



CBI PILOT EVALUATIONS: A PROTOTYPE STUDY

Technical Report No. 19

November 1983

TECHNICAL REPORT

**DIGITAL EQUIPMENT CORPORATION  
EDUCATIONAL SERVICES  
30 North Avenue  
Burlington, Massachusetts 01803**

**digital**

This Technical Report may be copied  
for non-commercial purposes with  
credit to the authors and Digital  
Equipment Corporation.

## CBI PILOT EVALUATIONS: A PROTOTYPE STUDY

Constance J. Seidner, Ph.D.

### ABSTRACT

The purpose of this report is to describe evaluation strategies that can be used in the pilot evaluation of CBI courses. The effectiveness, efficiency, and appeal of computer-based instructional materials are different, although related, elements that contribute to a good CBI course. Evaluators should be clear about the distinction between these elements and choose evaluation instruments and strategies that provide some indication of a course's performance on each criterion. Procedures in this report are based on a prototype study that was part of an ongoing effort to develop a standard set of guidelines for CBI pilot evaluations. The intent is to provide the novice evaluator with suggestions regarding the planning and administration of a small sample pilot study. Non-statistical techniques for discerning patterns in the data are suggested.

## RATIONALE FOR PILOT EVALUATIONS

Pilot evaluations are conducted to see how well a course is working before the course is released in its final form. The term "pilot evaluation", as it is used here, is similar to what is called a "formative evaluation" in educational evaluation research.

Ideally, you would like to do formative evaluations on parts of the course, for example a lesson or module, so that you can check out certain design strategies to see how well they work before you implement them in the entire course. At the very least, you should do a pilot evaluation on the course in its final or near final form.

The purpose of this report is to outline some suggestions for conducting pilot evaluations of CBI courses that are in the final stages of completion. The report is divided into three sections:

- Planning the Evaluation
- Evaluation Procedures
- Data Analysis

## PLANNING THE EVALUATION

There are at least three things that should be addressed in the planning stages:

- What do you want to assess in the pilot evaluation?
- How will you collect the data?
- Who will your subjects be?

### What to Assess

It is important that you be clear about the purposes of a pilot evaluation. There are different elements that contribute to a successful course. Among the quality dimensions that may be addressed are those suggested by Reigeluth et al. [1]:

- Effectiveness
- Efficiency
- Appeal

---

[1] Reigeluth, C.M., M.D. Merrill, B.G. Wilson, and R.T. Spiller. The elaboration theory of instruction: a model for sequencing and synthesizing instruction. Instructional Science 9:195-219, 1980.

While these three elements may be related, they are not precisely the same thing.

- Effectiveness refers to the student's achievement of course objectives. A course is effective if students learn what they are supposed to learn.
- Efficiency refers to the process by which students learn. An optimal process is efficient in terms of student time and effort. Courses that are well designed make it easy for students to learn. This implies that the course makes optimal use of techniques that are known to facilitate learning.
- Appeal refers to students' affective reactions to the course. Students' feelings about a course may be influenced by factors other than how much they learn in the course. For example, students tend to feel good about a course that is aesthetically appealing. Actual physical comfort or discomfort associated with the learning experience will also affect how students feel after they have completed a course. For example, if the course is so long that it caused undue tiredness or eye strain, students will not leave the course with a "warm and fuzzy" feeling. An appropriate level of dialogue between the program and the student will also increase a course's appeal. For example, feedback messages should not be intimidating or condescending. Instead, they should be designed to make students feel good about their accomplishments.

It is important that course evaluators keep these distinctions in mind so that they are clear about what is being measured in an evaluation. You would expect that the relationship between effectiveness, efficiency and appeal would be fairly strong. Students will tend to feel good about a course that teaches effectively and efficiently. And those aspects of appeal that relate to ease of use should contribute to the effectiveness of the course. For example, if students are not overly tired, and/or find screens attractive, they will be more likely to attend to the instruction that is being presented.

The relationship is not a perfect one, however. Courses can be effective in terms of student learning, and still lack appeal. Conversely, students may find the features of a course very appealing, even if the course does not teach effectively. A course can be effective and yet not equally efficient for all students because differences in student characteristics have not been provided for in the course design.

These three criteria -- effectiveness, efficiency, and appeal -- may be differentially weighted, depending on institutional restraints and the overall goals of a particular course. For

example, if a course is going to be used for some type of certification, course effectiveness will be of major importance. On the other hand, appeal can be very important if customer attitudes weigh heavily in the balance. The amount of development time required to produce courses that are responsive to individual students' characteristics may not be available, so compromises have to be made.

Why all of this concern about multiple criteria? Because it is important that an evaluator know what the relative importance of these three criteria are for a given course, so that appropriate evaluation strategies will be used. It is also a good idea if the several groups of people who make judgements about a course are in consensus about the relative importance of the criteria. Being specific about these criteria at the outset will increase the likelihood of this happening.

#### Data Collection Alternatives

Course Effectiveness - Course effectiveness is best evaluated by the administration of criterion measures that are matched to course objectives. For example, there may be an end-of-course test that is designed to measure course objectives.

Sometimes it is impossible or inappropriate to include criterion measures in the course. When this is the case, criterion measures should be constructed separately and administered during the pilot evaluation and field test. For example, it is often impossible to go from the course software into the operating system or applications software under program control (while still within the CBI environment) so that students can perform a software task in a "real" setting. In this case, tasks that correspond to course objectives should be constructed and administered separately after students complete the course.

The best measures of course effectiveness then, are direct measures of performance like pre- and posttests, and task performance. You may get indirect measures of course effectiveness by students' responses to questionnaires, rating scales and interviews, or by observing their performance in the course.

A sample questionnaire is included in this report as Appendix A [2]. A rating scale that was used in prototype evaluations is included as Appendix B. It is a confidence rating that asks students how confident they feel about their ability to perform a particular task or tasks (the criterion tasks). It should be remembered these these types of instruments can only get at what

-----  
[2] The Quality Assurance Group in Educational Services has now developed a Student Opinion Form for use in CBI pilot evaluations.

students think they have learned, and as such are not direct measures of learning.

Course Efficiency - Only the internal efficiency of a course can be assessed in a formative pilot evaluation. That is, you can only assess how efficiently this particular course teaches. You cannot compare the efficiency of this course to another course or another medium of instruction.

Internal efficiency of this sort can only be assessed in a general way. A person who is knowledgeable about instructional design can watch a student go through a course, noting areas of difficulty as well as sections of the course where students seem to progress with ease. From this type of observational data, inferences can be made about how effectively a course presents material and how readily students appear to learn the material. Students' appraisals of course efficiency also can be assessed indirectly through questionnaire and interview schedules.

Course Appeal - Students' spontaneous comments as they are going through a course are one indicator of the course's appeal, as are their responses to questionnaire and interview items.

### Selection of Subjects

A pilot evaluation, particularly if it involves observation of students taking a course, is a time-consuming process. Typically, only a few subjects are used for pilot evaluations -- perhaps 3 to 5 people. The people who participate in the pilot study should resemble the target audience as much as possible. If your target audience varies widely in terms of background or experience, your subjects should also vary on those characteristics that will impact their performance in the course.

## EVALUATION PROCEDURES

There is no one best procedure for doing a pilot evaluation. What you do will affect how you do it. The following set of guidelines should be interpreted with this in mind. It is important that you are clear about what you intend to do and why. Then be as consistent as possible in what you do with each subject.

### Setting

The setting for the pilot should be comfortable for the subject. You will probably want to choose a place that provides some privacy but does not give the impression of an antiseptic experimental setting.

Equipment for the subject will include such items as:

- Course materials (diskettes and whatever supporting materials are delivered with course, such as a Student Guide)
- Documentation (if appropriate)
- Paper and pencil for subject (if appropriate)

Equipment for the evaluator will include:

- Subject list
- Observation Form
- Questionnaire and/or Interview Schedule
- Pre- and Posttests (if used)
- Performance Task (if used)
- Copies of Confidence Rating (if used)

### Instructions to Subject

The instructions you give subjects will vary depending on the methods of data collection you use. It is a good idea to develop a rough script of what you are going to say so that you present a consistent stimulus condition to each subject. The following suggests some things you might include in your instructions to the subject.

1. The purpose of the pilot test is to find out how well the course is working.
2. What we want you to do is to take the course, just as you would if you were doing it on your own. However, since this is a pilot test, we want you to tell us about the course as you're going through it:
  - Note typos.
  - Tell us if an instruction isn't clear.



- Tell us anything that you think would make it easier for someone to learn the information in the course.
3. We will:
- Make a note of your observations.
  - Make independent observations of our own on things that we see as you are going through the course. These may have nothing to do with what you are doing, for example, we may make a note of the way a particular screen is designed. So pay no attention to us making notes, but be sure you tell us whatever you want us to know.
4. After you finish the course we would like you to:
- Practice what you have learned by completing the following task: (describe criterion task, if one is to be used).
  - Take a posttest on the material covered in the course (if used).
  - Fill out a questionnaire that asks how you feel about certain aspects of the course.
  - Spend some time talking with us informally about your responses to the questionnaire, and your reactions in general to the course.

Ask if there are any questions about the procedure.

#### Pre-Course Procedures

1. Ask background questions from observation form:
  - Previous experience that relates to the subject matter.
  - Job description.
  - Other appropriate information.
2. Administer pre-course confidence measure (if used).
3. Administer pretest (if used).
4. Point out available documentation, and pencil and paper for taking notes (if appropriate).
5. Give subject the course materials and ask the subject to begin.

### During-Course Procedures

1. Use a systematic method for recording your observations. A sample observation form is included as Appendix C.
2. Take notes on the subject's actions. Include behavior that is indicative of subject's emotional states, e.g., behavior that indicates the subject may be tired or frustrated, but report behavior as objectively as possible.
3. Try to record any verbal comments verbatim.
4. Keep track of the time students begin each major segment (module) of the course.
5. Use identifiers regularly (such as lesson number), so that your observations can be interpreted in context.
6. When noting a particular problem, be as specific as possible -- try to refer to the content of the frame.
7. Be sensitive to the subject's needs; suggest a break when appropriate.
8. Try to set an informal tone by your behavior and comments, establishing a balance a level of formality that lends objectivity to your data and a level of informality that makes the subject feel comfortable and not threatened by the situation.

### After-Course Procedures

1. Administer post-course confidence measure (if used).
2. Administer posttest (if used).
3. Administer post-course task (if used).
  - Explain that the purpose of the task is to give the student a chance to practice what has just been learned and to help us see how well the course teaches.
  - Let the subject read the task; ask if there are any questions related to the task (not how to do it).
  - Tell the subject that he or she can refer to documentation if this is appropriate.
  - Take notes on the subject's task performance.
4. Administer post-task confidence measure (if used).

5. Give the subject the questionnaires to fill out. You may want to leave the room so the subject can complete the form in private. You can also assure anonymity by telling subjects they do not need to put their names on the form and by supplying an envelope into which they may put their completed forms. Emphasize that you want their honest appraisals, so they should feel free to offer constructive criticism.
6. Conduct interview:
  - Explain that the purpose is to get general reactions to issues not covered in the questionnaire and/or to give the subject a chance to expand on some things covered in the questionnaire.
  - Ask questions, probe when necessary, take notes.
7. Thank subject for his or her help.
8. After subject leaves, put subject number on questionnaire and interview schedule. File under subject number.
9. Enter pre-course, post-course, and post-task confidence measure (if used) on subject's observation form.

#### DATA ANALYSIS

After the data is collected, it must be looked at in some systematic way so that patterns can be discerned and implications drawn. The method of analyzing the data varies with the type of data.

#### Performance Measures

Scores on Pre- and Posttests - Since samples will be small for pilot studies, statistical analysis is usually not warranted. A simple chart that gives pre- and posttest scores for each subject will usually be sufficient. From this it will be possible to see if scores tended to increase, stay the same, or decrease.

Task Performance - This usually can be summarized in terms of success (subjects were or were not able to complete the task), amount of time to complete the task, and difficulties encountered by the subject in carrying out the task.

### Questionnaires

A simple tabulation of results across subjects will usually be sufficient to see any patterns. This can be done by hand on a blank questionnaire. Software such as SPSS (Statistical Package for the Social Sciences) as well as specific software produced by Digital are available for the analysis of data from larger samples.

### Interview Schedules

Interview schedules are usually constructed to probe specific areas. For example, you may be interested in responses to menu structure, course organization, graphic illustrations, or special features. To see if there is a pattern to subjects' responses in each area, you can set up a simple matrix with area of concern on one axis and subject number on the other axis. Then you can enter subjects' responses (paraphrased and/or truncated) in each cell. Appendix D shows a sample Data Analysis Matrix.

### Observational Data

The technique used to analyze interview data can be used for observational data as well. However, you are less likely to know what the categories will be ahead of the time. You could begin by using one matrix for instructional design, one for screen design, one for bugs or editorial corrections, etc.

If you use a matrix to summarize the data, it is easier to see patterns. For example, you will see if more than one student had trouble with menu design or with a particular set of instructions. You will have to make judgements concerning the importance of the phenomena you observed. Not everything needs to be entered on the matrix. The matrix should summarize the most important things you observed. The instructional designer will always be able to refer to the complete observational files for each subject. However you should not necessarily exclude data from your matrix just because it relates to the performance of only one subject. If it seems important, include it.

### Summary Report

It is helpful to the developer if you summarize your findings in a final report. The following can serve as a suggested outline for a final reports.

#### I. Introduction

##### A. Goals of the Pilot Evaluation

- B. Evaluation procedure
  - C. Participants
  - D. Limitations and Qualifications
- II. Summary of Observational Data
- (sub-headings reflecting major topics)
- III. Performance Measures
- (pre- and posttests and/or task performance)
- IV. Summary of Questionnaire/Interview Data
- (sub-headings reflecting major points)
- V. Summary and Recommendations

#### ACKNOWLEDGEMENTS

These Guidelines grew out of a series of pilot evaluations done for three DECmate II CBI courses. I am grateful to the instructional designers whose courses were evaluated. John Hale, Marion Hamblett and Michele Fineblum exhibited a great deal of professionalism by volunteering their courses for evaluation, and by contributing to the development of these Guidelines. The Sample CBI Questionnaire in the Appendix is an amalgam of the work of several developers, including Betsey Henkels, Linda Marsh and Jesse Heines. Pat Billingsley shared her ideas on methods and instrumentation in the early stages of the process. The support and suggestions of Lisa Ehrlich provided impetus and direction to the project. The participants in the pilot evaluations, all of whom were Digital employees, were very helpful and cooperative. Without them, we could have not completed the project.

Appendix A

PROTOTYPE QUESTIONNAIRE

This questionnaire addresses your reactions to this course. Please use the following scale to respond to each question.

SA = Strongly agree  
A = Agree  
N = Neither Agree nor Disagree  
D = Disagree  
SD = Strongly disagree  
NA = Not applicable

Please CIRCLE the appropriate response for each question.

|   |    |   |   |   |    |    |
|---|----|---|---|---|----|----|
| I knew how to get in and out of the course.   | SA | A | N | D | SD | NA |
| I often felt lost or disoriented while taking the course.   | SA | A | N | D | SD | NA |
| There was usually too much information on the screen at any one time.   | SA | A | N | D | SD | NA |
| The graphics used in the course were useful in getting the point across.  | SA | A | N | D | SD | NA |
| The language used in the course was easy to understand.   | SA | A | N | D | SD | NA |
| I could easily distinguish material that was presented as instruction from material that simulated the real system. | SA | A | N | D | SD | NA |
| There were too many technical words that were not clearly defined.  | SA | A | N | D | SD | NA |
| I found it useful when the course asked me to respond to questions or to enter information.                         | SA | A | N | D | SD | NA |
| The feedback messages that appeared with an incorrect response were helpful in understanding the material.          | SA | A | N | D | SD | NA |
| I was able to proceed through this course at my own pace.   | SA | A | N | D | SD | NA |
| The examples shown in the course helped me learn the material.  | SA | A | N | D | SD | NA |

|   |    |   |   |   |    |    |
|---|----|---|---|---|----|----|
| There were enough examples.   | SA | A | N | D | SD | NA |
| The menus were easy to understand and use.                            | SA | A | N | D | SD | NA |
| The lessons followed a logical sequence.                              | SA | A | N | D | SD | NA |
| The information in this course will be useful for my job.             | SA | A | N | D | SD | NA |
| The summaries were helpful.   | SA | A | N | D | SD | NA |
| The course MAP helped me understand the way the course was organized. | SA | A | N | D | SD | NA |
| The HELP facility was a useful feature of the course.                 | SA | A | N | D | SD | NA |

How often did you use the following HELP features?

|          | Often | A Few Times | Never |
|----------|-------|-------------|-------|
| BACKUP   | _____ | _____       | _____ |
| CONTINUE | _____ | _____       | _____ |
| GLOSSARY | _____ | _____       | _____ |
| EXIT     | _____ | _____       | _____ |

What parts of the course did you particularly like?

Were there any parts of the course that were confusing, unnecessary, or boring?

Would you prefer taking a CBI course like this to learning about WPS features by reading a manual?

Appendix B

CONFIDENCE RATING

On a scale of 1 to 5, how confident do you feel right now that you could ...

(describe the task here)

Please use the following codes:

1. I don't think I could do it, even with documentation
2. I probably could do it, with documentation
3. I definitely could do it, with documentation
4. I probably could do it, even without documentation
5. I definitely could do it, even without documentation

YOUR RESPONSE: \_\_\_\_\_

NOTE

The points on this scale reflect a software task for which documentation (that is not a part of the CBI course) is available. The points on a scale should reflect the type of task the subject is asked to perform.





Appendix D

SAMPLE DATA ANALYSIS MATRIX

|               | Menu<br>Negotiation | Course<br>Structure | Student<br>Errors | Feedback<br>Messages |
|---------------|---------------------|---------------------|-------------------|----------------------|
| Subject<br>01 |                     |                     |                   |                      |
| Subject<br>02 |                     |                     |                   |                      |
| Subject<br>03 |                     |                     |                   |                      |
| Subject<br>04 |                     |                     |                   |                      |
| Subject<br>05 |                     |                     |                   |                      |