). In the instructional system, evaluation is an on-going and constantly evolving process designed to provide feedback as to whether the training program or materials, as designed and implemented, were successful.

According to Seels & Richey (1994), the domain of evaluation consists of four areas of study which include: (1) problem analysis, (2) criterion-referenced measurement, (3) formative evaluation and (4) summative evaluation. Confirmative evaluation is a fifth area of study within the evaluation domain recently added by Morrison, Ross and Kemp in 2007.

1. 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). Front-end analysis, training needs assessment (TNA) and needs assessment are frequently used alternatives for the term problem analysis. All terms are used to describe a systematic process of collecting and analyzing data for making decisions regarding the worth or value of an instructional program and whether the program has achieved its intended instructional goals.

Allison Rossett's Training Needs Assessment model (1987) is often used when conducting an analysis regarding training needs. The model addresses the use of techniques and tools for gathering information needed to determine optimal and actual levels of performance. Subtracting the actual from the optimal will reveal the gap in performance and the focus of the instructional needs. Instructional designers should also take into consideration the feelings of the workers or learners in order to determine causes and potential solutions (Rossett, 1987).

2. Criterion-Reference Measurement

Criterion-reference measurement involves "techniques for determining learner mastery of pre-specified content" (Seels & Richey, 1994, p. 56). An instructional technologist uses criterion-referenced test items to determine the degree to which learners have met the desired objectives, which parts of the instruction were effective and allow learners to evaluate their own level of

performance. Each test item correlates directly to a written performance objective (Dick, Carey & Carey, 2005, pp. 145-146).

Entry behaviors tests, pretests, practice tests and posttests are four major types of criterion-reference tests an instructional designer may create (Dick, Carey & Carey, 2005, p. 146-148). Entry behaviors tests are given to learners prior to instruction. These tests measure whether the learner has the required pre-requisite skills and knowledge to begin the instruction.

Pretests are administered to learners prior to instruction to determine the level of instruction needed within a unit. For example, if learners score high on a pre-test, maybe only a review of that unit is needed. A low score would mean the unit of instruction should be covered in more detail. "A pretest is valuable only when it is likely that some of the learners will have partial knowledge of the content" (Dick, Carey & Carey, 2005, p. 147).

Practice tests allow learners to apply their knowledge and skills and receive feedback regarding their level of understanding (Dick, Carey & Carey, 2005, p. 148). Practice tests also provide feedback to the instructional designer as to whether the instruction was effective or whether more review is needed to meet the instructional goal.

Posttests are administered after the instruction has been delivered and help the instructional designer identify areas of the instruction which may need further attention or review (Dick, Carey & Carey, 2005, p. 148).

3. Formative Evaluation

Formative evaluation is the "gathering of information on adequacy and using this information as a basis for further development" (Seels & Richey, 1994, p. 57). The purpose of formative evaluation is to collect and analyze data, then use this information to revise instructional materials" (Dick, Carey & Carey, 2005, p. 278). Its function is to inform the instructional designer and team as to how well the instructional materials are meeting the stated objectives as the development progresses (Morrison, Ross & Kemp, 2007, p. 236).

There are three phases of formative evaluation where the designer collects information from learners: (1) one-to-one evaluation, (2) small group evaluations, and (3) field trials. The purpose of the one-to-one evaluation is to identify obvious errors in the instruction, obtain initial information regarding learner performance and learner reaction to the material. The purpose of the small group evaluation is to determine the effectiveness of changes made to the materials during the one-to-one evaluation and to identify any further problems learners are having. The final stage, the field trial, again determines whether the initial two evaluation processes were effective and whether the materials can be used in the setting, or context, for which they are intended (Dick, Carey & Carey, 2005, pp. 282, 288,290).

4. Summative Evaluation

Summative evaluation is the "gathering of information on adequacy and using this information to make decisions about utilization" (Seels & Richey, 1994, p. 57). The purpose of the summative

evaluation is to measure the effectiveness of the instructional program and materials. The summative evaluation also measures the efficiency of the learning, the cost of the development of the program, any continuing expenses, learner reactions toward the program, and long-term benefits of the program (Morrison, Ross & Kemp, 2007, p. 240).

A commonly used summative evaluation model used by business and industry is Donald Kirkpatrick's Four-Level Model (1994). This model was originally developed in 1959 and has gone through several revisions. The four levels of this model are listed and each level builds upon information obtained in the previous level.

- Reaction evaluation – a measurement of the feelings of participants to a training program
- Learning evaluation – a measurement of the knowledge and skills gained from the instruction
- Transfer evaluation – a measurement of improvement of performance before and after the training/instruction.
- Results evaluation – a measurement of the results of the training on the organization.
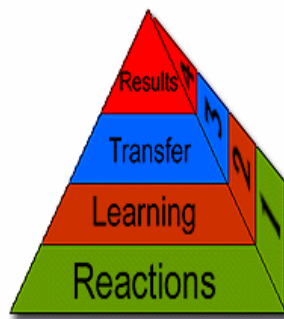

Figure 6: Kirkpatrick's Four-Level Evaluation
Adopted from: http://coe.sdsu.edu/eet/Articles/k4levels/index.htm

Corporations are increasingly looking at the need to not only evaluate the effectives of training in solving performance problems, but also the cost-benefit of the training.   In his 1996 article, *Measuring ROI: The Fifth Level of Evaluation*, author Jack Phillips stated that to obtain a true return on investment evaluation, "the monetary benefits of the program should be compared to the cost of implementation in order to value the investment." Phillips extended Kirkpatrick's evaluation model to include a fifth level, designed to determine the cost effectiveness of a training initiative.

5. Confirmative Evaluation

In the 2007 edition of their book, *Designing Effective Instruction*, Morrison, Kemp and Ross have extended the evaluation process to include confirmative evaluation. Confirmative evaluation is defined as "a continuous form of evaluation that comes after summative evaluation

used to determine whether a course is still effective" (Morrison, Kemp and Ross, 2007, p. 430). Confirmative evaluation can be conducted weeks or even months after the completion of a training program or instruction and is used to evaluate the continuing effectiveness of the new skills and knowledge in the workplace as a result of a training initiative. Data is gathered through the use of multiple information gathering tools, including questionnaires, performance assessments and interviews.