The Domain of Evaluation

Needs Assessment • Criterion-Referenced measurement • Formative Evaluation • Summative Evaluation

Seels and Richey (1994) defines evaluation as "the process of determining the adequacy of instruction and learning" (p. 54). In the field of instructional technology evaluation is an ongoing process throughout the entire project. It involves the sequent use of four types of evaluation: needs assessment, criterion-referenced testing, formative evaluation, and summative evaluation. They constitute a systematic approach to evaluation to determine the adequacy of an instructional technology program or project from initiation of it to full into use of it.

Needs Assessment

Rossett (1987) defines needs assessment as "the systematic study of a problem or innovation, incorporating data and opinions from varied sources, in order to make effective decisions or recommendations about what should happen next" (p.3). A needs assessment, also known as "front-end analysis," is intended to formulate the problem and determine the best way to try and solve it in any given situation. Going ahead with the needs assessment is the first step for an instructional designer when called on the stage. In the step an instructional designer gather a variety of data related to the problem, from the perspective of the various sources and tools, to make decisions about priorities and recommendations on the solutions. The results of the needs assessment indicate whether or not the problem is performance-based and can be solved by instruction. There is a number of models for the instructional designer' use in the front-end analysis. Allison Rossett's Training Needs Assessment (TNA) model and Roger Kaufman's system development model (1993). In the Rossett's model, the instructional designer goes through the process of :
- determine the purpose
- identify the sources
- select the data-collection tools
- gather data fulfilling the five purposes of TNA (Optimals, Actuals, Feelings, Cause(s), Solutions.)
- make decisions using findings.

Roger's model is often used to conduct the front-end analysis in the large operating system. The principle procedural steps in the model are:
- sharpen the problem statement
- analyze the operating system (needs analysis, contextual analysis)
- determine the causes (performance analysis)
- identify solutions
- choose the best solution.

In practice such data-gathering tools are commonly used by the instructional designer when conducting the front-end analysis. They include interviews, surveys, questionnaire, and observations. The use of them depends on the purpose that they serve and the particular situation. For instance, observation, if used for gathering data about performance, is the best tool although time-demanding and labor-intensive. But if observation is used for gathering data about attitude and feeling, it is not cost-effective. Interview, like observation, is labor-intensive for the designer, and but it allows the designer to obtain in-depth information about the system, processes and attitudes of the group being asked to perform the task.

Criterion-Referenced Measurement

Criterion-referenced measurement refers to measuring if the learner has mastered pre-specified learning objectives through specified objective-based test items. Learning objectives and test items are the keys for criterion-referenced measurement. The writing of learning objectives and

development of test items begins with task analysis. The objective of conducting task analysis is to identify the subtasks, information flow, and inputs required to perform a task and form a flowchart that revealing the relationships among task elements. Using the flowchart of task analysis, the instructional designer can translate task, subtasks, and relationships among task elements into observable and measurable goal and enabling objectives. Based on these enabling objectives, criterion-referenced test items are developed accordingly to evaluate the learners' process against the course or curriculum.

Formative Evaluation

Formative evaluation involves gathering information on adequacy and occurs under the development of instruction or product. Information on adequacy includes test data, comments and annotations, attitude data, testing time, and reactions from Subject Matter Experts (SME) and the clients. The instructional designer uses information gathered to identify specific errors in the instructional materials and its use, and correct them.

There are three phases for the instructional designer to follow or choose to accomplish formative evaluation: one-on-one, small group, and field trial. In one-to-one phase, the instructional designer works with individual learners to gather data to revise the materials. A small-group evaluation is to determine whether the learners can instructional materials or instruction without interacting with the instructors. The emphasis in the third phase of the field trial is on the testing of the procedure required for the installation of the instruction in a situation as close to the "real world" as possible. The three phases are conducted successively, and the prior phase provides information for the use of the next phase. The deeper the instructional designer goes into along the phases, the more adequate information gathered is to allow for many changes for adequacy of instructional materials and its use. Going through one or more phases in formative evaluation depends on specific situation, including time, money, and resources. After final revisions are made to the instructional material and its use, the instructional material is ready to put into use in the real world.

Summative Evaluation

Comparing to formative evaluation, summative evaluation is conducted after implementation and institutionalization by external evaluators. In summative evaluation qualitative and quantitative data are gathered and analyzed to d ocument strengths and weaknesses in instruction in order to decide whether to maintain or adopt it. Also, there are a number of models for the instructional designer's use. Such examples are Dick, Carey and Carey's model (2005) and Donald Kirkpatrick's four-level evaluation model (1998).

According to Dick, Carey and Carey's model (2005), there are two phases in summative evaluation: Expert judgment and field trial. Expert judgment is used to decide if the instructional material or instruction has the potential for meeting the organization's needs. In the expert judgment phase, the decisions to be made include (1) the congruence between the organization's needs and goals and those of the instructional material or instruction; (2) the completeness and accuracy of the instructional material or instruction; (3) the inclusion of the instructional strategy; (4) the utility of the instructional material; and (5) the current users' satisfaction. The field trail has two components: outcome analysis and management analysis. It is designed to gather data about learner performance and attitudes, implementer attitudes, and procedures and resources required by the implementation of the candidate instructional material or instruction. The main purpose of the field trial is to locate, analyze, and document both the strengths and weakness of the candidate instructional material or instruction. Both the expert judgment and field trial can be focused on one set of instructional materials or on completing sets of materials.

Donald Kirkpatrick's model has been often used to assess the effectiveness of the training in the business or industry setting. Today, it also is adopted by the practitioners in the educational setting. The four levels refer to reactions, learning, transfer, and results. According to the model,

evaluation should always begin with level one. As time and budget allows, Evaluation should then move sequentially through levels two, three, and four. Information from each prior level serves as a base for the next level's evaluation. Thus, each successive level represents a more precise measure of the effectiveness of the training program, but at the same time requires a more rigorous and time-consuming analysis. The level one measures the rough impressions of the learner about the instructional technology program, for example, satisfaction, relativeness. The 'learning' evaluation attempts to assess the learners' performance including skills, knowledge, or attitude. The 'transfer' level is related to the degree the learners apply knowledge, skills, and attitude acquired to other like real situations. The last level, results, is the most complex and time-consuming. It aims at assessing the impacts the utilization of the instructional technology program will have had on many aspects inside and outside of the organization, such as increased production, improved quality, decreased costs, reduced frequency of accidents, increased sales, and even higher profits or return on investment. The level evaluation, I believe, covers to some extent the fifth level-ROI Jack Phillips (1996) introduced into Kirkpatrick's model. The ROI evaluation has some priority with specific purpose and detailed procedure if decisions about results in financial terms are really in need in an organization.

The evaluation domain is a pretty key domain in the field, going throughout the research and practice of the field and then completing the systematic characteristics of the field as a professional field.