Domain of Evaluation

The process of evaluation is instrumental in the field of instructional technology. Evaluation provides instructional designers with data that drives design decisions. Seels and Richey (1994) define evaluation as "the process of determining the adequacy of instruction and learning" (p. 54). Evaluation occurs within each stage of the instructional systems design process. There are three forms of evaluation that instructional designers get engaged in during the instructional design process. The first type of evaluation is a needs assessment where instructional designers evaluate the gap between current performance and optimal performance to determine the needs of an organization. After instructional designers determine the need for instruction, instructional designers design instruction and develop assessment items that are matched to learning goals and objectives. Instructional designers utilize criterion-referenced measurements as an assessment of learning technique. Prior to the full-scale implementation of instruction within an organization, instructional designers then go through the process of formatively evaluating the instruction to determine mistakes or weaknesses in the instruction. Formative evaluation allows instructional designers to determine ways to modify the instruction as needed based on feedback from learners. After implementation, instructional designers summatively evaluate the instruction or training to determine the efficacy of the final product. Instructional designers use the findings from the summative evaluation to generate a report for stakeholders; so that the stakeholders can determine whether or not to adopt and institutionalize the innovation.

Problem Analysis

As an instructional designer, determining whether or not instruction needs to occur is based on factors that are analyzed at the beginning of the systematic process of instructional design. During the analysis phase, an instructional designer completes a front-end analysis. Front-end analyses are sometimes referred to as pre-training analysis, problem analysis or needs assessment (Rossett, 1987). The information that is gleaned from the front-end analysis helps instructional designers determine whether instruction is the best solution; if it is, then instructional designers analyze the needs so that instructional goals can be identified. According to Allison Rossett (1987), there are four types of causes to problems that instructional designers are faced with: an absence of skill or knowledge, an absence of incentive or improper incentive, absence of environmental support, or absence of motivation. Therefore, instruction is an intervention that presents information in a way that provides examples, practice, and provides feedback to teach something that people have never been taught, never learned, or have forgotten (Rossett, 1987). There are five purposes for completing a problem analysis:

- determine optimal performance levels or knowledge
- determine actual performance levels or knowledge
- determine feelings of all significant individuals
- determine the causes of the performance problem from different perspectives
- determine possible solutions of the performance problem from different perspectives (Rossett, 1987).

Each purpose of the problem analysis process is important because it sheds light on how to design and develop instruction that meets the specified needs of the client. Determining optimal performance levels or knowledge informs the instructional designer of the required skills and knowledge needed to complete a task well. Determining actual performance levels or knowledge

informs the instructional designer of how tasks are currently being completed. The gap between the optimal and actual performance levels indicates the performance problem or need. If the need of the organization can be resolved by an instructional solution, then the process of systematically designing instruction begins.

Assessment of Learning Outcomes: Criterion-Referenced Measurement

"A criterion-referenced measure provides information about a person's mastery of knowledge, attitudes or skills relative to the objective (Seels & Richey, 1994). Instructional designers utilize information from the front-end analyses, instructional goals and objectives, and task analysis to design and develop assessment items. The assessment items that are developed can be either norm-referenced or criterion-referenced. According to Linda Bond (1996), "These two assessment strategies differ in their intended purposes, the way in which content is selected, and the scoring process which defines how the assessment results must be interpreted."

Norm-referenced assessments items compare learners' performance to a norm group and rank learners based on their knowledge. The content for norm-referenced assessment items are constructed by subject matter experts and are standardized using a norm group (Fairtest, 2007). Criterion-referenced assessments, on the other hand, compare learners' performance to specific instructional expectations and determine whether or not a student has mastered specified skills or knowledge. In other words, criterion-referenced assessment items are derived from content that is specific to learning outcomes that correspond with the instructional goals and objectives of the intended instruction and measure learner performance without discriminating among learners (Linn & Gronlund, 2000). Instructional designers utilize criterion-referenced measurements more often than norm-referenced assessments because the findings of the criterion-referenced assessments point out whether or not learners have mastered specific knowledge or skills; if not, instructional designers can alter the instruction to ensure that it entails the desired learning outcomes.

Formative Evaluation

By completing formative evaluations, instructional designers can determine the weaknesses or problems that exist within the instruction (Dick, Carey & Carey, 2005). Instructional designers, acting as internal evaluators, often conduct three different types of formative evaluations: one-to-one, focus group, and field trial (Dick, Carey & Carey, 2005). During the one-to-one evaluations, instructional designers work with subject matter experts and potential learners to inform their revision decisions. Focus group formative evaluations require multiple learners from the target audience to study the instructional materials and then assess whether or not they understood the materials. Lastly, a field trial is conducted where the instruction is provided to learners in a real world context. The field trial should involve more learners than the focus group. Formative evaluations can use various assessment tools to monitor the effectiveness of instruction. These tools could include criterion-referenced measurements of learning outcomes, observation of learners as they complete the instruction, attitude surveys, or short interviews.

Instructional designers utilize the feedback from the formative evaluations to make any needed revisions to the instruction before it is ready for full-scale implementation. By revising the instruction, instructional designers can save the project from going over budget or getting

behind schedule. Otherwise, if instructional designers skipped over the formative evaluation stage, the instruction may be flawed during implementation; then it would cost more time and money to resolve the problems. It is only after the instructional designer goes through the revisions and makes sure that all possible problems have been fixed in the instruction and the product is ready for implementation.

Summative Evaluation

Summative evaluations occur after instruction has been implemented within an organization. The purpose of a summative evaluation is to measure long-term effects of the instruction and to determine whether or not the instructional performance problem has been resolved. Often during the summative evaluation process, a third party evaluator is hired to complete the evaluation in order to make sure that the findings of the evaluation are unbiased. There are two distinct reasons why instructional designers systematically complete the summative evaluation process.

The first reason is to determine the effectiveness of the instruction. During the evaluation phase, instructional designers often utilize Douglas Kirkpatrick's Four Levels of Evaluation Model (1998) for determining training effectiveness. Kirkpatrick's model is hierarchical. The different levels of Kirkpatrick's model guide instructional designers through the process of gathering information; then the designer builds on the information through the remaining levels. The first level of evaluation is reactions. At the reactions level, instructional designers determine how the learners felt about the training, the instructors, the environment, and it assesses learner attitudes towards the instruction (Kirkpatrick, 1998). When gathering reactions, instructional designers can figure out what worked and what did not work for the learners. Then instructional designers move to the learning level to determine the amount of knowledge that the learners have retained from the instruction/training. Often instructional designers utilize pre- and post- tests to assess learner achievement. Next, instructional designers move to the transfer level. The transfer level of evaluation helps the instructional design see whether or not the training provided the learners with information that made it possible for applying the knowledge in real world situations (Kirkpatrick, 1998). Lastly, you get to the results level and instructional designers can determine whether or not the training was in fact effective. During the last level of evaluation, instructional designers can determine the return on investment for the instructional solution. Return on investment findings help to justify the cost, time, and resources utilized to implement and utilize the instruction in the organization (Schwalbe, 2007).

Secondly, summative evaluations are utilized by instructional designers to gather information that will determine whether or not stakeholders will want to adopt an innovation for a larger scale implementation. The findings from the summative evaluation help to conclude the effectiveness of the innovation. After the instructional designer has completed the summative evaluation, a report can be generated. A summative evaluation report presents all the findings in an organized manner so that stakeholders can assess whether or not the implemented innovation met their needs. Therefore, it is important that the summative evaluation report is disseminated to the various stakeholders. After reviewing the summative evaluation report, stakeholders can determine whether or not they want to adopt the innovation. Summative evaluation findings are therefore vital to successful full-scale implementation and institutionalization of an innovation.

Appendix

Kirkpatrick's Four Levels of Evaluation Model (1998)

Kirkpatrick designed a model that was hierarchical and represented in a pyramidal shape. The are four different levels of Kirkpatrick's model that guide instructional designers through the process of gathering information then the designer builds on the information through the remaining levels.

- Reactions level - Determine how the learners felt about the training, the instructors, the environment, and assess learner attitudes towards the instruction.
- Learning level - Determine the amount of knowledge that the learners have retained from the training.
- Transfer level - Determine whether or not the training provided the learners with information that made it possible for applying the knowledge in real world situations
- Results level - Determine whether or not the training was effective by determining the return on investment for the instructional solution.
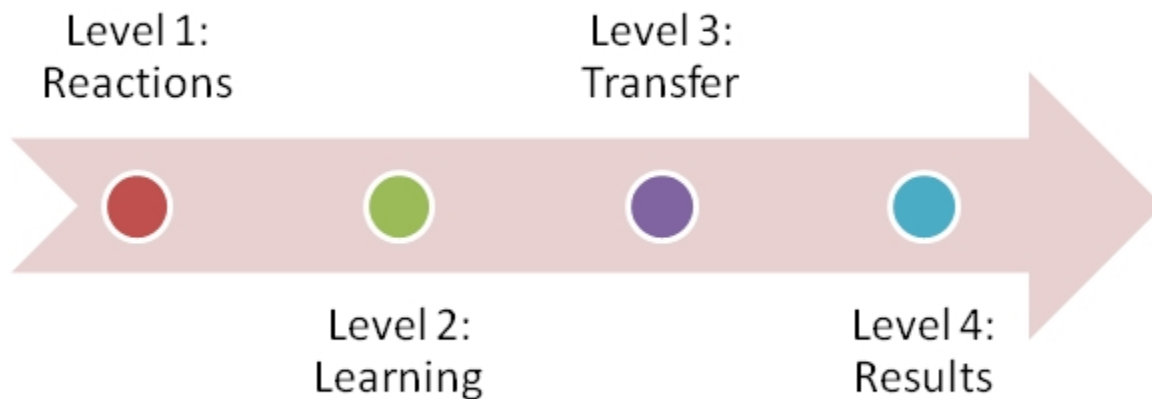
Kirkpatrick's Four Levels of Evaluation Model (adapted from D. Kirkpatrick, 1998)

References

Bond, Linda A. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research & Evaluation*, *5*(2). Retrieved April 15, 2010 from http://PAREonline.net/getvn.asp?v=5&n=2

Dick, W., Carey, L., & Carey, J.O. (2005). *The systematic design of instruction (6th ed*.). Boston, MA: Pearson.

Fairtest. (2007, August 17) *Norm-referenced achievement tests*. Retrieved April, 16, 2010 from Fairtest website: http://www.fairtest.org/norm-referenced-achievement-tests

Kirkpatrick, D.L. (1998). *Another look at evaluating training programs*. Alexandria, VA: American Society for Training & Development.

Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Prentice Hall.

Rossett, A. (1987). *Training needs assessment*. Englewood Cliffs, NJ: Educational Technology Publications.

Schwalbe, Kathy. (2007). *Information technology project management*. Boston, MA: Thomson Course Technology.

Seels, B.B. & Richey, R.C. (1994). *Instructional technology: The definitions and domains of the field*. Washington, D.C. Association for Educational Communications and Technology.