A Model for Evaluating Online Programs

Amy J. Nelson
Minnesota State College—Southeast Technical, USA

INTRODUCTION

According to the Joint Committee on Standards for Educational Evaluation (1994), "Program evaluation is a systematic investigation of the worth of an ongoing or continuing distance education activity" (Simonson, 1997, p. 88). As such, this article addresses the issue of evaluating programs rather than courses. Although it is true that content, instructional design, and delivery greatly affect the quality of the program, course evaluation is a topic in and of itself. Frydenberg (2002) noted that program evaluation was frequently listed as a separate item in standards documentation: "While assessment of student achievement is normally described as part of instructional design and tied to specific course objectives, program evaluation is an activity that incorporates all the aspects of the e-learning experience" (p. 7). High-level aspects of course design are, however, built into program evaluation as you will see because it is impossible to evaluate an educational program without looking at courses.

RESEARCH ON EVALUATING ONLINE PROGRAMS

"The literature on e-learning program evaluation is naturally skimpy, since few fully developed programs have arrived at a stage where summative evaluation is possible" (Frydenberg, 2002, p. 11). This does not prevent, however, the ability to develop a formative and summative evaluation structure because the process for evaluating a distance learning program should be founded on the principles of evaluating any educational program. According to the *Program Evaluation Standards* established by 16 professional associations, sound evaluations (of educational programs, projects, and materials) in a variety of settings should have the following four attributes (ERIC, 1995):

1. **Utility:** These standards are intended to ensure that an evaluation will serve the information needs of intended users. These include identifying all stakeholders, selecting a trustworthy and competent evaluator, collecting information that addresses pertinent questions and meets the stakeholders' needs, clarifying value judgments, describing the program being evaluated clearly, distributing significant interim findings and evaluation reports in a timely fashion, and reporting findings in ways that encourage follow-through by stakeholders.
2. **Feasibility:** These standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal. These include developing practical procedures to keep disruption to a minimum, anticipating stakeholders' views to gain their cooperation and curtail bias or interference, and conducting an efficient evaluation that keeps costs down while producing information of value.
3. **Propriety:** These standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results. These standards include designing evaluations that effectively serve the needs of the targeted participants, getting obligations agreed to in writing, respecting and protecting the rights and welfare of human subjects, respecting human dignity and worth in interactions with others, conducting a complete and fair examination that includes the program's strengths and weaknesses, disclosing findings to all stakeholders, addressing any conflicts openly and honestly, and applying sound accountability procedures.
4. **Accuracy:** These standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth of merit of the program being evaluated. These include documenting the program clearly and accurately, examining the context in which the program exists, describing the purposes and procedures of the evaluation

in enough detail so they can be identified and defensibly assessed, using valid and reliable information, reviewing quantitative and qualitative information systematically to find errors and effectively answer evaluation questions, justifying conclusions, avoiding distortion caused by personal feelings and bias, and evaluating the program formatively and summatively.

There are several types of program evaluation. Three common types used in non-profit and for-profit organizations alike are goals-based, process-based, and outcomes-based evaluations. Evaluations may often include a mix of these methods. The type or types used is determined by answering several key questions (McNamara, 1998):

1. What are the purposes for conducting an evaluation?
2. Who are the audiences for the information from the evaluation?
3. What kinds of information are needed to make the decision you need to make and/or enlighten your intended audiences?
4. From what sources should the information be collected?
5. How can that information be collected in a reasonable fashion?
6. When is the information needed?
7. What resources are available to collect the information?

Goals-based evaluations are used to determine the extent to which programs are achieving their overall, predetermined objectives, which are often described in the original program plans. Questions may include (McNamara, 1998): How were the program goals established? Was the process effective? Do personnel have adequate resources (money, equipment, facilities, training, etc.) to achieve the goals?

Process-based evaluations are used to understand how programs really work, as well as their strengths and weaknesses, which are useful in long-standing programs, those that have changed over the years, those coupled with stakeholder complaints or questions, or those inundated with inefficiencies. Questions may include (McNamara, 1998): On what basis do customers decide what products or services are needed? How do customers or clients come into the program? What typical complaints are heard from employees and/or customers?

Outcomes-based evaluations are used to determine the extent to which the program benefits its clients, which often include enhanced learning (e.g., knowledge, attitudes, skills) or conditions (e.g., increased literacy, self-reliance, etc.). Questions may include: What percent of graduates found related work or transferred? How does program retention compare to the norm? How does the GPA of participants compare to the norm?

Several organizations have developed standards for distance education institutions, programs, and courses. Most of these are from accrediting agencies and are very similar in content. In a summary of published quality standards in the US, Frydenberg (2002) found nine common themes: institutional commitment, technology, student services, instructional design and course development, instruction and instructors, delivery, finances, regulatory and legal compliance, and evaluation (not assessment). The Institute for Higher Education Policy (IHEP), for example, identified the following three *Evaluation and Assessment Benchmarks* (IHEP, 2000, p. 26):

1. The program's educational effectiveness and teaching/learning process is assessed through an evaluation process that uses several methods and applies specific standards.
2. Data on enrollment, costs, and successful/innovative uses of technology are used to evaluate program effectiveness.
3. Intended learning outcomes are reviewed regularly to ensure clarity, utility, and appropriateness.

In a similar effort to establish over-arching standards, the eight regional commissions that govern accreditation of higher learning in the US adopted and implemented common standards for distance learning called the *Principles of Good Practice in Electronically Offered Academic Degree and Certificate Programs,* originally developed by the Western Cooperative for Educational Telecommunications (WCET) and more recently detailed to fully explain each standard. Among these standards are those pertaining to program effectiveness (WCET, 2001, p. 15):

1. The extent to which student learning matches intended outcomes, including for degree programs