A practical framework for evaluating online distance education programs

Alfred P. Rovai*

School of Education, Regent University, 1000 Regent University Drive, Virginia Beach, VA 23464-9800, USA

Abstract

The case is presented for distance education administrators that evaluation is an essential element of successful distance education programs. A synthesis of the program evaluation and distance education research literature is used to form a framework for conducting evaluations of online programs. Evaluators should assess student performance, determine program and cost effectiveness, monitor quality to include technology and support services, evaluate course design and instruction, and ascertain teacher and student satisfaction. Strategies tailored to obtain such information are described within the context of an open-systems approach. An inventory of potential evaluation questions for input, process, output, and impact evaluations that respond to the potential needs of internal and external stakeholders are listed along with quantitative and qualitative data requirements that can be helpful in responding to these questions.
© 2003 Elsevier Science Inc. All rights reserved.

Keywords: Program evaluation; Distance education; Higher education; Program effectiveness

1. Introduction

There is a concern among some educators that distance education is compromising the quality of education (e.g., Nissenbaum & Walker, 1998; Trinkle, 1999). These educators are worried that technology will cheapen traditional education and destroy the special relationships

---

* Tel.: +1-757-226-4861; fax: +1-757-226-4857.

E-mail address: alfrrov@regent.edu (A.P. Rovai).

instructors have with their students. They cite research evidence suggesting that some courses taken at a distance are impersonal, superficial, misdirected, and potentially dehumanizing and depressing, and that they disrupt the interactions that create a learning community. On the other hand, substantial research evidence exists suggesting that the course-delivery medium is rarely the determining factor for a variety of educational outcomes, including student satisfaction, perceptions, and learning (e.g., Russell, 1999), and that strong feelings of community can be developed in distant learning environments (e.g., Rovai, 2001).

However, the validity of much of the research involving comparisons of online and traditional media (Russell, 1999) has been questioned by researchers who claim that *no significant difference* is an inconclusive result. In particular, Lockee, Moore, and Burton (2001, p. 62) believe that "a finding of *no significant difference* between face-to-face instruction and distance-delivered instruction does not mean they are equally good or bad." Moreover, such studies often violate the assumption of ceteris paribus, that is, all factors are the same except for the conditions that are manipulated. One solution offered by these researchers is that research should move away from comparison studies and focus on evaluations of specific distance education programs, a view promoted by this article.

Statistics reveal that not all distance education programs are equally effective. For example, Carr (2000) reported significant variation in distance education dropout rates among schools, with some postsecondary schools reporting course-completion rates of more than 80% and others finding that fewer than 50% of students finished their distance education courses. Moore, Thompson, Quigley, Clark, and Goff (1990) and Verduin and Clark (1991) suggested that teaching and studying at a distance can be as effective as traditional instruction provided: (a) the method and technologies used are appropriate to the instructional tasks, (b) there is student-to-student interaction, and (c) there is timely teacher-to-student feedback. Quality of course design and instruction are of primary importance to the success of distance education programs, as they are for traditional programs, and course design and instructional methods effectively adapted to the technology are more important than the types of technology used to deliver the program.

Most distance education administrators realize that they are in a highly competitive marketplace where quality and service are strongly related to success. Consequently, schools offering programs at a distance must compete with each other as well as with traditional programs to attract students. Schools must offer quality programs at a competitive price to be successful. Carnevale (2001) reported that the American Federation of Teachers requested that colleges not only acknowledge, but also assume the standards and collective bargaining agreements to protect the quality of distance education.

Evaluation is an essential component of program improvement and renewal and long-term success. Willis (1993, p. 70) cautioned that "even the best designed or adapted distance delivered course will likely require revision." Moreover, distance education program evaluation results can be extrapolated to improve other distance education programs. According to Verduin and Clark (1991, p. 184):

> Good evaluation will also assist distance educators in thinking about what they are trying to do and achieve as they implement programs and activities. Continued development in distance education is essential and evaluation can reveal what is effective and what is not.

Evaluations resulting in program refinements are highly important if quality is to be achieved and maintained. Galbraith, Sisco, and Guglielmino (1997, p. 113) stated, "Evaluation is an essential element in effective programs for any target group."

Scriven (1981) reported that a common use of program evaluation is for accountability purposes. He also identified other uses to include (a) determining program effectiveness, (b) identifying program weaknesses to enable administrators to improve program effectiveness, (c) providing evidence of effectiveness to doubters, and (d) providing information that can be used for program renewal. Keegan (1996) recommended that the evaluation of any distance education program should focus on four aspects: (a) the quantity of the learning (i.e., enrollment, new learner markets, and course-completion rates), (b) the quality of the learning (i.e., effectiveness of courses or program to enable desired learner outcomes), (c) the status of the learning (i.e., transferability of coursework and employer recognition of degrees and certificates), and (d) the relative cost of the learning (i.e., institutional cost effectiveness and cost benefits).

Measurement and judgment both have a legitimate role in program evaluations provided the evaluator distinguishes between both types of information. A comprehensive evaluation of the effectiveness of distance education programs should be based on multiple sources of evidence and the convergence of different measures. However, there has been a tendency in evaluations of distance education programs to be less rigorous and to rely on the use of limited self-reports and qualitative evaluations (Campbell, Lison, Borsook, Hoover, & Arnold, 1995). For example, learners may only be asked to rate how much they liked specific aspects of the distance learning system, and course evaluations may focus on how they liked the course design and the instructional methods used by the instructor. Fewer evaluators use the data from a comparison group (e.g., from another section of the course) and only a very few attempt evaluations that use experimental designs.

The purpose of this article is to draw from the program evaluation and distance education professional literature to synthesize a framework for evaluating online distance education programs. Such programs are delivered via the Internet and typically use e-learning software such as the Blackboard Learning System, Lotus LearningSpace, and WebCT. Added to this synthesis are the experiences of the author both as a practitioner of distance education and program evaluation for over a decade. The framework provided by this article is meant to assist those individuals who are involved in the direction, planning, or conduct of an evaluation by outlining a methodology and list of potential evaluation questions that can be used to help inform a comprehensive program evaluation.

Fig. 1 outlines the evaluation framework described in this article. The process starts with identification of the purpose of the program evaluation. For example, why is the evaluation necessary and what types of decisions will be made about the program? Should emphasis be placed on input, process, outcome, or impact issues? Once the purpose of the evaluation is understood, the type of evaluation can be determined. Then, evaluation strategies and questions can be developed in parallel. The specific evaluation questions to be used for any program evaluation should be based on the information needs of the evaluation's users and decision-makers, and the strategies used to respond to these questions should be selected

Purpose of the evaluation

↓

Determine type(s) of evaluation
Input
Process
Output
Impact

↙   ↘

Determine evaluation strategies   Develop evaluation questions
Objectives-oriented   Input
Management-oriented   ←→   Process
Consumer-oriented   Output
Expertise-oriented   Impact
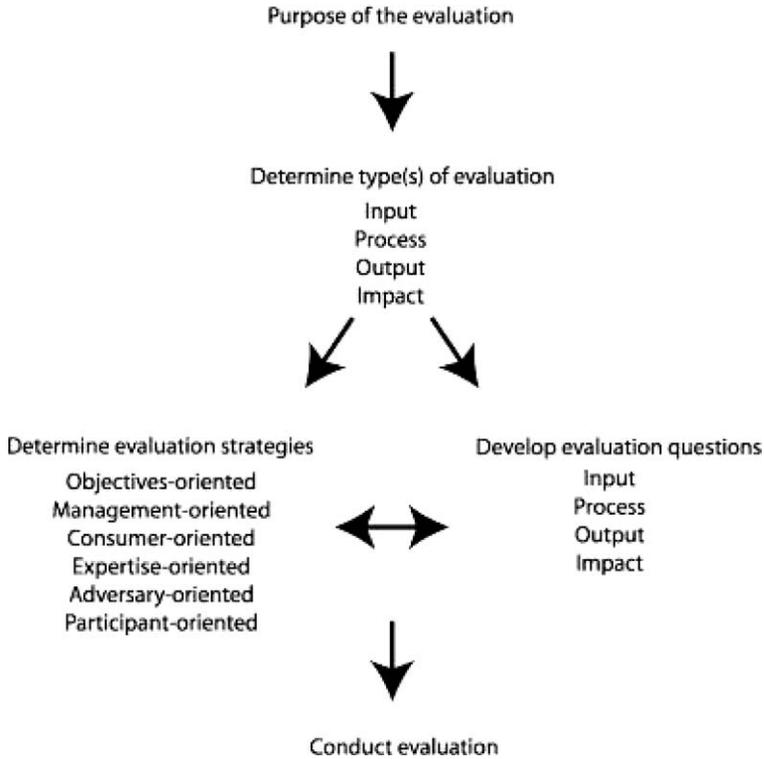Adversary-oriented
Participant-oriented

↓

Conduct evaluation

Fig. 1. A program evaluation framework that synthesizes the systems model of evaluation by Rossi et al. (1999), with the evaluation strategies identified by Worthen et al. (1997).

based on an understanding of the broader context in which the program operates. For example, there may be a variety of complex issues involving the distance education program that requires inclusion of an adversary-oriented strategy to ensure that all views are presented and considered during the evaluation process. A description of the types of evaluations and supporting strategies followed by an inventory of potential questions for use in an evaluation of online programs are presented below.

2. Types of program evaluation

According to Posavac and Carey (2002), *evaluation* is a collection of methods, skills, and sensitivities necessary to determine whether a human service is needed and likely to be used, whether it is conducted as planned, and whether the human service actually does help people. Evaluations of human service programs are conducted to answer questions and address issues that are raised by stakeholders. The process of identifying and selecting evaluation questions is central to the evaluation and normally represents an early step in the program evaluation planning process. Once selected, the evaluation questions provide a direction for the program

evaluation. Plans are then developed to gather evidence that allows the evaluator to answer these questions.

A common method of categorizing program evaluations is by whether they are formative or summative or a combination of the two (Scriven, 1981). Formative evaluations focus more on process and feedback to determine the extent to which the program is operating as intended. The idea behind the formative evaluation is to ensure that the course or instructional product is meeting its stated goals as efficiently and effectively as possible and that the program is being implemented as planned. The focus, therefore, is on academic quality management. Summative evaluations, on the other hand, seek to discover if the program made a difference. That is, the emphasis is on determining the results or outcomes of the program. Summative evaluations are often used for accountability purposes.

The systems model is an alternative approach to program evaluation (Rossi, Freeman, & Lipsey, 1999). Here, one can categorize an evaluation by type, as an input, process, output, and/or impact evaluation. Frequently, all types are included in the same evaluation. This approach is similar to Stufflebeam's (1971) CIPP model (i.e., context, input, process, and product), where *context* focuses on planning decisions, *input* on structuring decisions, *process* on implementation processes, and *product* on outcome attainment. Moore and Kearsley (1996, p. 5) suggested that a systems approach is very helpful to understanding distance education and that "the systems model provides a tool that not only helps us recognize many of the issues that separate distance education from conventional education, but also helps us distinguish good distance education from bad."

The systems view of program evaluation represents a departure from, and refinement of, the view of categorizing evaluations as formative or summative. However, the approaches are overlapping and can be viewed as different paths to the same goals. A key consideration in using the systems approach is the acknowledgement that a distance education program is an open system, and consequently, feedback from a variety of internal and external stakeholders is essential for system adjustment and improvement. Stakeholders may include learners and instructors, staff, administrators, policymakers, boards, vendors, consultants, community groups, accrediting agencies, government organizations, businesses, and employers.

The systems approach to program evaluation is highly compatible with strategic planning, which has the goal of strengthening the management decision-making process by having it recognize and address key internal and external factors that affect the organization. Since programs delivered at a distance consist of multiple components, e.g., the e-learning software, academic and technical support, presentation of content, and interaction, evaluators must recognize that all components of the program must work together in an efficient manner if the entire system is to be effective. Consequently, it is important to evaluate distance education programs by how they work as a whole rather than by evaluating individual components without regard to overall program effectiveness. By way of analogy, there is no need to have an expensive, high-performance carburetor in a motorbike if the rider rarely revs the engine past 5000 rpm.

2.1. Input evaluation

An input evaluation identifies and evaluates system capabilities to include equipment and technical expertise, alternative program strategies, and the designs used to meet the target audience and satisfy their needs. The purpose is to provide information on the quality of resources used by the program and to determine how to best use these resources to achieve program objectives. According to Moore and Kearsley (1996), important inputs that should be evaluated include student characteristics, instructor/tutor experience, competence of administrative staff, efficiency of course development, and institutional cooperation/support. A systematic examination of alternative economical models may also be required to predict and compare the expenditure (and potential revenue) of different models for each program component as well as the effectiveness of the overall program (Phipps & Wellman, 2001).

When evaluating program inputs, it is important to examine student needs, the number of students served, and program costs. According to Willis (1993), a needs assessment can provide the external data to verify the need for the program, identify the factors that led to the instructional need, and provide evidence that the instruction being planned can effectively meet this need. An input evaluation can also help forecast future costs and the sustainability of the program being evaluated, validate program and course objectives, and provide information for refining and positioning the program in competitive distance education markets.

It is also important to identify the needs of online instructors and to evaluate the extent to which these needs are satisfied. Distance education instructors must be trained to teach at a distance. According to Eaton (2000), president of the Council for Higher Education Accreditation in the United States, instructors are frequently thrust online without the proper training or time to manage effective online learning, thus, creating an input problem. Wilson (2001) reported that a study of distance education in Kentucky's higher education system revealed faculty willing to use the technology but needing more institutional support. In particular, she reported that the Kentucky faculty were (a) unsure of the instructional efficacy of distance education, (b) unconvinced about personal involvement in distance education, (c) underprepared in areas related to distance education, (d) under time pressure, (e) not rewarded for their work with distance education, and (f) feeling undersupported by the university infrastructure. Clay (1999) asserted that training for distance instructors is a continuous process and that instructor support should be ongoing. Such training programs should be periodically evaluated to determine whether or not to continue or modify the existing training.

2.2. Process evaluation

For a process evaluation, the evaluator must examine not only what is happening within the program as it is being implemented but also what should be happening and is not. It provides information about the state of all components of the program to include determining teacher and cost effectiveness. Cost effectiveness entails accomplishing the intended program goals by providing maximum values for limited expenditures. Program elements to be measured are applied against the cost, and the comparison of cost and effectiveness form the

basis of this analysis. Cost effectiveness analysis produces a dollar figure per unit of analysis, e.g., number of students graduating or increased enrollments for a school. Quantification might also focus on revenue or profit increases in the form of return on investment. A minimum threshold should be determined above which the program needs to perform to sustain support.

The evaluator must also analyze and evaluate the e-learning system as a whole, as well as by subsystems, to include instructor effectiveness. After all, technology is not self-implementing. In large measure, the teacher influences the effectiveness of any learning environment. There may be a great e-learning system that results in poor learning because of teacher ineffectiveness (a process problem). On the other hand, a marginal e-learning system (an input problem) can result in a superior learning experience because of the abilities of an exceptional instructor.

As part of process evaluation, one should also address the ongoing instructional efficiency of the teaching/learning process. Since learning is inherently a social process (e.g., Duffy & Cunningham, 1995), measurement of variables, such as sense of community and weekly rates and quality of student–student and instructor–student interactions, can be useful as proxy measures of learning. In particular, online students are more motivated when contacts with instructors are frequent, especially regarding feedback pertaining to how well they are performing on their assignments. They also value participation in small groups, particularly discussion groups.

2.3. Output evaluation

An output evaluation seeks to determine the immediate or direct effects of the program. It consists of collecting, analyzing, and judging results such as how much the program was used, how many people were reached and graduated, the extent to which program objectives were met, and changes in skills, knowledge, or attitudes. Graduation rates and learner achievement in the form of course grades or standardized test scores are frequently used as output evaluation measures. Moore and Kearsley (1996) also identified several outputs that provide valuable information of the overall effectiveness of distance education programs, e.g., student satisfaction ratings, completion rates, and staff turnover.

2.4. Impact evaluation

Impact or outcome evaluations address the longer-term results of the program and the extent to which the program reduced or eliminated student needs and the effects of the program on society at large. This type of evaluation typically involves tracking the performance of program graduates in their program-related job and often involves the use of graduate and employer surveys. It seeks evidence of the degree to which the program has reduced/eliminated student educational needs and the migration of need satisfaction towards more sophisticated ends, longer-term as well as unintended program effects, and evidence of program effectiveness (value) at the societal level (Worthen, Sanders, & Fitzpatrick, 1997). It also seeks to determine how staff and faculty attitudes, behaviors, and goals changed because of the program.

3. Evaluation strategies

Within the context of the type of evaluation (i.e., an input, process, output, or impact evaluation) described above, specific evaluation strategies are used to collect data. In selecting the strategy or combination of strategies, the evaluator must consider the interests of all stakeholders and how these interests can best be served. Worthen et al. (1997) identified six evaluation strategies that are frequently used, either singly or in some combination, to collect data for educational program evaluations. These six strategies are described below.

3.1. Objectives-oriented

Arguably, the most popular strategy is the objectives-oriented evaluation strategy. The distinguishing feature of this strategy is that the evaluation focuses on determining the extent to which program and instructional objectives have been met. Consequently, this approach is highly suited to distance education programs that have highly defined objectives, and the purpose of the evaluation is to determine if, and to what extent, these objectives have been met.

Major weaknesses often cited regarding this strategy include the difficulty of evaluators to operate in a program environment with ill-defined objectives, to identify unintended program outcomes, and to measure learning. Grades, often used to operationalize learning, can have little relationship to what students have learned as students may already know the material when they enroll, or their grades may be more related to class participation, or work turned in late, than to learning. Furthermore, grades may not be a reliable measure of learning, particularly for performance tests, as different teachers and even the same teacher over time will not likely assign grades in a consistent manner. Therefore, using grades as a measure of learning can be problematic.

3.2. Management-oriented

The management-oriented strategy is meant to serve decision-makers and is particularly useful for making decisions about the reallocation of funds. The rationale for using this strategy is that "evaluative information is an essential part of good decision making and that the evaluator can be most effective by serving administrators, policy makers, boards, practitioners, and others who need good evaluative information" (Worthen et al., 1997, p. 97). A weakness of this strategy is that it tends to reinforce the status quo of management rather than balancing the interests of management with those of other internal and external stakeholders. If management does not value distance education, evaluation results will likely reflect this bias. Woolcot (1997), for example, provided evidence that in some schools distance education is neither valued nor rewarded and is not seen as a means to faculty tenure or promotion.

3.3. Consumer-oriented

The market-driven, consumer-oriented strategy, typically summative, adheres to the type of evaluation used by the *Consumers Union*. Consequently, the central theme of this strategy is the

development of information on products for use by consumers (i.e., students). Such an approach has particular appeal for distance education programs because of the increasing competition among such programs. Limitations on the use of this model in a distance education context arise as the result of individual student differences. What appeals to one student may not appeal to another. Aptitudes (Ehrman, 1990), affective states (Westbrook, 1997), and learning styles (Dille & Mezack, 1991) are likely to interact in complex ways and respond differentially to distance education content, context, and preferred learning style (Coggins, 1988).

3.4. Expertise-oriented

The expertise-oriented approach to evaluation, widely used by accrediting agencies, depends primarily upon professional expertise to judge an educational program. The worth of a curriculum is evaluated by curriculum experts who observe the curriculum in action, examine its content and underlying learning theory, and render a judgment about its value. However, this strategy has a potential weakness in the evaluation of distance education programs. A subject matter expert may know the content area but may not know how to design a course that best uses the information technologies available to the school to present content. Moreover, theory and research has not kept pace with the accelerated and accelerating growth of information technologies (Simonson, Schlosser, & Hanson, 1999). Consequently, expert testimony will be affected by the limited body of research literature. A final weakness is the limited reliability of expert testimony. Different experts may not make the same judgments and recommendations regarding the program.

3.5. Adversary-oriented

The adversary-oriented strategy attempts to reduce bias by attempting to assure fairness by incorporating both positive and negative views into the evaluation itself. Several models have been used for adversary evaluations, to include structured public debates, such as town hall meetings, and the use of opposing evaluators that debate the issues. The idea of using this model is not so much to win a verdict as it is for all stakeholders and evaluators to acquire a better appreciation of the issues involved and to gain insights into other points of view (Worthen et al., 1997). For distance education, this strategy can be helpful if students and faculty members are identified who support and oppose the distance education program and are provided the opportunity to present their points of view to the evaluators.

3.6. Participant-oriented

Finally, the participant-oriented or naturalistic strategy involves all stakeholders and is used in qualitative research studies. Huxley (1959, p. 272) wrote an excellent metaphorical description that accurately captures the spirit of this strategy:

> The best way to find things out. . .is not to ask questions at all. If you fire off a question, it is like firing off a gun—bang it goes, and everything takes flight and runs for shelter. But

if you sit quite still and pretend not to be looking, all the little facts will come and peck round your feet, situations will venture forth from the thickets, and intentions will creep out and sun themselves on a stone and if you are patient, you will see and understand a great deal more than a man with a gun does.

A possible weakness to the participant-oriented approach is that each stakeholder is likely to have different criteria regarding program value and effectiveness. For example, some faculty may be opposed to the concept of distance education altogether. Tenured faculty who are heavily invested in traditional education may view change as a threat to their traditional roles (Beaudoin, 1990). They may also rebel against collective curriculum authorship or producing instructional materials that others will use (Rumble, 1989). Some faculty may even fear being exposed as poor teachers in a medium that broadcasts and records their work (Wilkes & Burnham, 1991). If the evaluator attempts to find common ground and to satisfy all stakeholders, the evaluation is likely to become ineffective, and those designing and conducting evaluations may focus on answering questions that are not relevant, but to which everyone agrees.

4. Evaluation questions

Drawing on the types and strategies of evaluation and the issues regarding distance education described above, a list of potential evaluation questions is presented below. These questions do not represent a comprehensive list. Instead, they respond to the potential areas of weakness of the programs identified in the professional literature, and outlined above, and from the personal experiences of the author in evaluating programs. No single plan can be used for all evaluations. The nature of the program and the requirements of decision-makers will influence the specific set of evaluation questions selected by the evaluator.

Posavac and Carey (2002) suggested that responses to the following questions often help inform the development of evaluation questions:

1. What program will be evaluated? What e-learning system is used to deliver the program?
2. Who will use the results of the evaluation?
3. What is the purpose of the evaluation? What issues have been raised, e.g., low persistence rates, poor student satisfaction, or low enrollment? How will the results be used? What decisions need to be made regarding the program, e.g., continuance, expansion, reducing costs?
4. Who is going to conduct the program evaluation?
5. How will the program evaluation be conducted? What method(s) should be used? What are the available resources? What are the constraints (e.g., time and money)?

Also provided below are possible evaluation strategies and quantitative and qualitative data requirements that will be helpful in responding to these questions. Common sources of data are school records, student and instructor surveys, course evaluations, expert reviews,

case studies, e-learning system statistical data, interviews, and focus groups. The evaluation questions presented below are organized by type of evaluation, starting with input evaluations.

4.1. Input evaluation questions

Is the program accredited? If so, is the accrediting agency recognized by the appropriate organizations, such as the Department of Education or the Council for Higher Education Accreditation? What are the fixed and variable costs associated with this program? What is the best way to use program resources?

1. Suggested strategy: management- and expertise-oriented.
2. Data requirements: information regarding program accreditation, the e-learning system, support services and funding levels, and costs.

Is the school's admission policy sufficiently rigorous? What types of students and instructors are attracted to the program? How many students are served by the program? Is the program reaching its intended audience? Why or why not? How effective was student preenrollment counseling? Does the school provide students with advertising, recruiting, and admissions information that adequately and accurately represent the program and services available?

1. Suggested strategies: consumer- and management-oriented.
2. Data requirements: student demographics, TOEFL scores for non-English-speaking students, standardized test scores, prior academic achievement, and employment status, instructor qualifications, entrance requirements, and other data as appropriate and available in school and program records.

What are the educational needs of students targeted by the program? Are program and course objectives sufficiently responsive to these needs? Are course materials current? How efficient is the course development process? How does the school ensure that students admitted possess the knowledge and equipment necessary to use the technology employed in the program? What evaluation and assessment methods does the school use to measure student learning? How does the program ensure the integrity of student work and the credibility of the degrees and credits awarded?

1. Suggested strategies: objectives-, management-, consumer-, expertise-, and participant-oriented.
2. Data requirements: program and course objectives, student ratings of the relative importance of each need (i.e., objective), availability of competing resources to satisfy needs, currency of course materials, identification of additional needs, the degree to which the curriculum responds to student needs (i.e., student need satisfaction), and student course evaluations.

How were faculty selected to teach at a distance? Are instructors qualified to teach the content of their courses? Are instructors qualified to teach online? What is the extent to which instructors control the content of their courses? What are the needs of instructors? What is the extent to which these needs are satisfied? What relevant professional development activities and support services are provided instructors? What is the administrative and teaching burden of instructors? Do the instructors feel adequately prepared to use the e-learning system? Are instructors sufficiently competent in designing courses for delivery using the e-learning system? Have instructors adopted the specialized teaching techniques considered appropriate for the e-learning system? What is the nature and extent of support services needed for instructor success?

1. Suggested strategies: management-, consumer-, expertise-, and participant-oriented.
2. Data requirements: instructor needs, instructor ratings of the relative importance of each need, extent to which each need is satisfied, availability of competing resources to satisfy each need, and instructor satisfaction (e.g., with compensation, workload, training, and teaching load).

Is the e-learning system adequate for the program? To what extent does the e-learning system meet program requirements? Are there any e-learning system integration issues and interoperability concerns, bandwidth and scalability problems, or lack of features and functionality? Is the e-learning system suitable in terms of cost, given the volume of student activity and the number of students involved over the length of the program? How easily can course materials be produced using this technology? Is the current technology still cost effective? Would other technologies allow for a more cost-effective program?

1. Suggested strategies: management-, expertise-, and participant-oriented.
2. Data requirements: technology availability and ease of use, quality of vendor support, student course evaluations, learner reactions to the learning environment (e.g., general level of acceptance), effect of the e-learning environment on the students' progress and performance, areas of weakness requiring remediation, and additional uses and value provided the school by the e-learning system infrastructure.

4.2. Process evaluation questions

How well are students satisfied with the program? How does student satisfaction compare with that of courses offered on-campus? What are the consistently most/least popular courses and why? Was the program equally effective for all participants? Why or why not? Are there characteristics that distinguish satisfied and dissatisfied students? Does the school apply this information to admission and recruiting policies and decisions?

1. Suggested strategies: management-, consumer-, and participant-oriented.
2. Data requirements: courses offered, course enrollments, enrollment trends, overall course persistence rates, student reasons for refunds and enrollment termination, percent of

students who take follow-on distance education courses, student progress (students should remain in the program long enough to meet their immediate educational goals and reenter and exit as their goals change), and course evaluations.

What is the nature and extent of support services needed for student and instructor success? Are these support services responsive to student and instructor needs? To what extent are these resources used? Does the school monitor the use and quality of support services?

1. Suggested strategies: expertise-, consumer-, and participant-oriented.
2. Data requirements: student and instructor support needs, scope of support services provided (e.g., online library databases, delivery of course materials, academic advising, counseling, program administrators, student placement, registrar's office, financial aid, etc.), and adequacy of support services.

Is the program implemented as intended? How efficiently is the program being implemented? What is the level of interaction between students and instructors and among students? Does the instructor provide timely feedback to students? What program components or activities are the most and least effective? To what extent is the learner involved in setting goals and in choosing the tasks, assessments, and standards to reach those goals? How is the integrity of student work assured?

1. Suggested strategies: objectives-, management-, consumer-, and participant-oriented.
2. Data requirements: management effectiveness, e-learning system cost structure, student enrollment and satisfaction, learner performance in the learning environment, adequacy of curriculum and instruction, adequacy of student assessment, instructor teaching loads, instructor effectiveness, timeliness of instructor feedback, instructor persistence rates, student time on task (deals with the amount of time a student spends in dealing with the content of the course), interaction rates, and cost effectiveness.

4.3. Output evaluation questions

What are the program results? Are courses transferable? Are certificates or degrees awarded by the program recognized by the appropriate profession? What structures or policies in the school or e-learning environment are supporting or hindering outcomes and overall program effectiveness? Were program and course objectives achieved? What program and course design components appeared to contribute the most and least to the attainment of these objectives? Do students feel that they gained the desired knowledge and skills?

1. Suggested strategies: objectives-, management-, expertise-, and participant-oriented.
2. Data requirements: transferability of coursework, recognition of degrees and certificates, access (i.e., identified students who enroll but who could not or would not have enrolled in a face-to-face program), persistence (i.e., program and course completion rates), relevance

(i.e., the extent to which the program corresponds to student perception of needs), gratification (i.e., the extent to which the program enhances the students' self-esteem and sense of integrity), skill development (e.g., educational gains), use of program opportunities, quality of student effort (e.g., grade trends), and graduation rates.

4.4. Impact evaluation questions

What are the effects of the program on graduates? As a result of completing the program did they receive increased pay, acquired professional certifications, received promotions, etc.? Did the program have any unintended impacts?

1. Suggested strategy: participant-oriented.
2. Data requirements: impact (the extent to which the program reduced/eliminated student needs), percent of graduates believing their instructional programs met their goals, attitudes of graduates concerning the program, changes (if any) in student educational expectations, level of employer satisfaction, benefits of the learning to an employer, changes in learner job performance, learner performance on the job (e.g., supervisor ratings), and return on investment in terms of cost and organizational impact.

5. Conclusion

The evaluator concludes the work by developing a holistic evaluation based on the analysis of information. Such an evaluation is not as simple as merging multiple viewpoints by performing cuts and pastes using one's word processor. Viewpoints may well have interaction effects with each other, e.g., one viewpoint may be more beneficial for some students and not for others. In such cases, trade-off decisions must be made. The goal, at this point of the evaluation, is to identify factors that decrease costs and improve overall learning and organizational impact. The economic key is to spread development and revision costs over large numbers of learners and to drive down ongoing costs. One way this can be accomplished is by automating much of the learning support in course readers, tutorials, simulations, and online responses to frequently asked questions.

Programs that are well defined will include norms (i.e., program standards and expectations) that evaluators use to judge the worth of the program. One source of such norms is the Council for Higher Education Accreditation, the national coordinating body for national, regional, and specialized accreditation. However, in some cases no such norms exist, or are so vague that they are of little use to the evaluator. In such cases, the program evaluation should include a norm-setting phase in which all relevant stakeholders are included. A major issue will be how high to set the bar. Benchmark targets can be set by using outcomes obtained from the school's comparable traditional programs, such as persistence rates, satisfaction, and learning, or by comparing current outcomes to previous classes with the expectation that a specified amount of yearly growth should be achieved. Alternatively, an expert panel of

educators can be convened to decide on the norms used to judge various aspects of the program. Additionally, the outcomes documented in the evaluations of comparable online programs could be adopted. Whatever method is used to establish norms, the norms should be agreed to before the start of the evaluation to avoid the situation where standards are set based on actual program outcomes.

The approach to evaluation presented in this article draws heavily on open-systems theory (Rossi et al., 1999). This theory provides a simple, familiar, and intuitively logical model that facilitates an appreciation for a variety of program processes including marketing, admissions, advising, course design, teaching, learning, and support services. Program evaluation is essential to distance education when it is defined in open-systems terms. When the structure and functions of a program change, systems theorists call the process *morphogenesis*. Applied to distance education programs, morphogenesis changes the program's procedures, services, and behavior to better serve the mutual interests of the school, the program, and its internal and external stakeholders. The results of periodic program evaluations provide the impetus for this change. Without periodic renewal, programs tend to be drawn toward disorganization and eventual demise (Katz & Kahn, 1978). The challenge facing evaluators is to compile enough evaluation information to articulate the place of technology and distance education in student and teacher learning.

References

Beaudoin, M. (1990). The instructor's changing role in distance education. *American Journal of Distance Education*, *4*(2), 21–29.

Campbell, J. O., Lison, C. A., Borsook, T. K., Hoover, J. A., & Arnold, P. (1995). Using computer and video technologies to develop interpersonal skills. *Computers in Human Behavior*, *11*(2), 223–239.

Carnevale, D. (2001, February 2). Union offers standards for distance education. *The Chronicle of Higher Education*, A23 (Available at: http://chronicle.com/free/v47/i21/21a03304.htm. Accessed: February 1, 2003).

Carr, S. (2000). As distance education comes of age, the challenge is keeping the students. *The Chronicle of Higher Education*, *46*(23), A39–A41.

Clay, M. (1999). Development of training and support programs for distance education instructors. *Online Journal of Distance Education Administration*, *3*(3) (Available at: http://www.westga.edu/~distance/clay23.html. Accessed: February 1, 2003).

Coggins, C. C. (1988). Preferred learning styles and their impact on completion of external degree programs. *American Journal of Distance Education*, *2*(1), 25–37.

Dille, B., & Mezack, M. (1991). Identifying predictors of high risk among community college telecourse students. *American Journal of Distance Education*, *5*(1), 24–35.

Duffy, T. M., & Cunningham, D. C. (1995). Constructivism: implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Handbook for research on educational communications and technology* (pp. 170–198). New York: Simon & Schuster.

Eaton, J. S. (2000). FOCUS: assuring quality in distance learning. *The CHEA Chronicle*, *3*(3) (Available at: http://www.chea.org/untitled_folder/Chronicle/vol3/no3/focus.cfm. Accessed: February 1, 2003).

Ehrman, M. (1990). Psychological factors and distance education. *The American Journal of Distance Education*, *4*(1), 10–24.

Galbraith, M., Sisco, B., & Guglielmino, L. (1997). *Administering successful programs for adults*. Malabar, FL: Krieger.

Huxley, E. (1959). *The flame trees of Thika*. London: Chatto and Windus.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations* (2nd ed.). New York: Wiley.

Keegan, D. (1996). *Foundations of distance education* (3rd ed.). London: Routledge.

Lockee, B., Moore, M., & Burton, J. (2001). Old concerns with new distance education research. *Educause Quarterly*, *2*, 60–62.

Moore, M. G., & Kearsley, G. (1996). *Distance education: a systems approach*. Belmont, CA: Wadsworth/ITP.

Moore, M. G., Thompson, M. M., Quigley, A. B., Clark, G. C., & Goff, G. G. (1990). *The effects of distance learning: a summary of the literature.* Research Monograph No. 2 (ERIC Document Reproduction Service No. ED 330 321). University Park, PA: Pennsylvania State University, American Center for the Study of Distance Education.

Nissenbaum, H., & Walker, D. (1998). A grounded approach to social and ethical concerns about technology and education. *Journal of Educational Computing Research*, *19*(4), 411–432.

Phipps, R., & Wellman, J. (2001). Funding the infostructure: a guide to financing technology infrastructure in higher education. *Lumina Foundation for Education, New Agenda Series*, *3*(2), 1–30 (Washington, DC: Institute for Higher Education Policy).

Posavac, E. J., & Carey, R. G. (2002). *Program evaluation: methods and case studies* (6th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: a systematic approach*. Newbury Park, CA: Sage.

Rovai, A. P. (2001). Classroom community at a distance: a comparative analysis of two ALN-based university programs. *Internet and Higher Education*, *4*(2), 105–118.

Rumble, G. (1989). On defining distance education. *The American Journal of Distance Education*, *3*(2), 8–21.

Russell, T. L. (1999). *No significant difference phenomenon*. Raleigh, NC: North Carolina State University.

Scriven, M. S. (1981). *The logic of evaluation*. Inverness, CA: Edgepress.

Simonson, M., Schlosser, C., & Hanson, D. (1999). Theory and distance education: a new discussion. *The American Journal of Distance Education*, *13*(1), 60–75.

Stufflebeam, D. L. (1971). The relevance of the CIPP evaluation model for educational accountability. *Journal of Research and Development in Education*, *5*, 19–25.

Trinkle, D. A. (1999). Distance education: a means to an end, no more, no less. *The Chronicle of Higher Education*, *45*(48), 1.

Verduin, J. R., & Clark, T. A. (1991). *Distance education: the foundations of effective practice*. San Francisco, CA: Jossey-Bass.

Westbrook, T. S. (1997). Changes in students' attitudes toward graduate business instruction via interactive television. *The American Journal of Distance Education*, *11*(1), 55–69.

Wilkes, C. W., & Burnham, B. R. (1991). Adult learner motivations and electronic distance education. *The American Journal of Distance Education*, *5*(1), 43–50.

Willis, B. (1993). *Instructional development for distance education*. Syracuse, NY: Syracuse University.

Wilson, C. (2001). Faculty attitudes about distance learning. *Educause Quarterly*, *2*, 70–71.

Woolcot, L. L. (1997). Tenure, promotion, and distance education: examining the culture of faculty rewards. *American Journal of Distance Education*, *11*(2), 3–18.

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: alternative approaches and practical guidelines* (2nd ed.). New York: Longman.