Effective Strategies to Assess the Impact of e-Learning

Barbara M. Olds, Colorado School of Mines

Abstract

As part of the recent emphasis on outcomes assessment in higher education, careful assessment of e-learning becomes increasingly important both for accreditation and accountability purposes. This paper introduces a proven assessment process that is both thorough and flexible. The process makes use of an assessment matrix consisting of seven components — objectives, learning outcomes, performance criteria, implementation strategies, evaluation methods, timeline, and feedback. Each of these components is discussed in some detail. The paper concludes with lessons learned about assessment as well as speculation about the future of assessment in e-learning.

I. Introduction

A genuine revolution in higher education assessment is taking place. As part of this revolution, an emphasis on <u>inputs</u> (e.g., number of credits taken in a subject, "seat time," number of books in the library) has been replaced with a focus on <u>outcomes</u> (e.g., what students actually know and are able to do). In the U.S., one of the leaders in this effort has been the Accreditation Board for Engineering and Technology (ABET) [1]; regional accrediting agencies in the U.S. are also focusing on the assessment of learning outcomes as a major part of the accreditation process [2]. Given this new emphasis on assessment, e-learning practitioners have an opportunity to develop cutting edge assessments and serve as a model for the rest of the education community.

I believe that the theories and methods that have been applied successfully in assessing traditionally delivered courses and programs apply equally well to technology-enhanced learning and teaching. In fact, technology is already widely used in developing and/or adapting assessment tools such as electronic portfolios, on-line surveys, etc. [3]. In this paper I hope to provide a model of a successful and flexible assessment process, discuss a few of the common assessment methods and their adaptability for e-learning, and speculate about the future for assessment of e-learning. [*]

A UNITED ENGINEERING FOUNDATION CONFERENCE
Davos, Switzerland 11-16 August 2002
http://www.coe.gatech.edu/eTEE

II. Assessment, Accreditation and e-Learning

E-Learning is obviously becoming more widespread each year. A 2002 report entitled "Accreditation and Assuring Quality in Distance Learning," from the Council for Higher Education Accreditation (CHEA) [4] reports that 5,655 institutions are accredited by the 17 institutional accreditors, regional and national, in their study. Of these, 1,979 institutions offer a form of distance-delivered learning program or courses, some leading to degrees. CHEA asserts that "standards, guidelines, and policies to determine academic quality are in place for the scrutiny of distance learning." Accrediting organizations routinely review seven key areas of institutional activity when examining the quality of distance learning including institutional mission, organization, resources, curriculum & instruction, support for students and faculty and <u>student learning outcomes</u>. An example of guidelines for examining the quality of student learning outcomes quoted by CHEA comes from the Accrediting Council for Independent Colleges and Schools (ACICS): "Requirements for successful course completion must be similar to those of residential courses and programs. Assessment of student performance must demonstrate outcomes comparable to those for residential programs. The institution must document that it conducts course/program evaluations, including assessment of educational outcomes, student retention and placement, and student, faculty, and employer satisfaction." In brief, the use of e-learning is increasing among colleges and universities and accreditors are responding with standards designed to ensure quality, including an insistence on assessment of student learning outcomes. However, a recent study found that only 41 percent of engineering instructors who use the Internet for instruction report that they evaluate the Internet components of their courses [5]. Clearly, more must be done.

---

[*]In this paper I am using the term "e-learning" to apply in a broad sense to all technology-enhanced learning and teaching including email, simulations, on-line courses, and courseware management tools, as well as hybrids of all of these.

II. The Assessment Process

As I work with people who are developing assessment plans, I see two common mistakes: 1) making the process too complicated; and 2) skipping some of the essential steps in developing an effective process. Because of this, my colleague Ron Miller and I developed a matrix that provides faculty with the structure they need to develop an effective assessment plan but also with the flexibility to be adapted for a variety of settings and purposes, including e-learning [6]. The matrix has been used successfully for course and program assessment at several institutions; it can also easily be used for educational project evaluation or individual course assessment. Our goal in developing the matrix was to help demystify assessment for faculty and assuage some of their fears about assessment. I first present a brief overview on developing an assessment plan and then introduce details of the matrix.

A. Developing an Assessment Plan

The following steps have been found valuable in developing an effective assessment plan:*

- Identify course or program objectives consistent with institutional goals and the needs of internal and external stakeholders including accrediting agencies.

- Develop course or program outcomes and performance criteria for each objective.

- Decide what course and program curricular and co-curricular activities will address each outcome.

- Determine the best methods for assessing and evaluating each outcome and decide when assessment data will be collected.

- Report results to stakeholders and use feedback to improve the program and the assessment process itself.

Each of these steps will be discussed more fully in the next section.

B. The Assessment Matrix

A relatively easy way to begin developing a course or program evaluation plan is to use the assessment matrix, summarized in Table 1, which Ron Miller and I adapted and expanded from a similar matrix included in the National Science Foundation's *User Friendly Handbook for Project Evaluation* [7]. We are also indebted to the outcomes assessment guidelines developed by Gloria Rogers and Jean Sando and published in *Stepping Ahead: An Assessment Plan Development Guide* [8].

Table 1. Assessment Matrix.

Objectives
　What are the overall objectives of the course or program? How do they complement institutional and accreditation expectations?
Learning Outcomes
　What are the program's educational outcomes? What should your students know and be able to do?
Performance Criteria
　How will you know the outcomes have been achieved? What level of performance meets each outcome?
Implementation Strategies
　How will the outcomes be achieved? What program activities (curricular and co-curricular) help you to meet each outcome?
Evaluation Methods
　What assessment methods will you use to collect data? How will you interpret and evaluate the data?
Timeline
　When will you measure?
Feedback
　Who needs to know the results? How can you convince them the objectives were met? How can you improve your program and your assessment process?

The assessment matrix provides faculty members (especially ones with little assessment experience) with a structure for developing their assessment plan using a series of questions — as they answer the questions, they essentially articulate the plan. Thus, the matrix provides a "hands-on," concrete tool for guiding development of the assessment plan—it is not an abstract document that simply lists what needs to be done. As shown in Table 1, questions are posed in the matrix to help develop the following aspects of the plan: program objectives and outcomes; performance criteria; implementation strategy; evaluation methods; timeline; and feedback. Each of these components of the plan should be treated as iterative and fluid as the program's curriculum is taught, assessed, and revised. Additional details to help faculty members work through the planning process are discussed below.

1) *Objectives and Outcomes:* Developing clear objectives and outcomes is the key to the success of an assessment plan.

---

* In this paper I am using the assessment terminology used by ABET and commonly adopted by engineering programs in the U.S. It is my view that the assessment nominclature selected is not important as long as terms are understood by all of their users and they are employed consistently. For example, whether an overarching principle is called a "goal" or an "objective" is not important as long as there is agreement and consistency among its users. Many hours have been wasted arguing about which words to use.

Faculty often fail to spend the time necessary to articulate clear objectives and outcomes before they rush to develop, measure, and evaluate a course or curriculum. They need to begin by defining broad objectives and then answering such questions as "What should students know and be able to do when they complete the course or program?" Here we are defining "objective" as a broad statement of desired results such as "students who complete the program should be able to communicate effectively." An outcome is a "detailed statement which describes under what circumstances the goal will be achieved" [8]. Outcomes should be clear, precise, and measureable.

In general, it is helpful to write outcomes using quantifiable action verbs (e.g., apply, calculate, describe, determine, demonstrate, analyze, evaluate), rather than vague terms (e.g., know, learn, appreciate, understand). The choice of verbs also indicates the level at which students are expected to demonstrate mastery of a concept. Many faculty have found Bloom's taxonomy, which moves from skills such as knowledge, comprehension, and application to "higher level" abilities such as analysis, synthesis, and evaluation, helpful in designing assessment outcomes [9]. For a complete list of the ABET criteria 3 outcomes including attributes and appropriate verbs at each Bloom level, see the website developed by a multi-institution group with NSF funding [10, 11].

2) *Performance Criteria:* Faculty should also articulate performance criteria for each objective to be evaluated. A performance criterion "defines the level of performance required to meet the objective" [8] and indicates the types of data that will be collected to provide supportive evidence. Once again, faculty must discuss and agree upon what performance levels they expect their students to achieve; this discussion will help make explicit faculty ideas and beliefs about satisfactory levels of student performance. The questions to be answered here are "How will you know the outcomes have been achieved?" and "What level of performance meets each outcome?"

3) *Implementation Strategy:* It is important to make ensure that learning outcomes, performance criteria, and implementation strategies mesh. For example, important questions such as "How will the outcomes be achieved?" and "Which course or program activities help to meet each outcome?" should be answered as the implementation strategy is developed. Many assessment plans include numerous lofty goals for student achievement between entry and graduation. However, the faculty developing these goals sometimes fail to allow sufficient opportunities in the curriculum for students to meet the goals. For example, if students are to learn the design process, or how to communicate effectively, or to gain an understanding of contemporary issues, they must have an opportunity within the curriculum and/or co-curriculum to learn, practice, and improve these skills and abilities. One benefit of developing an assessment plan for a program is that the process itself allows the faculty to examine the entire curriculum and to see how each faculty member's courses fit into the program's overall objectives and outcomes.

4) *Assessment Methods:* Once outcomes and an implementation strategy have been developed, general assessment methods and evaluation strategies should be selected. I am defining assessment as "collecting and analyzing data on student academic performance," and evaluation as "interpreting assessment data to draw conclusions about how well program goals and objectives are being met." [8] Good assessment allows faculty to draw sound conclusions about the course or program. The basic questions here are "What assessment methods will you use to collect data?" and "How will you interpret and evaluate the data?" The methods selected will depend on many factors including time and money available, but several rules of thumb apply:

- Explore a range of possible methods, qualitative and quantitative, formative and summative, depending on the course or program outcomes.
- Whenever possible, use more than one method—triangulate.
- Realize that for some program outcomes it may be difficult or impossible to obtain purely objective assessment results. However, methods exist to assess complex outcomes with a high degree of precision and reliability.

Many assessment techniques are available to interested faculty members. Prus and Johnson have developed a particularly useful compendium of common methods, many of which are easily adaptable to e-learning formats [12]:

- Commercial, norm referenced standardized exams
- Locally developed exams
- Oral examinations
- Performance appraisals
- Simulations
- Written surveys and questionnaires
- Exit interviews and other interviews
- Third party reports
- Portfolios
- "Stone" courses
- Archival data
- Behavioral observations

In their article, Prus and Johnson briefly define each method and then discuss its strengths and weaknesses, ways to overcome the weaknesses, and a "bottom line." A brief discussion of three of these methods — standardized tests, portfolios, and surveys — will give a sense of the variety of

assessment possibilities as well as some of the electronic applications of these methods that have been developed recently.

Standardized tests have several advantages including that little faculty time is required to develop or score them; scoring is objective; they have external validity; reference group comparisons are possible; and they may be beneficial where state or national standards exit. Disadvantages are that they limit what can be measured; they may eliminate a local goal-setting process; they are unlikely to measure a program or course's specific goals; and the results are easily misused. In selecting a standardized exam, faculty should also pay close attention to whether the exam is norm referenced (students are compared to other students) or criterion referenced (students' performance is measured by how well they do against set criteria, not other students). Disadvantages can be reduced by choosing tests carefully, reviewing reports carefully, using criterion-referenced tests if possible, and using cross-validation (checking results against other measures).

In addition to tests with which we are all familiar which have been moved on-line, such as the Graduate Record Exam (GRE), a number of creative standardized exams are being developed, many of them computer adaptive tests (CATs). Two examples cited in a recent article by Erwin and DeMars [13] include a conflict resolution assessment administered via computer and a branching test developed by the National Board of Medical Examiners where, as part of a licensure test, physicians are presented with basic facts about a patient's case and asked to provide treatment. They are evaluated on the effectiveness of their actions. Technologically delivered exams can also make creative use of multimedia. Erwin and DeMars cite a general education test in fine arts and humanities that has been developed using such media as slides of artwork, recorded theater performances, and musical performances.

Portfolios have been a popular type of assessment in the arts and humanities for many years, but they have recently been used more widely in engineering, technology, and science programs. Among their many advantages are that they can be used longitudinally; they can measure multiple parts of the curriculum at the same time; they arguably reflect student ability better than tests; the process of evaluating them provides opportunities for faculty development; and they increase student participation. The disadvantages of portfolios include their cost in terms of evaluator time and effort; management of the process; and (perhaps) the inability to compare with students at other institutions. Disadvantages may be reduced by using a representative sample of students rather than evaluating portfolios for all students; establishing inter-rather reliability and providing

training for evaluators; and cross-validating with more controlled student work samples.

In recent years, electronic portfolios have become increasingly used in higher education. For example, Kalamazoo College has perhaps the best-known general education electronic portfolio program[14]; Rose Hulman Institute of Technology has an electronic portfolio program for its engineering students [15]. In addition to the convenience of keeping portfolios electronically and avoiding rooms full of folders, it may even be possible soon to score portfolios and other written work electronically using the latent semantic analysis (LSA) process developed by Kintsch, Landauer, and their colleagues at the University of Colorado which purports to measure not only writing ability but textual meaning [16].

Surveys have a number of potential advantages including that they can cover a broad range of content areas within a brief period of time; results tend to be more easily understood by lay persons; can cover areas of learning and development which might be difficult or costly to assess more directly; can provide accessibility to individuals who otherwise would be difficult to include in assessment efforts (e.g., alumni, parents, employers). The biggest disadvantage of surveys is that results tend to be highly dependent on wording of items, salience of survey or questionnaire, and organization of instrument. Thus, good surveys and questionnaires are more difficult to construct than they appear. An excellent resource for faculty who which to develop a survey is Linda Suskie's book *Questionnaire Survey Research: What Works* [17]. In addition, surveys frequently rely on volunteer samples which tend to be biased; mail surveys tend to yield low response rates; require careful organization in order to facilitate data analysis via computer for large samples; commercially prepared surveys tend not to be entirely relevant to an individual institution and its students; forced response choices may not allow respondents to express their true opinions; results reflect *perceptions* which individuals are willing to report and thus tend to consist of indirect data. Disadvantages can be reduced by using only carefully constructed instruments that have been reviewed by survey experts; wording reports cautiously to reflect the fact that results represent perceptions and opinions respondents are willing to share publicly; using pilot or "try out" samples in local development of instruments and request formative feedback from respondents on content clarity, sensitivity, and format; and cross-validating results through other sources of data.

Electronic surveys have a variety of advantages over paper and pencil ones. For example, there is some evidence of higher return rates for electronic surveys and survey results can be easily compiled and analyzed on line.

Once the methods are selected, appropriate data should be collected and analyzed and results compared with the performance criteria established earlier. Note that the underlined evaluation of such data and decisions about how to use the results for program improvement are often complex; that is, the root cause of poor retention may be hard to identify. However, a high quality assessment process helps programs make decisions based on data, not hunches or anecdotal evidence.

5) Timeline: The important question here is "When will you measure?" Some outcomes require formative (midcourse) measurements; others are summative and can reasonably be put off until students reach the end of the curriculum. A combination of formative and summative assessment usually works best. For example, students' ability to work well in multidisciplinary teams is a skill that takes time to develop. Most programs would wish to measure teamwork ability at more than one key point in their students' education, perhaps at the end of each academic year. The same could be said for other developmental skills such as oral and written communication and critical thinking. Such formative assessment provides valuable feedback to both students and programs and encourages corrective action before it is too late. On the other hand, technical knowledge acquired in specific classes may be best assessed once at an appropriate place in the students' program. Although there are no clear-cut rules about when to collect data, most successful assessment programs prefer to maintain longitudinal data on their students so that progress toward meeting program objectives can be steadily monitored. In addition, the collection of longitudinal data allows programs to demonstrate "value added" by the curriculum, to show how students gain in knowledge and skills from one point, e.g., sophomore year, to another point, e.g., graduation. A single, summative, evaluation makes it difficult to measure any change in student outcomes.

6) Feedback: Here the key questions are "Who needs to know the results?", "How can you convince them the objectives were met?", and "How can you improve your program and your assessment process?" The stakeholders for a program or curriculum (e.g., faculty, students, other programs, accrediting agencies) should be identified and their needs analyzed. Different audiences clearly have different agendas and will need information presented in different ways to be informed that a program meets its objectives and outcomes. Evaluation reports should be customized to meet the needs of various audiences and delivered in time to be useful. Any data collection methods that turn out not to be useful should be discontinued.

III. Conclusion: Some Keys To Successful Assessment

Although all components of the assessment process are important, these three are, in my experience, the keys to a successful assessment:

a. Taking the time to select appropriate objectives and measureable outcomes;
b. Selecting appropriate assessment methods, preferably multiple measures; and
c. Making use of the assessment results for continuous improvement.

In addition, based on experience and observation, I have learned many lessons about effective assessment, including these:

- Avoid the temptation to start collecting materials before developing clear objectives, outcomes, and an assessment process. Before decisions are made about which materials to collect and assess, be sure to answer questions about what is being assessed, how the data will be analyzed, when materials will be collected, and who will receive the results.

- Be sure to promote stakeholder buy-in by involving as many constituencies as possible in the assessment development and implementation process.

- If you are a novice at assessment, look for campus resources to help get you started with assessment. Most institutions have some level of assessment expertise on campus — meet your colleagues in the education or psychology department.

- Remember that quality of results is more important than quantity. You do not have to measure every learning outcome in every course in the curriculum. Collect results that will be of most value in improving the learning and teaching process and, if it makes sense, use sampling techniques to collect a longitudinal snapshot of student achievement.

- Do not forget to assess and improve the assessment process itself. Few of us will get it right the first time, so revision and refinement is essential.

E-learning has both strengths and weaknesses when it comes to assessment. Many people, including many in the higher education community, are still skeptical about the efficacy of e-learning. It is therefore particularly important to carefully and professionally assess e-learning. Too much of the current work is unconvincing because it relies too heavily on self-reports from students rather than actual measures of their learning. Surveys

are easy to develop and administer (especially bad ones), but we need to do more in order to make our case.

I am excited by the opportunities technology affords for developing effective assessments, both formative and summative. On-line quizzes allow students to get instant feedback on their understanding of class materials; simulations allow faculty and students to test their knowledge in almost "real world" situations; new software may make measurement of complex cognitive processes much easier. The possibilities appear endless. I encourage developers of e-learning assessments to be creative in adapting tried assessment measures to the special circumstances of e-learning as well as in inventing assessments which have not even been conceived of yet.

References

[1] Accreditation Board for Engineering and Technology, <http://www.abet.org/images/Criteria/2002-03EACCriteria.pdf>.

[2] The Higher Learning Commission, <http://www.ncacihe.org/overview/2001HLCOverview.pdf>.

[3] See, for example, the regular "Web Corner" column in *Assessment Update* by Alec Testa, Ephraim Schechter, and Douglas Eder.

[4] Council for Higher Education Accreditation, "Accreditation and Assuring Quality in Distance Learning," CHEA Monograph Series 2002, Number 1.

[5] Nickles, G. and A. Pritchett, "Software for the Automated Evaluation of Web-Delivered Instruction," *Proceedings of the 2002 American Society for Engineering Education Annual Conference & Expostion*, Session 2258.

[6] Olds, B.M., and R.L. Miller, "An Assessment Matrix for Evaluating Engineering Programs," *Journal of Engineering Education*, 87, 2, April 1998, pp. 173-178.

[7] Stevens, F., F. Lawrence, and L. Sharp, *User-Friendly Handbook for Program Evaluation: Science, Mathematics, Engineering, and Technology Education*, J. Frechtling (ed.), NSF 93-152, National Science Foundation, 1996.

[8] Rogers, G.M., and J.K. Sando, *Stepping Ahead: An Assessment Plan Development Guide*, Rose-Hulman Institute of Technology, Terre Haute, Indiana, 1996.

[9] Bloom, B.S., M.D. Englehart, E. J. Furst, W. H. Hill, and D. R. Krathwohl, *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain*, New York: Longman, 1956.

[10] Engineering Education Assessment Methodologies and Curricula Innovation, National Science Foundation Action Agenda Grant EEC 9872498, <http://civeng1.civ.pitt.edu/~ec2000/ec2000_attributes.html>.

[11] Besterfield-Sacre, M., L. J. Shuman, H. Wolfe, C. J. Atman, J. McGourty, R. Miller, B. Olds, and G. Rogers, "Defining the Outcomes—A Bloom's Taxonomy Approach to EC 2000," *IEEE Transactions on Engineering Education*, April 2000.

[12] Prus, J. and R. Johnson, "A Critical Review of Student Assessment Options," in *Assessment & Testing, Myths & Realities*, (T.H. Bers and M.L. Mittler, eds.), New Directions for Community Colleges #88, Winter 1994, pp. 69-83, Jossey-Bass Publishers, San Francisco.

[13] Erwin, T. D. and C. DeMars, "Advancing Assessment: Why Not Computer-Based Assessment?" *Assessment Update*, March-April 2002, Volume 14, Number 2.

[14] <http://www.kzoo.edu/pfolio>.

[15] <http://www.rose-hulman.edu/users/groups/Assessment/Public/html/IRA/IRA/>.

[16] Rehder, B., M. E. Schreiner, M. B. Wolfe, D. Laham, T. K. Landauer, and W. Kintsch, "Using Latent Semantic Analysis to Assess Knowledge: Some Technical Considerations," *Discourse Processes*, 1998, 25, 337-354.

[17] Suskie, Linda A. *Questionnaire Survey Research: What Works*, 2nd Ed., Tallahassee: Association for Institutional Research, 1996.

Author's Biography

Barbara M. Olds is Associate Vice President for Academic Affairs and Professor of Liberal Arts and International Studies at the Colorado School of Mines where she has been a member of the faculty for the past eighteen years. She has also served as Principal Tutor of the McBride Honors Program in Public Affairs for Engineers, Director of EPICS (Engineering Practices Introductory Course Sequence) and chair of CSM's Assessment Committee. She has given numerous workshops and presentations on topics of interest to engineering educators. Dr. Olds has received the Brown Innovative Teaching Grant and Amoco Outstanding Teaching Award at CSM and was the CSM Faculty Senate Distinguished Lecturer for 1993-94. She also received the Helen Plants Award for Best Workshop at the 1992 Frontiers in Education national conference and was awarded a Fulbright fellowship to teach and conduct research in Sweden during the 1998-99 academic year. She is currently on the Board of Directors for the American Society for Engineering Education.