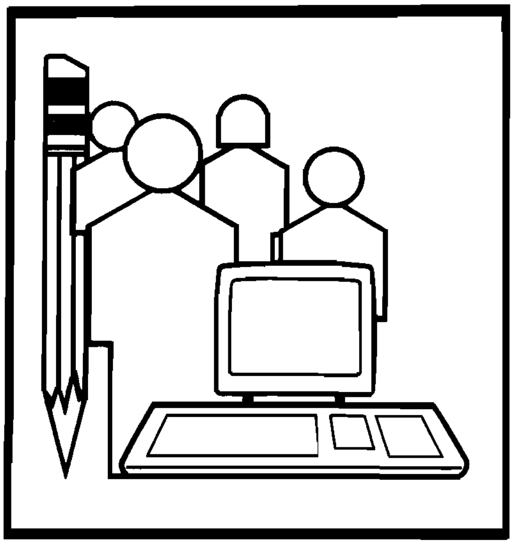

TECHNICAL REPORT

USING SIMULATIONS TO EVALUATE SKILLS AND COMPETENCIES

Technical Report No. 18

November 1983

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

USING SIMULATIONS TO EVALUATE
SKILLS AND COMPETENCIES

Lisa R. Ehrlich, Ph.D.

ABSTRACT

The technique of simulation is becoming increasingly popular in both training and educational environments. This approach places the learner in a mock or simulated setting which attempts to mirror the actual environment. This paper presents a brief overview of what simulations are and how they are used in training and educational settings. The major focus is on the use of simulation techniques to evaluate learners' skills and competencies, as opposed to the use of simulation to teach or train; evaluative simulation is distinctly different from instructional simulation. Since this paper discusses how simulations may be used to assess competencies, some discussion of basic measurement concerns and issues is presented. While this paper primarily addresses computer-based and paper and pencil presentation formats, many of the issues and concerns discussed relate to all currently used simulation formats.

This paper was presented at the 1982 Annual Meeting of the National Society for Performance and Instruction.

WHAT IS A SIMULATION?

Simulations allow learners the flexibility of trying alternative approaches to problems and situations presented to them. Students can be placed in a situation that is very similar to the actual environment for which they are being trained. In real or simulated time, students can see the results and ramifications of various choices they have made. This allows for a dynamic learning environment. Students can experiment within a safe environment and not be concerned with how the decisions made affect real circumstances and situations. They have the freedom to make mistakes without having to face the consequences of the same mistakes made in a real setting.

Simulations are very powerful teaching and evaluating tools that, if used wisely, can greatly enhance the learning or testing experience. The student is not a passive recipient of information. These instruments are interactive and allow the learner to participate actively in the learning experience by constantly being required to make decisions.

EVALUATIVE SIMULATION

Before discussing some instructional design concerns, it should be noted that this paper only addresses how the technique of simulation is used to evaluate students, not how to teach or train them. Evaluative simulation is essentially a test instrument used to obtain a measure of an individual's competency or skill level. Simulation as testing is appropriate when the intention is to assess how a person will function in the actual environment for which they have been trained.

An example of an evaluative simulation that could be used in a business training setting may help to clarify some questions and issues. Suppose that you train managers and supervisors, and that a group of lower level managers has just received an intensive three day training seminar on time management. An evaluative simulation instrument would be an excellent technique to measure what the managers learned from this seminar.

First, a realistic scenario must be created to which the managers can respond. In other words, you are "setting a scene" which will require the managers to use their newly acquired time management skills. Something like this could be the opening scene:

> You are the manager of ten people in a group. The group is working on eight concurrent projects. Four of these projects are on critical time lines. You also have the responsibility for setting up a formal train-

ing program for new hires within your group. This is also on a critical time schedule.

The student is then asked what s/he would first do to manage the time and resources allocated. The option choices are typically presented from a list. The student makes choices(s) and receives some form of feedback. This feedback may consist of ramifications of those choices, realistic "random" occurrences, or perhaps another option list. The feedback should be realistic for all option choices. The more contrived the simulation appears the less likely students will identify with it.

It is important to keep in mind that if the simulation is being used as a test of knowledge or competencies, the feedback should simply explain or state how the option choice affected the situation. This feedback should not instruct, coach or cue the student in any way. If it does, the instrument loses its ability to assess what the student knows or would probably do in the given situation.

The student progresses through the simulation making choices from presented option lists, receiving feedback and continuing to make option choices in light of the feedback. Allowing the students to continually choose what their next decision would be facilitates an interactive learning environment. In the example simulation, the manager has the flexibility to make time management decisions and see the ramifications of these choices.

The simulation allows each student the ability to create her/his own path or route as progress is made through the instrument. The routes are uniquely defined by the choices made.

Since timeliness and priorities change in the real world, these events can be built into the simulation as well. One of the most important considerations to keep in mind is the realism of the simulation. The more realistic the simulation, the better the estimate of what a student knows and what s/he can do in the real world.

DESIGN CONSIDERATIONS

The flexibility and realism of a simulation are determined by its design. Regardless of the ultimate presentation format of the simulation, some fundamental instructional design considerations must be addressed prior to any development.

The major stages that one goes through in designing evaluative simulations are:

1. Define the knowledge, skills, and/or competencies to be measured.

2. Define the simulation model, including its content, the relevant interactions between the main object system and its connecting super- and sub-systems, and the input and output dynamics.

3. Design the simulation's structure and flow, including the examinee interface with the model, the informative feedback, and pathways through the simulation.

4. Implement the simulation on a computer system.

5. Develop a scoring algorithm.

6. Assess the validity of the evaluation using external instruments.

7. Revise the evaluative simulation as needed.

The critical steps in this process are to identify exactly what aspects of competence are to be measured and to construct an environment which resembles, as closely as possible, the real life performance on which the examinee is being evaluated. Note that some aspects of reality cannot be simulated. These limitations translate into constraints placed on the model. It is critical to identify what aspects of reality are not incorporated into the model so that any generalization of examinee performance on the simulation can be prefaced by those aspects, and so that predictions of individual performance on the basis of responses to the simulation can be interpreted properly. Additional details on individual stages are provided in the paragraphs that follow.

Objectives

If is very important to identify which specific skills and competencies are to be measured. This is similar to defining terminal objectives. What is being measured must be very clear in the designer's mind.

Model Definition

The content or scene must be determined before any script development can occur. Once the content has been defined and delineated, an analysis of this system can begin. This step is very important, because it determines how similar the simulation will be to the "real world." All the static and dynamic aspects of the real system should be studied. The outcome of this design stage will define the simulation model.

Structure

The question of how the student will interface with the simulation instrument must be determined as part of the design phase. Two principle formats are computer-based or paper and pencil. Some of the issues that help determine the format are: economics, availablity of computer hardware, programming software and expertise and the purpose of the test instrument.

Most evaluative simulations use option lists as a means of eliciting student input. These individual options were predetermined by either a subject matter expert or by empirical means. With computer formats, a list is typically presented either directly on the terminal screen or in an accompanying booklet. For paper and pencil formats, a latent image technique is used. Here the student is presented an option list in the printed test booklet and given a special pen which, when applied to a treated area on the option page, reveals a response to the option choice. The construction of a paper and pencil test is similar to the programmed instruction "scrambled book" format.

Option Lists

The use of option lists is of primary concern to the design of these instruments. In order to design realism into the simulations, appropriate feedback responses need to be available. Thus, in order to design appropriate feedback into the simulations, the initial option choices must be known.

One of the criticisms against using list or menu options lies in the belief that the student's response choices are limited to the option list. The student is forced to choose from a finite option list, and is is not at liberty to compose an original response. If the option list is large enough then this restriction does not pose any great constraint; however, the question of what defines an exhaustive option list and yet still remains usable is not easily resolved.

Related to this concern is the issue of cueing the respondent. Presenting a list of choices not only forces the student's choice but may present options that the student had not considered on his/her own. This is of particular concern if the simulation instrument is part of a certifying examination. The cueing issue becomes less important if a validity estimate of sufficient magnitude has been determined for the instrument.

Design and Development

A detailed task analysis should be conducted. This analysis will enable the construction of option lists with the appropriate feedback. Ultimately a very detailed flowchart will result from the analysis. If done correctly, the actual materials development should be completely determined from the flowchart.

MEASUREMENT CONCERNS

The design of evaluative simulations requires a well-defined set of testing objectives. The test environment must be as similar as possible to the actual setting in order to obtain a measure or estimate of how the student will function in the real environment, as well as her/his level of performance.

At a fundamental level, the design of evaluative simulations is no different from that of any other test instrument. The test designer still needs to decide what knowledge and/or skill levels are to be measured, and, if the test is criterion referenced, what the predetermined passing level(s) need to be.

One of the major differences between evaluative simulations and multiple choice type tests is the manner in which the test items are constructed. More often then not, the items on a multiple choice exam are independent of one another. In other words, the way a student answers one item has nothing to do with the way the next item will be answered. This is not true of simulations. The choices (answers) that a student makes are interdependent; how a student decides to respond to a list of options determines what the next series of options will be.

Item Scoring

There is little empirical evidence to favor one scoring algorithm over another. The evaluative simulation literature is rich with many empirical and theoretical attempts to score individual items within these instruments (Ehrlich, 1981; Schultz, 1978; McGuire, 1976; Schumacher, 1974; Rimoldi, 1962).

There is much controversy over how the individual, but yet, interrelated test items should be weighted. Many evaluative simulations use an item weighting continuum from +5 to -5 (McGuire, 1976). A recent study investigated the use of three different scoring systems using a complex branching simulation. The results suggested that the use of a simple dichotomous (zero and one weightings) scoring systems performed as well as a complex differential weighting system (Ehrlich, 1981).

Validity

It is very important to obtain an estimate of the validity of the simulation instrument. One of the best ways of doing this is to have a student work through the simulation and then observe this student in a similar situation in the actual environment. The higher the correlation between the two performances, the more faith can be placed on the results of simulation performance. Take note that the designer is cautioned against interpreting the results from these instruments if its validity has not been estimated to be at an appropriate level.

The validity of these instruments can be determined in the same manner as validity estimates for other evaluative instruments are determined. What is ultimately at issue is whether the simulation is measuring the skills and competencies it purports to measure.

Reliability

Some factors which effect common parametric reliability estimation for simulations are as follows:

1. Data options are differentially weighted.

2. Data options are interdependent.

3. There are differential amounts of feedback obtained by the examinees which result in dissimilarity among the students with respect to the nature of the problem posed by any given data option.

4. An individual student can be denied the opportunity of responding to many of the data options because of the particular decisions s/he opts to make.

There is no one reliability index that is presently being used to estimate this test characteristic. The above issues effect the use of commonly used estimates. Much work needs to be done in this area. However, if the instrument is valid and the scoring system used is determined to be reliable, the simulation results obtained can be generalized with some amount of certainty.

The above measurement concerns and issues confound attempts to estimate student competencies and abilities. Yet, on a conceptual level, it seems that one of the best ways to measure what a student can do is to place them into a similar situation and see how they perform. While there are important psychometric issues surrounding the use of these instruments, this should not impede their use. The reader is encouraged to use simulations to assess skills and competencies in their particular training or educational environment.

REFERENCES CITED AND RELATED READINGS

ANDREW, B.J. An Approach to the Construction of Simulated Exercises in Clinical Problem Solving. Journal of Medical Education, 1972, 47, 952-958.

DONNELLY, M.B. Measuring Performance on patient Management Problems. Proceedings of the Fifteenth Annual Conference on Research in Medical Education, 1976, 161-164.

DOWALIBY, F.J., and B.J. Andrew. Relationships Between Clinical Competence Ratings and Examination Performance, Journal of Medical Education, 1976, 51, 181-188.

EHRLICH, L.R. An Investigation of Basic Psychometric Properties of Evaluative Medical Simulations. Doctoral Dissertation, University of Iowa, 1981.

GRACE, M., J. Hanson, S.M. Fincham, E.N. Skakon, and W.C. Taylor. A Scoring Technique for Computerized Patient Management Problems. Medical Education, 1977, 11, 335-340.

McGUIRE, C.H. Simulation Techniques in the Teaching and Testing of Problem Solving Skills. Journal of Research in Science Teaching, 1976, 13, 89-100.

PAGE, G.G., and D.W. Fielding. Performance on Patient Management Problems: Are They related? Journal of Medical Education, 1980, 55, 529-537.

PALVA, J.P., and V. Korhonen. Validity and the Use of Written Simulation Tests of Clinical Performance. Journal of Medical Education, 1976, 51, 657-661.

RIMOLDI, H.J.D. The Test of Diagnostic Skills. Journal of Medical Education, 1961, 36, 73-79.

RIMOLDI, H.J.D. Rationale and Applications of the Test of Diagnostic Skills. Journal of Medical Education, 1963, 38, 364-368.

RIMOLDI, H.J.D., J.R. Devane, and T.F. Grib. Testing Skills in Medical Diagnosis. Chicago: Loyola University Psychometric Laboratory, 1958.

RIMOLDI, H.J.D., J.V. Haley, and H. Fogliatto. The Test of Diagnostic Skills. Chicago: Loyola University Psychometric Laboratory, 1962, No. 25.

SCHULTZ, J.V. R.B. Friedman, R.S. Newsom, and S.M. Entine. Computer Based Examination Using a Simulated Patient-Physician Encounter. Paper Presented at the Annual Meeting of the American Educational Research Association, Toronto, 1978.

SCHUMACHER, C.F. A Comparative Study of Four Methods for Scoring Experimental Computer Based Examinations for Clinical Problem-Solving. _Proceedings of the Thirteenth Annual Conference on Research in Medical Education_, Chicago, 1974, 2-7.

SEDLACEK, W., and L. Nattress. A Technique for Determining the Validity of Patient Management Problems. _Journal of Medical Education_, 1972, 47, 263-266.

SENIOR, J.R. _Toward the Management of Clinical Competence_. Philadelphia, Rittenhouse, 1976.