

The Development of Effective Evaluation Methods for E-Learning: A Concept Paper and Action Plan

ELLEN B. MANDINACH

EDC Center for Children and Technology

E-learning is an emerging field as a promising instructional medium as well as a ripe arena in which to conduct research on its impact on teaching and learning activities. The fundamental nature of e-learning as an instructional medium differs substantially from face-to-face delivery, thereby requiring new and hybrid methods for evaluating its impact. This article examines the characteristics of e-learning that make it unique and traces some of the emerging trends in the field. The article then discusses evaluation methodologies that might be potentially informative in the examination of how e-learning is beginning to affect teaching and learning processes.

BACKGROUND

One of the contributions researchers can make to the field of e-learning is to provide enhanced methodologies by which to examine the process and impact of e-learning activities across all levels of education. E-learning is more than a new and emerging technology-based instructional delivery mechanism. It is a new form of teaching and learning that requires educators to rethink how the evaluation of process and outcomes should be conducted. New paradigms for teaching and learning call for new methods of assessment and evaluation.

E-learning creates new variables, constraints, and issues, making it fundamentally different from face-to-face learning environments. The roles of the professor, teacher, and student change. Requisite resources and infrastructure differ. Even the educational objectives may differ across students, professors, teachers, and institutions.

In such circumstances, tight experimental designs may not be practical or feasible. The environment contains many variables and little control, which makes using traditional experimental or even quasi-experimental methods difficult at best. That is, it would be difficult to create randomly assigned treatment and control groups, ensuring equivalent groups, because of the diverse and unknown characteristics of online student popu-

lations. Additionally, in an environment so open as online learning, it is exceedingly difficult to control for extraneous and confounding influences as required by experimental design.

Educational researchers, more generally, currently are facing similar challenges as new legislation, the No Child Left Behind Act (2001) mandates rigorous, controlled experiments in precollege educational settings to determine “what works.” There is substantial discussion about the feasibility of conducting such studies (Berliner, 2002; Erickson & Gutierrez, 2002; Feuer, Towne, & Shavelson, 2002; Pellegrino & Goldman, 2002; Shavelson & Towne, 2002; St. Pierre, 2002). There is thus a pressing need to develop new or hybrid ways to examine e-learning and its impact. Existing methods may need to be modified, expanded, enhanced, and made more flexible to be responsive to the specific constraints imposed by the environment.

The leitmotif of this article, focusing primarily on higher education, is that effective evaluation is only possible when the objectives of the educational intervention, in this case e-learning, are clear. Of course the need for such specification is true throughout education, but the assumption is mostly taken for granted in classroom learning environments. It helps the evaluator to understand how and why e-learning programs are being developed. Currently e-learning, however, is being conducted for many reasons and often without a clear specification of its educational objectives. Institutions of higher education know they should be joining the e-learning movement, but for what purposes and for what gains? Is its purpose to make money, save money, enhance learning, increase accessibility, improve instruction, or something else? Institutions are often lacking a clear rationale for why they are developing and providing e-learning. Once institutions have delineated their purposes for developing and providing e-learning, effective evaluation requires that they specify clearly their educational objectives for instituting the effort and that they make clear that stakeholders¹ in their schools agree that their purposes and objectives are appropriate. If these elements are not in place, the evaluator should help elicit them from the institution.

E-LEARNING ISSUES AND THEIR RELEVANCE TO THE RESEARCH COMMUNITY

Researchers can help stakeholders and institutions make sensible decisions about the questions that result from e-learning. Methodologies must be designed to examine how technology enables or facilitates teaching and learning activities in the context of e-learning. It would be remiss to look only at the medium and then make comparisons across various delivery mechanisms. Technology is only a tool. The technology becomes a necessary but not a sufficient condition. The focus of examination therefore becomes the interaction among several important levels of variables with the tech-

nology. The levels of variables include institutional infrastructure, pedagogical or teacher processes, and student learning processes. Researchers must capitalize on and begin to understand the impact of the unique characteristics of the technological environment, such as anywhere/anytime/any pace access, interactivity, and accessibility to enormous amounts of resources that heretofore have not been possible in higher education. E-learning should not be simply a current incarnation of correspondence courses if its implementation is to utilize fully the affordances of technology.

Lessons can be drawn from the implementation of computer-based testing. Simply translating a test from paper-and-pencil to a computer is not an effective use of the technology. For example, the computer version of the Graduate Record Examination decreases the number of items an examinee must answer based on algorithms that select successive items dependent upon response patterns. The computer determines which items are to be administered and adapts the test based on how a student has answered previous items. Thus, the examinee is not simply receiving a paper-and-pencil test delivered via the computer. In contrast, much of the early drill-and-practice applications contained exercises that were nothing more than textbook content delivered in a computer environment, so-called computer assisted instruction. Making tests adaptive capitalizes on the capacity of the technology. Similarly, e-learning should be more than a replication of an old mechanism of instructional delivery in a new medium. It must capitalize on the new capabilities that are made available by the technology.

It is difficult to envision effective evaluation strategies without distinguishing among different types of e-learning applications. How e-learning in a specific setting is implemented will directly impact the evaluation questions that can be asked and the types of evaluation designs that can be considered. One simple taxonomic classification that may be useful is to distinguish between e-learning applications that are used *in* a course from those used *as* a course. "As a course" means that the entire course is conducted online: There is no face-to-face contact between professor and students, and all communication is done electronically, with the possible exception of an on-site testing requirement.

"In a course" implementation is a more diverse category. It could entail several types of hybrid models. A traditional, brick-and-mortar institution course might include some online requirement in terms of communication or access to supplemental materials. Professors may ask students to communicate through threaded discussions, organized study groups, or simply by e-mail. These interactions might occur via various media, including audioconferencing, videoconferencing, e-mail, or telephone. They may conduct tutorials online rather than in a face-to-face manner. Students may be required to access supplemental materials online through digital libraries, rather than going to actual campus buildings. They may be asked to conduct

projects online, such as computer-based laboratories or simulations. Another way to classify e-learning applications is by the level of interactivity. Is the activity synchronous or asynchronous? Although real-time and interactivity capitalize on the medium and are preferable, many applications exist that are not interactive and tend to be the norm (e.g., chat rooms or list procs).

There is no question that e-learning is more than a passing trend. It will no doubt affect pedagogy and policy at multiple levels of the educational system. At the very least, it will impact institutions of higher education from the perspective of organizational change and infrastructure. It also will stimulate changes in terms of how teaching and learning occur. E-learning will increase institutions' capabilities to reach more students and different and evolving student markets and provide access to educational activities and infinite resources on an anytime/anywhere/any pace basis. E-learning could be the great equalizer, or it could further exacerbate the digital divide.

E-learning is not a new format of instructional delivery. It has its roots in distance learning and continuing education, trying to reach audiences that are not well served by traditional delivery mechanisms (Lyll, 1999; Olsen, 2002). The proliferation of e-learning, as we currently know it, began in the mid- to late 1990s, with both existing institutions of higher education and for-profit companies as well developing courses, e-learning divisions, and sometimes entire universities, attempting to establish a share of the market (Applebome, 1999; Carr, 2000). The goal was not only to reach the underserved, but to deliver courses at low cost with high return (Pohl, 2003; Traub, 2000). The for-profit model, however, did not fit well in many institutions of higher education (Carlson & Carnevale, 2001; Carr & Kiernan, 2000; Traub, 2000). Professors have been skeptical about their role in the for-profit model, expressing concern about their loss of decision-making power and the increase in time commitment (Applebome, 1999; Carr, 2000; Mangan, 2001; Maslen, 2001). There also has been a concern about the extent to which for-profit e-learning may negatively impact instructional quality (Altschuler, 2001; Olsen, 2002; Pohl, 2003; Tedeschi, 2001).

The investment in the infrastructure for e-learning is an expensive proposition, however, and it is not without substantial risks. Many institutions, both nonprofit and for-profit, accredited and unaccredited, have tried and failed. The *Chronicle of Higher Education* and other publications announce on a regular basis that yet another institution has decided to drop its e-learning programs. A recent *New York Times* article (Pohl, 2003) noted three new failures of online master's in business administration programs: Pensare, Quisic, and the State University of New York at Buffalo. Pensare, a collaboration among Duke, Wharton, and other institutions, filed for Chapter 11 bankruptcy when it failed to sell its business courses to corporations (Blumenstyk, 2001). Columbia's online learning venture was shut down (Carlson, 2003) as was NYUonline (Carson & Carnevale, 2001). Some

institutions have withdrawn their participation in collaboratives. For example, the University of Toronto resigned from Universitas 21, and the University of Michigan decided not to participate in a consortium of 18 research universities from 10 countries (Maslen, 2001). Questions have been raised about the motivation for establishing e-learning endeavors. Are institutions merely seizing a business opportunity, recognizing the potential of reaching a limitless pool of clients? Is the primary rationale return on investment? Or is e-learning viewed in terms of a higher educational mission—as a more altruistic and equitable solution to the provision of educational opportunities to segments of the populace that heretofore have not been well served? One perspective is that the success of a program depends on its coherence and connection to the institution (Carlson & Carnevale, 2001). The examination of institutional decisions leads to a wealth of policy questions ripe for research.

E-learning enables us to think beyond the traditional student and the traditional brick-and-mortar institutions. It provides opportunities to reach nontraditional students as well as an entrée into the arena of corporate training. It will play a fundamental role in education and training in the pursuit of continued professional development and lifelong learning, facilitating opportunities not only for learning but also for job certification and qualification. As has been noted by an international panel of educational and corporate training experts, individuals will need to be literate in terms of information and communication technologies (ICT) in order to function effectively in a knowledge society (ICT Literacy Panel, 2002). Fundamental to ICT literacy is the merging and application of cognitive and technical proficiency in a way that enables the transformation of knowledge and skills. E-learning is seen as a primary delivery medium in which such knowledge transformations can occur.

The examination of e-learning must be both deeply systemic and hierarchical. It calls for more sophisticated design and analysis methodologies. Impact will occur at many levels—student, pedagogical, and institutional processes—as well as with interactions crossing levels. Stakeholders will be found at different levels of the institutional hierarchy. They will have different questions, different agendas, and different views of what is important in terms of e-learning. In addition, the focus of examination is a moving target that is constantly evolving, in the hope that results will provide valuable feedback in terms of implementation and process as well as outcome.

EVALUATION ISSUES

There is much debate in the evaluation field as to the appropriate balance between highly quantitative versus qualitative methods, formative versus summative approaches, internal versus external validity, and experimental

versus more ethnographic designs. There is also debate among three movements in the evaluation community about the relative emphasis placed on learning, theory, and evidence as the primary reason for undertaking an evaluation (European Evaluation Society, 2002). The learning movement espouses the premise that stakeholders should learn from an evaluation. The theory-based movement maintains that evaluative information must be interpreted within a theoretical framework. Evidence-based evaluation tends to use experimental methods in an attempt to determine what works. Fundamental to any evaluation, however, is the identification of key questions reflective of the needs of the stakeholders, decisions on appropriate constituencies, methodology, measures, data collection methods, and designs. These components must be adaptable because evaluation should be iterative and cyclical in nature.

Areas of Disagreement and Agreement Among Evaluators

Evaluation methodology is not unique to any particular programmatic or learning environment. There are many commonalities based on the methodology, but there are also unique constraints imposed by specific learning environments. There are unique components in e-learning, as noted in the introduction, particularly the issue of face-to-face contact. Although such characteristics may create specific challenges to the conduct of research, they do not negate the need to balance some of the methodological and theoretical debates within the field of evaluation.

There are at least three major dimensions by which evaluations can be characterized: (a) internal versus external validity, (b) formative versus summative, and (c) qualitative versus quantitative. Evaluators stress different perspectives and different dimensions. Although there is overlap within the dichotomies, there certainly is not consensus. The extent to which evaluators place emphasis on internal versus external validity also can be seen as the distinction between basic and applied perspectives. Proponents of internal validity argue for tight experimental control from which valid explanations of the specific phenomenon can be made (Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002). A concern here is the extent to which experimental designs can be practically implemented in most applied and real-world settings. The more complex the setting, the more difficult it is, of course, to control for extraneous confounds, maintain random assignment, and exercise experimental control. The external validity perspective focuses more on the generalizability of the findings across settings (Cronbach, 1982; Cronbach et al., 1980). Replicability, however, may not be of interest to specific stakeholders. Evaluation should not be an abstraction; its results should have practical consequences.

Perhaps the one issue on which experts in evaluation agree is the belief that good designs are possible, but that there is no such thing as a perfect design. Even on this issue there is a degree of disagreement. Rossi and Freeman (1993), for example, take the position that an evaluation should begin with what they consider the strongest possible design, the randomized experiment, and then begin compromising design principles based on the constraints imposed by the circumstances. Guba and Lincoln (1989), on the other hand, strongly recommend qualitative studies that use multiple perspectives and are responsive to stakeholders. Weiss's (1998) belief is that "form follows function, design follows question" (p. 87).

From this point of view, design decisions are personal choices made by the evaluator in collaboration with relevant stakeholders. Design is seen as a continuing and necessarily flexible process that evolves along with the evaluation. Evaluators must adapt and adjust designs in a cyclical manner in response to the evolving constraints of the program. Methods must match the target of examination. If the components of the program are dynamic, it probably will be necessary to triangulate among multiple methods qualitatively to gain an understanding of the process. Conversely, if the unit of analysis is more static and clearly defined, it may be possible to use focused quantitative methods. Whereas qualitative methods tend to be less precise in terms of the measurement of outcomes, they can provide rich information on the details of the process. Evaluations must be meaningful and relevant to the stakeholders. According to Cronbach (1982), one goal could be to develop a clear understanding of the program and how to improve it, using multiple, small evaluations rather than one large study. Exploratory techniques that help to identify the important issues and questions can then lead to more controlled studies that will help to clarify assumptions by testing theory or hypotheses.

Especially if the goal of evaluation is to provide constructive feedback, it is essential to gain a comprehensive understanding of the processes that affect the implementation of an e-learning program and how improvements might be made. Thus, one perspective is that a primary focus should be on formative evaluation, focusing on multiple components of the program rather than investigating the entire program and triangulating multiple methods of data collection. This necessitates an open-ended form of inquiry that is sensitive to unexpected problems and issues (Cronbach, 1982). It also requires continuous interaction with stakeholders, as the evaluation feeds information back to the program under examination in an iterative manner. Another perspective is to embrace the program in its entirety and examine the extent to which it works. Many evaluators, even proponents of summative work, would balk at this notion, admitting that it is virtually impossible to examine the whole of complex phenomena. Both qualitative and quantitative techniques are appropriate, but for different purposes. Qual-

itative methods can be used to gain an understanding of the program, then the quantitative techniques can be used to validate the prior findings. The triangulation of methods on proximal outcomes and processes can then lead to more summative assessments. Hybrid or eclectic models enable the evaluator to determine appropriate methods and match them to specific questions.

Potential Methodologies

There are a number of methods that might be especially adaptable to the study of the implementation and outcomes of e-learning. Because e-learning is a relatively young and emerging instructional medium, it would be premature to look only at its outcomes. Much valuable formative information can be obtained from the examination of how programs are being implemented and the processes by which they are delivered. One potential technique that seems to be particularly informative for some e-learning evaluative questions is implementation analysis (Love, 2003). A goal of this method is to bridge the gap between process and outcome analyses by examining key elements of a program in an attempt to understand what variables affect implementation. Information gained from implementation analysis can be used to improve the functioning of the program during and after implementation. Another hybrid method that might be successfully adapted from classroom-based computer environments to e-learning is the formative experiment (Newman, 1990). The objective of a formative experiment is to observe how the technology is being implemented, given the specified goals of its use. Instead of the technology as the unit of analysis, the focus is on the environment, including the instruction, roles of the professor and student, the institution as an organization, and the technological infrastructure. Once an understanding of the phenomenon has been gained through the formative experiment, systematic research then can be planned and carried out to examine specific factors that contribute to successful or unsuccessful educational practices. Also fundamental to formative experiments is the role the researcher assumes, namely, that of a facilitator of innovation, rather than an impartial observer.

Another method that is promising for the examination of e-learning, as well as the impact of technology in classroom settings, is systems analysis (Cline & Mandinach, 2000; Mandinach & Cline, 1994). The concepts of interaction and feedback are fundamental to systems analysis. It focuses on dynamic rather than static phenomena. Systems analysis recognizes the need to examine phenomena in real-world contexts with multiple components and multiple impacts that interact across levels of analysis. The implementation of e-learning is a complex and dynamic phenomenon with

many interacting components set within the context of the organizational structures and systems of institutions of higher education. As part of any comprehensive evaluation of e-learning, systems analysis might be used to parse out how those components (e.g., students, faculty, infrastructure, etc.) interact. Systems analysis can be used purely as a heuristic device to explore relationships among the components of a program. It also can be used as a quantitative analytic device that requires the specification of numerical values for the components that then can be mathematically modeled. The quantification process, however, can be problematic, unscientific, and almost haphazard. Yet much valuable information about a program can be gained from systems analysis if one uses the technique solely for its heuristic value.

Given that the tools of systems analysis have traditionally been used in domains other than education, one can also consider using multidimensional scaling, a more standard technique, to quantify the relationships among elements in the system or program. Such analyses could be conducted in parallel with what Yin (1995) refers to as partial comparisons, namely, the focused examination of specific aspects of a program. To study just the specific entities, however, misses a wealth of important data. These sorts of formative activities then can be supported by quantitative data collection methods, including cost-benefit analyses, which can serve as a method of validation. The systems approach allows the evaluator to gain a deep understanding of the program in its entirety, focusing on its dynamic nature and the complexity of interacting issues. The qualitative approach then can serve to address specific theory-driven or empirically based questions or hypotheses that have been raised by the systems analysis. These methodologies can help the evaluator to gain both a qualitative and a quantitative understanding of the impact that the implementation of an e-learning program is having on various entities within the institution and how those components interrelate.

Another evaluation approach that can yield informative data about the effectiveness of a program is econometric modeling. Data can be collected to describe the characteristics and outcomes of a program using economic measures. It is important, however, to recognize the limitations to econometric modeling, as well as the limitations to any particular evaluation method. Such a technique can be a valuable inclusion in a multiple-methodology approach to evaluate e-learning. There are different ways of looking at evaluation and analysis, many of which are sufficiently compatible to form a multimethod approach to research. Although approaches have merit independently in their ability to answer different types of questions, strength often is found in merging their capabilities to examine phenomena from different but complementary perspectives. The selection of a method or multiple methods will depend on the nature of the program, the salient questions, and the needs of the stakeholders.

There are evaluation issues that are especially critical to the some stakeholders concerned about the impact of e-learning. For example, substantial pressure is being leveraged to conduct cost-benefit analyses to validate the expenditures of technology at all levels of the educational system, but particularly e-learning, in terms of hard and measurable outcomes. Stakeholders want to know how much bang they are getting for their buck. There is no question that this is a valid issue from the standpoint of funding. Econometric studies may help to answer some of the financial questions. The economic questions, however, only provide a small piece to a complex puzzle that is highly interactive, constantly evolving, and systemic in nature.

Directions for the Development of Evaluations for E-Learning

There are two possible directions that could be taken in terms of applying evaluation to e-learning. A first option is to develop new methodologies for data analysis, data collection, and design, capitalizing on the affordances of the e-learning environment. A second option is to adapt existing methodologies. It will be prudent for researchers to do both. Adapting existing methodologies might be viewed as a short-term strategy through which researchers can begin to explore a number of the issues that relate to e-learning. These methodologies must use a mixed model, combining both appropriate qualitative and quantitative techniques, given the specific questions we seek to examine. Although some experts believe it is not realistic to combine the methodologies because of epistemological differences (e.g., Guba & Lincoln, 1989), it is common practice to use qualitative methods to examine process and quantitative techniques to study outcomes. Further, it is also possible to combine basic research with evaluation where appropriate given the specific questions to be addressed. The development of new methodologies is a long-term goal.

Evaluation research needs to acknowledge the different objectives for e-learning and the differing perspectives of the stakeholders. A primary target is the higher education audience, where infrastructure is evolving quite rapidly in many locations. The infrastructure and use in precollege settings are not as evolved. Precollege settings may be a potentially fruitful market, however, in terms of assisting schools with implementation issues and training the enormous pool of teachers who will need to use e-learning. Perhaps the fastest-growing arena for e-learning is not for degrees, but for professional development and certification, particularly at the community college level, for corporations, and the military. These venues provide great possibilities for new research and development to assist educational institutions, corporations, and the military in their efforts to educate and train individuals, using e-learning as the medium of instruction.

A set of research questions that focus on essential issues in the implementation and evaluation of e-learning are included in Appendix A. These questions are roughly categorized into assessment of student learning, pedagogical and institutional evaluation issues, and broader policy issues and focus primarily on higher education. Many issues can be generalized to precollege implementations as well. They are the tip of the iceberg. The infrastructure on which e-learning is founded and the dynamic nature of the field make it fertile ground for research, evaluation, and development.

ACTION PLAN

ACCREDITATION AND CERTIFICATION

The accreditation of e-learning is controversial, as many wonder if e-learning institutions should be held to the same criteria as those used for brick-and-mortar universities (Carnevale, 2000; Olsen, 2002; Traub, 2000). One critical piece of evidence for evaluating an e-learning program, however, is whether it meets specified standards of accreditation from one of the six U.S. regional accrediting associations.² Two other agencies, both recognized by the U.S. Department of Education, are the Accrediting Council for Independent Colleges and Schools (ACICS, 2002) and the Distance Education and Training Council (DETC, 2002). It will be important to obtain detailed information from these agencies to gain a better understanding of their criteria for successful and unsuccessful programs. Until recently, the criteria have not been publicly delineated. The DETC Web site (2003a, 2003b) now contains an accreditation handbook and the processes by which accreditation is granted. It outlines a dozen standards that institutions must meet to receive accreditation (e.g., mission, objectives, educational services, financial responsibility). A goal may be for researchers to work with accreditation agencies to help develop the procedures by which to evaluate e-learning programs for accreditation. Such information also will enable researchers to be more effective in working directly with and evaluating specific institutions of higher education.

A second focus is in the area of certification. There are two principal kinds of e-learning certification. First, there is the certification of teachers and professors who are engaged in e-learning activities. Second, there is professional certification for the student that may result from the successful completion of e-learning courses.

For example, Thomas Edison State College in New Jersey provides a course for faculty that results in a Certificate in Distance Learning Program. The objective is to provide faculty with the skills sets necessary to develop and conduct online courses. According to the course description (Thomas

Edison State College, 2002), the objectives for the course are for “faculty to have skills in: (a) preparing course materials, (b) developing computer based presentations, and (c) interacting effectively with students using computer based communications utilities.” The critical issue here is to determine what skills are necessary to enable faculty to be successful in an e-learning environment. Researchers can contribute to the knowledge base of the field by providing information about the requisite skills and best practices for training faculty.

The second form of certification is related to meeting the demands of employment training and professional development. A large segment of the population enrolled in e-learning are adults who are taking courses not just for degrees, but for career advancement, continuing education credits, and possibly professional certification. E-learning provides opportunities for employees to enhance their work skills and helps employers maintain competitiveness in a global economy. Professional development opportunities and certifications are being offered by accredited and nonaccredited institutions as well as by internal training divisions of corporations (e.g., University of Wisconsin Extension, 2002). Certification must necessarily follow the identification and specification of requisite skills and knowledge for technology literacy. Many organizations have attempted to specify requisite skills (e.g., Committee on Information Technology Literacy, 1999; Information Technology Association of America, 2000). One such international project has identified the skills and knowledge needed for individuals to be considered literate in terms of information and communication technologies, ultimately leading to the development of productive employees and successful life-long learners (ICT Literacy Panel, 2002). In precollege education, standards have been written or prescribed as to the technology skills and knowledge students should demonstrate at specific grade levels (International Society for Technology in Education, 2000). As the technology continues to evolve, it will be possible for researchers to make unique contributions to the areas of skill identification, assessment, evaluation, and certification, building on the foundation of existing research and development projects.

SPECIFICATION OF GOALS FOR THE EVALUATION OF E-LEARNING

Several important goals for the evaluation of e-learning can be envisioned. Perhaps the ultimate educational objective for e-learning is to enhance student learning. It will be necessary to develop new forms of assessments to measure students’ knowledge, skills, and learning, while capitalizing on the technology of the e-learning environment. Assessment must be in the service of learning for the individual student or small groups of students. Parallel to these development activities should be an examination of current

practices and formative evaluation as new assessment techniques are designed and implemented. Evaluation should be a critical component here, in support of the basic research and development activities.

The student should also be the focus of other evaluation activities. Evaluations should examine the effects of e-learning on conative processes, student satisfaction, role changes, and other noncognitive variables. Although these variables often are not a primary curriculum objective, they are nonetheless important to a fuller understanding of how e-learning affects students.

A second general area for evaluation is the impact e-learning has on teachers, professors, and pedagogical processes. This is evaluation in the service of effective pedagogy. If one valuable lesson can be drawn from our work on systems thinking (Cline & Mandinach, 2000; Mandinach & Cline, 1994), it is that teachers are the linchpin in the educational system and should be a primary focus of inquiry. It may be premature to concentrate on student learning processes and outcomes without understanding the pedagogical skills teachers or professors need to bring to e-learning and the various systemic interactions that affect how they carry out their instructional activities. Focusing on how e-learning affects teachers and instruction could stimulate a series of research and development activities. Researchers need to examine what underlies pedagogical effectiveness in e-learning and then conduct comparative studies among various implementations of e-learning and between e-learning and traditional delivery.

One can envision job analyses of the skills and knowledge teachers need in an e-learning environment, surveys, observations, and interviews to establish best practices, less successful implementations, and evaluations of variables that influence pedagogical activities. Building on knowledge gained from the series of studies, it might be possible to develop products and services for teachers and professors who intend to enter the world of e-learning. Collaboration with teacher training institutions for preservice and schools for inservice teacher professional development opportunities will be critical.

A third area where evaluation can be useful is to provide information to educational institutions at the organizational level about the impact of e-learning. This is evaluation in the service of the enhancement of education. As was noted above, there are many issues as to why institutions choose to initiate e-learning programs and the commitments that must be made in terms of infrastructure and support. Policy issues provide input for formative and summative evaluations that can help to inform institutions about the key processes and outcomes on implementing e-learning programs.

These areas of focus cannot be considered independently. They are interacting components across levels of the educational system. Some research and evaluation effort must necessarily and logically precede others,

with the enhancement of student learning being a key goal. It is important, in my opinion, to focus evaluation efforts first on the teachers and professors, while continuing a parallel track of research and development activities on individual assessment and student learning, as well as the institutions or schools as learning organizations. The reason for this initial focus on the educators is that they are the key to change and student learning. Without their contributions to e-learning, and instruction more generally, the delivery of education and the students' ability to profit from instructional interactions would be fundamentally and structurally different. It is perhaps ironic that arguments made in this article about the implementation and evaluation of e-learning are not fundamentally different from the implementation and evaluation of any educational reform or innovation. This parallel is done purposefully. There necessarily will be uniqueness to any instructional medium, because of the so-called affordances of learning environments. There also will be many similarities. The key here will be to appreciate the similarities and learn from them, while developing hybrid methodologies to capture the uniqueness created by the affordances of the e-learning environment.

E-learning is one of many emerging technologies that can benefit from the development of new and improved research methodologies that are sensitive to the dynamic nature of the learning environment. The No Child Left Behind (2001) legislation, focusing on precollege educational outcomes, has forced educators and researchers to provide "scientific" (i.e., quantifiable) evidence of impact (e.g., Olson & Viadero, 2002; Shavelson & Towne, 2002). It is reasonable for politicians and policymakers to expect usable data that show return on investment. It is, however, often difficult and challenging, and many times impossible, to deliver the expected evidence when the target of investigation is a dynamic and constantly evolving phenomenon. The challenge that researchers must strive to address is the development of methodologies that are both responsive to the political climate and more importantly sensitive to the systemic complexities in which e-learning and other technology-based instruction are being implemented. Institutions of higher education and schools are set in complex systems with interacting components that are often in conflict. The implementation of a costly program such as e-learning that fundamentally alters the delivery mechanism and medium of instruction only adds to the complexity of the interacting components and the subsequent evaluation. To ignore the evolving and dynamic nature of educational innovation is naïve, thereby diminishing the utility of the research results.

The evaluation of e-learning, just like its implementation, no doubt will be challenging and sometimes problematic, but potentially effective and informative. It will be essential for evaluators to understand and work with stakeholders to reflect their need for both formative and summative infor-

mation. But as in any educational program, the promise to deliver definitive research results and outcomes is both premature and unfair to the institutions and programs, dooming them to less than informative data. Perhaps the most timely and effective contribution the field of evaluation can make in the area of e-learning right now is to help document the factors that either facilitate or impede the development and implementation of e-learning programs so that formative data can be fed back to the institutions for further improvement of implementation. With such formative information, taking into consideration contextual factors, differing objectives, and fiscal commitments, few programs will fail and the future of e-learning may be brighter and their instantiations more effective.

APPENDIX A

A GUIDE TO RESEARCH QUESTIONS FOR E-LEARNING EVALUATIONS

It is the essence of effective evaluations that they be based on a detailed understanding of a specific program, context, and stakeholders. It is possible, however, to delineate a set of research questions and issues. Research and evaluation are not differentiated here. Evaluation methodologies may be more appropriate for some issues and questions. Basic research methodologies may be better for others. The appropriate methodologies, however, should be determined by the specific questions and phenomena of interest.

As mentioned above, it is imperative that researchers ask questions about the impact of e-learning that are not only of scholarly interest, but also of interest to stakeholders such as the institutions and funding agencies. Any evaluation will begin by identifying the specific questions to be examined, then designing appropriate methodology to seek answers to those questions. Some stakeholders may not even know where to begin in terms of asking appropriate questions and will look to researchers to help identify salient issues and relevant questions.

ASSESSMENT OF STUDENT LEARNING

Assessment must be done in the service of learning. Although assessment per se focuses on the individual learner, the evaluation of assessment techniques and a needs analysis can contribute to a broader understanding of the area and feed directly into development activities. Surveys of existing best and least-effective practices should be conducted. These surveys should address the following issues:

- What kinds of assessment techniques can be used, adapted, or developed to maximize the capabilities of the emerging technologies to enhance teaching and learning in higher education? In precollege education?
- What are the innovative applications taking place now and how do we improve upon those activities?
- What kinds of new assessments can be developed so that students can demonstrate both cognitive and affective processes online?
- What kinds of assessment tools can be developed to take advantage of the affordances of the technology, not simply translating old techniques into a different medium?
- How can the assessments capitalize on the continuous and dynamic nature of learning to make the feedback loop between instruction and assessment more meaningful?

PEDAGOGICAL AND INSTITUTIONAL EVALUATION ISSUES

Evaluation of pedagogical and institutional issues can best be seen as focusing on questions and components of the institution embedded within an organization. It can provide a window to what is effective and what is not. It can identify best practices for institutional self-assessment and for dissemination to others. It also can provide essential information for the improvement of process and implementation, identifying additional questions about particular components or interactions within the e-learning system or infrastructure.

At least three types of evaluation issues can be considered. The first set of questions might survey what is being done, forms of implementation, and perceptions about e-learning from students, professors, administrators, and other relevant individuals. It seems logical that these sorts of questions should be asked initially to gain an understanding of the demands of e-learning environments.

- What are the technical/computer skills students need in e-learning?
- What are the technical/computer skills faculty need?
- What are the types of support and help needed for faculty?
- What are the types of implementation strategies and processes?
- What are the factors associated with “successful” implementation and what is “success”? What is considered best practice? And why? What are the factors associated with less successful implementation?

- Are there some courses that lend themselves to e-learning better than others? If so, what courses and what are their characteristics?
- How do students meet course objectives?
- What is the impact on affective processes? Motivation? Collaboration? Self-efficacy? Satisfaction? Self-regulation or metacognition? Self-directedness? What sorts of assistance must be given to less self-directed students to assist them in e-learning environments?
- What role changes, incentives, and disincentives are there for students?

The second type of question is comparative in nature and can be used as a follow-up to the initial set of queries. This type of evaluation either compares the relative efficacy of traditional courses versus e-learning or evaluates different forms of implementation of e-learning.

- Do e-learning courses differ in terms of instructional and assessment methodology? Comparing e-learning courses to traditional courses. Comparing differences among e-learning courses.
- What are the pedagogical opportunities from and implications of using new technology versus the potential problems?
- How does e-learning course delivery differ from traditional courses?
- Are there skills learned better in an e-learning environment? If so, what?
- E-learning uses the term “mortality” for students who fail to complete a course. Is there a differential mortality rate for e-learning versus traditional courses? Are there differential patterns of attrition, retention, completion, and graduation rates? What can be done to reduce the mortality rate?

A third type of evaluation question focuses on technology’s potential to transform teaching and learning from a passive environment to one that is active and learner-centered. E-learning provides “just-in-time access” that allows maximum flexibility for students. At the same time, e-learning creates an environment that makes the professor continuously available. Much to the dismay of many educators, in e-learning, they are constantly on-call. Thus, e-learning fosters fundamental changes in the manner of educational delivery and the roles of students and professors. Research is needed to examine the effects of such 24/7 availability for students and professors. Research is needed to examine how learner characteristics such as metacognition, ability, and motivation affect participation and interactions within e-learning. Similarly, research is needed to examine the pedagogical and

personal characteristics that either facilitate professors in or impede professors from working effectively in e-learning environments. Research also is needed to examine the impact of e-learning on special populations of students.

- What are the implications for e-learning teaching and learning activities for special populations of students (e.g., students with disabilities, nonnative speakers of English, low-socioeconomic-status students)?
- How does e-learning enhance teaching and learning activities while maximizing the use of technology?
- What are the expectations for students in terms of participation, work products, contact with professors and other students?
- Can e-learning effectively be done entirely from a distance with no personal contact? What is the impact of the lack of face-to-face communication for professors and students?
- What is the impact on faculty in terms of demands, workload, accountability, time, satisfaction, collaboration, and intellectual property?
- Many observers of e-learning have noted that professors must be on-call 24/7 to meet the needs of the students, given the asynchronous medium. How does this impact the functioning of academics? What role changes for professors will occur? What are the incentives and disincentives?
- Do students and professors use resources differently in an e-learning environment where there are immediate and infinitely vast resources available at all times?

BROADER POLICY ISSUES

A number of broader policy issues are likely to arise with the implementation of e-learning. Many of the policy questions focus on issues relevant to the level of the organization or institutional infrastructure. Perhaps the overarching policy question is, What makes a grade or degree achieved by e-learning credible or valid? And to whom? Are there differences, depending on the objective (i.e., degree, certification, etc.)?

- What is the role of e-learning in higher education, K–12, or professional training?
- What are the implications of e-learning for access and equity?
- Is there a perception that the worth of an e-learning degree means something different than a traditional degree? Is there differential va-

lidity? What are the perceptions of and implications for institutions and employers concerning the comparability of outcomes?

Does it even matter? Does it matter if the student obtains a degree as opposed to a certificate? An undergraduate degree versus a graduate degree?

- Is the e-learning degree considered equivalent to a traditional degree?
- For employers, what will it take for them to have confidence in the educational outcomes from e-learning?
- What impact on the admissions process for further education will there be for applicants who have e-learning degrees?

- Are there differences in the structure and functioning for nonprofit institutions versus for-profit e-learning ventures? In the validity of degrees and credentials for the two types of institutions?
- Many e-learning programs emphasize what a student has done experientially rather than what is accomplished in terms of educational experiences in class. They stress “getting it done” rather than learning for knowledge. What do stakeholders think should be the educational objectives of e-learning? How can e-learning programs make the instructional and assessment experience meaningful and engender lifelong learning?
- To what extent is there a synergy between e-learning and traditional course offerings on campus? Is there cross-fertilization? How is e-learning used? As a course? In a course?
- What is the rationale for the implementation of an e-learning program? Is the rationale primarily a business strategy to increase student enrollment, or is it a new delivery mechanism to enhance teaching, learning, and scholarship?
- What are the costs associated with the implementation of an e-learning program? Can the institution justify the expenditures on e-learning in terms of positive effects on traditional courses and infrastructure as well?
- What are the key barriers to e-learning and how can they be overcome? What are the potential pedagogical benefits and opportunities?
- What are some of the best practices, and how can they be characterized in terms of synchronous versus asynchronous, as a course/in a course/hybrid models of use?

- What is needed in terms of changes to the infrastructure of higher education institutions to incorporate e-learning technology?

The author wishes to acknowledge the suggestions and feedback made by Henry Braun, Hugh Cline, Carol Dwyer, Carol Myford, and Michael Rosenfeld on an earlier version of this manuscript. The author also wishes to thank Margaret Honey and Andrew Hess for their contributions to the revised document.

Notes

1 A stakeholder here is defined as an individual or individuals who have some vested interest in an evaluation or the program that is being evaluated.

2 The regional accreditation agencies are the New England Association of Schools and Colleges, the Middle States Association of Schools and Colleges, the North Central Association of Colleges and Schools, the Northwest Association of Colleges and Schools, the Southern Association of Colleges and Schools, and the Western Association of Schools and Colleges.

References

- Accrediting Council for Independent Colleges and Schools. (2002). [Web site]. Retrieved February 7, 2002, from <http://www.asics.org>
- Altschuler, G. C. (2001, August 5). College prep: The e-learning curve. *New York Times*. Retrieved April 22, 2003, from <http://query.nytimes.com/search/restricted/article?res=F70D15FF3F590C768CDDA10894D9404482>
- Applebome, P. (1999, April 4). Distance learning: Education.com. *New York Times*. Retrieved April 22, 2003, from <http://query.nytimes.com/search/restricted/article?res=20816F93A540C778CDDAD0894D1494D81>
- Berliner, D. C. (2002). Educational research: The hardest science of all. *Educational Researcher*, 31(8), 18–20.
- Blumenstyk, K. G. (2001, June 15). Online-course company files for bankruptcy. *Chronicle of Higher Education*, Retrieved April 26, 2003, from <http://chronicle.com/weekly/v47/i40/40a03201.htm>
- Carlson, S. (2003, January 17). After losing millions, Columbia U. will close online-learning venture. *Chronicle of Higher Education*. Retrieved April 26, 2003, from <http://chronicle.com/weekly/v49/i19/19a03003.htm>
- Carlson, S., & Carnevale, D. (2001, December 14). Debating the demise of NYUonline. *Chronicle of Higher Education*, Retrieved April 26, 2003, from <http://chronicle.com/weekly/v48/i16/16a03101.htm>
- Carnevale, D. (2000, February 28). Jones International U. to offer accredited online MBA. *Chronicle of Higher Education*. Retrieved April 26, 2003, from <http://chronicle.com/daily/2000/02/20000022802u.htm>
- Carr, S. (2000, June 9). Faculty members are wary of distance-education ventures. *Chronicle of Higher Education*. Retrieved April 26, 2003, from <http://chronicle.com/weekly/v46/i40/40a4101.htm>
- Carr, S., & Kiernan, V. (2000, April 14). For-profit Web venture seeks to replicate the university experience online. *Chronicle of Higher Education*. Retrieved April 26, 2003, from <http://chronicle.com/weekly/v46/i32/32a05901.htm>
- Cline, H. F., & Mandinach, E. B. (2000). The corruption of a research design: A case study of a curriculum innovation project. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 169–189). Mahwah, NJ: Erlbaum.

- Committee on Information Technology Literacy. (1999). *Being fluent with information technology*. Washington, DC: National Academy Press.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F., & Weiner, S. S. (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.
- Distance Education and Training Council. (2002). *The accrediting commission*. Retrieved February 5, 2002, from <http://www.detc.org/content/theaccrediting.html>
- Distance Education and Training Council. (2003a). *The DETC accrediting commission: Process of accreditation*. Retrieved October 6, 2003, from <http://www.detc.org/theaccrediting.html>
- Distance Education and Training Council. (2003b). *The DETC accreditation handbook—2003*. Retrieved October 6, 2003, from <http://www.detc.org/accredittHandbk.html>
- Erickson, R., & Gutierrez, K. (2002). Culture, rigor, and science in educational research. *Educational Researcher*, 31(8), 21–24.
- European Evaluation Society. (2002). *Three movements in contemporary evaluation: Learning, theory, and evidence*. Retrieved January 30, 2002, from http://www.europeanevaluation.org/general/ees_conferences.htm
- Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher*, 31(8), 4–14.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
- ICT Literacy Panel. (2002). *Digital transformation: A framework for ICT literacy*. Princeton, NJ: Educational Testing Service.
- Information Technology Association of America. (2000). *Bridging the gap: Information technology skills for a new millennium*. Arlington, VA: Author.
- International Society for Technology in Education. (2000). *National educational technology standards for students: Connecting curriculum and technology*. Eugene, OR: Author.
- Love, A. J. (2003, January 17–18). *Implementation analysis for feedback on program progress and results*. San Francisco: Evaluators' Institute.
- Lyll, S. (1999, April 4). Distance learning: The British are coming. *New York Times*. Retrieved April 22, 2003, from <http://query.nytimes.com/search/restricted/article?res=F40B16F93A540C778CDDAD0894D1494D81>
- Mandinach, E. B., & Cline, H. F. (1994). *Classroom dynamics: Implementing a technology-based learning environment*. Hillsdale, NJ: Erlbaum.
- Mangan, K. S. (2001, October 5). Expectations evaporate for online MBA programs. *Chronicle of Higher Education*, Retrieved April 26, 2003, from <http://chronicle.com/weekly/v48/i06/06a03101.htm>
- Maslen, G. (2001, May 25). Leery about use of their names, Michigan and Toronto opt out of Universitas 21. *Chronicle of Higher Education*, Retrieved April 26, 2003, from <http://chronicle.com/weekly/v47/i37/37a03802.htm>
- Newman, D. (1990). Opportunities for research on the organizational impact of school computers. *Educational Researcher*, 19(3), 8–13.
- No Child Left Behind Act of 2001. (2001). Retrieved January 30, 2002, from <http://www.nclb.gov>
- Olsen, F. (2002, November 1). Phoenix rises. *Chronicle of Higher Education*, pp. 29–31.
- Olson, L., & Viadero, D. (2002). Law mandates scientific base for research. *Education Week*, Retrieved January 30, 2002, from <http://www.edweek.org/ew/newstory.cfm?slug=20wjatwprls/h21>

- Pellegrino, J. W., & Goldman, S. R. (2002). Be careful what you wish for—You might get it: Educational research in the spotlight. *Educational Researcher*, 31(8), 15–17.
- Pohl, O. (2003, March 26). Universities exporting M.B.A. programs via the Internet. *New York Times*, p. D7. Retrieved April 22, 2003, from <http://query.nytimes.com/search/restricted/article?res=F40D10FC3B540C758EDDAA0894DB404482>
- Rossi, P. H., & Freeman, H. E. (1993). *Evaluation: A systematic approach* (5th ed.). Newbury Park, CA: Sage.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- St.Pierre, E. A. (2002). “Science” rejects postmodernism. *Educational Researcher*, 31(8), 25–27.
- Tedeschi, B. (2001, March 12). E-commerce: The idea of selling classes online has not caught on, but at least one company is pushing ahead. *New York Times*, Retrieved April 22, 2003, from <http://query.nytimes.com/search/restricted/article?res=FB0616FC385E0C718DDDA A0894D9404482>
- Thomas Edison State College. (2002). *Thomas Edison’s certificate in distance education program*. Retrieved February 6, 2002, from <http://www.tesc.edu/CDEP>
- Traub, J. (2000, November 19). This campus is being simulated. *New York Times*, pp. 88–93, 113–114, 118, 125–126.
- University of Wisconsin Extension. (2002). *Certificate programs*. Retrieved February 12, 2002, from <http://www.uwex.edu/disted/certificates.html>
- Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Yin, R. K. (1995). New methods for evaluating programs in NSF’s Division of Research, Evaluation, and Dissemination. In J. A. Frechtling (Ed.), *Footprints: Strategies for non-traditional program evaluation* (pp. 25–36). Arlington, VA: National Science Foundation.

ELLEN MANDINACH is the associate director for research for the EDC Center for Children and Technology. She received her Ph.D. in educational psychology from Stanford University in 1984. She has done extensive work in the field of educational technology. She was part of a team that wrote *Remaking the Concept of Aptitude: Extending the Legacy of Richard E. Snow*. She also was the primary author of *Digital Transformation: A Framework for ICT Literacy*, the product of an international panel of experts on information communication technology literacy. Her most recent work, with Margaret Honey and Katherine McMillan Culp, was *A Retrospective on Twenty Years of Education Technology Policy*, for the National Education Technology Plan.