LESSON 26 AND 27:
EVALUATING TRAINING AND RESULTS

Hi Friends

In this chapter we are going to discuss about the last step in training process. As it is said that no communication is complete without feedback similarly no training is complete without its evaluation.

In following lesson you will get exposure towards:

Perspective on Evaluating Training
Basic Suggestions for Evaluating Training

You will be able to :

1. Identify crucial factors pre and post training
2. Carry out evaluation of training
3. Explain the need for evaluation of training
4. Design various types of evaluation
5. Take a project on training evaluation in an organisation

Introduction

Evaluation includes getting ongoing feedback, e.g., from the learner, trainer and learner's supervisor, to improve the quality of the training and identify if the learner achieved the goals of the training.

Perspective on Evaluating Training

Evaluation is often looked at from four different levels (the "Kirkpatrick levels") listed below. Note that the farther down the list, the more valid the evaluation.

a. Reaction - What does the learner feel about the training?

b. Learning - What facts, knowledge, etc., did the learner gain?

c. Behaviors - What skills did the learner develop, that is, what new information is the learner using on the job?

d. Results or effectiveness - What results occurred, that is, did the learner apply the new skills to the necessary tasks in the organization and, if so, what results were achieved?

Although level 4, evaluating results and effectiveness, is the most desired result from training, it's usually the most difficult to accomplish. Evaluating effectiveness often involves the use of key performance measures - measures you can see, e.g., faster and more reliable output from the machine after the operator has been trained, higher ratings on employees' job satisfaction questionnaires from the trained supervisor, etc. This is where following sound principles of performance management is of great benefit.

Basic Suggestions for Evaluating Training

Typically, evaluators look for validity, accuracy and reliability in their evaluations. However, these goals may require more time, people and money than the organization has. Evaluators are also looking for evaluation approaches that are practical and relevant.

Training and development activities can be evaluated before, during and after the activities. Consider the following very basic suggestions:

Before the Implementation Phase

1. Will the selected training and development methods really result in the employee's learning the knowledge and skills needed to perform the task or carry out the role? Have other employee's used the methods and been successful?
2. Consider applying the methods to a highly skilled employee. Ask the employee of their impressions of the methods.
3. Do the methods conform to the employee's preferences and learning styles? Have the employee briefly review the methods, e.g., documentation, overheads, etc. Does the employee experience any difficulties understanding the methods?

During Implementation of Training

1. Ask the employee how they're doing. Do they understand what's being said?
2. Periodically conduct a short test, e.g., have the employee explain the main points of what was just described to him, e.g., in the lecture.
3. Is the employee enthusiastically taking part in the activities? Is he or she coming late and leaving early. It's surprising how often learners will leave a course or workshop and immediately complain that it was a complete waste of their time. Ask the employee to rate the activities from 1 to 5, with
5. Being the highest rating. If the employee gives a rating of anything less than 5, have the employee describe what could be done to get a 5.

After Completion of the Training

1. Give him or her a test before and after the training and development, and compare the results?
2. Interview him or her before and after, and compare results?
3. Watch him ore her perform the task or conduct the role?
4. Assign an expert evaluator from inside or outside the organization to evaluate the learner's knowledge and skills?

Measuring Training Effectiveness/Impact

Training can be Measured in a Variety of Ways Including

I - Prior to Training

- The number of people that say they need it during the needs assessment process.
- The number of people that sign up for it.

II - At the End of Training

- The number of people that attend the session.
- The number of people that paid to attend the session.
- Customer satisfaction (attendees) at end of training.
- Customer satisfaction at end of training when customers know the actual costs of the training.
- A measurable change in knowledge or skill at end of training.
- Ability to solve a "mock" problem at end of training.
- Willingness to try or intent to use the skill/ knowledge at end of training.

III - Delayed Impact (non-job)

- Customer satisfaction at X weeks after the end of training.
- Customer satisfaction at X weeks after the training when customers know the actual costs of the training.
- Retention of Knowledge at X weeks after the end of training.
- Ability to solve a "mock" problem at X weeks after end of training.
- Willingness to try (or intent to use) the skill/ knowledge at X weeks after the end of the training.

IV - On the Job Behavior Change

- Trained individuals that self-report that they changed their behavior / used the skill or knowledge on the job after the training (within X months).
- Trained individuals who's managers report that they changed their behavior / used the skill or knowledge on the job after the training (within X months).
- Trained individuals that actually are observed to change their behavior / use the skill or knowledge on the job after the training (within X months).

V - On the Job Performance Change

- Trained individuals that self-report that their actual job performance changed as a result of their changed behavior / skill (within X months).
- Trained individuals who's manager's report that their actual job performance changed as a result of their changed behavior / skill (within X months).
- Trained individuals who's manager's report that their job performance changed (as a result of their changed behavior / skill) either through improved performance appraisal scores or specific notations about the training on the performance appraisal form (within X months).
- Trained individuals that have observable / measurable (improved sales, quality, speed etc.) improvement in their actual job performance as a result of their changed behavior / skill (within X months).
- The performance of employees that are managed by (or are part of the same team with) individuals that went through the training.
- Departmental performance in departments with X % of employees that went through training ROI (Cost/Benefit ratio) of return on training dollar spent (compared to our competition, last year, other offered training, preset goals etc.).

Other Measures

- CEO / Top management knowledge of / approval of / or satisfaction with the training program.
- Rank of training seminar in forced ranking by managers of what factors (among miscellaneous staff functions) contributed most to productivity/ profitability improvement.
- Number (or %) of referrals to the training by those who have previously attended the training.
- Additional number of people who were trained (cross-trained) by those who have previously attended the training. And their change in skill/ behavior/ performance.
- Popularity (attendance or ranking) of the program compared to others (for voluntary training programs).

The Ten Rules for Perfect Evaluations

On Choosing Between Training Excellence and Great Evaluations

by Jay McNaught

(Originally published by Data Training Magazine in May of 1991)

Among trainers, Joe Rogers was legendary. You would hear his name whispered whenever trainers gathered to discuss evaluations. A trainer among trainers, they said. The instructor with perfect evaluations. They claimed that he had never received less than a perfect evaluation from any of his students.

As a new trainer, I had to know how anyone could be so good that he always scored perfect evaluations. During a business trip, I found myself in the town where he worked, so I decided to give him a call and ask if he could meet with me to give me some pointers.

He turned out to be more than a well-evaluated trainer. He was a generous one as well, and he invited me to sit in on one of the famous training sessions in which I could watch Mr. Rogers's Rules for Perfect Evaluations in action. He even promised to give me an in-depth explanation of what he had done after the session was over.

The day of class, I arrived very early. I didn't want to miss anything. Rogers was already there. From all appearances, he had been in the classroom for some time and was busy preparing. He obviously left nothing to chance. Thick manuals were placed at each seat. I introduced myself, and Rogers told me to have a seat and observe. He pointed out that the work of gaining perfect evaluations required preparation, preparation, preparation.

1. Perfect Order Makes Perfect Evaluations

Cardboard name tents were already placed neatly in front of the manuals. They were hand-lettered in tasteful calligraphy.

"This is incredible," I said. "Who does the lettering on these name tents?"

"Oh, I send those out," he said. "It costs a lot, but the effect is worth it." He placed each manual a precise distance from the

name tent, and each name tent was a precise distance in front of a color computer terminal.

That was what drew my attention to the terminals. I had never seen a clean computer terminal. Yet each terminal in this training room was spotless. The screens were free of dust, and the keyboards were missing the typical sludge that develops over years of use by greasy fingers. I spotted a bottle of cleaner and a rag, and I began to understand.

I sat down and began thumbing through a manual. I was amazed at the detail and was becoming engrossed in the depth of the material, when my thoughts were interrupted. "Please, don't be moving that manual now," said Mr. Rogers. He looked at me over the top of his horn-rimmed glasses and I felt as if I were back in grade school. I set the manual down and walked nervously to a corner of the room where I had spotted refreshments earlier. As I poured a cup of coffee, I noticed him moving the manual back to the precise location where it had been before I violated it.

Satisfied with the placement of the manual, he looked up at me. "Perfect evaluations require perfect attention to detail," he said simply. "When the students are asked if the classroom was neat and orderly, the effort of arranging these manuals will be rewarded."

At this point I noticed the refreshments. Not only was there coffee, but there was juice and soda pop. Also included on the lavish refreshment tray were donuts, Danish, fresh-baked cookies, fresh fruit, and rolls. This was nothing less than a complete breakfast.

2. Good Evaluations can be Bought.

He noticed the way I was staring at the refreshments. "The shortest path to a good evaluation is through the student's stomach," • • he said. "Never let a student sit down to an evaluation form with an empty stomach." Then he walked over to the clock on the wall, pulled it down, and began resetting it.

When he replaced the clock on the wall, I noticed that he had set it a full five minutes earlier than the true time.

When the students began arriving a moment later, they would look up at the clock and then hurry to their seats. You could hear them say things like, "Goodness, I didn't realize it was so late."

3. A little guilt Never Hurts.

I took him aside and asked him why he had altered the time. His response was straightforward. "Make it obvious when you are right," he said, "especially if the student is wrong. You'll notice when you see the evaluation form that one of the questions asks if the instructor began the class on time."

"I don't understand," I said. "Why not just start the class on time?"

"I used to always start classes on time. But students never noticed what time it was when I started, so invariably, one or two students would just assume that the class had started late and mark the evaluation accordingly. I have learned that your good work gains you very little if you don't call attention to it."

When about half of the students had arrived, he began teaching the class. "The clock on the wall says that it is eight o'clock, so let's go ahead and get started. My name is Joe Rogers, and this class is titled, 'Using the Inventory System.'"

At that moment, several other students walked in. Rogers stopped his remarks and stared as they entered. "Welcome to class. The class started at eight, so we went ahead and started without you, but you haven't missed much yet." The new students all seemed to turn the same shade of red.

Rogers continued his introduction. "I have been working with the new inventory system for about a year now. I was actually one of the founding members of the project team which developed the system. I have a master's degree in system development and a Ph.D. in inventory systems."

4. Evaluations Start with Student Impressions.

I was impressed, and I could see that the students were, too.

Rogers continued, "If you'll take the manual on your desks, I will give you a moment to familiarize yourselves with the extensive documentation of the system which I have prepared. Pay special attention to the chapter headings and the table of contents." The students began thumbing through the thick manuals. Rogers quit talking while they read. More students entered, and there was some hubbub as they took their seats. Rogers began strolling around the room. When he came past my chair, he whispered:

"Impressions count. Look at this manual. The impression is that it is very detailed and rich in content. I have been using this manual now for a year, and it has helped me to get perfect evaluations. Look closely. The chapter headings are all accurate. And the first paragraph of every chapter is authentic. But the rest of it is simply the text of those first paragraphs, repeated over and over again in different formats." He rolled his eyes upward. "Thank heaven for word processing."

"Doesn't anybody complain about the content?" I said.

"Oh, come now," he said. "Have you ever met anyone who has read a manual? Excuse me. I don't want to give them too much time to browse it."

5. Teach to the Evaluation.

Rogers turned his attention to the classroom. "Now that you have had a chance to get acquainted with the manual, let's take an opportunity to get to know each other. Would you mind introducing yourself one at a time, and telling what experience you have had with the inventory system?"

The introductions took several minutes. There were 20 students in the classroom, and as each student introduced himself, Rogers would move close to the student, stand only a few feet in front him, and stare intently at him while he spoke.

With the introductions complete, Rogers launched into a lengthy dissertation on the goals and objectives of the class. His explanation took over 15 minutes. I noticed that he repeated himself several times during this explanation. Evidently, the students noticed this as well. About the fifth time I heard him repeat, "So the primary goal for the class today is to make you very familiar with the manual and to inform you about the inventory system," it became apparent that the students were no longer paying any attention to him at all. Rogers recognized

this as well. “Does anyone have any questions concerning the goals and objectives for this course?”

His question was greeted with total silence.

”Please do me a favor and turn to page five in your manual.” He waited while the students obliged. “Now look about half way down the page and circle the section titled ‘Goals and objectives of the class.”’

The students all followed his instructions.

”Now, are there any questions concerning anything we have done so far?”

One student raised his hand, “Why does this class last only two hours? How are we going to learn all of this material in only two hours?”

Rogers took 10 minutes to answer the question. He went into theories on adult attention span and talked about the interactive nature of the inventory system. He discussed his theories on adult learning and told how the on-line inventory help facility was so powerful that they could no doubt use the system with no training at all. By the time Rogers had finished his response, the student had clearly forgotten the original question. “Did I answer your question?”

”Yes.

”Are you sure that I thoroughly answered your question?”

”Oh, yes, I’m sure.”

Much later, during our private, in-depth discussion, I was able to ask Rogers why he had begun the class in this fashion. His response was straightforward. “I’m not going to waste my time doing anything that won’t be directly reflected on my evaluation,” he said. “Question two on the evaluation asks ‘Were the goals and objectives clearly stated at the beginning of the class?’ Question three asks, ‘Were the students allowed to introduce themselves?’ Questions four asks, ‘Did the instructor have good eye contact?’ And question five asks, ‘Did the instructor adequately answer any student questions?’ The students in this class may learn absolutely nothing, but they will know that I had them introduce themselves, that I told them the goals of the class, that I had good eye contact, and that I thoroughly answered their questions.”

6. Good Breaks Lead to Good Evaluations.

After his lengthy answer to the student’s question, Rogers must have sensed that it was time for a break. He went into great detail explaining where the restrooms were as well as phones and even nearby fax machines. “Now, we still have a lot to cover, so let’s hurry back from break. I want to get started again promptly in 25 minutes.”

The students didn’t waste any time in leaving. Again, I was curious, and when Rogers and I were alone in the room, I asked him about it. “Isn’t a 25-minute break a bit exces- sive for a two-hour class?”

’The highlight of any class is the break,” he said. ‘From the moment the student first sits down, he is wondering when the break will be. I am convinced that the longer the break, the better the evaluations!”

Eventually the students returned.

7. Don’t Let the Learning Get in the Way.

”I want to begin the second half of the class by giving you a quick orientation to the classroom and showing you how to use some of the equipment,” said Rogers. “To begin with, you are sitting in special chairs designed to accommodate a variety of preferences and physical needs.” He showed them how to adjust the chairs for maximum comfort.

”I want to point out that these terminals are also specially designed to afford maximum comfort and total student control.” He pointed out the ergonomically correct keyboards and the special non-glare monitors. He showed them how to adjust the contrast for maximum eye comfort. “Now these terminals may be nothing like the terminals you have at your work location, but I do want you to be comfortable in class.” The students were very impressed.

”I also want you to note our state-of-art projection equipment.” He showed them how it worked, and how he could project a computer image on the wall that they could all see. No doubt this would help them in understanding the new inventory system, since they would be able to see the system demonstrated.

But then a student, whose name tent identified him as Sam, asked the question that I was wondering about, “Why are you wasting time showing us this equipment?”

Rogers launched into another of his flowery explanations, emphasizing the importance of using the right equipment in training, and how he was dedicated to quality instruction.

After the session, he told me the real reason. “There is a lot more at stake here than any of these students understand,” he said. “I can’t afford the risk of a lot of instruction when other things could affect the evaluation.”

’Risk?” I said.

”Every year my boss gives me a performance review,” he said, “and it is based entirely on how well I do on student evaluations. My boss doesn’t care if the students learned anything or not. He only cares that I do well on the evaluations, because those evaluations are what he reports to his boss, and they are the basis for my boss’s own performance reviews. If he gets a good performance review, then I get a good performance review, and we both get big raises. I can’t jeopardize the welfare of so many people by spending a lot of time on something that counts for very little in the evaluation process.”

8. Absent students don’t complain.

His explanation to Sam about quality instruction and good equipment was still hovering in the air when the telephone rang. Rogers apologized for the interruption and quickly picked up the receiver. He had a brisk and hushed discussion. “I see. I’ll tell him right away.”

As he hung up the phone, Rogers turned to Sam, “That was your office. They said something about an emergency project. They want to know if you can leave class early and get back to the office right away.”

Later, during our review discussion, 1 remarked on his good fortune. “It was lucky that Sam had to leave before you gave out

evaluations. I had the feeling he didn't like you and might have given you a bad evaluation out of spite."

Rogers grinned and shook his head. "A good instructor leaves nothing to chance."

It took a moment for the implication of his remark to sink in. "You mean you planned to have the student taken out of your class?"

"It's an easy thing to program my personal computer to ring the classroom telephone every day right before I hand out evaluations."

9. Timing is Everything.

"Well, we only have 20 minutes left," said Rogers after Sam left. "I am hoping that we can wrap this class up a little early, so I'm going to hand out the class evaluation forms now. This way we can make sure that you won't have to rush to finish them." He handed out the forms. As the students began marking the forms, Rogers continued talking. "After you return your class evaluation form to me, I will give you your plaque, stating that you have successfully completed this class."

As the students began filling out the evaluations forms, we were able to have another of our hushed conversations.

"Why now?" I whispered. "Why not wait and let the students do a class evaluation after they get back on the job? It would seem to me that they could better evaluate the leaming once they actually started using the inventory system."

Rogers looked like he was going to laugh out loud, but he caught himself and looked around at the students. "You're so hung up on leaming," he said. "If I really wanted to know if the students had learned something, I would wait at least three months before doing an evaluation."

That seemed reasonable to me.

"I don't want my success tied in with the student's ability to learn," he continued. "What if I did wait to do the evaluation and then discovered that a student couldn't do anything he had learned in the class? What if the student was a total moron and just plain could not learn anything anyway? It would reflect poorly on me. So I give the evaluation immediately after the class, while everything I did is still fresh in the student's mind. This way, my performance in the classroom is all that is being evaluated. You might have heard this kind of evaluation form called a smile sheet. Well, to the extent it proves to my boss what a great trainer I am, it makes me smile!"

10. Ask the Right Questions Right.

I watched the students fill out the evaluation forms. I was somewhat dazed by all that I had witnessed. I picked up an extra copy of the evaluation form and was enlightened concerning his tenth and final rule: Just ask the right questions and ask the questions right.

As I read over the evaluation form, I was sad to see that it mentioned nothing about what the student had learned. Of course, that was because the students had learned nothing. Then I remembered that Rogers said that I was too hung up on learning. I noted the way he had phrased the questions. There was no room here for ambiguity, no space provided for comments. In each case, the answer Rogers desired was the only one possible.

Later, I asked one final question. "Is that all there is to it?" I said. "Don't you do anything to measure performance?"

"Most definitely." he said emphatically. "After each class, I send a glowing letter to each student's supervisor. I tell how well the student did, and how confident I am for expert performance on the inventory system."

"Isn't that a lot of work?" I said.

"Sure it is," he said, "but it's worth it. A report like that puts the onus on the student to meet performance expectations."

I realized then that he was both determined and tireless in his pursuit of perfect evaluations, and he wasn't about to let student leaming stand in the way.

Class Evaluation Form

Class: Using the Inventory System

Instructor: Joe Rogers

Student Name:

1. Were the classroom facilities adequate? Yes/No
2. Were the goals and objectives clearly stated at the beginning of class? (Refer to page five in your manual.) Yes/No
3. Were the students allowed to introduce themselves? Yes/No
4. Did the instructor have good eye contact? Yes/No
5. Did the instructor adequately answer any student questions? Yes/No
6. Was the class too in-depth? Yes/No
7. Did the instructor begin the class on time? Yes/No
8. Did the instructor state his name at the beginning of class? Yes/No
9. Were the refreshments adequate? Yes/No
10. Were the handouts adequate and thorough? Yes/No
11. Was the classroom neat and orderly? Yes/No
12. Was the instructor knowledgeable about the subject? Yes/No

Evaluating Training

There is No "Cookbook" Approach

This is a close-to-the-original version of an article prepared for a 1992 ASTD Tool Kit edited by Karen Medsker and Don Roberts. The original version was published in three separate pieces. This one is more or less intact.

Evaluate What and Why?

Evaluate? Evaluate what? Training? What do we mean by training? What's to be evaluated? A particular training course? The trainees? The trainers? The training department? A certain set of training materials? Training in general?

More to the point, why evaluate it? Do we wish to gauge its effectiveness, that is, to see if it works? If so, what is it supposed to do? Change behavior? Shape attitudes? Improve job performance? Reduce defects? Increase sales? Enhance quality?

What about efficiency? How much time does the training consume? Can it be shortened? Can we make do with on-the-

job training or can we completely eliminate training by substituting job aids instead?

What does it cost? Whatever it costs, is it worth it? Who says? On what basis? What are we trying to find out? For whom?

The preceding questions illustrate the complexity of any effort to evaluate training and emphasize the importance of being clear about the purposes of and the audiences for any such evaluation.

It is the central thesis of this article that the evaluation of training poses a problem for many trainers, managers, executives, and other professionals with an interest in training. Further, it is my firm conviction that these problems are most productively addressed by examining their underlying structure. As Dewey (1910) wrote, "A difficulty clearly apprehended is likely to suggest its own solution (p. 94)". This article, then, will examine various elements in the structure of the problem of evaluating training.

The centerpiece for the collection of articles comprising the ASTD Tool Kit for which this paper was originally written is Donald Kirkpatrick's well-known framework for evaluating training, frequently referred to as "Level One," "Level Two," and so on. Much has changed since Kirkpatrick's framework first appeared and it might help to better understand and appreciate the truly seminal nature of his work if we attempt a very brief review of some of the major changes in the training and development world since then.

A Brief Historical Perspective: 1960-1990

Donald Kirkpatrick set forth his four-level approach to the evaluation of training in a series of articles appearing in the journal of what was then known as the American Society of Training Directors. The first of these four seminal articles was published in November of 1959. The remaining three articles were published in the succeeding three months, with the fourth and final article appearing in February of 1960. These articles can be found in Evaluating Training Programs, a collection of articles compiled by Kirkpatrick from the pages of the ASTD Journal and published by ASTD in 1975.

In 1959, when Kirkpatrick launched his views, the American Society of Training Directors (ASTD) was about as close-knit a "good old boys" network as one could find. Since its inception in the 1940s, ASTD membership had consisted primarily of training directors, known also as training officers. Even as late as 1969 (the year in which I took up the training profession), ASTD was still dominated by training directors. That the members of ASTD were in fact "old boys" is amply demonstrated by some figures from the 1969 ASTD national conference, which was held in Miami, Florida (Reith, 1970): Only nine percent of the attendees were 29 years of age or younger. Fully 59 percent were 40 years old or older. Only nine percent of the attendees were females. To elucidate the obvious, 91 percent were males. Any group consisting of more than 90 percent males past the age of 40 certainly seems vulnerable to charges of being a bunch of "good old boys."

Changes, however, were already evident. Of the 1,081 full-time attendees filling out the Miami conference feedback form, almost half or 49 percent were attending their first ASTD national conference. More than 77 percent had been in training assignments for more than three years and roughly 40 percent had been in training assignments for more than 10 years. But, at the same time, more than 50 percent of those attending had been in their present jobs for less than three years.

Elsewhere, the training business was stirring. The likes of Bob Mager, Susan Markle, Tom Gilbert, Geary Rummler, Joe Harless and Karen Brethower were shaking up the training establishment and would continue to do so for several more years. The development business was stirring too. Rensis Likert, Chris Argyris, Douglas McGregor, and George Odiorne were shaking up the management mindset and a new term had entered our vocabulary: "Organization Development (OD)."

The board of governors of the American Society of Training Directors, perhaps sensing some kind of shift in the tide of human and organizational affairs, changed the name of the society from the American Society of Training Directors to the American Society for Training and Development, and moved its headquarters from Madison, Wisconsin to the Washington, D.C. area (Alexandria, Virginia).

Other changes affecting the training and development worlds were taking place during this same time period. Behaviorism flowered for a while then wilted in the face of the shift to knowledge work. Peter Drucker, in book after book, beginning with Landmarks for Tomorrow (1959) and continuing through The New Realities (1989), kept reminding us that the center of gravity in the employed workforce was shifting from those who worked with their muscles to those who worked with their minds. By 1980, the shift to knowledge work was more or less complete and, three years later, I spelled out some of its consequences for training and trainers (Nickols, 1983).

As perceptions of the locus of working gradually and painfully shifted from the workers' muscles to their minds, the focus of managerial control over work and working shifted from the exercise of direct control over overt physical behavior to a search for ways and means of influencing covert mental processes. In short, the cognitive view gained sway (and it is likely to hold sway for the foreseeable future). Nevertheless, behaviorism, mostly through the efforts of Bob Mager, did give us this central question pertaining to the evaluation of training: "What is the trainee supposed to be able to do as a result of training?" — and the training business hasn't been the same since.

Programmed instruction blossomed for a while too, and was then displaced by its own progeny: self-instructional materials, job aids, and performance technology. Another society, the National Society for Programmed Instruction (NSPI), moved its headquarters from San Antonio, Texas to Washington, D.C., and changed its name to the National Society for Performance and Instruction. (It has most recently become the International Society for Performance Improvement.)

Systems concepts and the systems approach came rushing at us from two very different angles. We didn't stand a chance; we were overwhelmed by superior forces. Systems engineering, apparently obeying the biblical command to be fruitful and multiply, gave us the systems approach to this, that, and the other. Its primary legacy consists of (1) the instructional systems development (ISD) model originally developed in the

military and (2) the computer systems development process found throughout business and industry.

General systems theory (GST) was fertile and prolific too, mostly on the organizational side of things. The concepts of "open" and "socio-technical" systems came into vogue and stayed. "Systems thinking" is with us still, so pervasive now that we hardly give it a second thought. Human relations was a burgeoning movement in this same period. Years earlier, Elton Mayo had given us the "Hawthorne effect" and, in the 1960s and 1970s, his legatees gave us sensitivity training, T-groups, and organization development (OD). One of Mayo's philosophical descendants, Len Nadler, coined the term "human resources" and people haven't been looked upon as people since.

Technology was at the heart of much of what was going on from 1960 through 1990. For 10 of those years (1965 to 1975) a brief war was waged between "educational technology" and "instructional technology." It was a civil war, of course, and like a lot of recent wars it ended in a draw; there weren't any clear-cut winners, but at least the hostilities came to an end.

Donald Kirkpatrick's four-level evaluation framework has survived all this turbulence. One might even say that it has prospered. At the very least, one must acknowledge its staying power — and rightly so, for, although his framework might not be the last or latest word in the evaluation of training, it certainly comes close to being the first word on the subject.

Let us now shift our focus from the past to the present and begin our examination of the evaluation of training problem. Our starting point is with the structural relationship between training and the workplace.

Training and the Workplace

Most training takes place in an organizational setting, typically in support of skill and knowledge requirements originating in the workplace. This relationship between training and the workplace is illustrated in Figure 1.

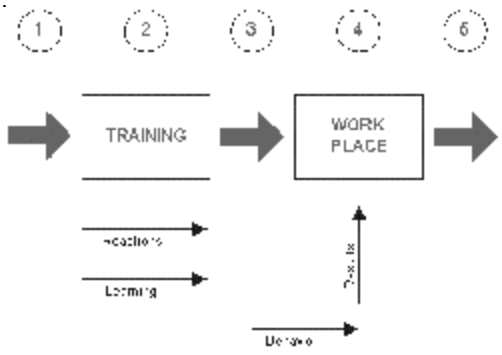


Figure 1 - The Structure of the Training Evaluation Problem

Using the diagram in Figure 1 as a structural framework, we can identify five basic points at which we might take measurements, conduct assessments, or reach judgments. These five points are indicated in the diagram by the numerals 1 through 5:

1. Before Training
2. During Training
3. After Training or Before Entry (Reentry)
4. In The Workplace
5. Upon Exiting The Workplace

The four elements of Kirkpatrick's framework, also shown in Figure 1, are defined below using Kirkpatrick's original definitions.

1. Reactions. "Reaction may best be defined as how well the trainees liked a particular training program." Reactions are typically measured at the end of training — at Point 3 in Figure 1. However, that is a summative or end-of-course assessment and reactions are also measured during the training, even if only informally in terms of the instructor's perceptions.

2. Learning. "What principles, facts, and techniques were understood and absorbed by the conferees?" What the trainees know or can do can be measured during and at the end of training but, in order to say that this knowledge or skill resulted from the training, the trainees' entering knowledge or skills levels must also be known or measured. Evaluating learning, then, requires measurements at Points 1, 2 and 3 — before, during and after training

3. Behavior. Changes in on-the-job behavior. Kirkpatrick did not originally offer a definition per se for this element in his framework, hence I have not enclosed this one in quotation marks. Nevertheless, the definition just presented is taken verbatim from Kirkpatrick's writings — the fourth and final article. Clearly, any evaluation of changes in on-the-job behavior must occur in the workplace itself — at Point 4 in Figure 1. It should be kept in mind, however, that behavior changes are acquired in training and they then transfer (or don't transfer) to the work place. It is deemed useful, therefore, to assess behavior changes at the end of training and in the workplace. Indeed, the origins of human performance technology can be traced to early investigations of disparities between behavior changes realized in training and those realized on the job. The seminal work in this regard is Karen Brethower's paper, "Maintenance: The Neglected Half of Behavior Change" (see the references section).

4. Results. Kirkpatrick did not offer a formal definition for this element of his framework either. Instead, he relied on a range of examples to make clear his meaning. Those examples are herewith repeated. "Reduction of costs; reduction of turnover and absenteeism; reduction of grievances; increase in quality and quantity or production; or improved morale which, it is hoped, will lead to some of the previously stated results." These factors are also measurable in the workplace — at Point 4 in Figure 1.

It is worth noting that there is a shifting of conceptual gears between the third and fourth elements in Kirkpatrick's framework. The first three elements center on the trainees; their reactions, their learning, and changes in their behavior. The fourth element shifts to a concern with organizational payoffs or business results. We will return to this shift in focus later on.

Thinking about the Evaluation of Training

The diagram shown in Figure 1 not only depicts Kirkpatrick's evaluation framework, it also indicates the points at which it takes measurements, collects data, and so forth. We can create

other possibilities for evaluating training by altering the points at which these same measures are taken.

Trainee reactions, for instance, could be assessed at Point 4, after the trainees have been on the job for a while, instead of so soon after the completion of training. In a slightly different vein, we could compare Points 2 and 4, which essentially amounts to comparing the training environment with the workplace environment. From such a comparison we might be able to gauge the "authenticity" of the training, that is, how closely the training environment matches or resembles the workplace environment and, from this, draw some conclusions about the likelihood of a phenomenon known as the "transfer of training."

We can "get outside the box," so to speak, and pick points not even shown on the diagram. Moving all the way to the left of Point 1, for instance, we can speculate that trainees arrive at Point 1 as a result of some kind of selection process. In the course of evaluating training, we (or someone else) might wish to measure the effect selection has on success in training. Moving all the way to the right, beyond Point 5, we can inquire as to where people go when they leave the workplace, perhaps at the end of the day or perhaps at the end of a career. One answer is that they go home. Another is that they reenter the larger community in which the organization is embedded and from whence they came. From this perspective, one might ask, "What good or harm comes to the community as a result of the organization's training and workplace practices?" Alternately, "Is the organization turning out skilled, self-supporting members of the community, or is it simply chewing up people and spitting out dull-eyed, unthinking, uncaring automatons who are of no further value to themselves or to society?" In short, by moving all the way to the right in Figure 1, we begin examining the societal impact of organizations — and of the training they provide — or don't provide, as the case may be.

Another way to make use of the structure depicted in Figure 1 is to change the time perspective being used. Kirkpatrick's "Reactions" element is a retrospective or after-the-fact view. The trainees are looking back at the training (to the left from Point 3). Why not substitute a perspective of looking forward? At Point 3, the notion of looking forward raises the possibility of asking the trainees to provide their predictions regarding the nature of the workplace they're about to enter. In other words, we might consider assessing the image of the company and the workplace that is communicated by the training experience.

As seen earlier, learning is typically assessed through before and after measures. This is a point-to-point measurement and comparison, it spans a "chunk" of the framework. By varying the points used, we can identify other "chunks" and come up with other evaluation issues. We could, for instance, create a span encompassing all of Figure 1 — Points 1 through 5 — and this might suggest larger learning issues that involve training and development in an integrated fashion. How do training and workplace developmental experiences dovetail, for instance, in mapping out career paths?

Create a span from Points 1 through 3, the same span used in gauging learning, but take the perspective of the manager of the people who are going through training. A couple of likely evaluation issues from this perspective can be expressed in two terse questions: "How long is it going to take? What is it going to cost?"

Let's pick yet a different audience for the evaluation of training: The professional training community. And let's use Point 2, the training process, as our focal point. It could well be the case that an evaluation for this audience at this point in the structure we are using would center on matters like adherence to standards for design and delivery, that is, the "professionalism" of the training.

Stay at Point 2 and adopt the trainees' perspective. Perhaps the chief evaluation issue in this case can be expressed in a single question: "How does all this (the training) relate to my job?"

Suppose we go to Point 1, adopt a looking forward (to the right perspective), and put on our executive's hat. What might we be interested in from that perspective? One quick answer is the results that can be expected in the workplace, at Point 4. Another is the resources required to achieve those results.

Training, like all organizational functions, must compete for resources. Moreover, resources must be allocated before any effort can be undertaken. From this it follows that resource allocation decisions must be made before the resources can be expended. Consequently, from the resource allocation perspective, the case to be made regarding the results of training must be made before the training is conducted, not after.

The preceding examples of evaluation possibilities were arrived at by varying elements of the structure of what might be termed "the evaluation of training problem." One of the elements varied was the point or span of points in the process at which measurements might be taken. Another element varied was the audience for the results of the evaluation. Yet a third element varied was the time perspective employed. Varying these elements, singly or in combination, permits us to identify some of the many purposes for evaluating training. In turn, the purposes for evaluating training are inextricably bound up with the purposes of the training being evaluated.

The Many Purposes of Training

Almost 20 years ago I wrote a brief article addressing what I saw as the need to adopt a "strategic view" of training (Nickols, 1981). My aim then, as now, was to point out that "training is a management tool, not the private domain of those who specialize in its development or delivery, nor of those who make its development and delivery contingent upon some other methodology." By "some other methodology," I mean performance technology, which seems to me to view training as little more than an occasionally useful remedy for skill or knowledge deficiencies.

As a management tool, training serves many masters and many purposes. In the article just mentioned, I presented and explained examples of three such purposes (the first three in the list below). Additional purposes for or uses of training are given in the list below. It is not my intent here to elaborate upon these many purposes. Instead, I wish merely to prompt you to think about how the evaluation of training might vary with the purpose or use of the training itself.

1. Focusing energy on issues.
2. Making work and issues visible.
3. Supporting other interventions.
4. Legitimizing issues.
5. Promoting change.
6. Reducing risk.
7. Creating a community based on some shared experience.
8. Building teams.
9. Indoctrinating new staff.
10. Communicating and disseminating knowledge and information.
11. Certifying and licensing.
12. Rewarding past performance.
13. Flagging "fast trackers."
14. Developing skills.

Given the diverse array of purposes listed above, it seems reasonable to conclude that the results sought from the training would also be diverse. And so they are. It is time now to return to the issue postponed earlier; namely, the fourth element in Kirkpatrick's framework, the results of training.

The Results of Training

When we speak of measuring the results of training — and we mean results beyond those of simply equipping people with the skills and knowledge necessary to carry out their assigned tasks and duties — we are redefining training as an intervention, as a solution to some problem other than equipping people to do their jobs.

In cases where skill and knowledge deficiencies are leading to mistakes, errors, defects, waste, and so on, one might argue (and many do) that training which eliminates these deficiencies and in turn reduces mistakes, errors, defects, and waste, is a solution to a performance problem. This argument is extended to assert that the reductions in mistakes, errors, defects, and waste, as well as the financial value of any such reductions constitute the "results" of training.

The logic of this argument has a certain superficial appeal but it is far from impeccable and even farther from compelling. In short, it does not withstand serious scrutiny. It is frequently pointless to ask "What business results were achieved as a result of training?" because the goal of training is generally one of preventing mistakes, errors, defects, and waste, not correcting them. Thus, by a strange twist of circumstances, the only way to prove that such training is successful is to shut down the training. As is the case with some other things, it is sometimes the case with training that the true measure of its value lies in its absence, not its presence, but shutting down training is hardly a practical way of testing that proposition.

At this point, it seems worthwhile to see if the evaluation of training problem can be cast in a more practical light. To accomplish this aim, we will use a completely fictitious, hypothetical, situation, one in which an equally fictitious executive, Lee Resnick, will play a central role. In short, let's pretend.

Let's Pretend

Pretend you are Lee Resnick, senior vice president for systems and operations at the Cowardly Lion Insurance Company. You are cutting over to a new, multi-million dollar insurance policy administration system in just a few months and your neck is on the line to the CEO for a "smooth, problem-free introduction" of the new system. You know that's a joke and so does the CEO — there's no such thing as a "problem-free introduction" of a new system — but the underlying message is also clear: If things get too screwed up, it'll be you that gets the ax, not the CEO.

The new system radically alters the way the clerical staff members do their jobs; indeed, the jobs themselves have been radically restructured. Obviously, the people need to be retrained. They need to know how the new system works and how to carry out the many new and different procedures they'll encounter. They'll also have to be sold on the new system, so as to reduce the friction at installation time. Moreover, you don't need some training consultant to tell you all this. You also know that, given enough time, the clerical staff wouldn't need much in the way of formal training at all. Sooner or later, they would figure out how to make the system do what it was supposed to do. In short, they would learn how to do the job even if they weren't trained how to do it. But you don't have time. And you can't afford to live with the financial and political costs of the error rates you'd encounter in a world where people are learning solely from their mistakes. You don't need to be told this, either. So, you know you're going to spend some money on training. The primary issue facing you is how much? How much money and for how much training?

Depending on the riskiness of the situation, your personal circumstances, your career ambitions, and a host of other factors, you might be inclined to go for the minimum amount of training and the minimum expenditure of cash or, conversely, the cost and length of the training might be no object. Which of these is the case is more or less immaterial because your choice, in either case, will be governed by what is essentially the same criterion: Of the options available to you, which seems most likely to serve your purpose?

When you follow up, which you're very likely to do, you're likely to make do with a few phone calls, a few questions, and a few answers. Formal, structured, and expensive after-the-fact evaluations are of little use and could even pose an inadvertent threat. What would you do, for instance, if you commissioned the kind of evaluation the training people are pressing for and it revealed that the money you spent on training was wasted? Now how's that going to look come performance appraisal time? (Fortunately, you can always hang the blame on the trainers.)

As Lee Resnick, you can probably relate very quickly to item six in the list of training purposes presented earlier: Reducing risk. Your primary motive in providing the training is simply to ensure that the lack of training doesn't create a problem during cutover. Training, in this case, is insurance; prevention as much or more than intervention.

Let's Pretend Some More

Suppose now that you are a new general manager and that your

department heads have a long history of isolation and compartmentalism, a history of not talking to one another. Further, suppose you decide to use some training sessions as a means of bringing them together and getting them started talking with one another. How would you evaluate this training?

Suppose instead that, historically, a deaf ear has been turned to laments and complaints about the company's performance appraisal system. A new CEO, charged with changing the corporate culture, is willing to modify it. How could training be used in support of this objective? Which of the purposes in the list above might this kind of training serve? How would you evaluate this training?

Suppose, finally, that the officers of the company are dissatisfied with the quality of their own training and education and decide to institute an advanced management program. First, they attend. Next, some but not all the of senior managers in the pool from which the officers are selected also attend. What's going on here? Which purposes are being served? How would you evaluate this training?

The root word of interest in this article is a verb: "Evaluate." To evaluate some thing is to determine its value, to find its strength or its worth. To evaluate training is to determine its value. Value is relative. What is of great value to one person is of little or no value to another. In evaluating training, then, it is important to know one's audience — the person or persons for whom the determination of value is to be made. As noted earlier, there are several possible audiences for evaluation results. These include the trainees, their managers, the trainers and their managers, the executives of the organization wherein the training is taking place, members of the training profession and even, as we saw at one point, members of the larger community in which the organization is embedded.

Because the definition and perception of value varies from person to person, so do the purposes of evaluation. Moreover, the various audiences for evaluation frequently act as their own evaluators. If you look carefully about you, or if you reflect upon your own experiences as a "trainee," you will quickly discover that training is being evaluated every day, but by trainees, managers, and executives — and in accordance with their criteria and purposes.

Conclusion

The concluding point to be made here is very, very simple and very, very important: There is no "cookbook" approach to the evaluation of training. To properly evaluate training requires one to think through the purposes of the training, the purposes of the evaluation, the audiences for the results of the evaluation, the points or spans of points at which measurements will be taken, the time perspective to be employed, and the overall framework to be utilized.

Evaluation Tools

Evaluation Matrix

- Although by all appearances, the "Evaluation Matrix" is a very simple tool, it has a powerful purpose. It helps you to consider a wider range of data collection methods than you might otherwise consider in relation to each of the questions addressed by your evaluation. Evaluators sometimes get into the habit of using one or other data collection method, e.g., an end-of-training questionnaire, without considering the advantages of alternative methods. This tool prompts you to consider each evaluation question and to decide which of the many data collection options have the greatest potential for providing the desired information.

Instructions For Use

1. The "Evaluation Matrix" tool will help you consider the most appropriate and feasible data collection method for each of the questions identified in your evaluation plan. (Remember that evaluation is different from assessment in that evaluation is focused on the effectiveness and worth of programs or products whereas assessment is focused on estimating student learning.)
2. List your questions on the vertical side of the matrix.
3. List the feasible data collection methods on the horizontal side of the matrix.
4. Consider each question carefully and choose the most appropriate data collection method.
5. To make your own matrix, copy the matrix below and paste it into a new ClarisWorks file using the "Drawing" option in the "New Document" dialog box.
6. In the "Drawing" option, the elements of the matrix can be edited.

Evaluation Matrix

Evaluation Questions	**Data Collection Methods**												
	Anecdotal Records	Expert Review	Implementation Logs	In-Basket Exercises	Instructor Interviews	Observations	On-Line Data	Portfolios	Tests	User Interviews	User Logs	User Questionnaires
a. What knowledge was learned by trainees?							X	X	X	X		
b. What skills were developed by trainees?						X	X	X	X	X		
c. What attitudes were formed by trainees?						X				X	X	X
d. What were trainee reactions to the IMM?	X									X		X
e. What were instructor reactions to the IMM?					X	X						

Anecdotal Record Form

- Evaluation data does not have to be reported as "cold hard statistics." Often you will want to tell the "human story"

involved in your development or implementation project. One way of capturing those important stories and critical incidents that provide the human story is the "Anecdotal Record Form." Participants in an interactive multimedia design project can use this instrument to describe a noteworthy event and to offer their own interpretation of its relevance. It is very important to try to complete an Anecdotal Record Form as soon as possible after a critical event has occurred so as not to forget critical information. It is equally important to separate your description of the incident from your interpretation of it!

Anecdotal Record Form

Evaluation data does not have to be reported as "cold hard statistics." Often you will want to tell the "human story" involved in your development or implementation project. One way of capturing those important stories and critical incidents that provide the human story is the "Anecdotal Record Form." Participants in an interactive multimedia design project can use this instrument to describe a noteworthy event and to offer their own interpretation of its relevance. It is very important to try to complete an Anecdotal Record Form as soon as possible after a critical event has occurred so as not to forget critical information. It is equally important to separate your description of the incident from your interpretation of it!

Instructions:

1. As a participant in an interactive multimedia design project, you will observe incidents or listen to reports of incidents which relate to the development and impact of the program. It is important that this kind of anecdotal information be systematically recorded so that the story of the development and outcomes of this project can be understood. Therefore, you should complete an Anecdotal Record Form whenever you witness or hear of a significant incident relating to the progress and accomplishments of project. An anecdotal record is a verbal account which exhibits these characteristics:

a. Each anecdote should be limited to a single incident.

b. It should contain a factual, non-inferential description of the observed or reported incident. (For example, "The trainees said 'I've never enjoyed using a computer before.' " instead of "The trainee expressed satisfaction with the training system.")

c. It should include a description of the situation in which the incident occurs so that the meaning of the behavior can be understood.

d. It should be written as soon as possible after witnessing or hearing about the incident so that all important details can be included.

e. It should include a separate section describing your interpretation of or feelings about the anecdote. Your personal evaluation is important because your judgments about the project are valued highly.

2. A copy of a blank Anecdotal Record Form as well as a sample completed form appears below.

Blank Anecdotal Record Form

DATE: _____ PLACE: _____

NAME OF OBSERVER: _____

Description of the incident:

Interpretation:

Sample Anecdotal Record Form

DATE: July 23, 1992 PLACE: Beta Site 2

NAME OF OBSERVER: Lucy Schweitzer

Description of the incident: About two hours into the course, one of the trainees suddenly got up and left the class. I followed him out into the hall and asked if anything was wrong. He replied: "I can't waste my time sitting in the class because I don't intend to use the new system." I asked him why and he answered: "Computers don't work for me. As soon as I touch one, the program blows up. You'll be glad I won't use your system because it would just fall apart if I did. It's nothing against you or your course, I just know it won't work."

I tried to talk to him more, but he indicated that he had to make some phone calls and left.

Interpretation: The "(Insert name here.)" course training is innovative and user-friendly in our eyes, but in the eyes of a person with high anxiety about technology, it is just another threatening computer program. I suspect that this person strongly fears computers and that he has an unusually strong degree of "learned helplessness" with respect to them. It may be worthwhile to conduct some sort of a pre-assessment with respect to "techno-phobia" and makes special efforts to help those who express high anxiety. Also, this person indicated before the beginning of the course that he was only there because his boss insisted that he attend. We may need to clarify the enrollment procedures for this and other clients.

Expert Review Checklist

- Expert review is one of the primary evaluation strategies used in both formative (How can this multimedia program be improved?) and summative (What is the effectiveness and worth of this multimedia program?) evaluation. It is often a good idea to provide experts with some sort of instrument or guide to insure that they critique all of the important aspects of the IMM program that you want reviewed. This "Expert Review Checklist" has been designed for use by an instructional design expert. You would employ different sorts of Expert Review Checklists with different types of experts such as a content expert or a human computer interface expert.

Expert Review Checklist

Expert review is one of the primary evaluation strategies used in both formative (How can this multimedia program be improved?) and summative (What is the effectiveness and worth of this multimedia program?) evaluation. It is often a good idea to provide experts with some sort of instrument or guide to insure that they critique all of the important aspects of

the IMM program that you want reviewed. This "Expert Review Checklist" has been designed for use by an instructional design expert. You would employ different sorts of Expert Review Checklists with different types of experts such as a content expert or a human computer interface expert.

Instructions:

1. The "Expert Review Form" is a tool that will help assure that the experts reviewing your interactive program focus on the variables of most interest to you. (Of course, they will usually provide you with additional aspects of the program. That's why they are called experts!)
2. A sample "Expert Review Form" appears on the next page.

Expert Review Form

Expert Review Check List for Interactive Multimedia

REVIEWER: Dr. Lynn Knowitall **DUE DATE**: June 10, 1994

Please circle your rating and write comments on each aspect of the interactive multimedia (IMM) package. 1 represents the lowest and most negative impression on the scale, 3 represents an adequate impression, and 5 represents the highest and most positive impression. Choose N/A if the item is not appropriate or not applicable to this course. Use additional paper for comments.

NA=Not applicable 1=Strongly disagree 2=Disagree
3=Neither agree/nor disagree 4=Agree 5=Strongly agree

Area 1: Instructional Design Review

1. This IMM provides learners with a clear knowledge of the program objectives. N/A 1 2 3 4 5
2. The instructional interactions in this IMM are appropriate for the objectives. N/A 1 2 3 4 5
3. The instructional design of this IMM is based on sound learning theory and principles. N/A 1 2 3 4 5
4. The feedback in this IMM is clear. N/A 1 2 3 4 5
5. The pace of this IMM is appropriate. N/A 1 2 3 4 5
6. The difficulty level of this IMM is appropriate. N/A 1 2 3 4 5

Area 2: Cosmetic Design Review

7. The screen design of this IMM follows sound principles.
   N/A 1 2 3 4 5
8. Color is appropriately used in this IMM. N/A 1 2 3 4 5
9. The screen displays are easy to understand. N/A 1 2 3 4 5

Area 3: Program Functionality Review

10. This IMM operated flawlessly. N/A 1 2 3 4 5

Focus Group Protocol

- Focus groups are a powerful means of collecting data about learner or instructor reactions to a new interactive multimedia program. However, focus groups need to be carefully planned so that you get the kind and quality of information you are seeking. This "Focus Group Protocol" is a brief example of a list of questions that might be addressed during a focus group regarding an interactive multimedia program.

Focus Group Protocol

Focus groups are a powerful means of collecting data about learner or instructor reactions to a new interactive multimedia program. However, focus groups need to be carefully planned so that you get the kind and quality of information you are seeking. This "Focus Group Protocol" is a brief example of a list of questions that might be addressed during a focus group regarding an interactive multimedia program.

Instructions:

1. The tool below is merely a template. You should modify it as needed for your distinct purposes.
2. Using an focus group is a type of survey activity. Conducting a survey should be done systematically. The overall steps in the survey process are:

a. Organize the survey team.
b. Determine the survey goal.
c. Select a representative sample.
d. Generate the questions.
e. Construct the instrument (questionnaire, interview protocol, or
   focus group protocol).
f. Test the instrument.
g. Administer the instrument.
h. Analyze the data.
i. Share and use the results.

3. Consider collecting evaluation data with more than one method if time and resources allow. For example, a questionnaire can be used to collect information about global reactions to an interactive multimedia program. Then, either interviews or focus groups can be used to collect more detailed information. Alternatively, interviews or focus groups might be used to identify the most important evaluation issues that will be included in a questionnaire sent to many people.
4. A sample focus group protocol begins on the next page.

Evaluation Focus Group Protocol

of participants: _____Host: _____

Date: _____ Site: _____

1. What is your opinion of the interactive multimedia system used to deliver this information?
2. Was the interactive multimedia system available at times and places convenient to you?
3. What is the word on the shop floor about this interactive multimedia system?
4. What could be done to improve the interactive multimedia system?
5. What other types of information or training should be available via interactive multimedia?

Formative Review Log

- The "Formative Review Log" is a simple instrument that can be used by anyone you have asked to review your program in its formative stages. The instrument has three columns, the first for recording the screen or format sheet number that the person is reviewing, the second for writing down observations (e.g., errors, confusing points, or ideas), and the third for recording what actions have been taken in reaction to the feedback provided by members of the project team. Using an instrument like this with many different types of users will probably have the greatest pay-off for formative evaluation throughout the life of the project.

Formative Review Log

The "Formative Review Log" is a simple instrument that can be used by anyone you have asked to review your program in its formative stages. The instrument has three columns, the first for recording the screen or format sheet number that the person is reviewing, the second for writing down observations (e.g., errors, confusing points, or ideas), and the third for recording what actions have been taken in reaction to the feedback provided by members of the project team. Using an instrument like this with many different types of users will probably have the greatest pay-off for formative evaluation throughout the life of the project.

Instructions:

1. In addition to space to record who is doing the evaluation, when, etc., there are three main columns that should be in the Formative Review Log:

- A column for indicating which part of the program is being review (e.g., a screen number, format sheet number, script version, etc.);
- A column for recording the reviewer's reactions, questions, errors, etc., and
- A column for recording what was done as a consequence of the feedback provided by the user.

2. Keep copies of formative reviews that have been done in the project notebook or diary.
3. As an interactive multimedia program nears completion, it is sometimes useful to watch a user and fill out the log for the user so that the user can concentrate on interacting with the program.
4. A sample "Formative Review Log" appears on the next page.

Formative Review Log

_____ _____ _____

(IMM Module) (Reviewer) (Date)

SCREEN	COMMENTS & SUGGESTIONS	ACTIONS TAKEN

Implementation Log

- It is one thing to plan and develop a good interactive multimedia program. It is entirely another thing to implement it as planned. Many training innovations have failed because implementation factors (such as instructor motivation) were not considered. It is essential to make every effort to collect information regarding the actual use of an interactive multimedia program as compared to the planned use. The "Implementation Log" tool has been designed to make that comparison a little more systematic.

Implementation Log

It is one thing to plan and develop a good interactive multimedia program. It is entirely another thing to implement it as planned. Many training innovations have failed because implementation factors (such as instructor motivation) were not considered. It is essential to make every effort to collect information regarding the actual use of an interactive multimedia program as compared to the planned use. The "Implementation Log" tool has been designed to make that comparison a little more systematic.

Instructions:

1. In addition to space for indicating what program is being implementing and other types of site-specific information, an implementation log should include the following:

- a column for describing what is planned to happen during the implementation,
- a column for describing what actually happened,
- a column for commenting on the differences between planned and actual activities (if any), and
- spaces for additional questions that might be important in the context.

2. A sample implementation log appears on the next page.

Implementation Log

DATE: 10/2-3/94 **PLACE:** Chicago **TIME:** 9:00 - 5:00

TRAINER: Larry R Jones **NUMBER OF TRAINEES:** 15

Time	Recommended Activities	Actual Activities	Comments
09:00-09:15	Introductions of participants and review of agenda. (Leader-led)	No changes.	
09:15-09:30	Overview of "New Course." (Leader-led)	No changes.	
09:30-12:00	Trainees begin Module 1. (multimedia systems)	One system failed to function because someone removed system files.	I need to check all systems personally before course begins.
12:00-01:00	Lunch Break	Two trainees chose toskip lunch and keep working.	
01:00-05:00	Trainees continue working through the modules.	No changes.	Frequently had to refer trainees to help routine.

1. What training activities would you like to modify the next time you conduct this course?

   I will personally check each one 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.

Interview Protocol

- Interviews are a powerful means of collecting data about learner or instructor reactions to a new interactive multimedia program. However, interviews need to be carefully planned so that you get the kind and quality of information you are seeking. This "Interview Protocol" is a brief example of a list of questions that might be addressed during an interview regarding an interactive multimedia program.

Interview Protocol

Interviews are a powerful means of collecting data about learner or instructor reactions to a new interactive multimedia program. However, interviews need to be carefully planned so that you get the kind and quality of information you are seeking. This "Interview Protocol" is a brief example of a list of questions that might be addressed during an interview regarding an interactive multimedia program.

Instructions:

1. The tool below is merely a template. You should modify it as needed for your distinct purposes.
2. Using an interview protocol is a type of survey activity. Conducting a survey should be done systematically. The overall steps in the survey process are:

a. Organize the survey team.
b. Determine the survey goal.
c. Select a representative sample.
d. Generate the questions.
e. Construct the instrument (questionnaire, interview protocol, or focus group protocol).
f. Test the instrument.
g. Administer the instrument.
h. Analyze the data.
i. Share and use the results.

3. Consider collecting evaluation data with more than one method if time and resources allow. For example, a questionnaire can be used to collect information about global reactions to an interactive multimedia program. Then, either interviews or focus groups can be used to collect more detailed information. Alternatively, interviews or focus groups might be used to identify the most important evaluation issues that will be included in a questionnaire sent to many people.
4. A sample interview protocol begins on the next page.

Evaluation Interview Protocol

Name: _____ Interviewer: _____
Date: _____

1. What is your specialty?
2. How many years and months in present position?
   _____years _____months
3. How many years experience with this company? _____years _____months
4. Please describe your use of the XYZ since the "XYZ Training" IMM course?
5. Please describe your first reactions to "XYZ" IMM course.

6. Please describe your present opinions of "XYZ" IMM course.
7. Do you need additional training for the "XYZ?"
8. To what degree did you accomplish the performance objectives established for the "XYZ" IMM course?
9. What would you tell another person about to take the "XYZ" IMM course for the first time?
10. What kinds of successes have you experienced with the "XYZ" since the training?
11. What kinds of problems have you experienced with the "XYZ" since the training?
12. Please describe the areas in which you feel most competent concerning use of the "XYZ."
13. Please describe the areas in which you feel least competent concerning use of the "XYZ."
14. What improvements would you recommend for the "XYZ Training" IMM course overall?
15. What improvements would you recommend for the "XYZ" IMM course training manual?
16. What is your opinion of the interactive multimedia system

Questionnaire

- Questionnaires are undoubtedly the single most frequently used type of evaluation instrument. Poorly designed questionnaires are often administered at the close of a course or training session as a "smilometer" or "happiness indicator." They are also often distributed to users of interactive multimedia programs. If the only thing you find out about your interactive multimedia program with a questionnaire is whether the trainees liked it, you are not making good use of this strategy. As shown in the "Questionnaire," a wealth of information can be provided by a well-designed instrument.

Questionnaire

Questionnaires are undoubtedly the single most frequently used type of evaluation instrument. Poorly designed questionnaires are often administered at the close of a course or training session as a "smilometer" or "happiness indicator." They are also often distributed to users of interactive multimedia programs. If the only thing you find out about your interactive multimedia program with a questionnaire is whether the trainees liked it, you are not making good use of this strategy. As shown in the "Questionnaire," a wealth of information can be provided by a well-designed instrument.

Instructions:

1. Questionnaires can take many forms, e.g., checklists, rating scales, multiple-choice questions, open-ended questions, and so forth. Most instruments include a combination of several different types of questions or items. You should design an evaluation questionnaire very carefully so that you get the information you need without requiring the persons completing the questionnaire to spend too much of their time.
2. A sample evaluation questionnaire appears on the next page.

Evaluation Questionnaire

Course Name: _____ Training Location: _____
Participant Name (optional): _____ Date: _____

Job Title: _____Years in present position? <1 1-3 3-5 5+

Instructions

Please circle your response to the items. Rate aspects of the course on a 1 to 5 scale 1 equals "strongly disagree" and 5 equals "strongly agree." 1 represents the lowest and most negative impression on the scale, 3 represents an adequate impression, and 5 represents the highest and most positive impression. Choose N/A if the item is not appropriate or not applicable to this course. Your feedback is sincerely appreciated. Thank you.

Course Content (Circle your response to each item.)
NA=Not applicable 1=Strongly disagree 2=Disagree 3=Neither agree/nor disagree 4=Agree 5=Strongly agree

1. I was aware of the prerequisites for this course.
   N/A 1 2 3 4 5
2. I had the prerequisite knowledge and skills
   N/A 1 2 3 4 5 for this course.
3. I was well informed about the objectives
   N/A 1 2 3 4 5 of this course.
4. This course lived up to my expectations.
   N/A 1 2 3 4 5
5. The content is relevant to my job.
   N/A 1 2 3 4 5

Course Design (Circle your response to each item.)

6. The course objectives are clear to me.
   N/A 1 2 3 4 5
7. The course activities stimulated my learning.
   N/A 1 2 3 4 5
8. Interactive multimedia was essential in the course.
   N/A 1 2 3 4 5
9. The activities in this course gave me sufficient
   N/A 1 2 3 4 5 practice and feedback.
10. The test(s) in this course were accurate and fair.
    N/A 1 2 3 4 5
11. The difficulty level of this course is appropriate.
    N/A 1 2 3 4 5
12. The pace of this course is appropriate.
    N/A 1 2 3 4 5

NA=Not applicable 1=Strongly disagree 2=Disagree 3=Neither agree/nor disagree 4=Agree 5=Strongly agree

Course Instructor (Facilitator) (Circle your response to each item.)

13. The instructor was well prepared.
    N/A 1 2 3 4 5
14. The instructor was helpful.
    N/A 1 2 3 4 5

Course Environment (Circle your response to each item.)

15. The training facility at this site was comfortable.

N/A 1 2 3 4 5

16. The training facility at this site provided

N/A 1 2 3 4 5 everything I needed to learn.

Course Results (Circle your response to each item.)

17. I accomplished the objectives of this course.

N/A 1 2 3 4 5

18. I will be able to use what I learned in this course.

N/A 1 2 3 4 5

Self-paced Delivery (Circle your response to each item.)

19. IMM was a good way for me to learn this content.

N/A 1 2 3 4 5

20. Video is an important aspect of the course.

N/A 1 2 3 4 5

21. How would you improve this course?

(Check all that apply.)

___Provide better information before course. ___Clarify the course objectives.

___Reduce content covered in course. ___Increase content covered in course.

___Update content covered in course. ___Improve the instructional methods.

___Make course activities more stimulating.___Improve course organization.

___Make the course less difficult. ___Make the course more difficult.

___Slow down the pace of the course. ___Speed up the pace of the course.

___Allot more time for the course. ___Shorten the time for the course.

___Improve the tests used in the course. ___Add more video to the course.

22. What other improvements would you recommend in this course?

23. What is least valuable about this course?

24. What is most valuable about this course?

User Interface Rating Form

- The "User Interface" of an interactive instructional product, e.g, a multimedia program, is a critical element of the product that must be carefully evaluated. If the user interface is not well-designed, learners will have little opportunity to learn from the program. This rating form includes ten major criteria for assessing the user interface for an interactive program, such as "ease of use" and "screen design." Not all of the criteria may be relevant to the particular program you are evaluating, but most of them will. You may need to add additional criteria to the list. Novice users of interactive instructional products are generally not good candidates for using this form. The people rating the user interface should be experienced users of the type of program you are asking them to rate. Even better, they could be experienced designers of interactive programs.

User Interface Rating Form

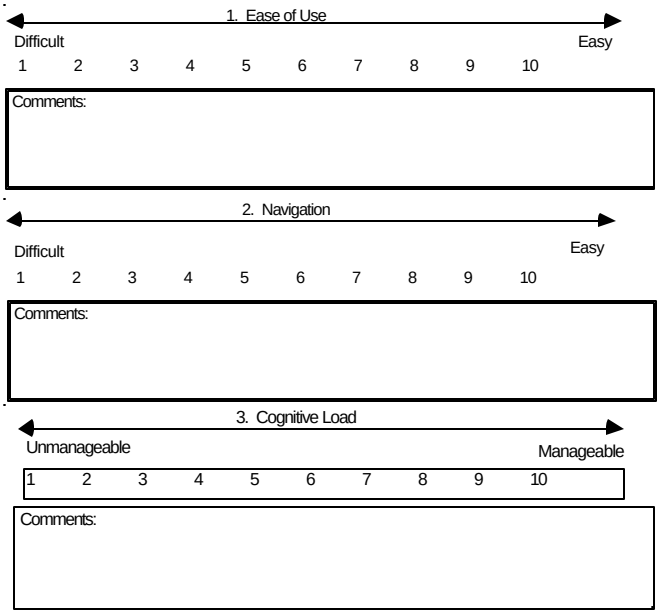
The "User Interface" of an interactive instructional product, e.g, a multimedia program, is a critical element of the product that must be carefully evaluated. If the user interface is not well-designed, learners will have little opportunity to learn from the program. This rating form includes ten major criteria for assessing the user interface for an interactive program, such as "ease of use" and "screen design." Not all of the criteria may be relevant to the particular program you are evaluating, but most of them will. You may need to add additional criteria to the list. Novice users of interactive instructional products are generally not good candidates for using this form. The people rating the user interface should be experienced users of the type of program you are asking them to rate. Even better, they could be experienced designers of interactive programs.

Instructions:

1. The "User Interface Rating Form" should be used by experienced interactive multimedia designers or users to rate the interface of a new program or one under development.
2. The ten criteria used in the "User Interface Rating Form" are explained in detail at the end of the instrument itself.
3. Some criteria may not be relevant in every IMM program.

User Interface Rating Tool for Interactive Multimedia

Instructions: For each of ten user interface dimensions illustrated below, rate the program you have reviewed on a one to ten scale by circling the appropriate number under the dimension. (Accompanying this tool are definitions for each of the ten user interface dimensions.) Please add any comments that may help to clarify or explain your rating. If a specific dimension does not seem appropriate to the interactive program you are reviewing, do not circle any numbers on the scale for that dimension and add a brief comment to explain your response.



4. Mapping

None — Powerful

1 2 3 4 5 6 7 8 9 10

Comments:

5. Screen Design

Violates Principles — Follows Principles

1 2 3 4 5 6 7 8 9 10

Comments:

6. Knowledge Space Compatibility

Incompatible — Compatible

1 2 3 4 5 6 7 8 9 10

Comments:

7. Information Presentation

Obtuse — Clear

1 2 3 4 5 6 7 8 9 10

Comments:

8. Media Integration

Uncoordinated — Coordinated

1 2 3 4 5 6 7 8 9 10

Comments:

9. Aesthetics

Displeasing — Pleasing

Comments:

10. Overall Functionality

Dysfunctional — Functional

1 2 3 4 5 6 7 8 9 10

Comments:

Please add other comments related to the user interface of this program below:

Definitions for User Interface Rating Tool

User Interface Dimension 1 – Ease of Use

"Ease of Use" is concerned with the perceived facility with which a user interacts with an interactive multimedia program. Figure 1 illustrates a dimension of such a program ranging from the perception that the program is very difficult to use to one that is perceived as being very easy to use. Like many of the dimensions described in this tool, ease of use is both an aggregate and individual dimension. For example, in the aggregate sense, the Windows interface is generally perceived as easier to use than the command interface of the Microsoft disk operating system (MS DOS). However, in the individual sense, some people may perceive the MS-DOS interface to be easier to use because of their own unique experiences and attributes.

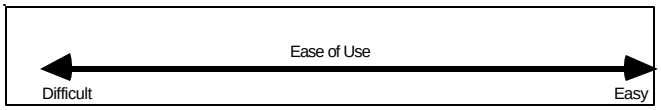


Figure 1. "Ease of Use" dimension of user interface.

User interface dimensions may be highly correlated with how well users enjoy using a specific program. Whether users like a program may be more or less important, depending on the intent of the program and the context for its use. Certainly, not liking an interactive program that is intended to be highly motivating is a major problem, whereas users' affect for a program may be less important in a training context in which strong extrinsic motivational factors exist. Nonetheless, in the long run, improving the user interface dimensions of multimedia, such as "ease of use," is a highly desirable goal, regardless of context.

User Interface Dimension 2 – Navigation

"Navigation" is concerned with the perceived ability to move through the contents of an interactive program in an intentional manner. Figure 2 illustrates a dimension of interactive multimedia ranging from the perception that a program is difficult to navigate to one that is perceived as being easy to navigate. An important aspect of navigation is orientation, i.e., the degree to which a user feels that he/she knows where he/she is in a program and how to go to another part of it. This is a critical variable because users frequently complain of being lost in a interactive program (Utting & Yankelovitch, 1989). Designers use several ways of supporting navigation and maintaining orientation. A popular approach to navigation is the WIMP (window-icons-mouse-pointing) interface.

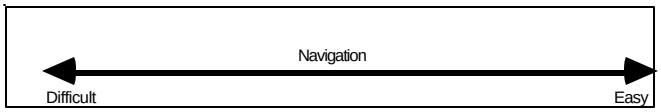


Figure 2. "Navigation" dimension of user interface.

User Interface Dimension 3 – Cognitive Load

Using an interactive program requires different mental efforts than performing tasks via print or other media. In order to make any meaningful response to an interactive program, users must cope with and integrate at least three cognitive loads or demands, i.e., (a) the content of the program, (b) its structure, and (c) the response options available. To use interactive

programs, users must perceive options, conceptualize a choice, and make some physical action, all while mentally coordinating the demands of these three cognitive loads. The user interface is the vehicle that allows perceptual, conceptual, and physical contacts with the interactive program. In terms of "cognitive load," the user interface can seem unmanageable (i.e., confusing) at one end of the continuum and easily manageable (i.e., intuitive) at the other end (see Figure 3).

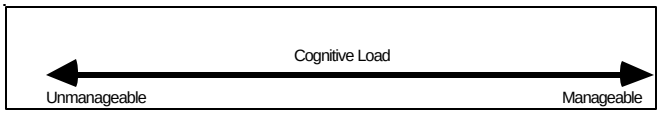


Figure 3. "Cognitive Load" dimension of user interface.

Learners acquire and structure information delivered via interfaces, conduct mental operations, and accomplish physical activities during their interactions with interactive multimedia. The limited capacity of working memory to hold only five to nine chunks of information simultaneously makes it difficult for users of complexity structured programs to reason when numerous cognitive load factors must be handled simultaneously. Users may feel overwhelmed by numerous options that increase the cognitive load. The risks of confusion are especially high when users confront programs which by their very nature include many interactive options. The possibility of user disorientation is a major concern in the increasingly popular multimedia programs that feature a complex, flexible structure.

User Interface Dimension 4 – Mapping

"Mapping" refers to the program's ability to track and graphically represent to the user his or her path through the program. In complex, non-linear programs, user-disorientation can be alleviated if users can see what parts of the system they have already accessed. Utting and Yankelovitch (1989) discuss user disorientation as referring to, among other things, the user's not knowing "the boundaries of the information space." Having a detailed mapping system gives users an aid in understanding which parts and how much of the information space they have interacted with, and conversely which parts and how much of it they haven't. Interactive programs fall in a continuum of containing no mapping function to an appropriately powerful mapping function (see Figure 4).

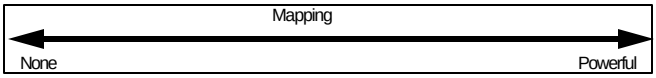


Figure 4. "Mapping" dimension of user interface.

The notion of an "appropriately powerful" mapping function requires some explanation. Just as it is important to possess a map of the most usable scale when taking a road trip, it is important for interactive programs to provide enough, but not too much, detail in showing user paths. A map that shows every piece of a program's knowledge space might prove to be so tedious or unwieldy as to be of as little value as an interactive program with no map.

User Interface Dimension 5 – Screen Design

"Screen Design" is a particularly complex dimension of interactive programs that can easily be broken down into many sub-dimensions related to text, icons, graphics, color, and other visual aspects of interactive programs. Shneiderman (1987) maintains that although certain design principles have been established, "screen design will always have elements of art and require invention" (p. 326). A separate dimension has been defined to deal with the artistic aspects of interactive programs (see Dimension 9 - Aesthetics below). We define "screen design" as a dimension ranging from substantial violations of principles of screen design to general adherence to principles of screen design (see Figure 5).

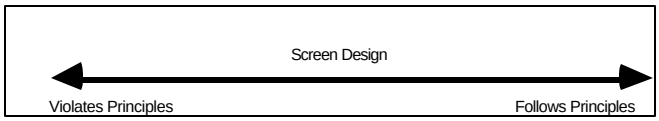


Figure 5. "Screen Design" dimension of user interface.

There are two problems with this dimension. First, screen design principles have not kept up with the rapidly changing nature of interactive technology. Second, creative designers may sometimes intentionally violate screen design principles for effect or to otherwise focus the user's attention. Nonetheless, we think that there exists enough knowledge about the principles of screen design that people, particularly experienced designers, can make meaningful distinctions among poorly and well designed screens in interactive programs.

User Interface Dimension 6 – Knowledge Space Compatibility

"Knowledge space" refers to the network of concepts and relationships that compose the mental schema a user possesses about a given phenomena, topic or process. Subject matter experts and/or designers of interactive programs are generally perceived as possessing an expert knowledge space with respect to the content included in the programs they create. This expertise usually is the basis for the structure of the knowledge or information presented in a program. Novice users, on the other hand, often possess an inadequate knowledge space with respect to the content of a program. The knowledge space of novices may be inadequate because of ignorance, misconceptions, or some blending of ignorance and misconceptions. When a novice user initiates a search for information in an interactive program, the interface should be powerful enough so that the user perceives the resulting information as compatible with his or her current knowledge space (see Figure 6). If the information received is not perceived as relevant to the search strategies used by the user, the system will be perceived as incompatible.

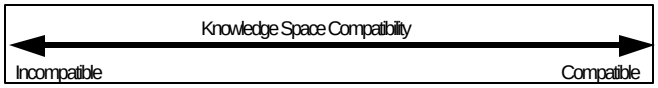


Figure 6. "Knowledge Space Compatibility" dimension of user interface.

Admittedly, this is a very difficult dimension to judge. However, if a user initiates a search for information about a topic, e.g., the procedures for installing new software, the resulting information should seem compatible with that search once the information is thoroughly explored. If the information seems arbitrary or irrelevant to the search that was initiated, the knowledge space representation should be judged as incompatible.

User Interface Dimension 7 – Information Presentation

The "Information Presentation" dimension is concerned with whether the information contained in the knowledge space of an interactive program is presented in an understandable form. The most elegantly designed user interface for an interactive program is useless if the information it is intended to present is incomprehensible to the user. Certainly the user might be able to find all of the information about a subject, but whether the user could then comprehend, understand, or learn that information is another matter. Imagine a complicated installation procedure presented in textual form, written in a stream of consciousness style reminiscent of James Joyce's Ulysses. Or consider a video presentation on sales techniques for ATMs, directed and produced by Andy Warhol. In each case the information requisite for understanding may be present, but would probably be difficult if not impossible to comprehend. Information presentation is defined as a dimension ranging from obtuse to clear (see Figure 7).

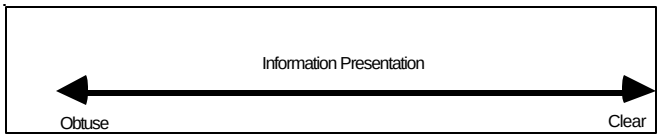


Figure 7. "Information Presentation" dimension of user interface.

User Interface Dimension 8 – Media Integration

The most important aspect of the media integration dimension refers to how well an interactive program combines different media to produce an effective whole. Do the various media (text, graphics, audio, video, etc.) work together to form one cohesive program, or is the program a hodgepodge of gratuitous media segments? Are the various media components necessary to the function of the program or would the program function equally as well without them? The media integration dimension is defined as ranging from uncoordinated to coordinated (see Figure 8).

Figure 8. "Media Integration" dimension of user interface.

User Interface Dimension 9 – Aesthetics

"Aesthetics" refers to the artistic aspects of interactive programs in the sense of possessing beauty or elegance. In the aggregate sense, many people may praise the aesthetics of an automobile design or the elegance of a bridal gown. However, in an individual sense, aesthetics are highly unique and one person's sense of the beautiful may seem grotesque to another. Eisner (1985) described the need to develop "connoisseurs" in evaluation of education and training, just as we have connoisseurs in the arts. Connoisseurs have refined tastes and a deep sensitivity to aesthetics that enable them to criticize phenomena (e.g., plays, paintings, musical scores, or interactive programs) in a manner that can be communicated to others. In turn, other people, perhaps less refined and less sensitive, may become more informed by "consuming" the expert reviews provided by the connoisseurs. In the absence of such connoisseurs, the aesthetics dimension of the user interface of an interactive multimedia program is defined as ranging from displeasing to pleasing (see Figure 9).

Figure 9. "Aesthetics" dimension of user interface.

User Interface Dimension 10 – Overall Functionality

"Overall Functionality" is an aspect of interactive multimedia programs related to the perceived utility of the program. The perceived functionality of an interactive program is obviously highly related to the intended use of the program. A given program may have multiple uses. Its overall functionality must be judged in relation to the specific intended use that exists in the mind of the users. Figure 10 illustrates a dimension of the user interface of interactive programs that ranges from dysfunctional to highly functional.

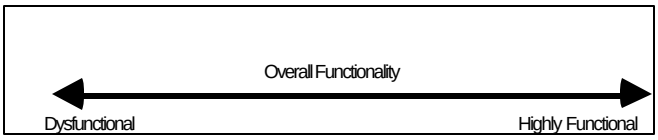


Figure 10. "Overall Functionality" dimension of user interface.

Evaluation Report Sample

- The "Evaluation Report Sample" presents one way of structuring an evaluation report. Evaluation reports are notorious for being weighty volumes that few people read. Not surprisingly, lengthy reports have little effect on decision-makers. This tool illustrates a strategy for dividing an evaluation report into two-page sections that each include four parts:

1. an attention-getting headline,
2. a description of the major issues related to the headline,
3. a presentation of data related to the issues, and
4. a bottom-line recommendation or summary of the findings.

- People who receive a report in this format can take two or three sections at a time and make them agenda items for their team meetings. In this way, the evaluation findings are much more likely to have an impact on practical decisions

Evaluation Report Sample

The "Evaluation Report Sample" presents one way of structuring an evaluation report. Evaluation reports are notorious for being weighty volumes that few people read. Not surprisingly, lengthy reports have little effect on decision-makers. This tool illustrates a strategy for dividing an evaluation report into two-page sections that each include four parts:

1. an attention-getting headline,
2. a description of the major issues related to the headline,
3. a presentation of data related to the issues, and
4. a bottom-line recommendation or summary of the findings.

People who receive a report in this format can take two or three sections at a time and make them agenda items for their team meetings. In this way, the evaluation findings are much more likely to have an impact on practical decisions.

Instructions:

1. Evaluators often write long reports of evaluations that few people bother to read. Needless to say, these lengthy reports are unlikely to have much impact on decision-making if people don't even read them. One way of increasing the likelihood that your evaluation report will be read and acted upon is to report it in easy-to-consume "chunks" of information.
2. One strategy of "chunking" an evaluation report is to have no section of your evaluation report longer than two pages. An illustration of this format is presented below. The example is divided into four parts: 1) an attention-getting headline, 2) a description of the issues involved in that part of the evaluation, 3) evidence that relates to the issue(s), and 4) a bottom line recommendation based upon the evidence.

Evaluation Report Sample

Although students value the self-paced nature of the L2 Algebra I multimedia program, it is not always implemented.

• Issues

Self-paced learning is an academic ideal held forth by many, but few curricula have implemented this principle on a large scale. The L2 Algebra I course incorporates genuine self-paced learning, but it is not implemented in each school. Hence, student satisfaction with the self-paced nature of the L2 Algebra I course varies considerably. Overall, almost 50% of the 1,400 L2 students surveyed agreed with the statement that "The pace of this course is just right." However, in those schools where teacher-paced guidelines have been most strictly enforced, as many as 90% disagree with the statement! Overall student satisfaction with the L2 Algebra I course is much lower in those schools that have abandoned the principle of self-paced learning. Reasons for instituting teacher-pacing may stem from factors outside the teachers' control (e.g., the inability to issue incomplete grades or no provisions for summer school) to differences in individual teacher styles (an inability to tolerate ambiguity or a need for more control). Regardless of the source, teacher-pacing seems to be a major violation of the L2 program's design.

• Evidence

Example comments of students in schools where self-pacing is implemented include:

Most of Gary's comments centered around how much he liked the self-paced nature of the instruction. He told me this: "If you don't understand something, the computer will take you back, if you had a teacher and you didn't understand, you would just have to go on."

Sheri likes L2 primarily because she enjoys the self-paced instruction. She also prefers working with a computer to listening to a teacher because she hates "teachers that sit there and talk to you for the whole period."

Aretha failed algebra last year in a traditional book-based, teacher-led class. She appreciates the opportunity that L2 has given her to be independent. "I'm on my own. Here you go, teach yourself. No excuses."

.....he [Don] likes it [L2] even more than he thought he would. Primarily, he likes the self-paced method of instruction. He told me "with a teacher, if you don't understand and the rest of the class does, they're going to go on without you."

He [Jeffrey] would recommend the class to his friends because of the self-paced, individualized nature of L2. "You go at your own pace and the computer pertains to you better than a teacher," he said.

Charles really likes the self-paced learning in L2. "I can't think of a better way to learn," he said. He says that in a book class, the teacher goes too fast. He says that with L2, "you work your butt off" but that in the book class "everyone is failing."

Example comments of students in schools where teacher-pacing exists include:

Her initial reaction to L2 was that it was "a lot better than text-based classes! It's self paced!" Her favorite aspect of L2 is that it shows you what to do step by step. She especially disliked the fact that after a short period of self paced use, the rules were changed so that L2 time deadlines were placed on students by the teacher. "We must be to chapter 8 by the fourth six weeks!"

The worst thing about L2 is that now they have a teacher imposed pace that if not equaled will result in INCOMPLETES at the end of 6 week grading periods. He [Edward] has already missed one honor roll because of this policy and was peeved.

Farrah initially liked L2, "...it seemed easier than the teacher. It was self paced, and had individualized instructions. It was self paced!" She expressed definite resentment about the current policy of teacher set pacing. When asked for her opinion of L2 now Farrah says, "I don't like it. Really - there are only so many days to complete X number of topics. If you don't finish them - you get an incomplete. I have some incompletes. A lot of people do. What are they (the administration) going to do at the end of the year?"

She [Angela] also expressed her feeling that they had been misled regarding the self paced concept. "Now we have to be at certain points at the end of each grading period. I had to come after school to keep up."

Matt initially was very excited about the L2 program. The self paced approach really appealed to him. By the time the L2 program was discontinued seven weeks ago his attitude toward the program had drastically changed. Matt's comments reflect this; "When we stopped L2 I disliked the program– a lot. I especially disagree with the time requirements and pass/fail grading policy. You should be given credit for increments in progress. If you haven't finished a chapter by the due date you get an 'F' even if you were about 80% finished. This is not right!"

He [Paul] likes the self-paced aspects of the program, and he said he strongly resents the pressure that the teacher puts on him to make more progress. He says that when he rushes through a section, he "really doesn't learn anything."

• Recommendation

The National Science Center should work with school administrators, teachers, and parents to assure that the principle of self-paced learning can be implemented within the total educational environment of the local school system. In those schools where this is not possible, the L2 Algebra I course may be reduced to a supplementary role or eliminated entirely.