CHAPTER 8

Effectiveness Evaluation

Objectives

After reading Chapter Eight, you should be able to:

• identify decisions involved in evaluating the effectiveness of interactive learning systems;
• specify questions that should be answered before making these decisions;
• identify the information needed to answer questions about the effectiveness of interactive learning systems; and
• decide how to collect and report information so that the effectiveness of an interactive learning system can be demonstrated.

Why should you conduct effectiveness evaluation?

The overall purpose of effectiveness evaluation is to determine whether an interactive learning system accomplishes its objectives within the immediate or short-term context of its implementation. As illustrated in Figure 8.1, different types of decisions must be made concerning the dissemination, adoption, and/or implementation of interactive learning systems, each of which is related to one or more specific questions that can be addressed through the strategies used in effectiveness evaluations.

Think back on an education or training program you recently completed. Perhaps you had to be trained on how to handle hazardous chemical spills in your work place. Or you may have been trained in using a complicated new telephone system your company recently installed. Alternatively, you may have recently completed a graduate degree in Instructional Technology. Regardless of the instructional methods used in the education or training program, e.g., teacher-led, interactive multimedia, lecture, video, etc., hopefully there was some degree of difference between the knowledge, skills, and attitudes (KSAs) you possessed before the instruction and the KSAs you had after the instruction (Kirkpatrick, 1998). Determining that difference is the focus of effectiveness.
evaluation. Evaluating whether you actually apply those KSAs in the work place or in your life is the focus of another facet of our evaluation model, i.e., impact evaluation, which is presented in Chapter Nine.

<table>
<thead>
<tr>
<th>Decisions</th>
<th>Example Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>How should the interactive system be marketed?</td>
<td>Do learners achieve the objectives of the program?</td>
</tr>
<tr>
<td>Should this program be adopted?</td>
<td>What are implementation requirements?</td>
</tr>
<tr>
<td></td>
<td>What are the outcomes of the different implementations of this program?</td>
</tr>
<tr>
<td>How should instructors be trained to implement this</td>
<td>To what degree do instructors implement the program as designed?</td>
</tr>
<tr>
<td>interactive learning system?</td>
<td>What creative adaptations of this program have instructors made?</td>
</tr>
<tr>
<td>What price should be charged for this program?</td>
<td>How do the results achieved with this program compare to alternatives?</td>
</tr>
</tbody>
</table>

**Figure 8.1.** Typical decisions and questions in an effectiveness evaluation.

When should you conduct effectiveness evaluation?

If what you are concerned about ultimately is changing the performance of people in the work place or helping students become lifelong learners, making an effort to measure learning in the immediate context of education or training may seem like a waste of time. After all, why not focus on impact evaluation, i.e., evaluating the long term effects of your interactive learning system? There are several problems with this notion. To begin, suppose you look for impact and find none. Should you then conclude that your interactive learning system was ineffective? Not necessarily! It could be that people learned new knowledge, skills, and
attitudes with your e-learning program, but the larger environment did not allow them to put the new KSAs into practice. This occurs more often than you think. Employees are given training to learn new skills, but when they return to their jobs, they find that their managers or peers force them to keep doing things “the old way” or “our way.”

Suppose instead that you conduct an impact evaluation that actually reveals a great deal of evidence supporting the impact of your interactive learning system, e.g., sales went up, accidents decreased, or customer satisfaction ratings soared. But should all (or any) of the credit for these improvements be attributed to the training if you don’t have evidence of immediate effectiveness? Others factors in the larger environment may provide better explanations for the increased sales (a new advertising campaign), reduced accidents (elimination of hazards in the work place), or higher customer satisfaction (lower prices). Unless you present evidence that your program was effective in the short-term context of the training per se, you will not be able to build a convincing case for impact.

Finally, as will be explained in the next chapter, evaluating impact is extremely difficult, even impossible, in many situations. After all, how long is long enough to detect your impact on the lifelong learning habits of your students? Gathering short-term effectiveness data may be the best you can do with respect to evaluating the outcomes of your interactive learning system.

What kinds of decisions can you anticipate?

Evidence of the effectiveness of interactive learning systems is often desirable (if not required) for marketing. Whether you are developing training or education programs for internal clients or commercial markets, you will need to make decisions about how to market your interactive program. Although an idealized marketing strategy involves providing hard evidence that the program accomplishes what it is intended to accomplish, interactive learning systems are more often marketed on the basis of other criteria such as novelty (This detergent is new and improved!), endorsements (Ninety-two percent of dentists surveyed approve of our gum.), or veiled threat (Don’t leave home without it.).

How will you decide to market your interactive learning system? Should you provide evidence of the knowledge, skills, and attitudes resulting from the program? Should you use learner or instructor endorsements? Should you stress the “bells and whistles” of the new multimedia delivery system used to deliver the instruction or the powerful pedagogical dimensions the program enables you to implement?
(Hopefully, unlike many “charlatans” in the software industry, you’ll decide to emphasize the pedagogy over “bells and whistles”!)  

In addition to marketing decisions, effectiveness data are useful in making decisions about program adoption, adaptation, and implementation. For example, you may find that your interactive learning system works as designed with experienced corporate employees, but new hires need a simpler version or a modified implementation of the program. Or you may find that an interactive learning system that worked well at beta sites during your formative evaluation fails miserably at customer sites where no one has taken responsibility for insuring that the program is installed properly. Effectiveness evaluations provide valuable data in making decisions about training, help, and other forms of user support.

Evidence of effectiveness is also likely to influence decisions about future development efforts, including deciding whether or how to modify your development process, whom to hire or fire for your development team, how to prepare a budget for the next project proposal, and what delivery and authoring systems to adopt. The business of developing interactive learning systems is so volatile that no company, university, or government agency can rest on the laurels of past accomplishments for very long. Using effectiveness information to guide “business decisions” is essential.

**What questions should be answered before making decisions?**

Before making your marketing, training, implementation, and/or support decisions, you will want reliable, valid answers to questions about the appeal, effectiveness, and efficiency of your interactive learning system. What are learner reactions to the program? What are instructor reactions? What knowledge, skills, and attitudes do learners derive from your program? Does your e-learning system accomplish its objectives in less time than competing programs? Or does your program allow more people to be trained or educated within a given amount of time?

Unfortunately, there is a tendency within business circles to boil the issue of effectiveness down to one simple question: “Is the interactive learning system more effective than the training program previously used (or some competing training program)?” This simplified approach to evaluation questions may be due to the widespread influence of Kirkpatrick’s four levels of training evaluation model, which was first introduced in the 1950s and continues to be promoted today (cf. Kirkpatrick, 1998). Although Kirkpatrick’s model represents an important contribution to the development of evaluation theory and methodology, it is not sufficient to deal with the complexities of training and education today.
Kirkpatrick (1998) includes four levels of evaluation as described in Figure 8.2. “Level 2 – Learning” of Kirkpatrick’s model corresponds most closely to what we mean by effectiveness evaluation, but the actual strategies he recommends for evaluating learning are too simplistic. Kirkpatrick presents the following guidelines for evaluating learning:

1. Use a control group if practical.
2. Evaluate knowledge, skills, and/or attitudes before and after the program.
3. Use a paper-and-pencil test to measure knowledge and attitudes.
4. Use a performance test to measure skills.
5. Get a 100 percent response.
6. Use the results of the evaluation to take appropriate action.

(p. 40)

<table>
<thead>
<tr>
<th>Level</th>
<th>Label</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reaction</td>
<td>What levels of satisfaction do trainees express about the training program?</td>
</tr>
<tr>
<td>2</td>
<td>Learning</td>
<td>How were the knowledge, skills, and attitudes of the trainees changed by the training?</td>
</tr>
<tr>
<td>3</td>
<td>Behavior</td>
<td>How does the behavior of trainees change in the workplace?</td>
</tr>
<tr>
<td>4</td>
<td>Results</td>
<td>What is the impact of training on results such as production, sales, profits, accidents, etc.?</td>
</tr>
</tbody>
</table>

Figure 8.2. Kirkpatrick’s four levels of training evaluation.

Why do we consider Kirkpatrick’s approach to be overly simplistic? First, there are few instances when control groups are feasible within education or training contexts, and even when a control group can be included, there are numerous problems with such comparisons, as will be described later in this chapter. Second, the recommendation to use paper-and-pencil tests is unrealistic in light of the lack of reliable and valid instruments to measure most of the KSAs learners are intended to develop as a result of using interactive learning systems. Third, with the possible exception of a military base, where are evaluators likely to get 100% representation in an evaluation? In our experience, effectiveness evaluations almost always entail major compromises between ideal and real conditions for sampling and data collection. Fourth, although the advice to use the results for action is sound, Kirkpatrick neglects to acknowledge the reality that most evaluations are only one source of
Influence on decision making, and that a great deal of negotiating and dealing with local politics are necessary before any actions result from effectiveness evaluations (Patton, 1997).

In our opinion, effectiveness is a difficult beast to pin down, and instead of relying on simplistic designs and tests, you should plan your evaluation activities and procedures in ways that will yield as much evidence as possible to build a case that your interactive learning system is effective. As noted in Chapter One, this approach to evaluation has more in common with the work of an attorney trying to build a case for a client’s innocence than it does to the experimental scientist measuring the behaviors of rats in a laboratory.

What kinds of information do you need to answer questions?

Effectiveness data take many forms. The most widely used type of effectiveness data is based on testing, although all too often the tests used have questionable reliability and validity. Effectiveness data can also be derived from questionnaires, interviews, focus groups, observations, ratings, simulations, and other methods. At a minimum, you will want to consider gathering information about the effectiveness of an interactive learning system in the following categories:

- knowledge
- skills
- attitudes
- appeal
- implementation

Knowledge takes many forms. Much knowledge is conceptual in nature. For example, a computer company may offer training to increase the “mind share” that their computers hold in the minds of salespersons who sell their systems as well as competing machines. Knowledge may take the form of specific rules, e.g., a bank teller knowing the protocol for signaling a robbery in progress. Often, knowledge is quite diffuse, involving a complex network of associations among concepts and rules, e.g., principles of perception and cognition learned in an introductory psychology class. The meaning of knowledge is changing in Western society as a result of the so-called “cognitive revolution” in learning theory, and what we thought of as knowledge in the past (facts and figures) is changing to a conceptualization of knowledge as the capacity to learn (Simon, 1987). These changes have enormous implications for how we conduct effectiveness evaluations.
Skills are also quite diverse. They can be as discrete as knowing how to forward a telephone call to another office or as general as the ability to make an effective diagnosis of patient illness. In training, as opposed to education, skills generally are regarded as being more important than knowledge. Of course, educators have long stressed the “basic skills” of reading, writing, and mathematics, but they are also beginning to realize the critical importance of “learning to learn” skills. Contemporary cognitive scientists have come up with various taxonomies of learning. One we find especially useful was developed by Kyllonen and Shute (1989) at the U.S. Air Force Armstrong Laboratory in San Antonio, Texas. Their taxonomy represents the spectrum of internal states with which cognitive psychologists are concerned, beginning with simple propositions (e.g., stating that Japan sells more electronic products than any other nation), proceeding through schema, rules, general rules, skills, general skills, automatic skills, and finally, mental models (e.g., estimating the potential of a trade war between Japan and the United States based on an analysis of balance-of-trade trends). The latter type of knowledge seems particularly important because mental models are the basis for generalizable problem-solving abilities (Halford, 1993).

Attitudes are perhaps the instructional outcomes most difficult to measure. Attitudes include such factors as motivation, morale, values, and prejudices. These are frequently measured with self-report instruments, but the differences between self-reported attitudes and actual attitudes are often considerable. Sometimes people are purposefully deceitful in reporting attitudes, but more often they just don’t know how they feel or are unable to express their attitudes.

Appeal of interactive learning systems for education and training is related to understanding effectiveness because you generally hope that learners perceive using these systems as desirable, even enjoyable, experiences. If learners enjoy using an interactive learning system for education and training, the enjoyment generally creates internal states that support learning, such as alertness, motivation, and stimulation.

Implementation of an interactive learning system refers to the program as it actually occurred as opposed to the way it was planned. Learners as well as teachers have “bad days.” Power failures, fire drills, and other unplanned disturbances disrupt the best-laid implementation plans for interactive learning. An instructor, programmer, or someone else might decide to add “something new” to the interactive learning system at the last moment with the result that it crashes when learners access it. (We both have had this happen!) Deviations from the interactive learning system as designed can be either positive or negative, but it is essential to evaluate the kind and degree of divergence as much as possible. Other-
wise, you won’t know if the outcomes you measure resulted from the system as designed or some alternative version of it.

How should effectiveness information be collected?

You are strongly advised to prepare a careful evaluation plan before you undertake an effectiveness evaluation (or any other facet of evaluation). As described in Chapter Four, an evaluation plan is a document that spells out in considerable detail the whys and hows of an evaluation effort. Planning and negotiating an evaluation will often represent the major part of the effort you will invest in evaluating an interactive learning system. In fact, once you have a well-designed plan in hand, accompanied by a set of reliable and valid evaluation instruments, it is often possible to turn much of the actual implementation of the plan over to other people. Figure 4.1 in Chapter Four presents a list of the major components of an evaluation plan.

One advantage of an evaluation plan is that it is a public document that your clients can review and approve. You are advised to have your evaluation plan thoroughly reviewed by clients and have them indicate their approval using a sign-off form (see Figure 4.18 in Chapter Four). Another useful tool in planning an effectiveness evaluation is a data collection matrix, as illustrated in Figure 4.9 in Chapter Four. Down one side of the matrix, you should list your evaluation methods, and across the other side, you should list your specific evaluation questions. Then you can consider the acceptability, feasibility, reliability, and validity of each data collection method in reference to each question.

Acceptability is one of the criteria you must consider in selecting data collection methods. Some methods may not be acceptable in certain training contexts, e.g., if union rules do not permit employees to be tested. Feasibility is another of the criteria you must consider. Although you might consider participant observation as a powerful indicator of effectiveness, you may lack the resources to use this method in a specific evaluation context. Reliability is the criterion concerned with the consistency of a measure. For example, whenever you use multiple experts to review the effectiveness of a program, you have to be concerned with their inter-rater reliability. Finally, validity is the criterion that addresses the issue of whether you are measuring what you really intended to measure. You are probably familiar with training situations in which a test was administered that seemed to be assessing KSAs that had very little to do with what you actually learned in the training. The test itself may have been a well-constructed one, but it was invalid for that purpose.
Determining the reliability and validity of outcome measures is an often ignored aspect of effectiveness evaluation. Few evaluators have sufficient psychometric expertise to conduct reliability and validity studies. In addition, clients are inclined to accept test results as somehow sacrosanct, viewing a test score as something valuable in and of itself. We regard establishing the reliability and validity of evaluation instruments as a professional responsibility. There are books that can help (Gronlund, Linn, & Davis, 1999; Loewenthal, 2000; Morris, Fitz-Gibbon, & Lindheim, 1987; Wiggins, 1998), but testing consultants are often required.

**Implementation logs**

Implementation logs are useful in effectiveness evaluation because unless you know how the interactive learning system was actually implemented, you will have little basis for interpreting the outcomes you measure. Implementation logs are used to assess the fidelity of the implementation. Sometimes, there may be major differences between an interactive learning system as designed and the program as implemented. For example, consider an interactive program that assumes that pairs of learners will complete the program. Occasionally, that same program might be used by a larger group than was intended or even by an instructor with a whole class. Although these divergent uses may be unavoidable, it is necessary to have some record of the actual use. Figure 8.3 provides an example of a completed implementation log.

The log in Figure 8.3 indicates precisely what should occur during the training day. It also has space for noting any discrepancies between the planned implementation and what actually took place. The form is designed to encourage the evaluator to think about how problems that occurred might be avoided in the future. Finally, it has space for noting creative changes to the program that might be recommended. Deviations from the planned implementation are not always negative, and you will want to record ideas for improving your strategies.

Whereas implementation logs are usually maintained by an instructor or evaluator, trainees may also keep logs during their use of an interactive learning system. A version of the formative review log presented in Figure 7.4 in Chapter Seven may be used for this type of data collection. You should estimate the degree to which keeping such a log may interfere with the learning process, and if the interference seems too great, consider other methods of data collection. Some interactive learning systems are designed with online logs that learners can use to record their feedback to the developers and administrators of a program as they use it. Or it may be appropriate to save learners’ work automatically and then collect work samples for analysis at other times.
IMPLEMENTATION LOG

DATE: October 2-3           PLACE: Chicago           TIME: 9:00 - 5:00

TRAINER: Larry Jones       NUMBER OF TRAINEES: 15

<table>
<thead>
<tr>
<th>Time</th>
<th>Recommended Activities</th>
<th>Actual Activities</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00-09:15</td>
<td>Introductions of participants and review of agenda. (Leader-led)</td>
<td>No changes.</td>
<td></td>
</tr>
<tr>
<td>09:15-09:30</td>
<td>Overview of new course. (Leader-led)</td>
<td>No changes.</td>
<td></td>
</tr>
<tr>
<td>09:30-12:00</td>
<td>Trainees begin Module 1. (multimedia systems)</td>
<td>One system failed to function because someone removed system files.</td>
<td>I need to check systems personally before course begins.</td>
</tr>
<tr>
<td>12:00-01:00</td>
<td>Lunch Break.</td>
<td>Two trainees chose to skip lunch and keep working.</td>
<td></td>
</tr>
<tr>
<td>01:00-05:00</td>
<td>Trainees continue working through the modules.</td>
<td>No changes.</td>
<td>Frequently had to refer trainees to help.</td>
</tr>
</tbody>
</table>

1. What activities would you like to modify the next time you conduct this course?
   I will personally check each of the multimedia systems to insure that the modules function as designed.

2. How can the training materials used in this course be enhanced?
   Trainees pointed out several errors in the data communications module. See attached list.

Figure 8.3. Sample implementation log.

Questionnaires

Questionnaires are one of the most frequently used methods of collecting effectiveness data. As noted above, the acceptability and feasibility of measurement instruments are major concerns in selecting any type of data collection method. Collecting data about learner reactions to instruction, which is Level 1 in Kirkpatrick’s (1998) classic evaluation model, is widely acceptable and usually quite feasible in any context. This explains the popularity of the post-instruction questionnaire that you and virtually every other learner complete at some point. These instruments have been variously called “evaluation forms,” “happiness indexes,” and “smi-lometers.” It is tempting to assemble evaluation questionnaires hastily, but you should invest considerable planning and revision time in these instruments. Figure 8.4 presents a sample questionnaire used in the evaluation of interactive multimedia in a university course. Several
different types of closed and open-ended questions are posed in this instrument.

---

**ECOLOGY 100 – Stream Lab Data Analysis Program**

We need your feedback to help us develop better interactive tools for the Ecology 100 course. Your responses are anonymous. Thank you for your time.

Today’s Date? _________________ Your T.A.’s Name? __________________

Your Age? ____ Years Class? __FR __SO __JR __SR __ Other ________

Gender? __M __F Major? ____________________________________________

1. Please circle your responses to the following statements using the scale below.

<table>
<thead>
<tr>
<th>1 = Strongly Disagree</th>
<th>2 = Disagree</th>
<th>3 = Neither Agree or Disagree</th>
<th>4 = Agree</th>
<th>5 = Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. The “Stream Lab” field work was a valuable learning activity.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Using the “Stream Lab” multimedia program was a valuable learning activity.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. The “Watershed” animation helped me to understand what a watershed is.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. The “Effects of Land Use” simulation helped me to understand how differences in land use affect stream quality.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. The “State Botanical Gardens Walk” helped me to understand the land uses in the State Botanical Gardens.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. The “Tanyard Branch Walk” helped me to understand the land uses surrounding Tanyard Branch.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. The “Bioassessment Pamphlet” helped me to understand why we sampled macroinvertebrates in the stream.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. The data analysis portion of the “Macroinvertebrate Bioassessment” was an effective learning tool.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. The overall multimedia treatment or design of this “Stream Lab” is appealing and easy to use.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. What is the best thing about the “Stream” program? (Use back if needed.)

3. How would you improve the “Stream” program? (Use back if needed.)

4. Please write any other comments about this course on the back of this sheet.

---

*Figure 8.4. Sample student evaluation questionnaire.*
Developing questionnaires is the subject of many workshops and books (Converse & Presser, 1986; Fink & Kosecoff, 1998; Fowler, 1995), and you are advised to avail yourself of the opportunity to develop advanced survey skills if questionnaires are going to be a major tool that you will use in effectiveness evaluations. Alternatively, you can hire a survey consultant to develop your instruments. Questionnaires should be reviewed carefully by other members of the development team as well as by any external clients to make sure that all of the salient issues are addressed. It is also important to pilot a questionnaire with individuals who are as much like the eventual respondents as possible. Even the most carefully crafted questionnaires will have items that are confusing for some and misunderstood by others. Eliminating as many of these problems as possible involves several rounds of testing the instrument with typical respondents.

**Interviews**

Interviews are powerful tools for collecting data about the effectiveness of interactive learning systems. Interviews provide participants in an evaluation more opportunities to speak in their own voice instead of merely responding to the categories of questions that others have defined for them. It is important to plan your interview protocol carefully. This will often involve several trial interviews and subsequent revisions of the questions. Figure 8.5 presents a typical post-training interview protocol.

Most interview protocols allow adjustments in questioning strategies to fit the nature of individual interviewees. You should always be prepared to ask questions in areas that may not have been anticipated by the designers of the interview protocol.

Recording interview data presents challenges. Audio or video recorders may be used, but these devices may intimidate interviewees so much that their responses are skewed. Practice taking brief notes during an interview, and subsequently recording a more detailed transcription of the interview immediately after its conclusion. If possible, it is advisable to do a “mind dump” (writing everything you can recall) concerning one interview before starting to conduct another. Otherwise, you are liable to confuse the responses from one person with those from someone else.

It is sometimes difficult to strike a balance between reporting a highly subjective interpretation of the results of interviews and a verbatim report of all the interviewees’ responses. Most decision makers will reject the former as being too biased and will reject the latter for being too long. The following tips may help:

- Have several different people read interview transcripts and then use group brainstorming techniques to tease out the major findings.
• Look for exemplary quotes which capture in well-spoken words the feelings and beliefs of a significant portion of the interviewees.

• Provide appendices of complete transcripts of the interviews so that readers can spot-check your interpretations.

Another useful strategy is to have interviewees themselves review your reports, adding clarification as needed.

INTERVIEW PROTOCOL

A subset of the personnel completing the Learner Questionnaire will be interviewed either in person at the work site or via telephone. The following interview protocol is likely to be modified during use. All interviews will be audio-recorded and transcribed for analysis.

1) What previous experience have you had with interactive multimedia programs?
   • What kinds of computer games have you played?
   • What kinds of multimedia training have you experienced?
   • What other kinds of multimedia have you used?

2) How does the experience of using this program compare with your earlier experiences?
   • How does it compare with computer games?
   • How does it compare with other multimedia training programs?
   • How does it compare with other kinds of multimedia you have used?

3) Describe your experiences with using the World Wide Web (WWW).
   • How have you used the WWW to find information?
   • How have you used the WWW to purchase things?
   • How have you used the WWW for entertainment?
   • How have you used the WWW to learn?

4) What would you say to other workers with jobs like yours about this program?

5) What would you say to the company managers about this type of training?

6) What hardware and software would you need to use this type of program at home?
   • What kind of personal computer would you need?
   • What kind of software would you need?
   • What kind of Internet access would you need?

7) What areas of training and education would be suitable for this multimedia approach?

8) What is your preferred way of using this program, alone or with someone else?
   • What are the advantages and disadvantages of working with someone else?
   • What are the advantages and disadvantages of working alone?

9) Should this type of multimedia training program be used more widely within this company? Why or why not?

10) What other feedback can you provide concerning this program?

Figure 8.5. Sample interview protocol.
Observations
Observations are useful as measures of the implementation of interactive learning systems and as relatively unobtrusive measures of effectiveness. Observations require carefully designed protocols to provide useful data (see Figure 7.4 in Chapter Four). Observations are somewhat time consuming and labor intensive compared to other data collection methods, but the rich nature of the data obtained often justifies the extra work required to observe a program in use. Depending on the complexity of the interactive program, you may want to videotape your observations or simply keep notes as you observe. Participant observation is an especially powerful variant of observational methods. Having development team members actually participate in a training session using the interactive learning system they developed can be an extremely enlightening experience for all, especially the team members.

Automated user tracking
Computers are usually essential components of the delivery system for interactive learning systems, and because this is so, you automatically have access to a reliable and tireless electronic observer. The computer can record every key stroke, mouse click, screen touch, or other interaction with the interactive program while providing precise times down to the microsecond for each action. This results in what can be called an “audit trail” of the learner’s progress through an interactive learning system (Schwier & Misanchuk, 1993). Of course, you will soon be drowning in an empirical swamp of data unless you carefully plan ahead for the data you want and how you are going to analyze it.

Consider a typical e-learning program that includes a number of discrete modules, with each module containing several different types of instructional options, e.g., tutorials, quizzes, practice exercises, and mastery tests. You can record trainee paths through the modules, their choices among the instructional options, and their quiz and test scores. Most learning management systems (LMS) include audit trail functions (Hall, 2001). You could even make your online data collection proactive by displaying questions on the screen that inquire about the trainee’s reactions to various aspects of the interactive program at pre-specified intervals or points.

The analysis of audit trail data within complex multimedia or hypermedia programs is especially challenging (Salter, 1995). When learners can go wherever they want in any sequence, the possibility of detecting interpretable paths without the input of the learners becomes almost impossible. This problem is even more complex when the World Wide Web is used for the delivery of interactive learning. It is technically quite easy to track wherever a user goes on the Web, and in fact, commonly used
browsers automatically keep track of path data. However, the interpretation of such data involves a great deal of inference. Even more problematic is the issue of the ethics of tracing where someone goes on the Web. In most cases, you are not advised to collect audit trail data unless complete anonymity can be guaranteed, and even then, learners must be informed that such data are being tracked.

Tests

Although testing is widely accepted as an effectiveness evaluation strategy, you should be wary of tests unless you have evidence of their reliability and validity. Paper-and-pencil tests using multiple choice test items are pervasive as evaluation methods, but the reliability and validity of these instruments are often suspect. There are many different types of test questions (e.g., short answer, true-false, simple multiple choice, complex multiple choice, and essay), and each has advantages and disadvantages in terms of acceptability, feasibility, reliability, validity, and security. You are advised to consult a test construction manual before attempting to assemble a test for effectiveness evaluation (Haladyna, 1999; Morris, Fitz-Gibbon, & Lindheim, 1987; Westgaard, 1999).

Angelo and Cross (1993) provide an excellent introduction to alternative assessment strategies. Viable alternatives to traditional paper-and-pencil tests are performance tests or mini-simulations of the ultimate performance that your interactive learning system is designed to teach. Suppose your interactive learning system trains healthcare workers how to dispose of syringes and other sharp medical instruments without risk of cutting themselves or others. A paper-and-pencil test item might ask them to recite rules for handling this situation, but a simulation or performance test can be set up that actually requires the trainees to demonstrate the knowledge, skills, and/or attitudes being trained. Of course, it is often more difficult and expensive to use performance tests, but the validity of this method is usually high.

Expert review

Expert review can be used for effectiveness evaluation as well as formative evaluation. The immediate supervisors of trainees may provide you with an insightful perspective on the effectiveness of your interactive program if they are given the opportunity to observe or actually participate in the training. Depending on the type of technology you are using, you may also want to utilize instructional experts to review your product’s effectiveness. Content experts are more frequently used for formative evaluation, but they can also provide you with important information about the perceived effectiveness of interactive learning systems.
A few final caveats must be issued regarding effectiveness evaluation. First, you should not rely on only one or two measures of the effectiveness of your interactive learning system. Instead, attempt to “triangulate” effectiveness by measuring it with as wide a variety of approaches as your resources allow. Second, it is as important to evaluate implementation of a program as it is to assess effectiveness because, without knowing how it was actually used, you can’t judge outcomes fairly. Third, rarely trust test data in isolation. Reliable, valid tests are exceedingly difficult to construct, and their security is frequently compromised in education and training contexts.

How should effectiveness studies be designed?

Don’t be surprised if your managers or clients want you to set up a comparative evaluation wherein the effects of the new interactive learning system will be compared with the effects of a competing instructional approach. This type of “horse race” evaluation is frequently conducted, but these kinds of studies usually reveal “no significant differences” in effectiveness between the alternatives. This unfortunate outcome can often be traced to weaknesses in the comparative evaluation design itself and/or a lack of understanding of the dimensions that are common to both programs (Reeves, 1993a, b).

If you must do a comparative evaluation, you are advised to expand the criteria for comparison beyond traditional effectiveness measures to include such dimensions as the experiential value, costs, safety, flexibility, and efficiency of the competing programs (Reeves, 1988). Despite the hype surrounding multimedia and the World Wide Web, interactive learning systems are not always more effective than traditional methods of education and training. However, the implementation of interactive learning systems can often be justified by other criteria. Consider experiential value. There is considerable evidence that experiential learning is more motivating, and if an interactive learning system can be designed to provide more realistic and challenging scenarios than traditional lectures or linear media, learners may be more motivated to learn. The award-winning medical simulations designed by Dr. Joe Henderson (1991) at Dartmouth Medical School are excellent examples of programs with high experiential value.

Another important criteria on which to base a comparative evaluation between interactive learning systems and alternative methods is costs. Although many uninformed people automatically assume that technology-based instruction will always be more expensive than teacher-centered instruction, this isn’t necessarily the case. Gustafson, Reeves,
and Laffey (1990) describe the results of *Macintosh Fundamentals*, an interactive videodisc-based training system developed for Apple Computer, Inc. (This program was described in more detail in Chapter Three.) Although the program cost a million dollars to develop in the late 1980s, 10,000 people completed an average of over nine hours of training via the system during its first year of use alone. The cost per hour of this training compared quite favorably to the cost of any instructor-led approach.

Safety is another important factor in education and training. There are numerous examples of interactive learning systems in science education where hazardous laboratory experiments have been simulated with high fidelity. Also, in an age of rampant litigation, teachers are much less likely to be able to take their students outside the school for field research, especially in dangerous habitats such as lakes and rivers. Fortunately, ground-breaking innovative CD-ROMs like *Investigating Lake Iluka* and *Exploring the Nardoo* enable high school students to carry out research in simulated environments (Hedberg *et al.*, 1994; Wright, Harper, & Hedberg, 1999). In some situations, including military and industrial training, an evaluation based on safety alone may justify the adoption of interactive learning systems even if effectiveness is not greater and costs are higher. Airlines have recognized this for years.

Flexibility is an issue that is increasingly important for learners at a distance, or even for those learners who may be physically near traditional education and training institutions, but who prefer the flexibility provided by interactive learning systems that can be used anytime, anywhere. Comparative evaluations of most distance learning systems with traditional classroom instruction indicate equivalent effectiveness, but the criteria of great flexibility and accessibility tip the scales for the distance approaches (Moore & Kearsley, 1996; Schlosser & Anderson, 1994).

Yet another important criteria for comparative evaluation is efficiency. A consistent finding from the research literature on interactive learning systems has been a 25 percent or more savings in the amount of time learners take to complete a given set of instructional objectives using interactive learning systems over traditional classroom approaches (Hall, 1997; Reeves, 1988; Rosenberg, 2000). These results make sense if you think about group pacing versus individualized pacing. In the former, all learners proceed at the rate determined by the instructor, whereas in interactive approaches, learners often proceed at their own speed. Thirty learners in an hour-long lecture take an average of 60 minutes each to complete instruction. The same learners using computer-based training may take from 20 to 80 minutes to complete instruction, but their average time may only be 45 minutes.
As noted above, evaluators have often relied upon traditional empirical evaluation methods that compare the effectiveness of an instructional innovation with some other approach, but the results of these types of studies have been disappointing (Clark, 1992). A major weakness in comparative approaches to evaluation is that the systems being compared (e.g., interactive multimedia versus classroom instruction) are assumed to be cohesive, holistic entities with meaningful differences. Berman and McLaughlin (1978) as well as Cooley and Lohnes (1976) have illustrated the fallacy of assuming that there are meaningful differences between two programs just because they have different names. Given the great variance in implementation of various instructional methods, there may be more differences among methods with the same name than between differently labeled methods. It is imperative to open up the “black boxes” of alternative learning systems and reveal the relevant pedagogical dimensions they express if an evaluation is to be meaningful and useful.

Reeves and Harmon (1994) describe a set of 14 pedagogical dimensions that can be used to analyze multimedia as an interactive learning system. Figure 8.6 is an example of a rating form using 10 of those dimensions that may be used to analyze the pedagogical differences between two different instructional approaches within a comparative evaluation. It is important to note that this instrument is a descriptive rather than a prescriptive one. Each of the ten dimensions in this instrument is presented as a two-ended continuum with contrasting values at either end. Of course, the world is rarely dichotomous, and there is more complexity involved in learning than any one of these dimensions represents. The individual dimensions themselves are not as important as the arrays across the ten dimensions that represent the instructional designs of various interactive learning systems. A particular array of ratings is not guaranteed to yield an effective interactive learning system. This instrument is to be used to identify the pedagogical similarities and differences between two programs you intend to subject to a comparative evaluation. A description of each dimension follows.

**Pedagogical philosophy**

The debate between instructivist and constructivist approaches to teaching and learning continues throughout education and training (Duffy & Jonassen, 1992; Kafai & Resnick, 1996). At the risk of over-simplifying this issue, the “pedagogical philosophy” dimension ranges from a strict instructivist structure to a radical constructivist one. Instructivists stress the importance of objectives that exist apart from the learner. Once
objectives are identified, they are sequenced into learning hierarchies, generally representing a progression from lower to higher order learning. Then, direct instruction is designed to address each of the objectives in sequence. Little emphasis is put on learners per se, who are viewed as passive recipients of instruction or treated as relatively “empty vessels” to be filled with learning. (Or so constructivists claim.)

**Figure 8.6.** Pedagogical dimensions of interactive learning systems.

Instructivists espouse an objectivist epistemology that defines knowledge as separate from knowing. They believe that reality exists regardless of the existence of sentient beings; humans acquire knowledge about this reality in an objective manner through the senses; learning consists of acquiring truth; and learning can be measured precisely with tests. Examples of interactive learning systems grounded in an instructivist pedagogy abound, especially in the form of computer-based training used in industry and popular computer-based tutorials used in schools.
By contrast, constructivists emphasize the primacy of each individual learner’s intentions, experiences, and cognitive strategies. According to constructivists, any given learner constructs different cognitive structures based upon his or her previous knowledge and what he or she experiences in different learning environments. It is paramount for constructivists that learning environments be as rich and diverse as possible. The learner is regarded as an individual replete with pre-existing knowledge, aptitudes, motivations, and other characteristics that are difficult to assess, much less accommodate. Direct instruction is replaced with tasks to be accomplished or problems to be solved that have personal relevance for each learner.

With regard to epistemology, constructivists believe that knowledge does not exist outside the minds of humans, and that what we know of “reality” is individually and socially constructed based on prior experience. Rather than truth, learning consists of acquiring viable strategies that meet one’s objectives, and at best, learning can be estimated only through observation and dialogue. Wilson (1996) describes numerous examples of interactive learning environments that function primarily as resources for learners engaged in constructing their own knowledge representations.

**Learning theory**

The design of interactive learning systems should be based upon sound learning theories. Although there are other useful learning theories, two that dominate instructional design are behavioral and cognitive psychology. These two theories are often juxtaposed, and thus the “learning theory” dimension has behavioral psychology at one end of the continuum and cognitive psychology at the other. Behavioral psychology continues to underlie most interactive learning systems. According to behaviorists, the critical factors in learning are not internal states, but observable behavior. Behaviorist instruction involves shaping desirable behaviors through the arrangement of stimuli, responses, feedback, and reinforcement. A stimulus is provided, usually in the form of a short presentation of content. Next, a response is demanded, often via a question. Feedback is given as to the accuracy of the response, and positive reinforcement is given for accurate responses. Inaccurate responses result in a repetition of the original stimulus or a modified (often simpler) version of it, and the cycle begins again.

Cognitive psychologists, by contrast, place more emphasis on internal mental states than on behavior. As described above, a typical cognitive taxonomy of internal learning states includes simple propositions, schema, rules, general rules, skills, general skills, automatic skills, and mental models (Kyllonen & Shute, 1989). Cognitivists claim that a
variety of learning strategies, including memorization, direct instruction, deduction, drill and practice, and induction, are required in any learning environment depending upon the type of knowledge to be constructed by the learner. Interactive learning systems can be grounded in behavioral learning theory, cognitive learning theory, or a blend of both.

Goal orientation
The goals for education and training can range from sharply focused ones (e.g., following medical trauma protocols) to general, higher order ones (e.g., developing patient rapport). Similarly, the goal orientation dimension of an interactive learning system can vary in degree of focus from sharp to broad. Cole (1992) maintains that some knowledge “has undergone extensive social negotiation of meaning and might most efficiently and effectively be presented more directly to the learner” (p. 29). In such cases, direct instruction, perhaps in the form of an interactive tutorial, may suffice. Other knowledge is so tenuous, creative, or of a higher level that a constructivist learning environment supporting inductive learning is much more appropriate (Wilson, 1996). Advanced learning environments might include a blend of direct instruction with opportunities to use technology as a cognitive tool (Jonassen & Reeves, 1996).

Task orientation
A basic tenet of adult learning theory (andragogy) is that the context of learning is enormously important to adults (Beder, 1989; Merriam, 1993). Contemporary cognitive learning theories also emphasize the importance of context (cf. Brown, Collins, & Duguid, 1989). The “task orientation” dimension has academic tasks at one end and authentic tasks at the other. Most interactive learning systems still employ academic tasks, but interactive learning can be designed to focus on authentic tasks relevant to learners. Consider the design of educational software for adult literacy. An academic design would depend heavily on having the learners complete traditional academic exercises such as identifying parts of sentences. By contrast, an authentic design would engage the adults in practical activities such as preparing job applications, thereby situating practice and feedback within realistic scenarios. Cognitive learning theory indicates that the ways in which knowledge, skills, and attitudes are initially learned affect the degree to which these abilities can be used in other contexts. If knowledge, skills, and attitudes are learned in a context of use, they will be used in that and similar contexts. Otherwise, it is left up to learners to generate connections between problems and solutions. Whenever transfer to real-world contexts is of primary concern, situating interactive learning in relevant contexts may be advantageous, although the research support for this is still weak.
Source of motivation
Motivation is a primary factor in any theory or model of learning. Every new educational technology promises to be intrinsically motivating, and interactive learning systems are no exception. Consider the hype about the Internet as a vehicle for interactive learning (cf., Perelman, 1992). The “source of motivation” dimension ranges from extrinsic (i.e., outside the learning environment) to intrinsic (i.e., integral to the learning environment) (Malone, 1981). Intrinsically motivating instruction is elusive regardless of the delivery system, but some proponents seem convinced that multimedia motivates learners automatically, simply because of the integration of music, voice, graphics, text, animation, video, and a user-friendly interface. Actually, multimedia studies indicate that some learners tire of these media elements (Reeves, 1993a). It is obvious that motivational aspects must be consciously designed into an interactive learning system as rigorously as any other pedagogical dimension.

Teacher role
Interactive learning systems can be designed to support different roles for teachers, e.g., the traditional didactic role of the instructor as “sage on the stage” or the facilitative role as “guide on the side.” Accordingly, the “teacher role” continuum ranges from didactic to facilitative. Over a quarter century ago, Carroll (1968) told us that “By far the largest amount of teaching activity in educational settings involves telling things to students...” (p. 4). Little has changed since then, despite much discussion of a shift in the teacher’s role from a didactic one to that of a facilitator. Part of the problem is that educational technology research continues to be focused on how the computer can be used to present information and judge learner input (neither of which computers do well) while asking learners to memorize information and later recall it on tests (which computers do with far greater speed and accuracy than humans). It is time to assign cognitive responsibility to the part of the learning system that does it best, i.e., the learner (Jonassen & Reeves, 1996).

Well-designed interactive learning systems may be the ideal vehicle for this transformation. The learner should be responsible for recognizing and judging patterns of information, organizing data, constructing alternative perspectives, and representing new knowledge in meaningful ways, while the computer should perform calculations, store information, and retrieve it upon the learner’s command. For example, when multimedia is used by learners as a cognitive tool for representing their own knowledge, the teacher can be a coach or even a collaborator in the knowledge construction process.
**Metacognitive support**

Metacognition refers to a learner’s awareness of objectives, ability to plan and evaluate learning strategies, and capacity to monitor progress and adjust learning behaviors to accommodate needs (Flavell, 1979). In short, metacognitive skills are the skills one has in “learning to learn.” The “metacognitive support” dimension has unsupported at one end of the continuum and integrated at the other. Imagine an interactive learning system designed to challenge learners to solve complex problems such as troubleshooting electrical circuit boards. Metacognitive support integrated into such a program could provide learners with recapitulations of their troubleshooting strategies at any point in the problem-solving process. Much research and development remains before interactive learning systems regularly include sophisticated metacognitive support, but the potential is enormous.

**Collaborative learning**

Interactive learning systems can be designed to ignore or promote collaborative learning, i.e., some programs require cooperative learning whereas others make no provision for its support. The “collaborative learning” dimension ranges from a complete lack of support for collaboration to the inclusion of collaborative learning as an integral feature. Cooperative and collaborative learning refer to instructional methods in which learners work together in pairs or small groups to accomplish shared goals. When interactive learning is structured to foster collaborative learning, learners can benefit both instructionally and socially. Given an appropriate instructional design, two or more learners working together may accomplish more than a single learner because the interactions between the learners may be just as important for learning as the interactions between the learners and the interactive program.

**Cultural sensitivity**

All instructional systems have cultural implications. For example, whereas constructivist pedagogy advocates questioning on the part of learners, “why?” questions may be inappropriate in some cultures. Although it is unlikely that an interactive learning system can be designed to adapt to every cultural norm, programs should be designed to be as culturally sensitive as possible. The “cultural sensitivity” dimension ranges from insensitive to respectful. Few interactive learning systems have been developed in which cultural sensitivity is integral to their design, whereas more than a few are culturally insensitive. For example, multimedia that uses icons such as a pointing hand to indicate direction may violate a cultural taboo by representing a dismembered body part in certain African cultures. At the very least, interactive learning
systems should accommodate the diverse ethnic and cultural backgrounds among the learners expected to use it. Better yet, these systems should build upon the diversity in the populations where these programs will be used so that the overall learning environment is enhanced.

**Structural flexibility**

Interactive learning environments can be “fixed” or “open” with respect to time and/or place. “Fixed” systems, still dominant in U.S. education, are usually limited to specific places, e.g., a classroom or laboratory, at specific times, e.g., the ubiquitous 50-minute class period. “Open” systems can be used by the learner independent of time and/or place constraints, e.g., print-based independent study materials mailed to learners. Interactive learning systems can provide opportunities for more open (asynchronous) learning. Major questions remain about whether or how interactive learning systems can be used to break the hegemony of fixed instructional modes that limit pedagogical innovation in traditional academic settings. To date, most of the interactive learning systems developed for education and training simply supplement traditional “fixed” approaches to teaching and learning. However, the rapid growth of the Internet and the high bandwidth capabilities of the World Wide Web mean that interactive learning can be designed for delivery anytime, anywhere to anyone with a personal computer and a high-speed modem.

**Applying the Pedagogical Dimensions Analysis Tool**

The *Jasper Woodbury Problem Solving Series* (CTGV, 1992) was created in an academic environment within the context of a research and development program. Its use was initially confined to a few dozen schools in the southeastern U.S., but it is now commercially available. The *Jasper Series* (delivered via interactive videodisc) is regarded as an example of a well-designed interactive learning system because it provides students with opportunities to develop mathematical problem-solving skills through high-interest simulated adventures. Students discover the need to develop mathematical skills within the context of flying ultra-light planes and operating motor boats to solve challenging dilemmas. Numerous studies have been conducted using the *Jasper* series of programs (cf., Bransford *et al.*, 1990).

The *Jasper* series is an example of what Hannafin (1992) calls a “generative” learning environment, i.e., a program that requires students to construct or generate their own knowledge as opposed to one that requires them to select knowledge from prepackaged options. Knowledge constructed in generative environments is more likely to transfer than the inert knowledge that is acquired in traditional passive learning environments (CTGV, 1992). Figure 8.7 presents a profile of the *Jasper*
programs using the 10 pedagogical dimensions described above. These ratings are based on seeing demonstrations of the programs at professional conferences, reading several extensive reports about the programs (cf., Bransford, Vye, Goldman, Hasselbring, & Pellegrino, 1991), and limited first-hand experience with the programs themselves. This analysis reveals that the Jasper programs are grounded in constructivist and cognitivist foundations. Teachers are integral facilitators in implementing Jasper, and they are encouraged to modify it according to their local needs. Collaborative learning is strongly supported in these programs. Special attention appears to have been paid to cultural aspects of the populations likely to use these programs.

![Pedagogical Philosophy Diagram](image)

**Figure 8.7.** Pedagogical analysis of Jasper Woodbury programs.

If we were asked to conduct a comparative evaluation of the Jasper programs with an alternative approach to teaching mathematical problem-solving skills, one of the first things we would want to do is to analyze the pedagogical similarities and differences between the two programs using a version of the pedagogical analysis tool. This would be our first
step because we know that it is the mix of pedagogical dimensions, i.e.,
the instructional design of the programs, that is more important than the
particular delivery systems used to implement them, electronic or other-
wise. The failure to focus on pedagogical differences and the tendency to
over-emphasize technological ones are perhaps the greatest flaws we see
in attempts to evaluate (as well as to conduct research with) interactive
learning systems.

What are some examples of actual effectiveness evaluations?

We have conducted numerous effectiveness evaluations, some quantita-
tive and others more qualitative. Most often, our designs are eclectic,
using a blend of quantitative, qualitative, and even critical methods.

Military example

In the mid-80s, the U. S. Navy began to investigate the use of interactive
videodisc technologies to deliver emergency care training to Naval
Corpsmen (medics) through the CAMIS (Computer Aided Medical
Instruction System) program. A prototype interactive learning system
was developed to address several emergency care procedures that had
been taught via four hours of lecture and demonstration supplemented by
an illustrated text within the context of the overall Corpsman training
program. We were asked to conduct a comparative evaluation of the two
approaches. Accordingly, a quantitative, experimental approach was
adopted whereby 500 sailors were randomly assigned to use the new
CAMIS materials, and another 500 sailors were assigned to attend the
traditional lecture/demonstration classes. Standardized tests were used to
compare the outcomes of the two groups, and the results indicated that
the trainees using the interactive learning system scored an average of
84%, whereas the traditionally-trained corpsmen scored an average of
80%. These differences were statistically significant, and the Navy
officers promoting the new technology were quite pleased with the
findings (Reeves, Marlino, & Strub, 1989).

However, we had concerns about the small size of the gain, regardless of
its statistical significance. We wondered if it was educationally signifi-
cant, especially in relation to the extra costs involved in the new technol-
gy. Fortunately, we supplemented our quantitative design by
interviewing samples of corpsmen in both treatments, in person immedi-
ately after the training and via telephone three months after graduation
from the Corpsmen school. These interviews provided data, albeit “soft”
in the eyes of the Navy officers, that the interactive videodisc scenarios
used in the interactive learning system had been very motivating, and that
the procedures learned through those simulations seemed to be more
durable than those learned in the lectures. We also interviewed instruc-
tors at the Corpsman school and were able to report that they were very
much in favor of the new interactive simulations.

Corporate example

The effectiveness of the Macintosh Fundamentals interactive videodisc
training system mentioned earlier in this chapter was evaluated several
different ways (Gustafson, Reeves, & Laffey, 1990). For example, with
the cooperation of Apple’s Training Support Group in Cupertino,
California, several graduate students from The University of Georgia
enrolled in two-day training sessions at selected training sites, unbe-
knownst to the local trainers. As participant observers, these students
were able to provide a firsthand perspective on the implementation of the
programs. They were also able to talk about the perceived effectiveness
of the interactive modules with other trainees during breaks and meals.

At other training sites, we used both paper-and-pencil tests and perform-
ance exercises as more direct measures of effectiveness. For the latter,
after a trainee had completed a particular module, e.g., one on linking
peripheral devices with a Macintosh, we would invite the trainee into
another room where she would be asked to assemble a typical computer
system including several peripherals such as printers and external hard
drives right out of the shipping boxes. We also conducted exit interviews
with numerous trainees once they had completed Macintosh Funda-
mentals. We were able to report to the Training Support Group at Apple
Computer that MacFun (as it came to be known) was seen as an effective
training system by both trainees and trainers, and that test results indi-
cated significant improvements in learning knowledge and skills before
and after the training. Because of the commitment Apple had already
made to interactive training using their own computers, it did not make
sense to conduct a traditional comparative evaluation between the interac-
tive videodisc training and the previously used leader-led course.

How should effectiveness evaluations be reported?

The results of effectiveness evaluations can be presented in a number of
formats, including formal and informal written reports, multimedia
presentations, online databases, or Web sites. The type of report(s) you
produce will depend on the information needs of your clients and the
other relevant audiences that will be given access to the effectiveness
findings. It is generally advisable to provide multiple reports using
different formats so that your effectiveness data can be widely shared. Of
course, this may not always be possible, depending on the nature of your relationship with your clients. We have found that using a focus group approach wherein we present the results of the evaluation to a group of clients in person, and then solicit their reactions is an especially powerful way of communicating the results of effectiveness evaluations. The discussion that evolves in these focus groups helps to ensure that the results will actually influence decision making.

Summary

The bottom line with respect to effectiveness evaluation can be summed up as follows:

- Never use a test unless you are certain it is reliable and valid.
- Evaluate the implementation of an interactive learning system as carefully as outcomes.
- Triangulate effectiveness, studying it from many different perspectives.

In the next chapter, we present strategies for evaluating the impact of interactive learning environments.

References


