Book Review

Formative Evaluation for Educational Technologies
by Barbara N. Flagg

Hard cover, $34.50, 259 pages, ISBN 0-8058-0127-8

Reviewed by Betty Collis
University of Twente, The Netherlands

Formative evaluation is not a new concept. It is an activity carried out—generally in an unsystematic and intuitive way—by any thoughtful developer of educational materials. One of its major purposes is to improve the eventual effectiveness of the educational materials while they are still in the design and development phase. While this sounds like an eminently reasonable pursuit, in practice such evaluation is often left at the unsystematic and intuitive level. In particular, relatively little appears in print with respect to specific issues, methodologies, and experiences with formative evaluation during the design and development of educational materials that "use electronic technologies as their means of communication"—such as television, computers, interactive video-disks, compact disks, and teletext. Flagg's book is a welcome contribution to this undeveloped area.

Flagg limits her consideration in this book to formative evaluation used to provide feedback during various stages of the design and development process of educational materials in the above categories, where the purpose of the feedback is to improve the eventual product. She divides her book into three sections. The first provides an introduction, conceptual framework, and analysis of the contribution of systematic formative evaluations. The second section contains five detailed case studies (written by other authors), and the third section contains a closer look at methodological tools and issues involved in this type of formative evaluation. Throughout the book, the development of educational materials is referred to as a four-phase process—planning, design, production, and implementation—and specific examples of formative evaluation relative to each of these phases are discussed and illustrated.

The many examples and illustrations make the book particularly helpful. After her overview of the book and definition of terms in Chapter 1, Flagg begins, in Chapter 2, with three interesting examples: (a) a discussion of the use of
formative evaluation by the staff of the Children's Television Workshop involved in the "Sesame Street" programming as they must make a decision about how to deal with a sensitive issue (the real-life death of one of their long-time characters), (b) the use of formative evaluation by a two-person production team making an interactive videodisk for training purposes so that they could test the user friendliness of the program before the costly process of mastering the videodisk was completed, and (c) the use of formative evaluation to help decide how to adapt an educational software program for an age group different from that for whom the program had been originally designed.

In Chapter 3, Flagg discusses the application of formative evaluation during the development of the Voyage of the Mimi multimedia materials. The case studies (written by persons involved in the different evaluations) in Chapters 4 through 8, are well described by the chapter titles:

- "Formative Evaluation in the Design of Educational Software at the Children's Television Workshop"
- "The Program Development Process of the Agency for Instructional Technology" (involving computer/video materials)
- "The 'Palenque' Project: Formative Evaluation in the Design and Development of an Optical Disk Prototype"
- "Formative Evaluation of Interactive Training Materials"
- "In Search of a Methodology: Evaluating 'SCOOP,' the WGBH Teletext Magazine for High School Students."

The final case study leads into the last section of the book, which relates more generally to methodological approaches and issues. Chapter 9 is an excellent reworking of classic program evaluation issues, reconsidered specifically for the formative evaluation of electronic learning products during their development cycles. There are four basic questions addressed in Chapter 9:

1. Which purposes will the evaluation serve and for whom?
2. What are the evaluation questions?
3. Which methods of inquiry are appropriate?
4. What measures will be used with whom under what conditions?

These are well-known questions in the general framework of formative program evaluation, but Flagg does a good job of discussing them in terms of electronic learning materials and in bringing in references from the preceding examples and case studies. As an example, with respect to Question 1, Flagg references 13 approaches to educational evaluation, but then chooses four of them for particular comment with respect to electronic learning materials. These four approaches are: "connoisseur-based," "decision-oriented," "objectives-based," and "public relations." She illustrates the first three approaches in terms of the case studies and examples, and comments that the fourth approach points out the importance of the formative evaluator knowing who will be using his or her evaluation and for what purposes.
Chapters 10 and 11 are more specific. Chapter 10 focuses on methodologies for the formative evaluation of the eventual user friendliness of an electronic product. Chapter 11, somewhat unexpectedly, focuses on anticipating the user’s affective response to the learning product, in terms of “attention, appeal, and excitement.” The importance of these aspects to the eventual educational value of the materials is defended: “The first step in the comprehension or persuasion process is to attract and maintain the learner’s attention and interest... The program must compete with other stimuli to be positively received by the learner... Reception of a program is a necessary but not sufficient condition of learning” (p. 196). Nonetheless, devoting an entire chapter to this topic leaves the reader wondering if other points of consideration are more important.

The final chapter deals, somewhat briefly, with a major problem—what is meant by the effectiveness of electronic learning materials and how can this effectiveness be measured? Flagg identifies two of the thorniest aspects of this issue—the fact that many projects involving electronic learning products have vague or unarticulated goals, and even when there are goals, these are often long term. She also identifies the value of helping the project team agree on observable “precursor behaviors,” although in this chapter she gives fewer examples relative to these points than she has given throughout the rest of the book.

This book has many strengths. It is up to date, both in its examples and in its extensive reference lists at the end of every chapter. It is written by someone who obviously has had firsthand experience of what she is talking about. It is easy to read and therefore accessible to relative beginners in the area of formative evaluation, but the book could also be a challenging text for a graduate course or be used in the “real world.” It offers convincing evidence that applied formative evaluation research is not only an interesting and useful exercise in creative problem solving but also a theoretical and scholarly activity.

What would I add to the text? Perhaps two things. One would be more discussion of problems confronting the implementation of this sort of formative evaluation orientation. Flagg explicitly does this on pages 25-29 and 241-242, and what she says is realistic, but the sort of examples she gives throughout the rest of the book very infrequently occur in practice. In my experience, a major reason why systematic formative evaluation does not often occur during the development of electronic learning materials relates to the willingness of the design/development team to have outside criticism, or to think they can benefit from it. Flagg mentions this, and perhaps her book can indirectly help to better inform the field of the value of on-going openness to change and improvement, but the receptiveness of all the people involved in her examples and case studies to the formative evaluation process might present an unduly “rosy” view of the evaluative process to someone who has not been involved in it.

Another issue Flagg does not address is that of who should be responsible for the formative evaluation—an external evaluator? The project team itself? The project manager? Also, the time and resource costs of the field-involvement described in each of Flagg’s examples is considerable. Again, she briefly notes this, but it would be useful to analyze these requirements more specifically. What does the formative evaluator need to anticipate in terms of time and support and
organization if he or she wishes to incorporate various rounds of teacher, expert, and student evaluation into the design and development project? Again, from my experience, organizing groups of students to “try something out” or groups of experts to comment on a walk-through of a program storyboard requires a lengthy period of preparation, support for the released time of those involved, probably institutional support and perhaps parental agreement for the student involvement. It is an expensive, time-consuming, and labor-intensive business. A more explicit description of these organizational aspects would help the formative evaluator better plan for his or her work.

The second thing I would add to Flagg’s book is more of a summary. A final chapter, in which each of the four phases of program development introduced in Chapter 1 is recapitulated with some sort of summary statement of useful methodologies and evaluation issues (and perhaps referenced to the case studies and examples in the book), would help the reader sort out the overall picture of these phases relative to the many times they are mentioned throughout the book. The text seems to end somewhat abruptly without this, and I was surprised when I turned the page after the end of Chapter 12 and found I was at the author index for the book.

However, despite these comments, I recommend this book strongly and am looking for ways to use it as a text in my own classes as well as a reference in my writing and an exciting source of new citations and studies in an area which, as Flagg says, is largely characterized by “fugitive” in-house reports, if reports are even available. It is a good book from both theoretical and practice-oriented perspectives, and I hope it leads the way for other books in the complicated area of the formative evaluation of new technological instrumentation.