DOCUMENT RESUME

ED 348 021 IR 015 758

AUTHOR	Savenye, Wilhelmina C.
TITLE	Alternate Methods for Conducting Formative Evaluations of Interactive Instructional Technologies.
PUB DATE	Feb 92
NOTE	30p.; In: Proceedings of Selected Research and Development Presentations at the Convention of the Association for Educational Communications and Technology and Sponsored by the Research and Theory Division; see IR 015 706.
PUB TYPE	Information Analyses (070) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Computer Assisted Instruction; *Evaluation Research; *Formative Evaluation; Hypermedia; Instructional Development; *Instructional Effectiveness; *Interactive Video; Literature Reviews; Microcomputers; Models; Multimedia Instruction; *Research Methodology

ABSTRACT

This paper argues that formative evaluation of instruction, which is generally agreed to be critical for instruction in any medium, is even more crucial when the instruction is to be delivered by interactive technologies such as computers, interactive video, hypermedia, or the various forms of interactive multimedia systems. It begins by discussing formative evaluation as a formal step in instructional development models, noting that the models rarely specify where in the process such evaluation should take place. The foundational assumptions and biases of the paper are then discussed, including the current controversy over qualitative and quantitative research and various issues involved in selecting the research methods to be used. Several types of data collection and analysis methods that can be used to answer important questions concerned with interactive instructional technologies are considered, and the use of a method that is appropriate to answer the particular evaluation questions involved is advocated. A discussion of the benefits of considering alternate methods of formative evaluation introduces a review of the results of evaluations of the overall effectiveness of interactive technology-based instructional programs, primarily computer assisted instruction and interactive video. An overview of planning and conducting formative evaluations as an on-going process through all phases of design and development is then presented. Multiple methods for collecting and analyzing data are also reviewed, with emphasis on the selection of appropriate methods. Suggestions for reporting the results and a summary of some of the major considerations in conducting formative evaluations conclude this paper. (63 references) (BBM)

Title:

Alternate Methods for Conducting Formative Evaluations of Interactive Instructional Technologies

Author:

Wilhelmina C. Savenye

Alternate Methods for Conducting
Formative Evaluations of
Interactive Instructional Technologies

Paper Presented at the National Conference of
The Association for
Educational Communications and Technology
Washington, D.C.
February, 1992

Wilhelmina C. Savenye
Learning & Instructional Technology
Division of Psychology in Education
Arizona State University
Tempe, AZ 85287-0611
(602-965-4963)
BITNET - ATWCS @ ASUACAD

Interactive technologies such as computer-based instruction, interactive video, hypertext systems, and the broad category of interactive multimedia systems are increasingly making an impact in educational and training settings. Many teachers, trainers, decision-makers, community members, and business and government leaders contend that these technologies will change the face of education and training (Ambron & Hooper, 1990, 1988; Interactive Multimedia Association; Lambert & Sallis, 1987; Schwartz, 1987; Schwier, 1987; U.S. Congress, OTA, 1988). Of concern is the claim by some that educational technologists are no longer the leaders in developing these technologies. Some are concerned that educational technologists are being left behind. One reason may be that technologists have neglected to prove the value of conducting evaluations of these technologies and thus cannot always show data to prove that systematically designing technological innovations for education makes a difference. One common feature of models for systematically designing instructional materials is that draft versions of these materials be tested with representative learners to ensure that the materials are effective (Andrews and Goodson, 1979) in the process called formative evaluation. Other terms, including developmental testing, pilot testing, field testing and validation, are occasionally used for this process. Most designers also draw a distinction between formative and summative evaluation. Generally, formative evaluation is conducted for the purpose of improving the instructional program through revision (Dick & Carey, 1990; Gagne, Briggs, & Wager, 1988; Morris &

Fitzgibbon, 1978). Summative evaluation is usually conducted to determine the overall value of a program, such as to make a "go" or "no go" decision about a completed program, often comparing it with other programs or approaches (Dick & Carey, 1990; Geis, 1987). The focus of this paper is on formative evaluation of interactive technologies.

While educational technologists acknowledge that we should be conducting more research and more research on formative evaluation specifically, the value of formative evaluation in enhancing the learning effectiveness of instruction has often been shown. One example is the recent meta-analysis conducted by Fuchs and Fuchs (1986). These researchers analyzed the results of 21 studies which investigated the effects of formative evaluation of materials developed for mildly handicapped students. They found that systematic formative evaluation significantly increased students' achievement when students used the resulting materials.

When to conduct formative evaluation is not always clear in instructional design models. Simplified graphics used to illustrate some models show formative evaluation being conducted using an almost-final draft of the instruction at the end of the design and development process. This may be appropriate for simple, print-based instruction, however most models show that formative evaluation is an ongoing process conducted throughout development. For example, Dick & Carey (1990) in their widely-used model recommend that formative evaluation include frequent reviews of materials, several 1:1 tryouts with learners who represent several segments of the target population, at least one small-group tryout, and a field test in the actual learning setting. On-going formative evaluation is particularly important in developing costly and labor-intensive interactive technology-based instruction. In fact, many small but critical design aspects of the interactive instruction, such as user-interface features like icons, menus, and navigational tools, are evaluated continuously in small segments as the project evolves.

While it is generally agreed that formative evaluation of instruction developed in any medium is critical, it is the premise of this paper that it is even more crucial when the instruction is to be delivered via interactive technologies, such as computers, interactive video or the various forms of interactive multimedia systems. For example, most design models for developing computer-based instruction (Gagne, Wager, & Rojas, 1981; Hannafin & Peck, 1988; Smith & Boyce, 1984) as well as interactive video for instruction (Kearsley & Frost, 1985; Savenye, 1990) call for conducting formative evaluations. There is a tendency, however, for evaluation to be neglected during development due to constraints of budget, personnel, and scheduling (Brenneman, 1989). This is, unfortunately, especially true in large-scale technology-based projects, in which costs are already high. Patterson and Bloch (1987) contend that formative evaluation is often not done during development of interactive instruction using computers. They mention that one reason for this may be that decision-makers in education and industry hold a negative attitude toward formative evaluation. These authors contend that educational technologists should recognize this fact, and help such constituents see the value of formative evaluation.

At the same time as formative evaluation procedures, and sometimes systematic instructional design itself, is sometimes assailed by developers of interactive instruction who have backgrounds other than instructional design, there are increased calls by funding sources and educators to use "qualitative" or "naturalistic" research methods in studying school and training processes and projects (Bosco, 1986; Clark, 1983). These methods are considered alternatives to more traditional approaches, such as using scores on paper-and-pencil achievement tests in experimental and quasi-experimental comparisons of programs. Sadly, at times producers of interactive educational materials have responded by collecting data on inappropriately small samples of learners, watching learners without first determining the evaluation questions, collecting too much data to effectively analyze later, and/or focusing on how well students like instructional programs, rather than on whether students learn through the programs.

Similar to the contentions about use of formative evaluation during development of interactive instruction, researchers have noted that there is little research being conducted on formative evaluation (Chinien, 1990; Geis, 1987; Patterson & Bloch, 1987). Thus it is likely that, just when we should be improving the ways we conduct formative evaluations, especially when developing interactive instruction, we are not conducting the research necessary to develop and test these improvements.

The purpose of this paper is to present methods for planning, conducting, and using the results of formative evaluations of interactive technology-based instruction. The focus is on practical considerations in making evaluation decisions, with an emphasis on alternate methods in formative evaluation.

Foundational Assumptions and Biases of This Paper

Lest the reader believe this author is contending that all these approaches to evaluation are new, we need only look at Markle's 1989 "ancient history of formative evaluation" to remind ourselves that versions of these processes have been used by designers for many years, although often informally. Markle contends that even in the early, more "behavioralist", days of instructional design, developers listened to their learners, watched them carefully, and humbly incorporated what learners taught them into their drafts of instructional materials. Similarly, what recent authors, especially computer scientists, are calling testing in "software engineering (Chen & Shen, 1989; "prototype evaluation" (Smith & Wedman, 1988), "prototype testing", quality assurance" (McLean, 1989), or "quality control" (Darabi & Dempsey, 1989-90) is clearly formative evaluation by another name.

A controversy swirls in education and in our field regarding the relative value of "quantitative" and "qualitative" investigations. "Quantitative" usually means experimental or quasi-experimental research studies; in evaluation, these studies often compare one approach or technology with another. Some technologists have called for the abandonment of quantitative comparison studies (cf. Reeves, 1986), claiming they answer the wrong questions in limited ways.

Use of the term "qualitative" research is less clear. It usually refers to studies using anthropological methods such as interviews and observations to yield less numerical descriptive data. Unfortunately the resulting studies sometimes employ less than sound research methods. Such studies have given the term "qualitative research" a bad name in some circles, notably among those who are strong advocates of the sole use of quantitative methods.

It is the view of this author that when planning evaluations of interactive technologies the debate is not useful. Most practical educational developers have for many years used a blend of quantitative and qualitative methods in evaluation. "Quantitatively", there is, for example, a long tradition of using pretests and posttests to compare the performance of learners in a control group with those who have used a new educational technology program. "Qualitatively", evaluators have long collected attitude data using surveys, interviews and sometimes observations. Alternate research methods allow for collecting more types of qualitative data to answer the new questions which emerge in evaluating new technologies.

A more fruitful approach to the issue of which types of research methods to use is to select whatever methods are appropriate to answer the particular evaluation questions. Such an approach is in line with the recommendation of Clark (1983) that we reconsider our study of media. This approach is also similar to the ROPES guidelines developed by Hannafin and his associates (Hannafin & Rieber, 1989; Hooper & Hannafin. 1988) which blend the best of behavioralism and cognitivism in what they call "applied cognitivism". Finally, selecting methods based on questions supports Driscoll's (1991) suggestion that we select overall research paradigms based on the most urgent questions. Driscoll adds that instructional technology is a developing science in which "numerous paradigms may vie for acceptibility and dominance" (p. 310).

A final fundamental bias of this paper is that the most important question to ask during formative evaluation remains, "How well did the learners learn what we intended them to learn." This paper presents several types of data collection and analysis methods to answer important questions which emerge when interactive instructional technlogies are involved. Yet if we answer these questions and neglect whether students learn we will not know whether the technological innovation has any value.

Benefits of Considering Alternate Methods of Formative Evaluation

While, as noted earlier, the ideas are not strictly new, there are several reasons for a new and deeper look at formative evaluation when interactive technologies are involved. At one level developing instruction using technologies such as computers adds complexity to what can go wrong and what needs to be attended to, because there are hardware and software issues involved (Patterson & Bloch, 1987). For example, interactive systems are often multimedia systems, so formative evaluation questions often include how efffective graphics, animations, photographs, audio, text and video are in any lesson segment.

A second reason a new look at formative evaluation methods is warranted is that interactive technologies now allow developers to collect data about learners and learning that could not technologically be collected before. We can thus look at learning in new ways, and answer questions we may have wanted to answer before. For example, computer-based lessons can be programmed to record every keypress a learner makes. Developers can thus determine how many times a student attempts to answer a question, what choices they make, and what paths they follow through hypermedia-based knowledge bases. One danger, of course, is that developers can become "lost in data", collecting data without regard to evaluation questions and what to do with the data.

A third reason for study of formative evaluation methods is that with the recent renewed emphasis on qualitative research in education, have come increased numbers of good studies using alternate research methods. It is fortuitous that these methods, many borrowed from other fields, particularly anthropology and sociology, are being tested and results reported at a time when developers of interactive technologies are looking for new ways to measure how much our technologies help learners learn.

A final reason to expand our views of evaluation methods is to "push the envelope" of useful knowledge in our own field of educational technology, as many researchers and developers have called for. Reigeluth (1989) states that our field is now at a crossroads with considerable debate taking place regarding what we should study and how. Winn (1989) calls for researchers to conduct descriptive studies yielding more information about learning and instruction. In his often-cited article, Clark (1989) agrees with Winn, and states that researchers should conduct planned series of studies, selecting methods based on sound literature reviews. His recommendation that we conduct prescriptive studies to answer why instructional design methods work can especially be followed by evaluators using alternate research methods.

Results of Evaluations of the Overall Effectiveness of Interactive Technology-Based Instructional Programs

Computer-based Instruction

Recently several researchers have reported the results of meta-analyses of general evaluations of the effectiveness of various types of interactive technology-based programs. Evaluations of computer-based instruction (CBI) and interactive video will be presented. Although Ambrose (1991) has presented a literature review regarding the potential of hypermedia, there has not to date been a meta-analysis of research studies indicating the effects of these newer multimedia systems on learning. It is hoped that such meta-analyses may be conducted on these technologies in the future.

Several researchers have conducted meta-analyses to study the overall effects of CAI on student learning. For example, Kulik, Kulik and Cohen (1980) reviewed 59 evaluations of computer-based college teaching. They found that college students who learned using computer-based instruction (CBI) generally performed better on their exams than students who learned using traditional instruction, although the differences were not great. For example, they reported that the average exam score in CBI classes was 60.6 percent, while the average score in traditional classes was 57.6. While only eleven of the studies they reviewed reported attitude data, these researchers also reported that students who learned using CBI had a slightly more positive attitude toward learning using computers, and toward the subject matter. The most significant finding in this meta-analysis was that in the eight studies which investigated effects of CBI on instructional time, students learned more quickly using computers than in conventional classes. This finding of this study has often been cited as one powerful reason to use CBI in college classes.

Kulik and several other researchers (Kulik, Bangert, & Williams, 1983) also conducted a meta-analysis on the effects of computer-based teaching on learning of secondary school students. In this study they reviewed 51 evaluations that compared the final examination scores of students who had learned using CBI with scores of students who had learned using conventional methods. The results of this study indicated that learning using computers may be even more effective for younger students than for the older students described earlier. Again, students in CBI classes performed better. The average effect size in the CBI classes was .32 standard deviations higher than in conventional classes. Another way to describe this difference would be that students in the CBI classes performed at the 63rd percentile, while those in conventional classes performed at the 50th percentile.

In this meta-analysis, the researchers also investigated the effects of CBI on student attitudes. There was a small significantly positive effect of CBI on attitudes toward subject matter, computers and instruction. Only two of the studies they reviewed investigated instructional time, and in both students learned more quickly using computers. Thus, recent studies have indicated the effectiveness of CAI and CBI on student learning.

Interactive Video

Savenye (1990) presented findings of general reviews of evaluations of the effectiveness of interactive video as well as specific types of multimedia studies which have been conducted. To summarize her results, this researcher found that in the evaluations (Bosco, 1986; DeBloois, 1988; Slee, 1989) interactive video generally helped students learn better than they did through traditional instruction. She cautioned, however, as did Bosco, that when studies used statistical analyses differences tended to be smaller. In addition, this researcher reported that learners usually have positive attitudes towards learning through interactive technologies. As in the studies on CAI, researchers often found that learners learn faster using interactive video.

McNeil and Nelson (1991) conducted a meta-analysis of studies which evaluated cognitive achievement from interactive video instruction. These researchers used criteria including presence of learning measures, use of experimental or quasi-experimental design, and sound methodology to select 63 studies from an initial list of 367. One strength of their meta-analysis is that many studies had not been published, thus avoiding the bias toward significance of published studies noted by some authors.

Similar to Kulik, et al.'s results of meta-analyses on CBI, McNeil and Nelson found an overall positive effect size (.530, corrected for outliers) showing that interactive video is effective for instruction.

These researchers conducted several types of analyses in an admirable effort to isolate instructional design factors which contribute to the effect size. These analyses revealed that the effect sizes were homogenous, "but the selected independent variables did not explain the achievement effect," (p. 5). They did, however, note that there were some significant teacher effects indicating that interactive video was somewhat more effective when used in groups rather than individually. The authors remind us of the important role of the teacher in interactive instruction. In addition, similar to results noted by Hannafin (1985), as well as Steinberg (1989) for some types of learners, program control appeared to be more effective than learner control.

These researchers explained their results by noting that interactive video instruction consists of a complex set of interrelated factors. Reeves (1986) concurs. It will be a continuing challenge to researchers studying interactive instruction to isolate factors crucial to the success of innovative technologies.

Planning Formative Evaluations of Interactive Technology-Based Instruction

The following sections of this paper will present an overview of planning and conducting formative evaluations of interactive instructional programs. As noted earlier, it is assumed that formative evaluation is an on-going process, with activities conducted throughout all phases of design and development.

Begin Early

One key to conducting cost-effective and useful formative evaluations is to begin planning early, ideally from project inception. By beginning early, the goal of the formative evaluation is determined early as well. The subsequent processes and methods can be carefully selected and planned to collect the most important information. Stakeholders, managers, reviewers, instructors and learners can be identifed early, thereby limiting delays during development. Similarly, members of the development team who will assist in data collection can be identified, enlisted and briefed early. Early planning, in fact, can enable developers to collect data that would be impossible to collect if not identified early, because the systems (such as computer programs) to collect these data might not be developed when needed, or at all.

Determine Main Evaluation Goal

It is most useful for communication and efficiency purposes for one clear goal to be determined for the formative evaluation. Although developers may want to investigate many questions, the evaluation goal is usually some variation of how effective the interactive instruction is with learners and instructors. The 80/20 rule applies as much when conducting evaluations as it does in most other activities, that is, 80% of the benefits are derived from 20% of the effort. As development progresses keeping the evaluation goal in mind will yield maximum results and avoid team members wasting time on less important details. Maintaining a focus on one clear evaluation goal thus enables developers to keep a view of the forest, rather than getting lost in the trees.

Determine Major Evaluation Questions and Sub-questions

The major evaluation questions are derived from the evaluation goal. In formative evaluations of interactive instruction, as in evaluations of instruction using other media, there are typically, three major evaluation questions:

1) How well does the instruction help the students learn (an achievement question)?

2) How do the learners, instructors, and other users or constituents feel about the instruction (an attitude question)?

3) How is the instruction implemented (a "use" question)?

(Higgins & Sullivan, 1982; Morris & Fitzgibbons, 1979; Sullivan & Higgins, 1983).

To answer each question, evaluators and developers determine data to be collected, select data collection methods, develop instruments and procedures, and determine how data will be analyzed. One way to plan the evaluations to both keep the focus clear and make procedures most efficient is to develop a matrix to guide the evaluation and development team. Under each major evaluation question can be listed the related subquestions. Beside each question, as headings across the matrix would be "data sources" (instructors, learners, administrators, expert reviewers, etc.), "data collection methods" (measures of initial learning, measures of learning transfer, attitude surveys, interviews, observations, etc.), and "instruments" (pretests and posttests of achievement, instructor questionnaires, observational checklists, etc.) (cf. Savenye, 1986a).

Interactive technologies both allow for, and call for, different subquestions related to the three major types of evaluation questons. They also call for expanded views of what data can be collected and analyzed and how these data can be used. Developers and evaluators of interactive programs have a responsibility to add new methods to their evaluation "toolkits", to maintain an open view with regard to questions which need to be answered, as well as to report the results of their evaluations to benefit their colleagues who develop interactive technology-based instruction in all settings. The latter responsibility, in particular, has been noted by many authors (cf. Clark, 1989; Patterson & Bloch, 1987; Reigeluth, 1989; Winn, 1989).

Alternate methods of conducting evaluations are most useful, in fact may be critical, in answering the third major type of evaluation question - how is the interactive instruction implemented or used. Flexible, open views with regard to "what is really happening" when innovative approaches and technologies are used can result in finding that a critical component of instructor training, for example, had been left out of the initial design, or that learners are using the technology in ways developers never anticipated; in fact, they may be using it in better ways. This can yield what Newman (1989) calls answers to how the learning environment is affecting the instructional technology. Newman elucidates: "How a new piece of educational technology gets used in a particular environment cannot always be anticipated ahead of time. It can be argued that what the environment does with the technology provides critical information to guide design process" (p. 1). He adds, "It is seldom the case that the technology can be inserted into a classroom without changing other aspects of the environment," (p. 3), a fact often noted by instructional systems designers.

When such questions are not brought up and investigated, an instructional innovation can fail, as those who developed "programmed learning" in the sixties, or who have implemented educational technologies in other cultural settings without getting participant "buy-in" have learned. In other words, selecting evaluation methods with a critical eye toward the realities of what can be happening when we use new technologies is called for.

In addition, with the prospect of continued lack of support for "basic research" in educational technology, developers can contribute to the knowledge base in our field by conducting "applied research" in the form of rigorous high-quality formative evaluations and publishing their methods and results. Alternate research methods can be used carefully to answer "open questions" related to implementation of technology, such as how youngsters make decisions as they proceed through a simulation, or how teachers use an interactive videodisc program with whole classes. Results reported by an evaluator in one study can yield instructional design and implementation guidelines that developers can use and test. No less important to the continued improvement in our knowledge is that researchers can use results of "naturalistic methods" to formulate questions and isolate factors which can subsequently be investigated using experimental method, yielding causal interpretations.

The following section of this paper will present a discussion of multiple methods for conducting formative evaluations of interactive instruction, with particular attention to selecting appropriate methods.

Data Collection and Analysis Methods

While the goal of formative evaluation is to improve the learning effectiveness of the programs, the choice of methods for conducting evaluations is not clear-cut. As recommended by Jacob (1987) in her review of qualitative research traditions, methods should be chosen based on the research questions to be answered. In the case of evaluation, where resources are limited and the value of the process is not always clear to constituents, selecting methods should be driven by evaluation questions.

In addition, it is important that developers and evaluators contribute to our knowledge of effects of instructional design factors. As noted by many researchers (McNeil & Nelson, 1991; Reeves, 1986), instruction based on interactive technologies relies on many individual factors for its success, and each program is often unique in its approach, use of media, etc. The challenge to determine what factors make a difference in learning is great.

In the discussion below will be interwoven the utility of various alternate research methods with traditional methods. The methods will be presented with relation to major areas of evaluation types.

Instructional Design Reviews

Most instructional design models include the recommendation that draft versions of instructional materials be reviewed periodically during development. It is particularly important that aspects of interactive programs be reviewed at many stages during development. Geis (1987) recommends that materials might be reviewed by subject matter experts, instructional designers, technical designers such as graphic artists, instructors, individuals who have special knowledge of the target audience, influential community leaders, project sponsors, previous students and project editors. In a large-scale school science project, for example, initial objectives might be sent with brief descriptions of video and computer treatments for lesson segments to scientists, instructional designers and teachers using other materials developed by the organization. Subsequent reviews might elicit responses to depictions of computer menus, descriptions of branching options and simulation and game segments, as well as video storyboards.

While use of drafts of print materials, scripts and storyboards for reviews is traditional in formative evaluation reviews, it should be noted that computer programs, interactive video lessons, and interactive multimedia presentations are often too complex for many reviewers to evaluate in print form. Many evaluators, therefore, submit prototype versions of aspects of lessons, such as crucial menus, operational draft segments of simulations, or selected lessons to reviewers. Reviews are typically sclicited early during formative evaluation activities to answer format, style, and content questions, and reviews continue on an on-going basis.

Determining Learning Achievement

Paper-based Tests. Traditional measures of achievement are still appropriate for use in determining how well learners perform after completing interactive lessons. Such measures are usually forms of paper-and-pencil tests. What is critical is that the test items match the learning objectives developed during design (Dick & Carey, 1990; Higgins & Rice, 1991; Sullivan & Higgins, 1983). Without such a match the test is often not useful, and, unfortunately this can often be the case in evaluating interactive programs in which developers let technical "bells and whistles" drive the design process. The decision to use paper-and-pencil tests is often made based on practical considerations, such as the fact that there may not be enough delivery systems for each student in a class, or that tests must be taken after students have left the training setting, or due to time limitations in accessing equipment. There is a danger, however, in using paper-and-pencil tests when learners received their practice in lessons through the computer or other technology. The "conditions" of the performance in the test may no longer match that of the objectives. Even a difference such as having computer graphics with text in practice activities with text only in the paper-based test can invalidate the test items. Technology-based tests to match the lesson objectives, format, and practice are thus somewhat preferable, unless the paper-and-pencil tests and technology-based practice are carefully matched.

Technology-based Tests. Computer-based achievement tests offer other advantages to paper-based tests. A test-item bank can be developed to allow administering multiple forms of the test. Data collection can be greatly simplified, in that computer programs can be written to transfer performance data directly to files for data analysis. (Of course, use of optically-scannable answer sheets with paper tests also increases efficiency). Adaptive tests might be developed that present specified items based on performance, and, once a student has begun to fail items based on a knowledge or skill hierarchy, for example, save testing time by not administering more items for advanced skills.

On-The-Job or Real-World Observations of Performance. A critical issue in evaluating learning is often how well students perform in their real-world settings. Although most instructional developers have traditionally recommended evaluating on-the-job performance, the efficiency of using less-realistic measures often ensures that paper-and-pencil or computer-based tests are used. Observations of learner performance in any work or life setting can, however, be conducted using methods adapted from ethnographic research.

Should evaluators decide to conduct observations, several decisions must be made. The team should determine who will conduct the observations, how the observers will be trained to ensure consistency, on what performances they will collect data, how observations will be recorded, how inter-observer reliability will be determined, how the data will be analyzed and how the results will be reported. For example, if the learned task is primarily procedural, it may be a simple matter to develop a checklist for recording how closely a student follows the required procedural steps in an assessment situation. In contrast, if the learned task was a more "fuzzy" type of skill, such as how to conduct an employment interview, the observational procedures, checklists, etc., would be more complex, and reliability of observations could be a trickier issue, due to subjectivity of what observers might be recording. Conducting observations of behaviors will be discussed further in a section concerned with evaluating implementation of interactive systems in real-world learning environments.

It might be noted that when it is not practical to observe student performance on the job or out of school, it may still be practical to conduct observations of students in "classroom" settings who are engaged in formal role-plays or simulations of the skills they learned through interactive instruction. The considerations described above would also be relevant in evaluating learning through such simulations.

Products/Portfolios. As noted by Linn, Baker and Dunbar (1991), there is increased concern among educators that traditional assessment methods shortchange evaluation of complex performance-based learning. In school settings, for example in programs to determine which students to include in gifted and talented programs, it is becoming common to include portfolios of student writing and other samples of the products of students' work. Interactive technology-based systems are often developed specifically to teach complex sets of behaviors and problem-solving skills through simulations. It is to be expected that instructional developers of interactive learning systems would collect products of student work to directly measure achievement of complex objectives.

For example, a student who learned to repair equipment by experiencing computer-and-videodisc simulations could be expected to demonstrate learning achievement by repairing an actual piece of malfunctioning machinery. The repaired equipment would thus be a product. Here again, as recommended by Dick and Carey (1990), Sullivan and Higgins (1983), and most other instructional developers, an evaluation checklist would be developed to determine mastery of the skill as demonstrated by the quality of the product.

Similarly, if a student learned to create an art piece by participating in a videodisc-based interactive lesson about a particular type of art, the evaluation would logically involve determining the quality of the student's creation, according to criteria established in the lesson.

One caution that applies in all types of evaluations of student products and portfolios relates to the alignment between practice and assessment activities. Developers and evaluators cannot expect learners to move directly from doing practice in a technology-based simulation to performing the skill in the real-world setting. As noted earlier, the conditions of the practice and assessment in this situation would not match. Developers would do well to ensure that learners engaged in learning using their instructional system receive some type of practice on the actual equipment or in the real-world setting, or producing the real product, before they are tested in the latter situations.

Time Measures. For some types of learning, mastery is measured by the quality or frequency of student performance within given time parameters. This would be the case, for example, for keyboarding skills in which mastery is demonstrated both by accuracy and speed. Interactive technology-based systems can easily record how quickly or frequently learners perform. In addition, it has often been noted that the "claim to fame" for technology is that it helps learners progress to mastery levels more quickly than through traditional instruction. Many evaluations of interactive instruction therefore include measures of time to mastery. It is likely that evaluations of evolving systems will continue to include collecting time data.

Self-Evaluation. In some instances, particularly in adult or recreational learning settings, collecting data regarding learners' perceptions of their own achievement of skills is desirable. Such data can be collected using straightforward questions on survey instruments, such as "How would you rate your skill in ________ now?" Such self-report data is often biased, however, and so it is usually more useful to collect data which directly measures student learning. However, at times, developers may also be concerned with learners' perceptions of their learning, perhaps for political reasons, and these may be useful depending on the evaluation questions.

Interviews. Although not typically used to collect achievement data, there are a few instances in which interviews might be useful. Interviews may be conducted to collect self-evaluation data. In addition, with very young students or those who are not literate interviews may really be oral tests conducted to measure learning achievement.

Occasionnally, especially in training settings, interviews are conducted with managers to determine how well they believe employees learned the skills practiced through interactive instruction, and how well managers believe employees are now performing in their jobs. Collecting these data sometimes has the side benefit of contributing to managers' "buy-in" of the interactive training, as they reflect on what their employees learned through the training.

Documentary Data. In some settings, evaluators of interactive technologies will secure access to data already existing in the organization. In educational settings these data might be end-of-course grades. For example, the final grades of college students' who completed a course delivered via interactive video might be compared with those of students who completed the course in a traditional manner. In schools districts, evaluators may secure access to student performance on yearly standardized tests. In both these cases, these data would be more relevant for whole courses which used interactive technology than for those courses which employed technology in a supplementary and limited manner.

In training settings, evaluators might review industrial documentary data regarding increased production, decreased loss due to error, decreased reports of health or safety violations, reduced customer complaints, increased efficiency, etc.

Issues Related to Transfer. A particularly sticky issue in education is how well learners can perform in real-world settings the skills they mastered in artificial settings such as classrooms. Many advocates of multimedia argue that interactive instructional programs can closely simulate the real-world, sometimes calling such systems learning environments. For example, one interactive video curriculum has been developed to train reserve soldiers to repair and maintain the M-1 tank (Savenye, 1986b). Yet in this project military trainers noted that troubleshooting a firing system malfunction based on video and audio displays, and then selecting a decision such as replacing a part from icons on a menu is not the same as actually performing these activities on a tank. While few evaluations have measured learning transfer to on-the-job or outside-school tasks, some studies have indicated learning through interactive media does help students learn to transfer their knowledge to other settings more quickly than learning through traditional instruction (DeBloois, 1988).

It will remain a responsibility of developers to use technology to build learning systems, especially simulations, that enhance learning transfer, and of evaluators to creatively measure such transfer.

Issues Related to Retention. Regardless of how learning is measured it is advisable to administer delayed versions of tests or other measures to determine how much learner retain of what they have learned. It is not difficult in on-going interactive curriculum materials to build in periodic tests which students might view as "reviews", but which developers could use to measure retention.

Answering Other Types of Learning Questions

Interactive technology-based instruction may be used in nontraditional educational settings, such as in museums and parks, or even for delivery of information, as opposed to instruction. In these cases the learning to be measured may be quite different from achievement of learning objectives and the evaluation questions, therefore, may differ.

An example of such a situation was presented by Hirumi, Allen and Savenye (1989). These authors discussed the development and evaluation of an interactive videodisc-based museum exhibit to introduce visitors to the plants and animals of the desert. In a museum setting visitors experience an exhibit in groups, with only one or two individuals actually making choices on the computer. Visitors spend little time with an exhibit, and are unlikely to be willing to take traditional tests. In this type of setting, if learning from the interactive exhibit is a concern, a limited number of nonthreatening learning achievement questions can be asked of samples of visitors in very brief interviews. If classes of children visit the museum it is often possible to ask them to complete short activities which they may perceive as fun, but which actually measure learning.

Technology-based learning environments may be developed with broader goals than to teach specific objectives. They may, for example, be based on an exploratory learning approach with the goal of enhancing student motivation to prepare them to participate in more structured learning activities later. Such exploratory systems may include a videodisc that shows students whatever aspects of a setting they may choose, for example selected parts of a town or archaeological site, or the flora and fauna of natural surroundings. These systems often are based on hypertext, and thus allow learners to branch in a network fashion from any bit of information in the database to any other. Evaluators investigating effects of such exploratory learning environments may, depending on the evaluation questions, collect computerized data on the pathways learners take through information, and what choices they make. If most learners bypass some parts of the information, for example, or always go through some parts, evaluators could conduct followup interviews to ask learners why they make the choices they do.

A technique of using read-think-aloud protocols could also be used (Smith & Wedman, 1988) to analyze learner tracking and choices. Using this technique, evaluators could ask learners to "talk through" their decisions as they go through a lesson. Evaluators could observe and listen as learners participate, or they could audiotape the learners and analyze the tapes later. In either case, the resulting verbal data must be coded and summarized to answer the evaluation questions. Techniques of protocol analysis (cf. Ericsson & Simon, 1984) should be determined and tested early in the evaluation process.

As described earlier, observations of actual or simulated performances may be called for, although in these nontraditional learning environments performance is not always as much a concern as is motivation.

In contrast to these nontraditional educational settings, in which performance is not always critical, business and industrial settings in which training is delivered on-line or on-demand, do hold learner performance to be of utmost concern. Yet these settings do not always allow for traditional testing. It is not difficult, however, to develop unobtrusive, objectives-based performance measures that are resident in the computer system, and which learners would not object to. For example, if an employee calls up a brief tutorial while attempting to use a software feature which is new to her, her subsequent performance using the feature could be measured and recorded by the computer system. In training settings issues of confidentiality of performance achievement may arise, so evaluators and managers together might determine whether and how employees would be informed that their performance would be tracked, and how those data would be used.

As in the earlier discussion, in these settings, data regarding learning time and self-evaluation of performance, as well as documentary and interview data could be collected.

Determining Attitudes/Perceptions

The traditional methods of collecting data on the attitudes and perceptions of learners and other stakeholders have included questionnaires and, less frequently, interviews. The traditional issues of sampling and how to compare results of instructional methods continue to apply when evaluating interactive instruction.

Questionnaires. Using quantitative methods evaluators may wish to compare the attitudes toward a subject of students who learned the subject through interactive instruction with students who participated in a traditional course. This approach was used by Savenye (1989) who found that students who participated in a full-year videdisc-based high school physical science curriculum generally held more positive attitudes toward science and how they learned science than students who took the course via traditional instruction. In this evaluation, as always, sufficient numbers of students needed to complete the surveys to yield reliable results. Additionally, care was taken to include students from various types of schools and communities, such as urban, rural and suburban, and from representative geographic areas and cultural groups in the evaluation.

If an interactive program had as its primary goal an improvement in attitudes, evaluators are likely to need to collect preinstructional and postinstructional attitude data. In a related example, Savenye, Davidson & Orr (1992) collected pre and post data, and reported that preservice teachers' attitudes toward computers were higher, and their anxiety lower, after they had participated in an intensive computer applications course.

Questionnaire items and directions should be clearly-written, and the questionnaire should be as short as possible. Questions should be directly based on the needs of the evaluation. Evaluators may wish to make most items forced-response, such as Likert-scale items, to speed data analysis. A case can be made, however, for including a few open-ended questions in every survey, to allow learners to bring up issues not anticipated by evaluators, who may be unfamiliar with the learners' needs and concerns and constraints of their learning environment.

Interviews. Attitudes can also be measured using interviews. Often evaluators supplement questionnaires by conducting one-on-one or small, focus-group type, interviews with a small sample of learners to ensure that all relevant data were collected and nothing was missed by using a questionnaire. When budget is limited, interviews may be the sole means of collected attitude data, primarily to verify and explain achievement results. For example, Nielsen (1990) incorporated interviews into his experimental study investigating achievement effects of informational feedback and second attempt in computer-aided learning. Nielsen found that some of his learners, who not coincidentally were highly motivated Air Force cadets, who received no feedback determined that their performance depended more on their own hard work and they took longer to study the lesson, while the cadets who received the extensive informational feedback soon figured out they would receive the answers anyway, and so spent less time on the practice items.

Usually it is desirable when conducting interviews, particularly when several evaluators will be interviewing learners, for a set of structured questions to be developed. Otherwise ideosyncratic data may accidentally be collected from each learner, and data will not be comparable. In addition, it is usually useful for interviewers to be given the freedom to probe further as the interviews progress, particularly when the evaluation involves a completely new interactive system, which may be causing many types of changes in the instructional setting.

Learner Notes. In traditional field tests, evaluators often collect attitude data by allowing learners to write comments on their materials. In interactive systems, a computer program can be written to easily allow learners to write notes and comments to developers.

Other Types of Data Collection Methods. As described earlier, it may be desirable to observe learners and collect incidental attitude data, provided observers have agreed what type of data they will record. Additional data can also come up in interviews or through collecting documentary data. One example of such data was the observation by teachers and evaluators in many schools which used a videodisc-based science curriculum that many more students were coming into the science classrooms to "play" with the science lessons during their free time than ever occurred when traditional science lessons were being used.

An exemplary study using alternate research methods to determine teacher perceptions of competency-based testing serves as an example of what can be done to measure attitudes and perceptions. Higgins and Rice (1991) conducted a three-phase study. They initially conducted relatively unstructured interviews with six teachers regarding the methods they used to assess their students, and in what situations they used the techniques. From these interviews the researchers constructed a taxonomy of assessments methods. These researchers the employed trained observers to collect data during ten hours of classroom observations egarding how teachers measured their students. bsequent interviews were conducted to ask teachers their perceptions of how they were using assessments during their classes, and to have teachers rank their perceptions of the utility and similarity of the types of assessments the teachers had described. The interview and observation data were coded and summarized. The rankings from the teacher interviews were used to perform multimensional scaling, which yielded a two-dimensional representation of the teachers' perceptions. Similar techniques could be adapted by evaluators to answer questions related to instructor and learner perceptions of their technology-based lessons, or their attitudes toward content and skills learned.

Evaluating Use/Implementation

It is in answering questions related to how the interactive instruction is being used in the various learning settings that evaluators can most profitably use alternate methodologies. The most efficient, and therefore, first methods to use to answer implementation questions are still questionnaires and interviews. However, sometimes the question when using a truly new technology is often, "What is really happening here," as opposed to what developers may plan to or hope to happen. Here we especially need answers that ring true, and here we sometimes do not know the right questions to ask. Using an anthropological approach, evaluators can go into their learning settings with an open mind.

Participant Observation. Participant observation is a technique derived from ethnographic studies. It involves intensive observation of partipants in a setting. Anthropologists may spend years "in the field" becoming in a sense members of a community, therefore participants, while they observe and record the patterns and interactions of people in that community.

Evaluators often cannot, nor do they need to, spend as much time in their instructional settings, as do anthropologists, yet the activity is extremely labor-intensive, and data collection is usually limited to those data which will answer the questions at hand. Still, evaluators would do well to remember that although they do not spend years observing the particular instructional community, they do quickly become participants. Their presence may influence results, and their experience may bias what they observe and record. In subsequent reports, therefore, this subjectivity can simply be honestly acknowledged.

Methods of collecting observational data may include writing down all that occurs, or recording using a limited checklist of behaviors. Observers can watch and write as they go, or data can be collected using videotapes or audiotapes. As mentioned earlier, analyzing qualitative data is problematic. Every behavior that instructors and students engage in could potentially be recorded and analyzed, but this can be costly in money and hours, and would most likely be useless for evaluation purposes. Evaluators should determine in advance what they need to find out.

For example, Savenye & Strand (1989) in the initial pilot test and Savenye (1989) in the subsequent larger field test of the science videodisc curriculum described earlier determined that what was of most concern during implementation was how teachers used the curriculum. Among other questions, developers were interested in how much teachers followed the teachers' guide, the types of questions they asked students when the system paused for class discussion, and what teachers added to or didn't use from the curriculum. A careful sample of classroom lessons was videotaped and the data coded. For example, teacher questions were coded according to a taxonomy based on Bloom's (1984), and results indicated that teachers typically used the system pauses to ask recall-level, rather than higher-level questions. Analysis of the coded behaviors for what teachers added indicated that most of the teachers in the sample added examples to the lessons that would add relevance to their own learners, and that almost all of the teachers added reviews of the previous lessons to the beginning of the new lesson. Some teachers seemed to feel they needed to continue to lecture their classes, therefore they duplicated the content presented in the interactive lessons. Developers used the results of these evaluations to make changes in the curriculum and in the teacher training that accompanied the curriculum. Of interest in this evaluation was a comparison of these varied teacher behaviors with the student achievement results. Borich (1989) found that learning achievement among students who used the interactive videodisc curriculum was significantly higher than among control students. Therefore teachers had a great degree of freedom in using the curriculum and the students still learned well.

If the student use of interactive lessons was the major concern, evaluators might videotape samples of students using an interactive lesson in cooperative groups, and code student statements and behaviors, as did Schmidt (1992).

Reporting Results

How the results of formative evaluations are reported depends on how the results are to be use. For example, if the report is for a funding source, or to ensure continuing support for a large project, the report might be quite formal and detailed. In contrast, if the results of the formative evaluations are for immediate use by the development team only, the reports may consist of informal summaries, memos and briefings.

The primary rule in reporting is to keep it simple. Long evaluation reports may not be read by those who most need them.

The organization of the report may best be accomplished by using the evaluation questions as headings and answering each question in the sequence the audience most likely would desire.

At a minimum the report should usually include sections on learning achievement, attitudes, and use/implementation. With regard to achievement, at least the major mean scores should be reported, with a summary table typically included. Results of any statistical comparisons may be reported. Finally other learning results or anecdoctal data related to performance, such as the results of interviews, observations, or analysis of products or documentary data should be reported here (cf. Dick & Carey, 1990).

When reporting attitudes, the primary findings related to the evaluation questions can be described. It may be desirable to summarize the results of survey items on a copy of the survey or of interviews on a copy of the interview protocol. Again, summaries of other types of data collected may be written, or presented in tables.

Reporting the results on use or implementation questions may be more difficult. Results of surveys and interviews can be done in a traditional manner, however, reporting results of observations and microanalyses of data can be done many different ways. Frequency tables can be developed for categories of coded behaviors. Although not an evaluation study, pre se, an example of the reporting of teacher perceptions and planning behaviors reported in a case study style is presented by Reiser and Mory (1991). Alternately, some evaluators build a type of story description, or scenario, of patterns they have observed. It may be useful for evaluators to turn to descriptions of qualitative research in social sciences for types of methods to try (cf. Bogdan & Biklen, 1982; Straus, 1987).

Conclusions/Recommendations

In conclusion, alternate methods of conducting formative evaluations may be particularly useful and crucial when dealing with highly innovative interactive technology-based instruction. One key to success is to ensure that evaluation questions drive the choice of methods for collecting data and reporting results. Another is to keep the evaluation focussed, thus simple and efficient. Another factor in success is to use rigorous techniques and methods while experimenting with new ways of conducting evaluations. Evaluators will learn more about how their innovative technology systems are being used if they are open to what is really occurring, but not overwhelmed to the point that they gather too much data, collect data haphazardly, or focus on data items which are so ideosyncratic that the results cannot be compared to any other data or results of any other studies.

As always, the main question is whether students learned using the interactive instruction, no matter how attractive the "bells and whistles."

References

Ambron, S., & Hooper, K. (Eds.). (1988). Interactive multimedia: visions of multimedia for developers, educators, & information providers. Redmond, WA: Microsoft Press.

Ambron, S., & Hooper, K. (Eds.). (1990). Learning with interactive multimedia: developing and using multimedia tools in education. Redmond, WA: Microsoft Press.

Ambrose, D. W. (1991). The effects of hypermedia on learning: a literature review. Educational Technology, 31 (12), 51-55.

Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. Journal of Instructional Development, 3 (4), 2-16.

Bloom, B. S. (1984). Taxonomy of educational objectives. Handbook 1: cognitive domain. New York: Longman.

Bogdan, R. C., & Biklen, S. K. (1982). Qualitative research for education: an introduction to theory and methods. Boston, MA: Allyn and Bacon.

Borich, G. D. (1989). Outcome evaluation report of the TLTG Physical Science Curriculum, 1988-89. Austin, TX: Texas Learning Technology Group of the Texas Association of School Boards.

Bosco, J. (1986). An analysis of evaluations of interactive video. Educational Technology, 16 (5), 7-17.

Brenneman, J. (March 1989). When you can't use a crowd: single-subject testing. Performance & Instruction, 22-25.

Chen, J. W., & Shen, C.-W. (September, 1989). Software engineering: a new component for instructional software development. Educational Technology, 9, 9-15.

Chinien, C. (October, 1990). Paradigms of inquiry for research in formative evaluation. Performance & Instruction, 26-29.

Clark, R. E. (1983). Reconsidering research on learning from media. Review of Educational Research, 53 (4), 445-459.

Clark, R. E. (1989). Current progress and future directions for research in instructional technology. Educational Technology Research and Development, 37 (1), 57-66.

Darabi, G. A., & Dempsey, J. V. (1989-90). A quality control system for curriculum-based CBI projects. Journal of Educational Technology Systems, 18 (1), 15-31.

DeBloois, M. L. (1988). Use and effectiveness of videodisc training: a status report. Falls Church, VA: Future Systems, Inc.

Dick, W., & Carey, L. (1990). The systematic design of instruction (3rd ed.). Glenview, Illinois: Scott, Foresman.

Driscoll, M. P. (1991). Paradigms for research in instructional systems. In G. Anglin (Ed.) Instructional technology past present and future. Englewood, CO: Libraries Unlimited.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis - verbal reports as data. Cambridge, MA: MIT Press.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: a meta-analysis. Exceptional Children, 53, (3), 199-208.

Gagne, R. M., Briggs, L. J., & Wager, W. W. (1988). Principles of instructional design. (3rd ed.). New York: Holt, Rinehart and Winston.

Gagne, R. M., Wager, W. W., & Rojas, A. (1981). Planning and authoring computer-assisted instruction lessons. Educational Technology, 21, 17-26.

Geis, G. L. (May/June, 1987). Formative evaluation: developmental testing and expert review. Performance & Instruction, 1-8.

Hannafin, J. J. (1985). Empirical issues in the study of computer-assisted interactive video. Educational Communications and Technology Journal, 33 (4), 235-247.

Hannafin, M. J., & Peck, K. L. (1988). The design, development, and evaluation of instructional software. New York: Macmillan Publishing Co.

Hannafin, M. J., & Rieber, L. P. (1989). Psychological foundations of instructional design for emerging computer-based instructional technologies: Part 1. Educational Technology Research and Development, 37 (2), 91-101.

Higgins, N., & Rice, E. (1991). Teachers' perspectives on competency-based testing. Educational Technology Research and Development, 39 (3), 59-69.

Higgins, N., & Sullivan, H. (1982). Preparing special education teachers for objectives-based instruction. Teacher Education and Special Education, 5 (4), 51-55.

Hirumi, A., Savenye, W., & Allen, B. (1989, February). Convergent technologies and public information: using videodiscs to interactivate museum exhibits. Paper presented at the annual meeting of the Association for Educational Communications and Technology, Dallas, Texas.

Hooper, S., & Hannafin, M. J. (1988). Learning the ROPES of instructional design: guidelines for emerging interactive technologies. Educational Technology, 28, 14-18.

Jacob, E. (1987). Qualitative research traditions: a review. Review of Educational Research, 57 (1), 1-50.

Kearsley, G. P., & Frost, J. (1985). Design factors for successful videodisc-based instruction. Educational Technology, 25 (3), 7-13.

Kulik, J. A., Bangert, R. L., & Williams, G. W. (1983). Effects of computer-based teaching on secondary school students. Journal of Educational Psychology, 75 (1), 19-26.

Kulik, J. A., Kulik, C., & Cohen, P. (1980). Effectiveness of computer-based college teaching: a meta-analysis of findings. Review of Educational Research, 50 (4), 525-544.

Lambert, S., & Sallis, J. (Eds.) (1987). CD-I and interactive videodisc technology. Indianapolis, IN: Howard W. Sams.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: expectations and validation criteria. Educational Researcher, 20 (8), 15-21.

Markle, S. M. (August, 1989). The ancient history of formative evaluation. Performance & Instruction, 27-29.

McLean, R. S. (1989). Megatrends in computing and educational software development. Education & Computing, 5, 55-60.

McNeil, B. J., & Neson, K. R. (1991). Meta-analysis of interactive video instruction: a 10 year review of achievement effects. Journal of Computer-Based Instruction, 18 (1), 1-6.

Morris, L. L., & Fitz-Gibbon, C. T. (1978) Evaluator's Handbook. Beverly Hills: Sage Publications.

Newman, D. (March, 1989). Formative experiments on technologies that change the organization of instruction. Paper presented at the annual conference of the American Educational Research Association, San Fransisco, CA.

Nielsen, M. C. (1990). The impact of informational feedback and a second attempt at practice questions on concept learning in computer-aided learning. (Unpublished doctoral dissertation, The University of Texas at Austin).

Patterson, A. C., & Bloch, B. (November, 1987). Formative evaluation: a process required in computer-assisted instruction. Educational Technology, 27, 26-30.

Reeves, T. C. (1986). Research and evaluation models for the study of interactive video. Journal of Computer-Based Instruction, 13 (4), 102-106.

Reigeluth, C. M. (1989). Educational technology at the crossroads: new mindsets and new directions. Educational Technology Research and Development, 37 (1), 67-80.

Reiser, R. A., & Mory, E. H. (1991). An examination of the systematic planning techniques of two experienced teachers. Educational Technology Research and Development, 39 (3), 71-82.

Savenye, W. C. (1986a). A design for the evaluation of comptuer-based software. Austin, Texas: Instructional Technology Area: The University of Texas at Austin.

Savenye, W. C. (1986b, February). Instructional design considerations in development and production of an interactive videodisc for computer-based instruction. Paper presented at the annual meeting of the Association for Educational Communications and Technology, Las Vegas, NV.

Savenye, W. C. (1989) Field test year evaluation of the TLTG interactive videodisc science curriculum: effects on student and teacher attitude and classroom implementation. Austin, TX: Texas Learning Technology Group of the Texas Association of School Boards.

Savenye, W. C. (1990). Instructional interactive video - what works?. Young Academic Monograph. Washington, DC: National Society for Performance and Instruction.

Savenye, W. C., Davidson, G. V., & Orr, K. B. (1992, in press). Effects of an educational computing course on preservice teachers' attitudes and anxiety toward computers. The Journal of Computing in Childhood Education.

Savenye, W. C., & Strand, E. (1989, February). Teaching science using interactive videodisc: results of the pilot year evaluation of the Texas Learning Technology Group Project. In M. R. Simonson, & D. Frey (Eds.), Eleventh Annual Proceedings of Selected Research Paper Presentations at the 1989 Annual Convention of the Association for Educational Communications and Technology in Dallas, Texas. Ames, IA: Iowa State University.

Schmidt, K. J. (1992). At-risk eighth grade student feelings, perceptions, and insights of computer-assisted interactive video. (Unpublished doctoral dissertation, The University of Texas at Austin).

Schwartz, M. (1982). Designing interactive video programs. In S. Floyd & B. Floyd Handbook of interactive video. White Plains, NY: Knowledge Industry Publications.

Schwier, R. (1987) Interactive video. Englewood Cliffs, NJ: Educational Technology Publications.

Slee, E. J. (1989). A review of research on interactive video. In M. R. Simonson, & D. Frey (Eds.), Eleventh Annual Proceedings of Selected Research Paper Presentations at the 1989 Annual Convention of the Association for Educational Communications and Technology in Dallas, Texas (pp. 150-166). Ames, IA: Iowa State University.

Smith, P. L., & Boyce, B. A. (1984). Instructional design considerations in the development of computer-assisted instruction. Educational Technology, 24, 5-11.

Smith, P. L., & Wedman, J. F. (1988). Read-think-aloud protocols: A new data-source for formative evaluation. Performance Improvement Quarterly, 1 (2), 13-22.

Steinberg, E. R. (1989). Cognition and learner control: a literature review, 1977-1988. Journal of Computer-Based Instruction, 16 (4), 117-121.

Straus, A. L. (1987). Qualitative analysis for social scientists. Cambridge: Cambridge University Press.

Sullivan, H., & Higgins, N. (1983). Teaching for Competence, New York: Teacher's College Press, Columbia University.

The power of multimedia: a guide to interactive technology in education and business. (1990) Washington, DC: Interactive Multimedia Association.

U. S. Congress, Office of Technology Assessment. (1988, September). Power On! New Tools for Teaching and Learning (OTA-SET-379). Washington, DC: U. S. Government Printing Office.

Winn, W. (1989). Toward a rationale and theoretical basis for educational technology. Educational Technology Research and Development, 37 (1), 35-46.