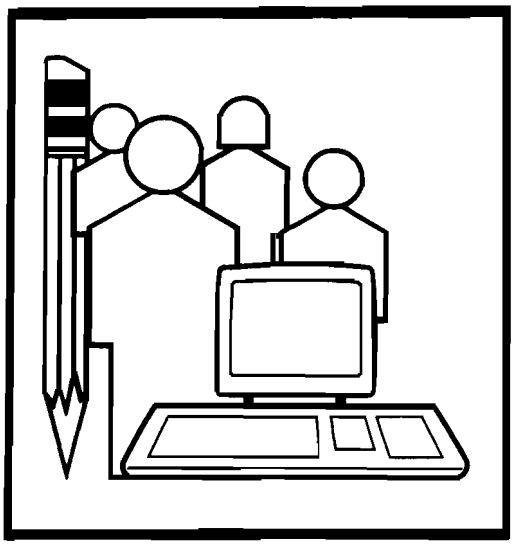

TECHNICAL REPORT

EVALUATING TECHNICAL COMPETENCIES: DESIGNING ASSESSMENT ITEMS USING A COMPUTER/VIDEODISC TRAINING SYSTEM

Technical Report No. 20

November 1983

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

digital

EVALUATING TECHNICAL COMPETENCIES: DESIGNING ASSESSMENT ITEMS USING A COMPUTER/VIDEODISC TRAINING SYSTEM

Lisa R. Ehrlich, Ph.D.

ABSTRACT

This paper begins with an investigation of reasons why computer/-videodisc systems may be used for testing. It presents a brief discussion of the characteristics of interactive computer/-videodisc training systems, and then focuses on the item design considerations faced by instructional designers and evaluators when using these delivery systems as a means of assessing learner comprehension and competency.

This paper was presented at the 1983 Annual Meeting of the Evaluation Research Society.

INTRODUCTION

The hardware and software integration of microcomputer and videodisc technology affords instructional designers a unique and challenging training medium. This individualized environment offers the capabilities of computer-based instruction and audio/-visual technologies within a single delivery system. Computer/-videodisc training systems may be used as assessment mediums for performance-based objectives requiring visual formats. The challenges of designing for this rich training environment lie ahead for evaluators and instructional designers.

Research investigating the use of audio/visual testing began in the mid 1940's with the use of motion films and slides using synchronized recorded sound. Thelen (1945) used slides and sound methods, which he believed could be used as a valid prediction of behaviors, to evaluate overt learner responses.

Gibson (1947), working for the Army Air Force Aviation Psychology program, identified the need for alternative testing methodologies to evaluate human skills and aptitudes requiring primarily pictorial rather than linguistic presentation:

> It is likely that there are types of human aptitudes and abilities which cannot be adequately measured by the relatively static problems and questions presented by ordinary test methods but which can be demanded by setting up tasks arising from the continuous flow of events portrayed on the motion picture screen. (Gibson, 1947, p. 98)

This paper begins with an investigation of reasons why computer/-videodisc systems may be used for testing. It presents a brief discussion of the characteristics of interactive computer/-videodisc training systems, and then focuses on the item design considerations faced by instructional designers and evaluators when using these delivery systems as a means of assessing learner comprehension and competency. It concludes with suggestions for further research investigations.

This paper does not argue for or against the use of visually-based test items; it does not investigate relationships which may exist between the use of visually-based test items and spatial visualization abilities. While it is critical that this relationship be explored, it is beyond the scope of this paper.

WHY USE COMPUTER/VIDEODISC SYSTEMS FOR TESTING?

One might question the use of computer/videodisc hardware/software systems to test learner comprehension and mastery of skills. What potential benefits are gained over the use of a paper and pencil test? Are the additional costs associated with the design and production warranted?

Both industry and government are concerned about skills training. Learners are taught functional literacy skills, equipment maintenance, programming, office skills, and so on. Much of this training is criterion-referenced, i.e., requiring content mastery. It is the author's belief that the retention of these skills will tend to increase as the testing environment closely approximates the instructional environment, assuming the assessment items are valid. This belief is indicated by the stimulus-generalization phenomenon, which states that "the amount of information acquired by students increases as the testing situation becomes similar to the situation in which students receive instruction" (Szabo, 1981, p. 179).

It is also the author's belief that the closer both the instruction and testing environments approximate reality, the easier skill transfer becomes. Most testing strategies currently used to evaluate mediated instruction are pencil and paper types which are highly verbal, rather than visual, in nature. The reading skills required to take verbal tests may exceed those skills actually required by the job itself. Learners deficient in reading ability may not achieve test scores which appropriately demonstrate their ability to perform on the job; one major reason for using audio/visual testing is that these tests are not influenced to a large degree by verbal ability.

Nugent (1982) tells us that the iconic system (that part of human cognitive structures primarily responsible for visual processing capabilities) is presumably specialized in processing concrete objects and events, thus providing an overall schema for the subject matter. Linguistic presentation is specialized in providing a specific focusing orientation but not good at providing an overall schema. The two can interact and work jointly to facilitate learning. Learners process pictorial and linguistic information through functionally independent, though interconnected, cognitive systems. A properly designed and presented training/evaluation package on a computer/videodisc delivery system can facilitate the flexibility of individualized storage, encoding and retrieval strategies.

Park (1972) believes that exams can be advanced two steps closer to "reality" by paying more attention to the increase in visual and audio sensory inputs. This provides an additional reason for the use of a computer/videodisc delivery systems for testing.

MEDIA CHARACTERISTICS OF COMPUTER/VIDEODISC TRAINING SYSTEMS

It is important to investigate the media attributes available with an interactive computer/videodisc system. The use of the computer as a training medium offers all the features of conventional computer-based instruction (CBI). Stand alone CBI may also use one of the audio tracks on the videodisc to enhance the primarily text and graphic presentation of the computer (see Figure 1). The videodisc player offers all of the traditional audio/visual training media attributes in differing combinations (depending on the hardware and software capabilities of the particular delivery system).

The integration of the media capabilities of a computer/videodisc training system are determined by appropriate instructional design strategies and visual design and message design techniques. Within the limitations imposed by the hardware and software capabilities of a particular delivery system, the instructional message should be dictated by sound instructional design strategies.

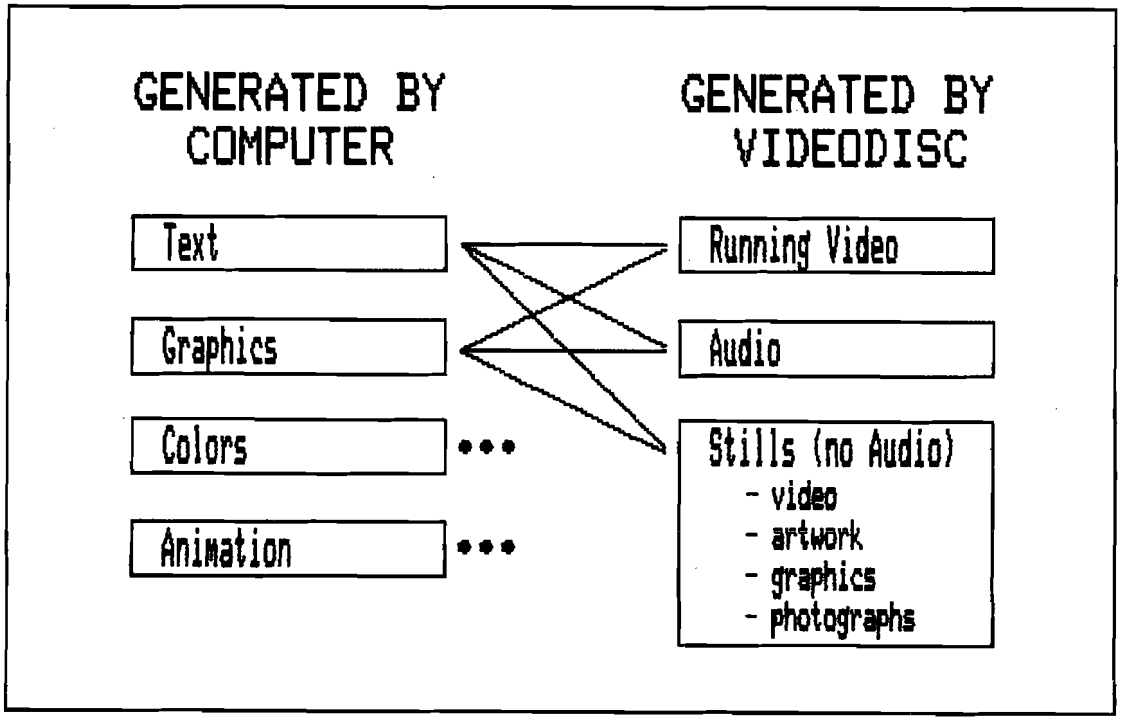

Figure 1

MEDIA ATTRIBUTES OF COMPUTER/VIDEODISC SYSTEMS

Each computer-generated attribute can be used in combination with any videodisc-generated attribute.

Some computer/videodisc systems use a two screen presentation system with one screen for computer-generated output and the other screen for the images from the videodisc player. With these systems, the software controls the output from the videodisc. Other systems use one screen (usually an NTSC monitor although high resolution graphic terminals are also used with more sophisticated systems) to display both the computer-generated output and the images from the videodisc. Some of the more advanced systems allow the computer-generated output to be overlayed directly on top of the visual images from the videodisc. In other words, the computer and videodisc output may appear concurrently on the screen. Other one screen systems only allow either computer-generated output or visuals from the videodisc on the screen at any one time.

The particular configuration of any one system will directly influence instructional strategies and visual design and message design decisions. Discussion of design and media selection decisions are beyond the scope of this paper.

DESIGN OF ASSESSMENT ITEMS

This section presents suggestions for test item design using a single screen interactive computer/videodisc system with a high resolution graphics monitor having overlaying capabilities. It discusses item design for assessment of various performance skill levels.

Visually-based test items are written from specific training objectives. The use of pictorial test items is warranted only in specific assessment situations which call for visual presentation. Visually-based items may be embedded within instructional sequences to ascertain learner comprehension of the instruction and they also may be used to assess learner competency in pre- and posttesting situations.

Psychometricians have recognized a spatial ability factor as a component part of human intelligence. There is evidence which suggests that a semantic test measures a factorially different ability than a visual test (Tinsley and Dawis, 1972). It should be emphasized that the author is not suggesting that textual-based testing be totally replaced by visually-based test items. This author is simply proposing that the same instructional decisions which identified an interactive computer/videodisc training system as the appropriate instructional delivery medium should also be the basis for deciding which skills are appropriately tested through a visual medium. Certainly, those skills which are better suited for verbal testing should be presented using a text-based format (on- or off-line).

Test item designs should approximate, as closely as possible, the actual skill performance environment. Item validation is very important if we are to feel confident that the learner can perform the actual skills taught. For example, a video test may place the learner in a passive evaluative role of watching someone else performing each task. Success in this passive role does not ensure success on the active role of performing the same task. Therefore, it is important that the test item match the actual performance objective.

The computer/videodisc system testing environment can have the learner enter responses via a keyboard, joystick, touch screen or voice if the responses are to be recorded on line (and the appropriate configuration exists). Test items can also be presented via the computer/videodisc system with the student responding on a paper response sheet.

The following is a list of test item designs for typical skill performance objectives, including a brief discussion of the potential use for each item type presented.

- If the performance objective is to assess the learner's understanding of a particular part removal and replacement:
  - The test item could consist of series of removal sequences asking for visual identification of the proper sequence.
  - This last item could be extended to exposing the learner to several complete removal sequences and asking for identification of the sequence which was performed in the correct removal order.
  - Isolated segments of removal sequences may be presented asking the learner to identify what the next task in the sequence would be.
- If the performance objective is to have the learner identify components the learner could be asked to identify:
  - A component as it is presented in isolation.
  - A component which is presented in the context of its actual operational environment.
  - Each component as it is removed from its environment.
- If the performance objective is to have the learner discriminate between components:
  - A picture could be presented containing a number of different components and the learner could be asked to find a particular piece.

- A component could be matched with either its associated function or its name.

- If the performance objective calls for interpretation of results:
  - A video sequence could be presented asking the learner to indicate if the equipment is functioning correctly. This question could further ask the learner to identify what the problem is and how it would be corrected.
  - A sequence could be presented containing common procedural errors. The item may ask for an identification of these procedural errors and also ask the learner to interpret the results of these errors.

- If the objective calls for the ability to interpret equipment readings:
  - A still frame picture could be presented asking the learner to read a particular dial or calibrator.
  - An associated question could ask the learner to specify what might be the cause for the reading or what the reading might indicate.

- If the objective is for the learner to be able to specify location or setup, the system might:
  - Present a still and have the learner specify erroneous or missing connections.
  - Show a picture of equipment and ask where a specific part or connection should be placed.
  - Put up a partial diagram and have the learner complete the diagram by choosing from a group of options or using free input.

- If the objective is to assess problem solving skills involved in troubleshooting, the system might:
  - Present a problem to the learner and ask for a diagnosis the problem.
  - Ask a learner to identify the component which is faulty.
  - Ask learner to define the next or a series of steps to be taken to correct the malfunction.
  - These evaluative problem solving simulations can become extremely complex using sophisticated programming and learner monitoring techniques similar to those used in

computer-assisted patient management problems in health science training. While it is certainly appropriate to use a computer/videodisc system for presenting evaluative simulations, the design and complexity of these problems are beyond the scope of this paper.

Associated with the item design are issues of screen design and formatting. Appropriate presentation of the item is critical. The manner in which the test item is produced may bias the item in some way, so it is important to pay close attention to screen layout. In addition, directions to the learner which explain specific item response are important. Due to screen design consideration, it may be appropriate to use audio to present these instructions and/or a visual identifier of some sort which cues the learner as to the kind of response expected.

There are number of related issues associated with the use of visually-based test items. If the test uses a fixed length format, the learner should be told how many items are in the test. There is no easy way of allowing the learner to scan through the test instrument, so it is important to give the learner a sense of test length. Of course, if the test uses an adaptive testing format the length of the test may vary for each learner. Nonetheless, the learner should be made aware of the scope of these adaptive testing formats in advance of taking these tests.

Questions using video segments are presented in fixed sequence. The learner should have access to replaying these sequences. This is analogous to "rereading" a question in a paper format. The learner should be allowed as much flexibility as possible in taking these test items.

A word of caution for test designers: It is crucial that all assessment items presented on-line using a computer/videodisc system be designed, scripted and storyboarded as part of the instructional design phase of the project. This calls for front end planning and design of all evaluation instruments. It is very difficult to "retrofit" a piece of instruction which was never meant to be evaluative into an assessment framework. An after-the-fact attempt at evaluation is very difficult unless a decision is made to go back into production to design and produce appropriate evaluation items, which is a costly endeavor. It is much to the benefit of the instruction and evaluation to plan and design appropriate assessment items as part of the project instructional design process.

ASSOCIATED RESEARCH AND MEASUREMENT CONSIDERATIONS

The growing availablity and use of interactive computer/videodisc training systems increases the feasibility of using this medium as a means of evaluating learner comprehension and competency. The use of this assessment medium paves the way for the investigation of various basic research questions.

Psychometric qualities requiring further investigation are in the areas of item validation and item scoring. How is a test instrument which uses motion presentation "different" in its measurement characteristics from a paper-based instrument? How should visually-based items be validated? How might the scoring of individual items differ from traditional dichotomous scoring methods? Are test items which are presented via video segments independent or are they interrelated? And, if so, how do these items affect the reliability of the test instrument?

Research investigating relationships existing between individual spatial abilities and success on visually-based test items is warranted. Are there certain spatial skills necessary for learners to retain visual information? Is is possible for learners to retain visual information and be unable to transfer these skills? Can visually-based test instruments be designed which are sensitive enough to assess a learner's ability to perform a specifically learned skill? What are the critical characteristics of visually-based test instruments as these items interact with spatial ability skills?

An important research issue requiring further investigation is a identification of specific abilities and aptitudes which might be appropriately assessed using visually based test items. What are these characteristics and what is the most appropriate means of presenting these items using a computer/videodisc system?

Now is the time for further investigation of these and related research areas. This author looks to practitioners and researchers alike in expanding our knowledge of visually-based testing using interactive computer/videodisc systems.

SUMMARY

The challenge of designing interactive computer/videodisc instructional courseware lies ahead. The merging of computer-based instruction and videodisc technologies invites designers to devise new instructional strategies and modifications to existing methodologies uniting these presentation mediums. A logical extension for the use of these training systems is the assessment of learner competency and comprehension.

In summary, visually-based test items may be used for learner assessment in skills training. As the testing environment more closely approximates the instructional environment skills retention may increase and become more easily transferred to actual operational environments. Visual-based testing is a more pictorial testing mode then conventional paper-based testing, which is presented using a primarily verbal format. It is suggested that test items presented using a verbal format may measure differing skills and abilities from test items presented using a visual format. The use of visually-based test items is warranted only in specific assessment situations requiring visual presentation. Test items should approximate, as closely as possible, the actual skill performance environment.

Additional research investigation is necessary. Item validation and item scoring issues require further inquiry. Research investigating relationships existing between individual spatial abilities and success on visually based test items is warranted. Another important research issue requiring further investigation is an identification of specific abilities and aptitudes which might be appropriately assessed using visually based test items.

REFERENCES CITED AND RELATED READINGS

Boyd, J.L., and B. Shimberg. Handbook of Performance Testing: A Practical Guide for Test Makers. Educational Testing Service, 1971. ERIC Document No. ED 052 220.

Brittain, C.V., and M.M. Brittain. Audiovisual vs. Paper and Pencil Testing of Occupational Competence. Paper presented at Annual Meeting of Eastern Educational Research Association, 1980. ERIC Document No. ED 211 572.

DeMelo, H.T., and F.M. Dwyer. A Multifactor Analysis of the Instructional Effects of Type of Instruction, Testing, Recall and Order of Testing. Paper presented at the Annual Meeting of Association for Educational Communication and Technology, 1982. ERIC Document No. ED 223 197.

Gibson, J.J. (ed.). Motion Picture Testing and Research. Report #7, AD-651-783, Army Air Force Aviation Psychology Program Research Reports, 1947.

Guay, R.B. Spatial Ability Measurement: A Critique and an Alternative. Paper Presented at Annual Meeting of American Educational Research Association, 1980. ERIC Document No. ED 189 166.

Hall, K.A., R.C. Comer, and J.A. Merrill. Taxonomy of Instructional Strategies for Computer-Based Education. Ohio State University, 1981. ERIC Document No. ED 217 867.

Hiscox, M.D. Integrating Testing and Instruction Using the Videodisc. Paper presented at Annual Conference of Society of Applied Learning Technology, 1981.

Levine, J.M., D. Schulman, R.E. Brahlek, and E. Fleishman. Trainability of Abilities: Training and Transfer of Spatial Visualization. Office of Naval Research, Report Number ARRO-3010-TR3, 1980. ERIC Document No. ED 183 954.

McKinley, M.B. A Comparison of Audio/Visual Testing Mode with Visual Test Mode as a Differential Variable in Test Performance, 1975. ERIC Document No. ED 120 247.

Merrill, P.F., and C.V. Bunderson. Guidelines for Employing Graphics in a Videodisc Training Delivery System. ISD for Videodisc Training Systems. First Annual Report. Vol. III. WICAT, Inc., 1979. ERIC Document No. ED 196 413.

Moses, B.E. The Relationship Between Visual Thinking Tasks and Problem-Solving Performance. Paper Presented at Annual Meeting of American Educational Research Association, 1980. ERIC Document No. ED 187 734.

Nugent, G.C. Pictures, Audio and Print: Symbolic Representation and Effect on Learning. Educational Communication and Technology Journal, 1982, 30, 162-174.

Park, J. The Use and Development of Videotape Tests to Assess Achievement of Educational Objectives: Application in Educational Psychology. Paper Presented at CAPE Conference, 1972. ERIC Document No. ED 069 754.

Rouse, W.B., S.H. Rouse, R.M. Hunt, W.B. Johnson, and S.J. Pelligrino. Human Decision-Making in Computer-Aided Fault Diagnosis. Army Research Institute for the Behavioral and Social Sciences, 1980. ERIC Document No. ED 192 743.

Shriver, E.L., J.F. Hayes, and W.R. Hufhand. Evaluating Maintenance Performance: A Video Approach to Symbolic Testing of Electronics Maintenance Tasks. Final Report. Air Force Human Resources Laboratory, Report # AFHRL-TR-74-57, 1974. ERIC Document No. ED 106 329.

Szabo, M., F.M. Dwyer, and H. DeMelo. Visual Testing -- Visual Literacy's Second Dimension. Educational Communication and Technology Journal, 1981, 29, 177-187.

Thelen, H.A. Testing by Means of Film Stills and Synchronized Recorded Sound. Educational and Psychological Measurement, 1945, 5, 33-48.

Tinsley, H.E., and R.V. Dawis. The Equivalence of Semantic and Figural Test Presentation of the Same Items. Office of Naval Re-

search, Report # TR-3004, 1972. ERIC Document No. ED 068 515.

Wildman, T.M., and J.K. Burton. Matching Learning Theory, Instructional Technology and Evaluation Design. Paper Presented at Annual Meeting of the Association for Educational Communication and Technology, 1980. ERIC Document No. ED 192 766.

Wissman, D.J. Audio/Visual Proficiency Testing: Annotated Bibliography. Interim Report. Air Force Human Resources Laboratory, Report # AFHRL-TR-77-54, 1977. ERIC Document No. ED 152 807.