

Assessment and Evaluation in Online Learning

Ross A. Perkins*

Humans are evaluative by nature. It is quite likely one of the essential characteristics of our species that has allowed us to persist for hundreds of thousands of years. Despite what might be considered our almost instinctual inclination to assess or evaluate, we do not always do it well. There are any number of examples of the wrong questions being asked, or the wrong data being collected, or the wrong analysis being conducted, or the wrong conclusions being drawn. An aphorism, perhaps especially well known to readers of this text, warns, “Don’t judge a book by its cover.” The maxim concerns assumptions about almost anything *except* books. It carries with it the notion that features other than surface ones need to be taken into account when making decisions about something—or someone. This chapter addresses how to evaluate and assess online learning in particular and how to do so in a way that is systemic and systematic. This chapter is not about how to measure student learning within an online course, as that is a separate topic entirely; however, any evaluation of online learning may well include data on student progress.

Although internet-based courses have existed for over thirty years, and though distance education programs are ubiquitous, the history and spread of this innovation do not mean that the fundamentals of instructional design, the sine qua non of any effective course, have always been applied. Because those developing online courses are assumed to be committed to quality, how then can one determine if courses bear the marks of quality instruction? While the measurement of quality is, to some degree, context-dependent, general principles exist that allow designers, instructors, directors—or whoever might

be a stakeholder—to both evaluate and assess online learning in a way that gives them confidence in their conclusions.

Evaluation and Assessment Defined and How They Compare to Research

The terms *evaluation* and *assessment* are sometimes used synonymously. At other times, a differentiation is made that specifies scale, target, or objective. Some may prefer to think of evaluation as large-scale, while assessment is small-scale. Others might assert that evaluation happens to people (in a job role), and assessment happens to programs or policies. However, in this present writing, *evaluate* will describe quantitative measures, and *assess* will describe qualitative measures. One should not assume that these distinctions apply whenever these terms are used, but they’ll allow for clarity in our discussion here.

From a practical perspective, an evaluation emphasizes the collection of numerical or survey data that might include the number of times or times of day that students access a course, student demographics, count of times students participate in discussion boards, grades on assignments, survey responses from questions with Likert-type answers (e.g., Strongly Agree, Disagree, etc.), grades on assignments, and so on. In many cases, such quantitative data becomes part of what has become known as *learning analytics* and can provide unique insights into how students are best supported in online learning. As a simple example, if an evaluation finds that students with a particular grade point average tend to have lower overall

* Ross A. Perkins, PhD, is an associate professor in the Department of Educational Technology at Boise State University in Boise, Idaho. He serves as a coordinator for the online EdS and EdD programs offered by the department. Perkins is the lead facilitator of the online master’s degree capstone course and has taught graduate classes on instructional design and evaluation. Perkins has been designing and teaching online courses since 2002. He earned his doctorate in curriculum and instruction, with an emphasis on instructional technology, at Virginia Tech.

course interactions after week 7 of a course, designers might create opportunities spurring involvement and, hopefully, success. It is common for advanced learning management systems (LMSs) to now have in place analytics systems that generate “smart reminders” for students who consistently lag behind in submitting assignments.

Assessment can happen alongside or independently of evaluation. Because the emphasis of assessment is qualitative (as we are defining it here), one focuses on collecting data such as content of posts on discussion boards, feedback that students give to each other on peer-reviewed assignments, open-ended responses to surveys about what users or instructors think about course assignments or alignment of goals to their learning needs, or transcripts of interviews with stakeholders about the online courses or programs. Both assessment and evaluation must be done—even if at different times and with different purposes—to help create a complete understanding of online learners or online courses or programs.

Evaluation or assessment, no matter how these are defined, should in most cases be thought of as different from research. To be sure, a well-designed evaluation or assessment can be part of research. The planning for either a research-driven measurement or one that is evaluation-centric includes carefully planned data collection and reporting. In the end, the motivation behind doing each one is different, and so is the end point. The purpose of evaluation and assessment is to make ongoing changes or to account for experiences after a course has been run; no other justification for the evaluation is needed. Dissemination tends to be internal, and the conclusions have practical implications. Research instead starts with a literature-based rationale for the questions and, in the end, relates what has been found back to those questions. It tries to align with, contradict, or help evolve theory. The readership of research is wider. An important caveat must be noted: a very good evaluation or assessment of online courses might be done in parallel with research goals. Given that all who produce or consume online courses need continuous examples of their production and implementation, it behooves librarians who are designing online instruction to think about how an evaluation might help inform a wider audience.

Systemic and Systematic Approaches

Librarians who are involved in course design should approach evaluation and assessment both systemically and systematically. A *systemic* approach refers to appreciating the fact that any online course is part of a system of people, tools, technologies, goals, and so on; all aspects are interrelated with varying levels of

connection. A *systematic* approach refers to approaching evaluation and assessment in a well-planned way that follows a series of steps to lead one toward the formation of useful questions, the collection of useful data, and analysis and reporting that take into careful consideration the process itself.

The Systems Perspective

A systemic approach takes into account as many elements that impact online learning (specific to one’s context) as possible. The perspective one must take is that any formal online learning opportunities are part of a system, which means a number of interrelated parts, processes, policies, and personnel are attached to the effort. In many cases, the online learning opportunity cannot take place or be sustained without the other elements functioning. In other cases, even if parts are not *dependent* on one another, changing or adjusting aspects of the online system affects the other parts. *Here’s the recast:* If one carefully takes the entire system into account, the impact of the evaluation or assessment may well be positive. If evaluation or assessment is done without planning, or based on pressures that do not take into account the system relationships, the impact can be irrelevant at best, and misleading or invalid at worst.

As an example, consider a series of online modules for high school students that teach them about library resources, makerspace policies, checkout procedures, citations, copyright laws, and so on. The modules have been set up via the school’s LMS with the intention that students can access the material on a school computer, on their home computer, or even on their mobile devices. However, due to scheduling, students almost never have time to use school computers to explore the modules, 30 percent of students are without consistent access to a computer or the internet at home, and although many use phones or tablets, the courses are not really designed to be mobile-friendly. On top of that, parents are not paying for data plans that allow students to download or stream instruction. As school personnel try to determine why the content is not being disseminated, an evaluation that examines all aspects of the system shows the logistical access challenges as a major cause.

Understanding system impacts is a major step toward conducting a good evaluation, but it is not enough to simply evaluate or assess the connected elements. Indeed, considering the system also requires one to be cognizant of the stakeholders as well and what impact a closer inspection might mean to them. People who have put a good deal of time and energy into creating an online learning experience are generally biased (understandably so), convinced that their product has many positive elements. While it may be true, the point of evaluation and assessment is not

merely to generate a report applauding the efforts, but to investigate what might need to be improved. How will that news be interpreted? If you are doing an evaluation, what data is available? If you are doing an assessment, is it possible to conduct interviews, and if so, who will be conducting them? Will the interviewer be perceived as someone who might impact the interviewees' grade or have influence over their workplace performance? Even if one finds what appears to be "the truth" through an evaluation, it is important to think about who will be reading the report and how it will be disseminated. This is not at all to suggest that an assessment should be avoided; rather, it is a caution that one must sometimes be "as wise as a serpent and as innocent as a dove" when navigating evaluation initiatives.

A Systematic Approach

Adopting a systems view and being wise when approaching an assessment or evaluation also means being systematic. Using a systematic approach means that one follows a carefully considered, reiterative plan, implementing research-based tools, to conduct evaluations or assessments of online learning. Being systematic is important whether one is looking very specifically at a single unit of content within a stand-alone course, or if one is trying to assess the impact of a multiple-course program such as a certificate or degree. It is instructive that the instructional design process itself, often described with the acronym *ADDIE*, begins with *assessment* and works toward *evaluation*. The assessment part of the process often relates to needs assessment, learner assessment, task assessment, context assessment, and so on. Although the last letter of *ADDIE* represents *evaluation*, it is by no means the last thing one does. In fact, evaluation should be among the first things planned when creating a course, a program, or a new policy. It is important to keep in mind that planning to measure the quality of online instruction is an activity that should happen at multiple points during a course or program and that data should be used for continuous improvement. Thus, assessment and evaluation are part of a reiterative cycle—not one-offs with information that never is used to ameliorate whatever has been examined.

To successfully conduct a systematic evaluation and assessment, those involved with the planning must consider the questions *why*, *who*, *what*, *where*, *when*, and *how*:

- *Why?* Likely the most critical question is the *why* of evaluation and assessment. To be sure, no online learning should go without a closer look into its reception, use, and impact. Yet, if the data is never going to be used, or if it is ignored altogether, is the energy involved in developing the

means to measure elements of a course worth the time? The question *why* must be answered *first*, rather than on a post-hoc basis. The answer may be very straightforward: "We are doing this evaluation because we want to know if students have used the course to achieve the following goals . . . ," or "At least one end-of-unit assessment will happen after each unit so that designers can better determine what is and is not clear to the learner." If one does not have clear answers about the *why*, then why evaluate at all?

- *Who?* As noted above, the *who* of the task includes the designer and instructors of the online course, but are they the best people to do the evaluation or assessment? It is often a good idea, if logistically possible, to have neutral parties involved (or at least anonymous surveys) because the type of information one gathers may well depend on who is doing the gathering and how the participants feel their responses might be used. If a librarian asks a student in an online course, "Tell me about how you use the library," a nonanonymous user might extol the "nice" things about the library, while leaving out feedback that could make the person gathering the data feel uncomfortable.
- *What?* The *what* of evaluation and assessment entails asking good questions—well before instruction begins—about aspects of the online course about which one wishes to know more. The questions might relate to one specific part—for example, how an activity in a single unit is perceived or how the assignments follow (or do not follow) the instructions or examples provided. At the programmatic level, one typically examines the *what* of alignment of activities to certain standards of learning or performance outcomes. The data collected can be from usage statistics, results of student assignments, discussion board text, or feedback left on surveys or given in interviews.
- *Where? When?* The *where* (at what point in the instruction) and *when* of a systematic inquiry into learning might largely be the same; one must decide whether to use a formative approach (with data collected along the way to make incremental changes) or a summative approach (with data collected only at the end of the course or program).
- *How?* The *how* of evaluation or assessment of online courses requires a good deal of reading, quite honestly. Any number of texts and articles exist that guide one through a systematic process. These texts or websites may even include rubrics that pose important questions and list research-based criteria. For-profit entities exist that, for a subscription fee, provide training to institutional personnel on use of a proprietary evaluation systems. The advantage of using a commercial product is that the work of developing the rubric,

testing it, training on it, and so on has already been done. In other cases, organizations or institutions develop their own rubrics that guide people in the process of looking more closely at online learning. At the very least, the *how* of evaluation and assessment should include aspects such as a logic model to determine what questions will be asked, how the data will be collected, and how the data will be analyzed.

Evaluating the Evaluation

A final thought: As part of the planning process, it is also important to be “meta-,” by which we ask others to help take a critical perspective on the evaluation plan to see if it contains appropriate questions, data collection schemes, time line, analyses, implementation, and reporting approaches. Since the goal

is ultimately to be able to do far better than simply “judge the cover,” it is helpful to establish people, processes, and procedures to ensure that assessments and evaluations provide the full measure of worth possible. To get a deeper understanding of evaluation, here are some texts you might consider:

- John Boulmetis and Phyllis Dutwin, *The ABCs of Evaluation: Timeless Techniques for Program and Project Managers*, 3rd ed. (San Francisco: Jossey-Bass, 2011).
- J. Michael Spector and Allan H. K. Yuen, *Educational Technology Program and Project Evaluation* (New York: Routledge, 2016).
- Jody L. Fitzpatrick, James R. Sanders, and Blaine R. Worthen, *Program Evaluation: Alternative Approaches and Practical Guidelines*, 4th ed. (Upper Saddle River, NJ: Pearson, 2010).