Chapter 9
Evaluation: Before, During, and After

9.1 Introduction

Evaluation ensures that the learning design will accomplish, or has accomplished, the desired ends. The focus of the final *E—evaluation*—in the ADDIE model is on the learning design, not student achievement per se. Student achievement usually is an important element of design evaluation. But it is not the sole determinant of whether the learning design is successful.

In various formative and summative ways, evaluation is key to each step of the learning design process. In particular, the feedback loop of formative evaluation must be robust in order to maximize the match between the learners' needs and the intended goals of the lessons that compose the design. Chapter 6 introduced OPUS, an acronym for observe, probe, unify, and stage. These critical processes, as I mentioned in Chap. 8, are not for one-time use. They inform the entire ADDIE continuum, from analysis to evaluation. The formative information gleaned through these processes loops back to reshape the learning design as it unfolds.

When a learning design has been fully implemented and an end point has been reached, it is essential to allow time for active reflection. The term *active* is important, because it means that reflective summative evaluation should consist of more than sitting back and concluding, "That went well," or conversely, "That went badly." The arc of the gradual release of responsibility model—from "I do, you watch" to "You do, I watch"—often can frame the designer's reflection on the learning design. Summative evaluation ultimately not only determines the successes or failures of the specific learning design but also guides future iterations of that design and informs the creation of new designs.

D.R. Walling, *Designing Learning for Tablet Classrooms: Innovations in Instruction*,
DOI 10.1007/978-3-319-02420-2_9, © Springer International Publishing Switzerland 2014

9.2 Phase 5: Evaluation

Evaluation is a normal function of human existence. We look in a mirror and evaluate our appearance before we leave for work. We evaluate how well the vegetable garden is growing, whether the lawn needs mowing, and whether we should water or wait for rain. We evaluate our chances of crossing the street safely before we jaywalk. While we do not engage in a formal process, we evaluate *everything*.

Most of this evaluation is formative: reflect and respond. We don't like the way our tie looks with the jacket, so we change ties. We decide the radishes need to be thinned and the mower blade sharpened, but the rain prediction seems reliable enough that we won't bother watering. We consider the speed of the traffic and our aging knees and decide against jaywalking.

Why should it be different when it comes to learning design? The short answer is, it shouldn't. Formative evaluation should take place in each phase of the ADDIE model:

- Is the *analysis* accurate? Is it sufficient to guide the learning design?
- Is the *design* well matched to the analysis? Does it target the desired goals?
- Is the *development* consistent with the design goals? Does it incorporate appropriate resources and strategies that will foster student ownership of their learning?
- Is the *implementation* plan sufficient to match students' learning needs with lesson goals? Can adjustments be made or supports be provided to ensure maximum effectiveness as problems or issues arise?
- Is the *evaluation* consistent with the goals of the learning design? Does it provide sufficient information to guide future iterations of the same design as well as the creation of new designs for learning?

If the answer to any of these questions is negative, then the phase can be modified. Teachers ask students to use a similar reflect-and-respond strategy when they plan and carry out assignments and projects. One strategy is K-W-L: What do you *know*? What do you *want* to know? What did you *learn*? K-W-L was coined by Donna Ogle (1986) as a reading strategy, but it is widely applicable across the curriculum—and it ramps up to the learning design level as a way to guide teachers' active reflection.

In a very real sense, formative evaluation is the comprehensive, self-reflective counterpart to formative feedback, which was detailed in Chap. 12. Formative feedback involved the teacher/learning designer observing students' activities during the implementation phase and providing supports to modify implementation. Formative evaluation incorporates opportunities throughout the ADDIE sequence for the designer to evaluate—that is, to reflect on what has happened, is happening, or is likely to happen—and to respond accordingly. The response may be to do nothing because the design is unfolding as conceived and everything is A-OK. Or the response may be a course correction—a new support, a different approach, an alternative resource, and so on.

The questions above are illustrative. Potential evaluative questions vary over a considerable range, depending on the actual learning design. Let's go back to the lesson idea in Chap 6: to locate, read, and summarize a nonfiction work by Mark Twain using a tablet computer. In Chap. 8, when this example was used during the implementation phase, I suggested several possible points of entry for this lesson. One might be to help students learn how to do an effective Internet search if they have had little experience with that function. Formative feedback might indicate a need for one or two learning supports, for instance, the teacher modeling an Internet search or structuring a small-group brainstorming session to generate potential keywords.

Formative evaluation requires the teacher to reflect on whether this portion of the implementation phase worked: Were some students Internet-savvy and others not? If so, how was instruction differentiated? Once the teacher modeled Internet searching, were students able to search independently? Will additional modeling or guided practice be needed in the future? If so, could some students who have mastered Internet searching mentor other students who need additional modeling and practice?

Within a tablet environment—or any technology-mediated learning situation—formative evaluation must consider students' development of technology skills alongside content knowledge acquisition.

9.3 The Test as Summative Evaluation

Formative evaluation is used to shape, or reshape, the learning design as it unfolds. But what about summative evaluation? Isn't a test sufficient to judge whether the goals of a learning design have been reached?

Student achievement is an important indicator in the overall evaluation of a learning design. It is not the only indicator. And a test is not the only way to evaluate student achievement. Indeed, a poorly designed test can be a faulty way to evaluate both student learning and the learning design. Tests can be simple and easy to use, but just as often, they are simplistic and simply wrong.

Over the past two decades, a burgeoning testing industry has fostered a culture in which The Test—often a standardized test—is the sole or primary vehicle for evaluation. Many politicians and pundits, mostly non-educators, along with a surprising number of educators have gravitated toward this simplistic strategy that often misjudges students' accomplishments and is misused to judge other educational factors, including teacher competence and school or district success.

Harris, Smith, and Harris (2011) point out that "the inherent unfairness of allowing the scores on standardized tests to be our primary—in some cases, our only—way of judging school quality is one of the cruel ironies of the way public education in America has evolved" (p. 45). They are not alone in taking this view. Using tests to judge teacher effectiveness, for example, "assumes that student learning is measured well by a given test, is influenced by the teacher alone, and is independent from the growth of classmates and other aspects of the classroom context" (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012, p. 8). The same could be

said of using a test to determine the effectiveness of a learning design. It's simply too limiting, and the information derived through a test alone can be misleading. The alternative is a multifaceted evaluation.

9.4 Multifaceted Summative Evaluation

Clearly, the learning designer who wants to understand and fully evaluate a learning design must adopt a multifaceted approach. I would pose this challenge in terms of a new Three R's: reactions, responses, and results. These Three R's can be characterized as follows:

- What *reactions* did students have to the learning design? Did they like the learning process? Were they engaged? Were they motivated to go beyond the basic requirements of the lessons?
- What *responses* resulted from the learning design? How did students go about learning? What behaviors did they engage in? Were students able to learn independently? Cooperatively?
- What *results* were achieved by the learning design? What did students learn? Were the goals of the learning design accomplished? Can future learning designs build on this one?

The following are several ways to answer these types of evaluative questions:

Observations. Just as in the OPUS strategy, observing students can provide many indications of whether a learning design has been effective. While observation early on can lead to changes, such as adding supports, in summative evaluation, it is focused on the total arc of the design. Did students move from questioning or hesitant to capable and confident? Do they evidence a sense of accomplishment, a desire to move forward and learn more; or conversely, are they frustrated or confused, discouraged or bored? In either case, teachers can draw conclusions about the effectiveness of the learning design and glean information to help shape future designs.

Discussions. One-on-one, small-group, and whole-class debriefing sessions allow students opportunities to talk about their lesson experiences. What were the pluses and the minuses? Teachers who use active listening and probe for salient details can gather a wealth of information from such discussions. Moreover, these summative discussions also allow students to listen to one another, which has the potential for further enriching the learning experience. For example, if students have worked in small groups, one member might report on a learning strategy that their group used. Other groups, hearing about the experience, may respond with, "We didn't think of that, but here's what we did…." Discussions transform the summative evaluation from an instance of information gathering by the teacher into a shared learning experience for the teacher and the students.

Reflective Questions. K-W-L-type questions, mentioned previously, provide an effective basis for reflection by students on their learning experiences, which in turn can help the teacher evaluate the learning design, as well as to evaluate students' learning. The teacher might take the general knowledge question, for example, and make it specific: "What did you know about doing an Internet search before we started this lesson?" "What did you want to know about using the Internet to find information?" And, "What did you learn about Internet searching?" In positioning tablet technology as a toolbox for learning, these questions about students' comfort and competence with technological functions—using the Internet, using specific apps—are as important as questions about content knowledge and skills acquisition. Of course, the same strategy applies to content: "What did you know about Mark Twain's nonfiction before this unit of study?" "What did you want to learn about Mark Twain's work?" "What did you learn about Twain's nonfiction writings?"

Tests. While complex learning often can be better understood using observations, discussions, and reflective questions, teacher-made tests can be effective for examining students' basic knowledge of both content and technology. A summative test might ask questions such as the following: What is a keyword? How are fiction and nonfiction different? Technology tools can help. For example, teachers with Google accounts can create tests using Google Forms (http://www.google.com/drive/apps.html). Forms provide a template that facilitates creating tests with multiple-choice, text, and other types of responses. It also will aggregate responses in several ways. The program works on tablets and other computers.

Questionnaires. Students also can be asked to reflect on their learning experience by responding to a questionnaire. Reflective questions, as suggested above, can be answered verbally or in writing. The answers in either case tend to be discursive. Questionnaires may be more focused than open discussions but also can run the risk of providing few details. The contrast might be compared to the difference between a still photo and a video. Nonetheless, questionnaires can be productive, are usually easy for the teacher to construct, and the recorded responses that can be considered over time. A questionnaire might focus on learning strategies, involvement or engagement, affect (how students feel about the lessons), materials, or other aspects of the learning design. For instance, the teacher might use a combination of discussion and testing to evaluate whether the design achieved content-acquisition goals but use reflective questions and a questionnaire to better understand students' facility with and feelings about the technology they used. Google Forms, mentioned above, can facilitate questionnaire design, as can other programs, such as Survey Monkey (http://www.surveymonkey.com).

9.5 Summary

Robust formative and summative evaluation ensures that the learning design is accomplishing (during its unfolding) and has accomplished (at its conclusion) the intended goals. Student achievement is an important element of design evaluation, but it is not the sole determinant of whether a learning design is successful. Sometimes content knowledge acquisition is a secondary goal. It would be narrow and simplistic to think of learning only in terms of knowing more "stuff." Content acquisition is enhanced when teachers/learning designers and others involved in the educative process attend to integral factors, such as affect (liking the subject, the device), environment (comfortable surroundings, adequate resources), ethos (how teachers and students interact with one another), and processes (learning strategies, technological functions). Examining all of these factors is the full intent of the final *E* in the ADDIE model.

References

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan, 93*, 8–15.

Harris, P., Smith, B. M., & Harris, J. (2011). *The myths of standardized tests: Why they don't tell you what you think they do*. New York: Rowman and Littlefield.

Ogle, D. M. (1986). K-W-L: A teaching model that develops active reading of expository text. *Reading Teacher, 39*, 564–570.