

Evaluation of technology-enhanced learning

Introduction

Evaluation is an important aspect to be considered in the implementation of technology-enhanced learning. Experts' views on evaluation vary from simple 'attainment of goals' (Tyler, 1949), to a 'critical analysis of the quality of a program' (Eisner, 1985). Though there are certain differences in the numerous views expressed over the years, the main emphasis in any evaluation is making decisions about the quality of something. As Kifer (1997, p. 384) describes, evaluation is, 'a disciplined inquiry to determine the worth of things, where *things* may include programs, productions, procedures or objects'. In this Essential Reading, we shall discuss the key concepts associated with evaluation, particularly in a technology-enhanced learning environment.

When evaluating technology-enhanced learning, evaluation methodologies common to all types of evaluation can be used. The main concern of the evaluation would be obtaining accurate information to improve the quality of the technology-enhanced learning environment. As such, careful planning is essential in order to conduct a meaningful evaluation.

In this reading you will be looking at what evaluation is, purposes of evaluation, different approaches and strategies for evaluation, and finally, how to plan an evaluation in a technology-enhanced learning environment.

What is evaluation?

Evaluation has been defined in many ways, by different people, in different situations. For instance, evaluation is defined as: the determination of the degree to which objectives have been attained by a program (Tyler, 1942), judgment of a program's worth (Scriven, 1967), description of a program's inputs, processes and outcomes (Stake, 1967), the ethnographic investigation of a program (Guba & Lincoln, 1981), and the critical analysis of a program's quality (Eisner, 1985).

All these definitions, though there are certain differences in the ideas expressed, share a common emphasis. That is, making decisions about the *quality* of a program.

Although early definitions focused mainly upon program evaluation, now, many aspects to evaluation are existent, such as, product evaluation, process evaluation, personnel evaluation and performance evaluation. Consolidating many earlier thoughts, Reeves, (1991, p.85) defined evaluation as, "the process of providing information to enlighten decision-making that will improve the quality of life".

Examining the views discussed above, we can identify some key concerns in any evaluation:

- Collecting information
- Making decisions
- Determining the quality of a product, program, and/or people

Evaluation in a technology-enhanced learning environment will also need to focus upon these aspects. These concerns decide the steps to follow in an evaluation. An overview of the steps of a 'typical' evaluation is given in Figure 1.

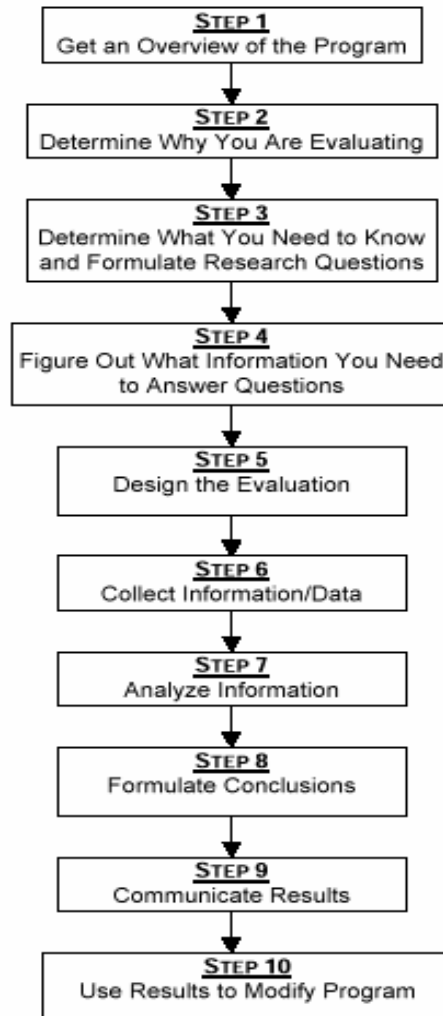


Figure 1 - Evaluation Overview

(Source: An Educator's Guide to evaluating the use of technology in schools and classrooms, 1998)

As Figure 1 indicates, in an evaluation, we collect data according to the determined purposes and then formulate conclusions regarding the program, in order to modify it if necessary. However, Scriven (1991) argues that evaluation is not the mere accumulation and summarizing of data for decision-making, but it also constitutes collecting, clarification and verifying relevant values and standards. This is an important fact to keep in mind when planning and conducting an evaluation.

Purposes of Evaluation

Before planning an evaluation procedure, first, it is important to identify the purposes of the evaluation. As definitions of evaluation indicate, the ultimate purpose of any evaluation would be to improve the quality of a product or a program. As Stufflebeam et al (1971) claimed, the purpose of evaluation is not to *prove*, but *improve*.

A single evaluation may have multiple purposes. For instance, evaluation may be performed to:

- Make decisions about individual learners (What are the learner needs? How should the instruction be planned or designed?)
- Make decisions about course improvement. (Decide what materials and methods are the most suitable, and how to revise the materials)
- Make administrative decisions about a program (Determine how effective the program is, how well individual learners or teachers function in specific situations and what impact the program/product is having on participants)

An evaluation will provide information to program personnel and others on aspects of the program that work well and potential problems. Identifying potential problems early in the program will be useful to correct them before more serious problems occur. It will also provide information on what technical assistance may be needed.

Further, an evaluation may be useful to bring in to light issues that need to be examined in greater detail. Hence, an initial evaluation of program implementation may be used, in part, to guide a later evaluation of long-term impact.

Approaches to evaluation

During the past few decades a large number of evaluation models have been developed and presented. Guba and Lincoln (1989) analyzed the changes taken in the approaches for evaluation, as four generations. According to them, the first generation evaluation was concerned with measurement of various attributes of students; the second generation evaluation was characterized by descriptions of patterns of strengths and weaknesses with respect to stated objectives (eg: Tylerian Model, 1940); and in the third generation evaluation, judgment was an integral part, where evaluator be the judge. [eg: Models of Stake (1967), Stufflebeam et al. (1971), Scriven (1973), Eisner (1979)]. Guba and Lincoln (1989) presented Responsive-Constructive approach to evaluation as the fourth generation evaluation. In this form of evaluation, the claims, concerns and issues of the stakeholders serve as the basis for determining what information is needed, where constructivist methodology is employed.

Examining the evaluation models presented throughout the years, Alexander and Hedberg (1994) categorized four main approaches to evaluation: objective-based, decision-based, values-based and naturalistic. The objective-based approach determines the degree of achievement of educational objectives (Tylerian Model, 1940). The decision-based approach looks at the program in terms of Context, Inputs, Process and Product (Stufflebeam et al. (1971). Values-based approaches are more concerned with overall merit or worth of products [Models of Stake (1967), Scriven (1973)]. In contrast to all these approaches which concentrate mainly on the product or outputs, the naturalistic approach (Guba & Lincoln, 1985) organize evaluation around the participants' concerns and issues.

Basic information about some of the models and their potential in the evaluation of technology-enhanced learning are summarized in Table 1.

Table 1: Models of evaluation and their potential in the evaluation of technology-enhanced learning

Evaluation Model	Key Ideas	Potential in evaluating technology-enhanced learning
Objectives-Based Evaluation Model Ralph W. Tyler (1940s)	<p>Evaluation consists of the measurement of whether the objectives of an educational program, project, or product are accomplished.</p> <p>Does not allow for unintended outcomes.</p>	<p>Objectives provide the basis for the development of measurement procedures and instruments that can be used to evaluate the effectiveness of a technology-enhanced innovation.</p>
Values-based Evaluation Model Michael Scriven (1960s)	<p>Is concerned with not only in achieving goals, but also about whether they are worth achieving.</p> <p>Distinguishes between formative and summative evaluation.</p> <p>Unintended outcomes are acknowledged.</p>	<p>Formative evaluation allows a technology-enhanced program to be improved at its developing stages, and summative evaluation can be used to make decisions regarding future use of it.</p>
Responsive Evaluation Model Robert Stake (1970s)	<p>Evaluation methods are negotiated by the actual “stakeholders” in the evaluation.</p> <p>The continuous nature of observations and reporting is an essential element.</p>	<p>Continuous feedback received from multiple audiences affected by a technology-enhanced program or product is valuable in making judgments on it.</p>
CIPP Model Daniel Stufflebeam (1970s)	<p>Facilitates decision-making at all stages of a development cycle.</p> <p>Four types of evaluation are used: Context, Input, Process & Product.</p>	<p>Useful to evaluate large-scale technology-enhanced program, through multiple levels.</p>
Naturalistic Evaluation Model Yvonne Lincoln & Egon Guba (1980s)	<p>Mainly focuses on the effect of a program on participants.</p> <p>Qualitative methodology is entirely used for data collection, and multiple methods are used.</p>	<p>Is a useful approach to gain an insight in to the real situation of a technology-enhanced learning process.</p>
Connoisseurship Evaluation Model Elliot W. Eisner (1980s)	<p>Education should have its own connoisseurs, individuals with refined tastes and sensitivity to educational phenomena.</p> <p>Educational connoisseurs or critics have the responsibility to convey their values to the public.</p>	<p>The use of expert reviewers as an evaluation strategy has much in common with educational connoisseurship and criticism.</p>
Multiple-Methods Evaluation Model	<p>Multiple methods are selected for investigating a complex program that cannot be adequately assessed with a single method.</p> <p>Involves using multiple measures to converge on a more accurate estimate of a variable. Triangulation is used.</p>	<p>As most technology-enhanced innovations will involve complexity, multiple methods of evaluation is useful. It can produce a more realistic representation of effectiveness and impact of the program.</p>

Types of Evaluation

Form and function of the evaluation process is referred to as types of evaluation. This is identified by the object or process being evaluated, and the purpose of the evaluation (Naidu, 2003). Evaluation may focus either on the instructional process or on the product of the instruction. Process evaluation examines the treatment or methodology used to reach the objectives of a program, whereas product evaluation looks at the functional artifacts or results of a program after its implementation (Scriven, 1991). In process evaluation, the evaluation is performed while learning is occurring, and examines the process itself. Product evaluation is usually performed after the instruction has taken place.

Often, two main types of evaluation are discussed: formative and summative. Formative evaluation is usually conducted during the time the instructional design is being developed and materials produced. It is, basically, testing an instructional method or material on learners, while there still is the possibility for modifying it. This concept was earlier called "developmental testing," "product tryout," and "learner verification." The primary purpose of formative evaluation is to improve the process or instructional methods and their products. Producers of instructional materials can improve their presentations by the use of formative evaluation procedures.

Summative evaluation is performed near the conclusion of the teaching/learning process to draw inferences or conclusions about the effectiveness of the program. Formative evaluation examines the outcomes and/or impacts of the process rather than the product, while summative evaluation focuses on the product itself. Thus, although both types of evaluation examine the learner, the teacher and the instructional design, their focus is different, and the data sources are also unique for each form of evaluation.

The instructional design is formatively examined for the suitability of the educational method, level of objectives, grading system, use of feedback to the student, media selection, etc. Summative questions regarding the instructional design address its costs, logistics requirements, ease of maintenance, and level of acceptance by students, teachers, and others concerned.

Process evaluation tends to focus on formative questions about what does and doesn't seem to need revision. However, summative process questions about time, cost, and acceptance can also be examined. Product evaluation tends to focus on summative questions, but can also examine whether the content is accurate or obsolete, which are formative questions.

In addition to the formative and summative evaluation, there are other types of evaluation found in the literature. Illuminative evaluation is described as an open-ended method associated with ethnography where information is gathered while associating with the participants. It is a systematic focus on discovering the 'unexpected'. Integrative evaluation (Draper, 1996). Integrative/Monitoring evaluation is an on-going process of data collection which is carried out as part of post-implementation phase of a program, in order to make improvements to the next iteration of the innovation, and seeks to assess the integration of the innovation (Naidu, 2003).

Designing an Evaluation

The design of an evaluation needs to be based upon the purposes of the evaluation. Hence, it is important when planning an evaluation, to be specific about what is to be evaluated. You have to decide if the evaluation will be formative, summative, or integrative, and also determine what strategies should be used. The main steps to follow in conducting an evaluation were looked at earlier. (See Figure 1) The evaluation plan should be made according to these steps.

An evaluation design should generally include:

- The evaluation question(s)
- The evaluation criteria
- The evidence which is to be used in the evaluation (the types of questions or test instruments which are going to be used)
- The role subjective judgment is to play in the evaluation (who is going to make which subjective judgments about the effectiveness of the program)
- A plan for implementing the evaluation design.

Reeves (1992) highlighted the importance of evaluating multimedia within the context of its use and the characteristics of users. An evaluation framework for educational multimedia proposed by Pham (1998) discusses three perspectives that are important to examine: the product itself, how it is used and the impacts it exerts. The quality of a product is determined by examining the knowledge content, the ways knowledge and tasks are represented and organized, and technical tools used for conveying and constructing knowledge.

Use of a multimedia product is mainly influenced by human-computer interface issues such as choice of media and proper use of media, and also navigation and links to information structures. Finally, the impacts of the products concern, whether the learners have achieved the intended objectives, performance of them in comparison with other similar products and any unintended outcomes.

Data collection techniques for evaluation

Evaluation involves the collection and use of information to make decisions about an instructional program. Collecting information for an evaluation can be accomplished in many ways. Selection of a method will depend mainly on the purposes of evaluation.

Some of the commonly used methods, that may be applicable in collecting data for evaluation of technology-enhanced learning environments, are briefly introduced below. .

- Achievement tests - These tests measure student knowledge over a wide spectrum of subject areas.
- Checklists - These are a structured lists where specific information about instructional courseware can be obtained
- Questionnaires - These are self-administered surveys consisting of sets of questions.
- Rating scales - These can be used for evaluation of individuals, events, or products. Learner attitudes, for example, can be rated on a five point scale from one to five.

- Ranking scales - A set of items are arranged into a hierarchy according to a value or preference.
- Semantic differentials - These may be used to measure attitudes and affect according to the indirect meanings of words. For instance, the format of a semantic differential question might be as follows:

Instructional Technology is -

Good --- --- --- --- Bad

Powerful --- --- --- ---Weak

Desirable --- --- --- --Undesirable

Effective --- --- --- ---Ineffective

- The Q-Sort - This permits individuals to rate items or statements by prioritizing them, for instance, from "very good" to "very bad".
- Diaries - Individuals are required to keep hourly, daily, or weekly accounts of specific activities, attitudes, thoughts or events.
- The critical incident technique - This requires the recording only of particularly important, unique or useful information. Emphasis is placed on recording or reporting those incidents or situations that seem to make a significant difference in system performance.
- Observations - These can be used to determine what instructional materials are being used. Eyewitness observation, self-completed checklist, rating scales, field notes, and summary reports are all examples of observation instruments.
- Interviews: These may be performed with individuals or a group. Interviews can be either unstructured or highly structured. A face-to-face interview permits the probing of sensitive issues like attitudes or values. Telephone interviews, although less sensitive to attitude and value information than face-to-face interviews, are less time-consuming.

- Performance tests - These require individuals to complete a task. Evaluation may focus on the performance itself, the end-product of the performance, or both. Rating scales may be used to evaluate the performance.
- Record reviews - This method can be used to access individual achievement, using score data, such as tests and grades.
- Self-reporting measures - These ask individuals to express their attitudes, beliefs, perceptions and feelings.

A technique that closely matches with the objectives of the evaluation should be selected and used. Generally, in practice, more than one type of strategies may be used. It is important to utilize close-ended as well as open-ended strategies to obtain information.

Conclusion

Evaluation is a key skill that any educational technologist must master. Evaluating technology-enhanced innovations need careful planning, in concert with evaluation theories. The evaluation design should be planned at the same time as the instructional design. Evaluation of both materials (products) and complete instructional process is a key to the success of any instructional activity.

References

- Alexander, S. & Hedberg, J.G. (1994). *Evaluating Technology-based learning: Which model?* In K. Beatie, C. McNaught & S. Willis (Eds.) *Interactive Multimedia in University Education: Designing for Change in Teaching and Learning*. Amsterdam: North Holland, Elsevier, pp 233-244.
- Draper, S. W. (1996). *Observing, Measuring or Evaluating Courseware*. [Online] Available URL: <<http://www.psy.gla.ac.uk/~steve/Eval.HE.html>> Accessed: 21.07.2005
- Kifer, E. (1995). *Evaluation – A general view*. In J.G. Anglin (Ed.) *Instructional Technology – past, present and future* (pp. 384 – 391) (2nd ed). Englewood, Colo: Libraries Unlimited
- Naidu, S. (2003). *E-learning: A guidebook of principles, procedures and practices*. New Delhi: CEMCA
- Pham, B. (1998). Quality evaluation of educational multimedia systems. *Australian Journal of Educational Technology*, 1998, 14(2), 107-121
- Reeves, T.C. (1991). Ten commandments for the evaluation of interactive multimedia in higher education. *Journal of Computing in Higher Education*. Spring, 1991. Vol. 2(2), 84-113.
- Scriven, M. (1991). *Evaluation Thesaurus*. (4th ed.). London: SAGE Publications.

Recommended Readings

Alexander, S. & Hedberg, J.G. (1994). *Evaluating Technology-based learning: Which model?* In K. Beatie, C. McNaught & S. Willis (Eds.) *Interactive Multimedia in University Education: Designing for Change in Teaching and Learning*. Amsterdam: North Holland, Elsevier, pp 233-244.

Available URL: <<http://ncode.uow.edu.au/info/orig/alexander.html>>

Lefoe, G. & Gunn, C. (Eds.) (2004) *Evaluation*

Available URL: <<http://ncode.uow.edu.au/evaluation.html>>

This is a useful collection of web links on different aspects related to Evaluation.

It Includes:

Articles on Evaluation <<http://ncode.uow.edu.au/evalart.html>>

Evaluation Case Studies <<http://ncode.uow.edu.au/evalcase.html>>

Evaluation Strategies and Methodologies
<<http://ncode.uow.edu.au/evalstrat.html>>

Evaluation Tools.<http://mime1.marc.gatech.edu/MM_Tools/evaluation.html>

Evaluation Cookbook

Available URL: <<http://www.icbl.hw.ac.uk/Itidi/cookbook/contents.html>>