Online Education Evaluation: What Should We Evaluate?

Khe Foon Hew
Shijuan Liu
Ray Martinez
Curt Bonk
Indiana University

Ji-Yeon Lee
University of South Carolina

Introduction

In recent years, Web technologies have helped expand distance education in higher education. Online programs have become widely and generally accepted in many countries (Moore & Anderson, 2003; Ratcliff et al., 2001). According to the U.S. Department of Education (1999), the number of Institutions of Higher Education (IHEs) offering distance education has increased by one-third since 1994-1995; and 44 percent of IHEs offer distance courses. The number of enrollments reached 1.3 million credit hours, and thirty percent of all distance education courses now use the Web (cited in Swail & Kamptis, 2001). A recently conducted survey found that the number of online learners in fall 2002 topped 1.6 million, with public institutions leading the way (having approximately 1.3 million online learners) (Allen & Seaman, 2003). Moe and Blodget (2000) further noted that due to the increasing demand and the limited access to high-quality postsecondary education in many parts of the world, there could be as many as 40 million online education learners by 2025. Given this great increase in online education programs, performing evaluation studies on such programs has become imperative (Thompson & Irele, 2003). The main purpose of this short paper is to describe the evaluation of current online education at three levels: (a) macro-level, referring to the evaluation of entire online programs; (b) meso-level, referring to the evaluation of individual online courses; and (c) micro-level evaluation, referring to the evaluation of online students' learning. At each level, relevant literature, pertinent evaluation issues, and questions are discussed. While many issues addressed in the paper are applicable to both educational and corporate settings, the paper is focused on the higher education setting.

Terminology

Online Education

Many terms have been found, both in research and practice, to refer to teaching and learning over the Internet. These include distance education and online education. Moore (2003) defined distance education as "all forms of education in which all or most of the teaching is conducted in a different space than the learning, with the effect that all or most of the communication between teachers and learners is through a communication technology" (p.xiv). Distance education, under this definition, then includes non-Internet technologies such as video or correspondence learning. Since the three evaluation levels being discussed (i.e., programs, courses, and student learning) are all Internet-based, the term online education will be used.

Evaluation

There have been many arguments about the correct use of the term evaluation versus the term assessment. One position is that evaluation is different from assessment. Some researchers, especially in the North America, use the term evaluation to refer to studies implemented to examine and report on the strengths and weaknesses of programs, policies, organizations, and the like to improve their effectiveness (American Evaluation Association, 1999); Assessment, on the other hand, is used to refer to "the formation of value judgments to determine the significance, the importance, the value of learning and knowing", and using "a variety of procedures to obtain information about individual's learning" (Delandshere, 2003). Another position regards assessment as a subset of evaluation and a valuable tool in the larger evaluation activity (Rowntree, 1992; Thompson & Irele, 2003). As Thompson & Irele (2003) argued, "Assessment asks 'How much?' whereas evaluation asks 'Is it good enough?' and 'If not, why not?'" In this paper, we used evaluation as a generic term to include evaluation done at the program level, the course level, and the student learning level.

Macro-level evaluation

Macro-level evaluation refers to the evaluation of the entire online education program. Macro-level evaluation helps ensure the accountability and quality of a program (Fitzpatrick, Sanders, & Worthen, 2004; Rutnam & Mowbray, 1983). Such evaluation is important for an online education program for several over-arching purposes which include: (a) justifying the investment of resources, (b) measuring progress toward program objectives, (c) measuring issues of quality and effectiveness, (d) providing a basis for improvement, and (e) informing institutional strategic planning and decision making (Thompson & Irele, 2003). From these general purposes, several pertinent evaluation questions can be drawn.

- Is the program consistent with the institutional mission (Law, Hawkes, & Murphy, 2002)?
- Does the program have clear goals and objectives, and are there agreed upon provisions and measures for program oversight, accountability, evaluation and assessment (Law, Hawkes, & Murphy, 2002)? Agreed upon methodologies for evaluation may be models such as I-P-O, CIPP, and AEIOU (Koessl, 2003).
- Are resources allocated for infrastructure and communication technologies, faculty training and support, and student recruitment, retention and support (Sherry, 2003)? Are criteria in place when deciding upon the allocation for competing interests?
- Is the program cost effective (Rovai, 2003; Lorenzo & Moore, 2002; Twigg, 2003)? Can a positive return on investment be shown for fiscal costs?
- Is the technology delivery system as reliable and secure as possible, include plans for backup and disaster recovery (IHEP, 2000).

Meso-level evaluation

Meso-level of evaluation refers to the evaluation of individual online education courses. Central to this level of evaluation is the question of what criteria should be used to evaluate the online courses and instruction.

Whereas the Seven Principles for Good Practice, developed by Chickering and Gamson (1987), were originally intended for undergraduate education, they have been used widely in the evaluation of online courses, including those taught at the graduate level (Moore & Anderson, 2003). In addition to the seven principles, Graham, Cagiltay, Craner, Lim, & Duffy (2000) added four principles related to human computer interface (HCI) design when evaluating four graduate level online courses. Based on a relatively comprehensive literature review of online course evaluation, Achtemeier, Morris, and Finnegan (2003) generated eleven evaluation questions and used them to examine course evaluation instruments that were employed with online courses or web-enhanced courses from thirteen institutions. In addition, other researchers and institutions have developed many guidelines and principles for the design and development of courses. Although the guidelines and principles were not developed (only) for the purposes of course evaluation, they shed great light on the selection of evaluation criteria (Achtemeire et al, 2003). Especially helpful resources are five principles for effective online instruction by Hacker & Neiderhauser (2000); thirteen faculty guidelines by Sherry (2003); and an emerging set of guiding principles and practices for the design and development of distance education by the Innovation in Distance Education Project (n.d.).

Based on the available literature and discussion with colleagues, the second author of this paper has identified and developed ten key questions (criteria) for the evaluation of online courses, The questions have been employed to evaluate an online course offered to undergraduate students (preservice teachers) (Liu, 2003). The ten key questions are:

- Are course objectives, instructor's expectations, and evaluation criteria of assignments well communicated to the learner?
- Does the course provide students sufficient support (including instructional and technical) for meeting the course objectives and other relevant needs of students?
- Does the class encourage students to be active learners?
- Does the online course encourage learner-instructor contact?
- Does the online course encourage interaction and collaboration between and among learners?
- Does the course encourage meta-cognitive skills?
- Is learning grounded in effective (contextual, authentic) examples?
- Is the instructor's feedback appropriate for supporting learner learning (e.g., not too much nor too little, just in time, etc.)?
- Does the class use multiple evaluation methods aligned with course objectives and designed activities?

- Does the course use technology effectively in supporting learning and teaching?

Further research on meso-level evaluation is still desired (Achtemeire et al, 2003; Bonk, 2003).

Micro-level evaluation

The focus of micro-level evaluation is on the individual online learner. A review of the relevant literature suggests that evaluation of the individual learner typically falls into one or more of the following three categories: (a) the learner's perception of online learning, (b) the learner's process of online learning, and (c) the learner's product from online learning.

Learner's Perception

When learners enroll in an online course, they enter into an educational experience very different from the usual classroom-based face-to-face environment. Current online education courses are usually mostly or entirely text-based, asynchronous, and have multiple discussion threads. As such, course administrators and instructors are usually interested to know how their learners "feel" about the course experience. Because of the difficulty of reaching online learners who are often separated by space or time, evaluation of learners' perception typically involves questionnaire surveys. Some of the common evaluation questions pertinent to this purpose include:

- Does the learner enjoy the whole course?
- What attitudes do the learners have towards online learning before, during, and after the whole course?

Learner's Process

While the learner's perception of online learning can be helpful and useful information, it is usually not sufficient to end just there. Most instructors would also want to know about their learners' engagement in online learning through many different processes such as collaboration, cognition, problem solving and others. One common method of determining such processes is through a content analysis of the learners' online discussion transcripts. Table 1 lists some of the content analysis models for evaluating these processes.

Table 1. *Content Analysis Models for Evaluating Student Processes in Online Education*

Processes	Content analysis model
Cognitive processes	Henri (1992)
	Newman, Johnson, Webb, & Cochrane (1997)
Meta-cognitive processes	Henri (1992)
Social construction of knowledge	Gunawardena, Lowe, & Anderson (1997)
Collaboration processes	Murphy (2004)
Problem solving processes	Murphy (2004)

The use of content analysis, however, is not without some problematic issues. Some of these issues include:
- What kind of unit of analysis should be used - entire message postings, thematic units, or sentences?
- How is high reliability insured in content analysis?
- How are "passive" learners evaluated? Passive learners, as found in a study by Sutton (2000), do not participate often in the discussion but consider themselves to have learned a lot from reading and reflecting on the comments and responses posted by others.
- How is the validity of the content analysis insured? Do the indicators of the content analysis models describe what they purport to describe (Krippendorf, 1980)?

Learner's Product

Typically the evaluation of the learner's product of online learning is used to determine how much knowledge or skills he or she has acquired at the conclusion of the course. This is typically administered in traditional ways such as end-of-course tests, final papers, and final projects. Alternative methods include portfolios and performances.

Summary

This paper describes three levels of evaluation (macro, meso, and micro) pertaining to online education in the higher education setting. Central to the macro-level evaluation is the question of whether the entire online education program is able to address five over-arching purposes which include (a) justifying the investment of resources, (b) measuring progress toward program objectives, (c) measuring issues of quality and effectiveness, (d) providing a basis for improvement, and (e) informing institutional strategic planning and decision making. Meso-level evaluation is related to the evaluation of individual online education courses. In this paper, we identified ten pertinent criteria that may be useful to practitioners who wish to evaluate at this level. Finally, micro-level evaluation focuses on the individual online learner. This level can be divided into three different areas: (a) the learner's perception of online learning, (b) the learner's process of online learning, and (c) the learner's product from online learning.

References

Achtemeier, S. D., Morris, L. V., and Finnegan, C. L. (2003). Considerations for developing evaluations of online courses. *Journal of Asynchronous Learning Networks*, *7*(1). Available: http://www.sloan-.org/publications/jaln/v7n1/pdf/v7n1_achtemeier.pdf

American Evaluation Association. (1995). Guiding principles for evaluators. *New Directions for Evaluation*, *66*, 19-26.

Bonk, C. J. (2003 Fall). Lectures given in the course *P600/R685 Online Learning Pedagogy and Evaluation*. School of Education, Indiana University: Bloomington.

Caffarella, R. S. (2002). *Planning programs for adult learners: A practical guide for educators, trainers, and staff developers*. San Francisco, CA: Jossey-Bass.

Delandshere, G. (2003 Fall). Class handout distributed in the course *Y527 Educational Assessment and Psychological Measurement*. Indiana University: Bloomington.

Institute for Higher Education Policy. (2000) *Quality on the line: Benchmarks for success in Internet-based distance education*. Washington, DC: The Institute for Higher Education Policy.

Graham, C., Cagiltay, K., Craner, J., Lim, B., and Duffy, T. M. (2000), *Teaching in a Web based distance learning environment: An evaluation based on four courses*. CRLT Technical Report No. 13-00. Available: http://www.crlt.indiana.edu/publications/workpapers/crlt00-13.pdf

Hacker, D. J.,& Neiderhauser, D.S. (2000). Promoting deep and durable learning. In R.E., Weiss, D., Knowlton, & B., Speck, (Eds.), Principles Of effective teaching in the online classroom. [Special issue]. New Directions for Teaching and Learning, 84.

The Innovations in Distance Education Project (n.d.). *An emerging set of guiding principles and practices for the design and development of distance education*. Retrieved December 10, 2003 fromhttp://www.worldcampus.psu.edu/ide/docs/guiding_principles.pdf

Koessl, D. (2003) *Review of research and research resources related to distance education and program evaluation*. Retrieved October 8, 2004 from http://cstld.colostate.edu/filesanddocument/CLTDBibliography1.pdf.

Krippendorf, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.

Law, J., Hawkes L., & Murphy, C. (2002). Assessing the on-line degree program. *New Directions for Teaching and Learning, 91*, 83-89.

Liu, S. (2003). *Evaluation of an online course*. An unpublished manuscript.

Lorenzo, G. & Moore, J. (2002). *The Sloan Consortium report to the nation: Five pillars of quality online education*. Needham, MA: Sloan Center.

Moe, M. T., & Blodget, H. (2000). *The knowledge web: Part 1. People power: Fuel for the new economy*. New York: Merrill Lynch.

Moore, M. (2003). This book in brief: Overview. In M. Moore & W. Anderson (Eds.). *Handbook of Distance Education* (pp.xiii-xxiii). Mahwah, NJ: Lawrence Erlbaum Associates.