DOCUMENT RESUME

D 164 614 TM 008 187

UTHOR Cockrell, John T.
ITLE Television as Stimulus Input : Synthetic Performance Testing. Technical Paper 303.
NSTITUTION Army Research Inst. for the Behavioral and Social Sciences, Arlington, Va.
UB DATE Aug 78
OTE 49p.

DRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
ESCRIPTORS Enlisted Personnel, *Feasibility Studies; Job Skills; *Performance Tests; Response Style (Tests); *Stim lation; *Television; *Testing; Test Validity; *Visual Measures; *Work Sample Tests
DENTIFIERS Paper and Pencil Tests

BSTRACT

As part of an effort to provide more economical ethods for job-skill evaluation in the Army Enlisted Personnel anagement System, a study investigated the use of television as a eans of presenting test items for a sample of tests from the job ield of tank crewman. The television scene provided the job setting or each item and then posed a question that required a real-time esponse from the examinee. All items were job-connected and herefore represented simulated skill items. The examinee responded irectly to the face of the television screen by means of an lectronic stylus or gun reticle. The test was compared with a aper-and-pencil test which covered the same items and a hands-on erformance test which covered many of the same items. The results ndicated ' at television testing is very acceptable to soldiers and easible. 'he validity of the prototype test could not be determined recisely ecause the criterion hands-on test was not usable; most xaminees made a perfect score on the hands-on test. (Author/GDC)

ED164614

Technical Paper 303

AD

TELEVISION AS STIMULUS INPUT IN SYNTHETIC PERFORMANCE TESTING

John T. Cockrell

ARI FIELD UNIT AT FORT KNOX, KENTUCKY

U. S. Army
Research Institute for the Behavioral and Social Sciences

August 1978

Approved for public release; distribution unlimited

SM008187

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL S ENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

WILLIAM L. HAUSER
Colonel, US Army
Commander

JOSEPH ZEIDNER
T hnical Director (Designate)

NOTICES

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
Technical Paper 303

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (*and Subtitle*)
TELEVISION AS STIMULUS INPUT IN SYNTHETIC PERFORMANCE TESTING

5. TYPE OF REPORT & PERIOD COVERED
--

6. PERFORMING ORG. REPORT NUMBER
--

7. AUTHOR(s)
John T. Cockrell

8. CONTRACT OR GRANT NUMBER(s)
--

9. PERFORMING ORGANIZATION NAME AND ADDRESS
U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Va. 22333

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
20763731A770

11. CONTROLLING OFFICE NAME AND ADDRESS
U.S. Army Training and Doctrine Command (TRADOC)
Fort Monroe, Va.

12. REPORT DATE
August 1978

13. NUMBER OF PAGES
38 pages

14. MONITORING AGENCY NAME & ADDRESS (*if different from Controlling Office*)
--

15. SECURITY CLASS. (*of this report*)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
--

16. DISTRIBUTION STATEMENT (*of this Report*)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (*of the abstract entered in Block 20, if different from Report*)
--

18. SUPPLEMENTARY NOTES
--

19. KEY WORDS (*Continue on reverse side if necessary and identify by block number*)
Simulated performance testing
Synthetic performance testing
Television testing
Job-skill testing

20. ABSTRACT (*Continue on reverse side if necessary and identify by block number*)
As part of an effort to provide more economical methods for job-skill evaluation in the Enlisted Personnel Management System (EPMS), a study investigated the use of television as means of providing test item presentation. The television scene provided the job setting for each item and then posed a question that required a real-time response from the examinee. All items were job-connected and therefore represented simulated skill items. The examinee responded directly to the face of the television screen by means of an electronic stylus or gun reticle. The test was

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

780

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*) compared with a paper-and-pencil test which covered the same items and a hands-on performance test which covered many of the same items. The results indicated that television testing is very acceptable to soldiers and feasible. The validity of the prototype test could not be determined precisely because the criterion hands-on test was not usable.

Tec ical Paper 303

TELEVISION AS STIMULUS INPUT IN SYNTHETIC PERFORMANCE TESTING

John T. Cockrell

and

Donald M. Kristiansen, Work Unit Leader

ARI FIELD UNIT AT FORT KNOX, KENTUCKY

Submitted as complete and technically accurate, by:
Donald F. Haggard
Field Unit Chief

Approvec

E. Ralph Dusek, Director
INDIVIDUAL TRAINING AND PERFORMANCE RESEARCH LABORATORY

Joseph Zeidner
TECHNICAL DIRECTOR (DESIGNATE)

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

August 1978

Army Project Number
2Q763731A77^

Performance-Oriented
Individual Training

ARI Research Reports and Technical Papers are intended for sponsors of R&D tasks and other research and military agencies. Any findings ready for implementation at the time of publication are presented in the latter part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

Research on improving performance both of groups of soldiers functioning as a crew and of individual soldiers forms a major program at the Army Research Institute for the Behavioral and Social Sciences (ARI).

The Performance-Oriented Individual Skill Development and Evaluation project is concerned with improving relevance, efficiency, and economy of individual enlisted training and evaluation. In a major move to improve the combat readiness of soldiers, the Army is implementing the Enlisted Personnel Management System (EPMS). This system requires restructuring the individual training and testing systems to make them job relevant.

Evaluation of the individual soldiers for career progression in the EPMS is based on critérion-reference performance testing of actual job skills rather than generalized knowledge. In an effort to achieve more economy in the large-scale testing required for the EPMS, a research program seeking to develop simulated performance tests has been initiated. One promising line of endeavor is the use of various audio-visual media to provide the stimulus input and job setting for the skill items. The present publication reports the results of a study investigating the use of television stimulus inputs in conjunction with an electronic responding vehicle which requires real-time decisions and responses.

This research was done at the ARI Field Unit at Fort Knox, Ky., in response to requirements of Army Project 2Q763731A770 and to special requirements of the Training and Doctrine Command (TRADOC), Fort Monroe, Va.

Joseph Zeidner

JOSEPH ZEIDNER
Technical Director (Designate)

Requirement:

To investigate the validity and feasibility of using television stimulus inputs in a synthetic performance test, and to determine if such tests can replace hand-on performance tests. This research is considered necessary because of the high cost of hands-on tests and the need to develop a less expensive, reasonably valid substitute.

Procedure:

A synthetic performance test using television as the stimulus input was developed and produced. The test was considered a performance test because the items covered the actual tasks the examinees were required to perform on the job. The test was administered to 70 soldier-trainees who had completed advanced training in the subject matter. Scores made by these same trainees on a hands-on performance test which had similar items were also obtained. The hands-on test was administered routinely by the Army to all trainees at the end of the advanced training. A parallel paper-and-pencil test was administered to 64 soldier-trainees, and hands-on scores were also obtained for these trainees.

Findings:

The results favored the feasibility of television testing. The test was produced and administered without difficulty, and the examinees had a very favorable attitude. The examinees had no trouble understanding and responding to the items. The examinees judged the test as "fair" (impartial) in terms of testing them on important tasks they should have mastered.

The validity of the results was inconclusive. The criterion scores for the hands-on test were unsatisfactory in that most examinees made a perfect score. The correlation between the television and hands-on tests was low-positive but nonsignificant. Comparison between the television and parallel paper-and-pencil tests also showed no overall difference, although there were significant differences between many items.

Utilization of Findings:

This study provides insufficient evidence to conclude that synthetic performance tests with television inputs can replace hands-on performance tests. To determine more precisely whether television testing has promise requires the development of a more satisfactory hands-on criterion test and a more thorough examination of those tasks and response components that appear most amenable to television testing.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

TELEVISION AS STIMULUS INPUT IN SYNTHETIC PERFORMANCE TESTING

INTRODUCTION

Over the past 10 years or so, the Army has tried to convert more of its testing to the "hands-on" performance mode, especially at training centers and at the beginning skill levels. Even more emphasis has been placed on performance testing in the last 2 or 3 years with the beginning of the Skill Qualification Testing (SQT) program. Performance testing is highly desirable because of its high face validity and high user acceptability; however, this type of testing is very costly, hard to standardize, and often not feasible.

The alternative to hands-on performance testing has generally been the standard, group-administered, knowledge-type, paper-and-pencil test. Although relatively easy to produce and administer, this type of test is generally considered to have low validity and low user acceptability.

Osborn (1970) has suggested that a compromise validity-feasibility tradeoff point might be reached by using synthetic performance tests. According to Osborn, the term "synthetic performance test" refers to any performance test that is less than a full hands-on test, but more than the group-administered, knowledge-type, paper-and-pencil test. Synthetic performance tests include all tests that use any type of simulated inputs or responses. Part-task tests, in which only one or a few response components of a task are measured, are also included under synthetic performance tests. The synthetic performance test is conceived as less costly than a hands-on test, but as a test that still has reasonable validity and user acceptability.

To support the Army's adoption of performance testing, the U.S. Army Research Institute has initiated a broad-based research program to investigate the possibilities of synthetic performance testing as a cost-effective alternative to the usual hands-on procedures. The goal of this research is to develop a psychometric base for both hands-on and synthetic methods.

The research focus has been on the use of audiovisual media to provide the simulated stimulus input. The reasoning behind this focus is that audiovisual media stand midway in the stimulus fidelity range, and at the same time, are at the medium to high end of the feasibility scale.[1] Thus, audiovisual media may represent a good fidelity-feasibility tradeoff

---

[1] Stimulus fidelity as used here refers to how closely the test stimulus resembles the real world, and stimulus feasibility refers to how much it costs to present the test stimulus in a testing situation (high feasibility equals low cost).

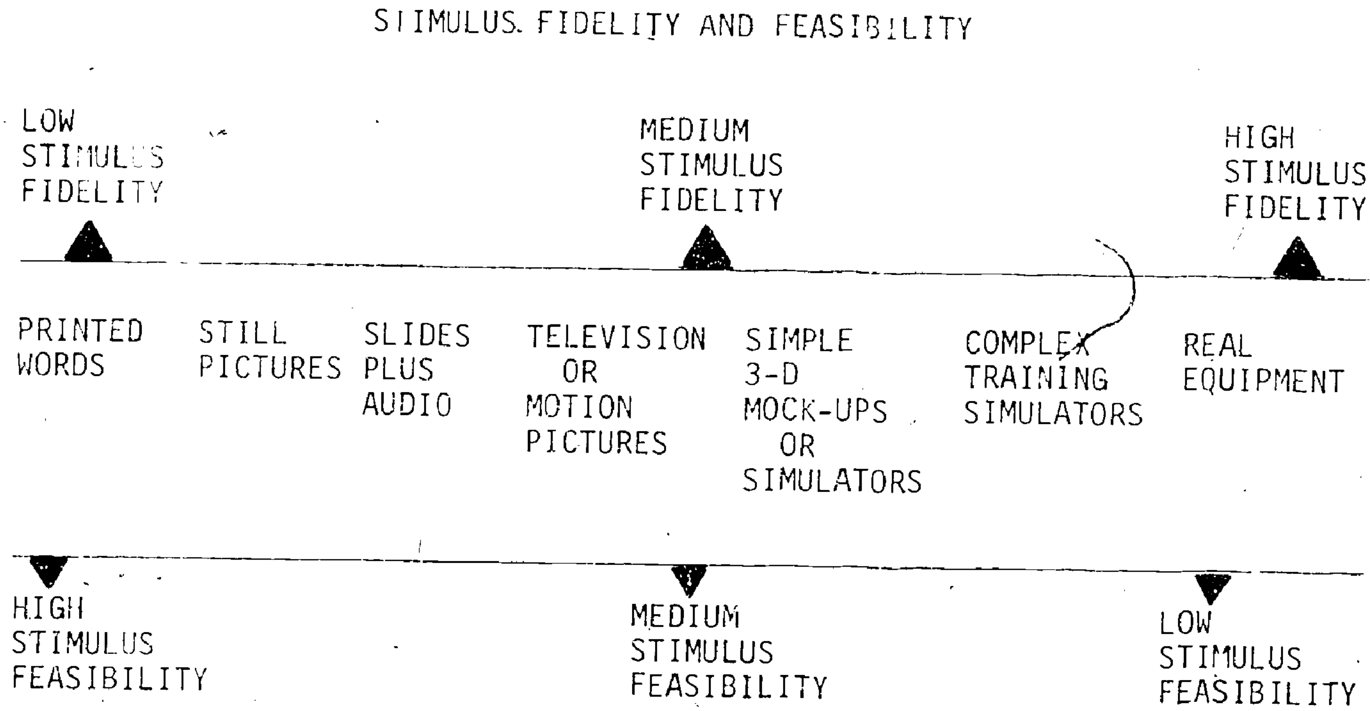


Figure 1. Conception of stimulus fidelity and feasibility tradeoff.

point insofar as stimulus input is concerned. Figure 1 shows a conception of this fidelity-feasibility tradeoff.

The overall research program has the following objectives:

1. To explore the parameters of the various audiovisual media to determine the media's applicability to synthetic performance testing.

2. To explore various responding modes and response devices that can be used with audiovisual stimulus inputs.

3. To determine whether those response components of a task that can be measured using audiovisual media are sufficient to yield an acceptable measure of the entire task.

4. To develop a task classification system that will enable a synthetic performance test developer to determine by analyzing the task (a) when audiovisual media should be used as the stimulus input, (b) which medium is advisable, and (c) which response components should be measured.

Several experiments in this research program are now in process using a number of different audiovisual media. This paper, which is concerned with television as the stimulus input, presents the results of the first of these experiments.

This first experiment was limited in nature and focused on the feasibility of using television as the stimulus input. As such it was concerned mostly with the first research objective--applicability of media to testing--with some exploration into the second and third objectives, responses to stimuli and test-task comparisons.

Background and Rationale

The impetus for this research stems from the Army's decision to substitute the Skill Qualification Testing (SQT) program for the current MOS testing program as a means of assessing the job skills of enlisted personnel. The SQT program is intended to be based on job-sample tests wherever practical, as contrasted to the current MOS paper-and-pencil knowledge test.

This change was brought about partially as a result of the research of a number of investigators (Engel, July 1970; Engel, October 1970; Engel & Rehder, 1970; Shirkey, 1965; Urry, Shirkey, & Nicewander, 1965) who questioned the validity of the MOS test for job skill assessment. In 1966 the Army convened a special board of inquiry (Brown Board) to survey the entire question of written MOS tests for assessing job skills and job knowledge. This board recommended that performance tests be substituted for written tests wherever practical (U.S. Army, 1966). Following the publication of the findings of the Brown Board, the Army has made substantial progress in implementing the recommendation (e.g., the Tank Crewman Advanced Individual Training performance tests administered in the form of a "county fair," with examinees moving from test to test around the examination area, during and at the end of each training cycle). However, due to high costs and difficulty in maintaining standardization, the performance test obviously is limited in terms of making up a substantial part of each SQT test. This is particularly true at the higher skill levels and for many hard-to-measure tasks. Occhialini (1972), for example, presents evidence that performance tests are extremely difficult to prepare and administer, and are of questionable validity. Engel and Rehder (1970) review the arguments against the use of performance tests for part or all of the SQT battery. Their general conclusion is that the exclusive use of performance tests in an SQT battery would be too costly and impractical.

Reacting to the pros and cons of paper-and-pencil vs. performance tests, several researchers have proposed compromises. Engel and Rehder (1970) advocate a mixture-of-measurement technique in each SQT test, combining work samples, simulated tests, peer ratings, and paper-and-pencil tests. They present evidence indicating that cognitive items can be measured adequately by paper-and-pencil tests; that motor-manipulative items require work sample or simulated tests; and that peer ratings can be used to judge social, leadership, and overall ability.

Osborn's (1970) approach is concerned with developing synthetic tests that it is hoped will eliminate some of the impracticality of administering performance tests, while reducing the verbal component and improving the validity of paper-and-pencil tests. Osborn visualizes a continuum bounded on one extreme by paper-and-pencil knowledge tests and on the other by job-sample skill tests. Within this continuum, a number of synthetic tests more or less removed from each extreme can be constructed. The continuum is conceived of as being scaled in psychological units and varies along the dimensions of stimulus fidelity and response fidelity (or a mixture of both).

In any combat situation, the stimulus dimension would be a large complex composed of visual, auditory, tactile, kinesthetic, olfactory, pain, and stress inputs. The response dimension would be an equally large complex of cognitive, motor-manipulative, and perceptual outputs. For the purposes of illustration, the stimulus and response fidelity dimensions for armor crewmen might be conceptualized as shown in Figure 2. Osborn maintains, in an analysis similar to the one shown in Figure 2, that one must pull away from each extreme of the continuum to develop synthetic tests that are both feasible and more valid than paper-and-pencil tests.

An important aspect of Osborn's conception is his reasoning with regard to <u>part-task</u> testing (Osborn & Ford, 1976). In this conception, each task is composed of a number of response components divided into cognitive, perceptual, and motor behaviors. Figure 3 shows a task broken

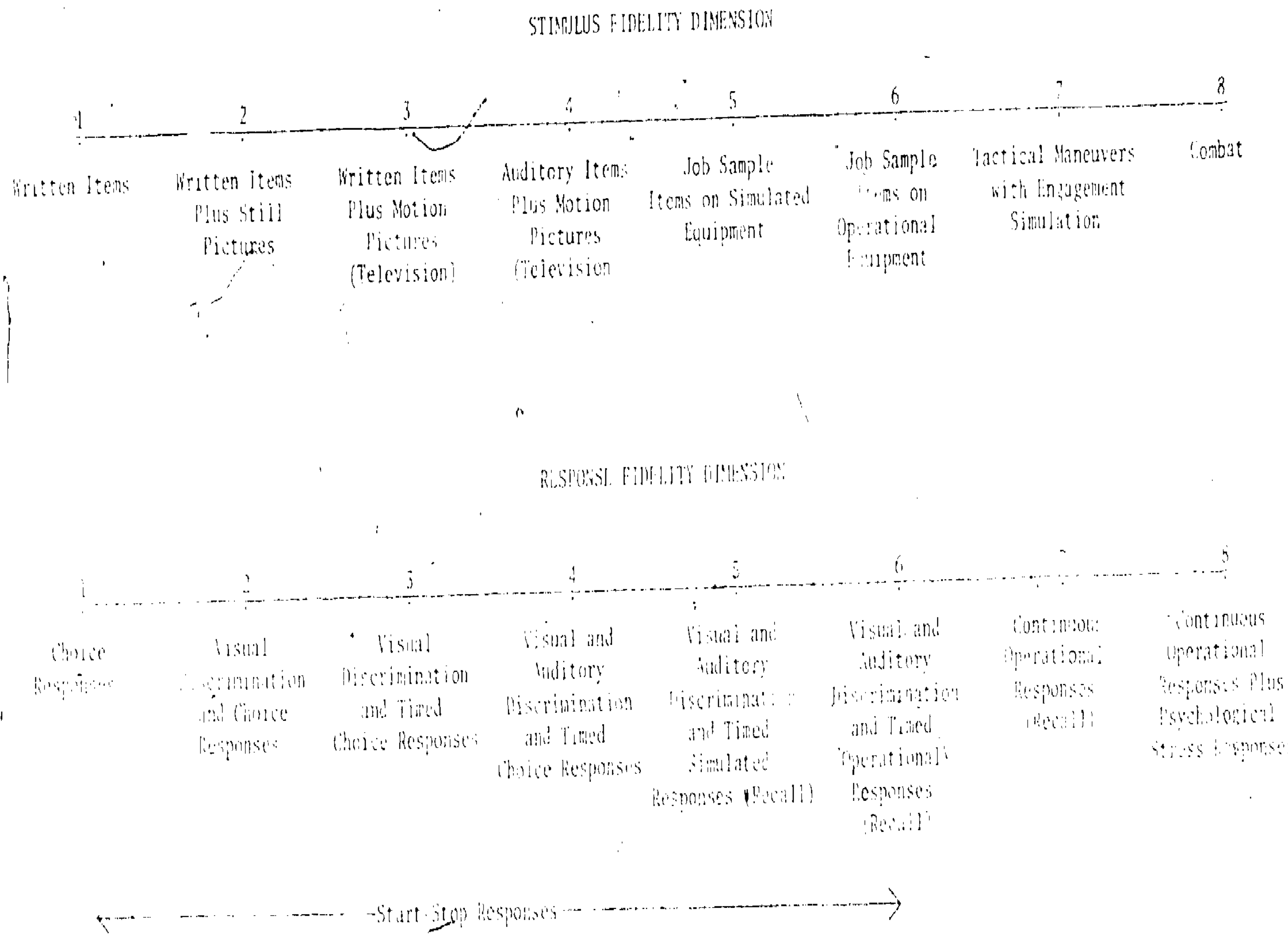


Figure 2. Stimulus and response fidelity dimensions.

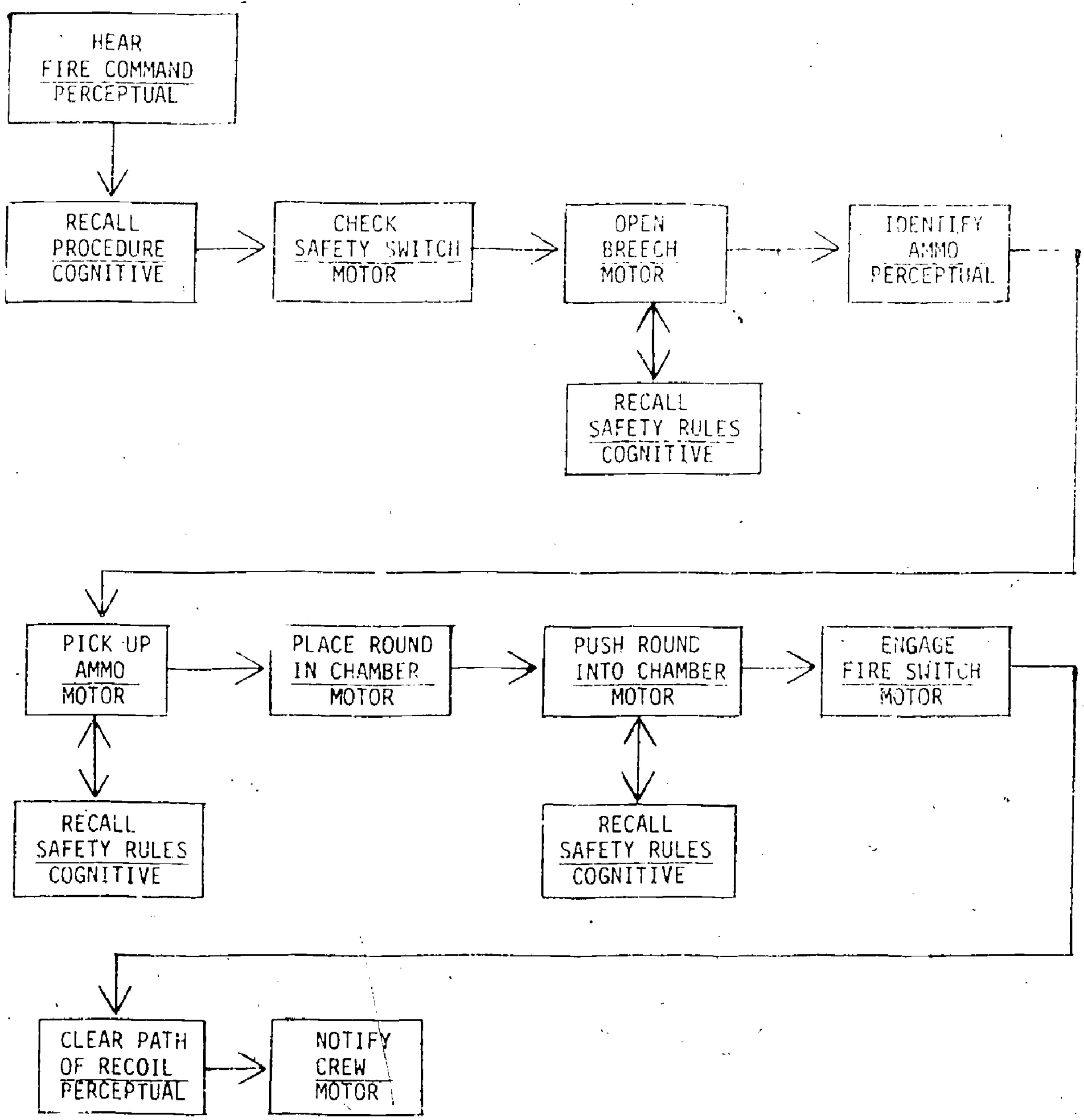


Figure 3. Response components for task: Loading round into main gun on an M60A1 tank.

down into response components. (This task is performed by the loader on an M60A1 tank.)

The reasoning behind part-task testing is that it may not be necessary to test every response component in a particular task in order to determine how well the whole task can be performed. It may be possible to get a good indication of whole-task performance by measuring only a few response components or perhaps measuring only one critical response component.

Part-task testing becomes crucial when audiovisual stimulus inputs are used because the nature of the medium precludes obtaining any measurements on most motor-response components. In order to obtain measurements on motor-response components one needs to test on real equipment or a hands-on simulator. Since the measurable response components in audiovisual simulation are limited to perceptual and cognitive ones, it follows that the usefulness of audiovisual stimulus inputs is dependent upon the validity of the part-task testing concept. One objective of the research program is to check the part-task testing concept by correlating scores made on part-tasks using audiovisual stimulus inputs with scores made on the corresponding whole task tested in the hands-on mode.

Use of Television in Testing. Television has been used in testing primarily as a recording medium (Cockrell, 1974; Hays & Pulliam, 1974). A study by Shriver (Shriver, Hayes, & Hufhand, 1974) explored the possibilities of using television as the stimulus input in a performance test. After developing the test, Shriver concluded that television did not offer much promise in terms of replacing hands-on testing. He listed eight disadvantages of the television medium and decided to abandon the method and not attempt a systematic comparison between the television test and hands-on performance tests. Some of the disadvantages mentioned follow:

1. Television tests place the subject in a passive role, watching someone else perform and evaluating the correctness of the performance. There is no reason to believe that success in this evaluation role will insure success in the active role of performing the task.

2. Television violates a major ground rule of criterion-referenced testing in that it emphasizes process measurement rather than product measurement.

3. Television costs are very high compared to those of slides or graphics because of the large amount of equipment needed and the large personnel time requirements.

Shriver's criticisms are informative, but they do not necessarily settle the case. The nature of the medium does include some practical difficulties both in producing the stimulus tapes and in administering the tests. However, these difficulties are minor compared to the complex task of administering hands-on performance tests. If television can show a useful correlation with job-sample tests and also show advantages over written and other audiovisual tests, it may well be worth the extra cost.

Objectives

The primary objectives of the present experiment were to appraise some of the practical difficulties in using television as the stimulus and to make a rough comparison among television, paper-and-pencil, and hands-on performance tests. The secondary objective was to conduct a checkout of a responding device (Telestrator) designed to permit examinees to respond directly to images on a screen (see Appendix C).

Specifically, the objectives were as follows:

1. Determine the feasibility of using television in testing. The items under consideration here were:

   a. Understandability of test items

   b. Ease of responding

   c. Time allotment for responding

   d. Difficulties and costs involved with administering television tests.

2. Determine the acceptability of television testing by examinees.

3. Compare the results made on the television test with those made on the paper-and-pencil and hands-on performance tests.

4. Conduct a checkout of the Telestrator response device.

METHOD

The overall method consisted of (1) producing a television test for a sample of tasks from the job field of tank crewman (11E MOS), (2) producing a parallel paper-and-pencil test covering the same items, and (3) comparing the results made on these two tests with the results made on an existing hands-on performance test that covered many of the same items.

The job field of tank crewman was selected because much prior research had been done in this field. A complete task analysis was available, and a hands-on performance test has been in use for the Tank Crewman Advanced Individual Training course for 2 years. This existing hands-on performance test was felt to be a good base against which to compare the television and paper-and-pencil tests.

Television Test

The first step in producing the television tape was to select the critical tasks in consultation with military experts. The selection criteria were set by the military and included such considerations as importance to fulfilling the mission, safety to the crewman, and safety to the equipment. The critical tasks selected were quite similar to the tasks covered in the Tank Crewman Advanced Individual Training course. After the critical tasks were selected, they were ordered according to skill level.[1]

For the final test, tasks were selected from skill levels 1, 2, and 3.[2] For the purposes of this experiment, the tasks can be considered to range from fairly easy to very difficult. Tasks were also selected such that each of the four positions (driver, loader, gunner, and tank commander) was covered, and a few tasks pertained to the crew at large.

In consultation with the military, each task was broken down into cognitive, perceptual, and motor components; and each response component was examined for its criticality to the task. Practical considerations such as overall test running time, time to televise each item, number of response components needed to cover a particular task, and achieving a balanced test (see Appendix A) eliminated many critical response components. For each of the remaining critical response components a television test item was conceived and a television shooting script was written. Each item was televised in a crude fashion with a handheld camera and a portable videotape recorder.

The raw footage was edited roughly into a prototype television test by the addition of narration and titles. The prototype tape was intended only as a model for a professional tape to be produced later and as a vehicle to check technical accuracy and television feasibility.[3]

Military experts checked the prototype tape for technical accuracy and understandability. A revised television script incorporated suggestions; a final television tape was produced using professional television personnel, cameras, and editing facilities. The shooting and editing of this final tape required approximately 30 calendar days (about 15 actual working days).

The final tape consisted of 47 test items plus 4 practice items and had a running time of 53 minutes. The items ranged in running time from

---

[1]There are five skill levels for each MOS ranging from skill level 1 (beginning) to skill level 5 (most advanced).

[2]Work on the preliminary television tape and the task selection required to produce it were done by Human Resources Research Organization under contract to the U.S. Army Research Institute.

Table [illegible]

[illegible]

[illegible]k and [illegible]	[illegible]	Item [illegible] [illegible]	[illegible]	Type [illegible]	Crew [illegible]	[illegible]	[illegible] response [illegible]
[illegible]rt [illegible]	[illegible] transmission	Yes	Yes	Words	Driver 1	Cognitive	Multiple choice
	[illegible] steering [illegible]	Yes	Yes	Words	Driver	Cognitive	Multiple choice
	[illegible] safety rule-starter [illegible]	Yes	Yes	Words	Driver 1	Cognitive	Error detection
[illegible]rgency [illegible]cedures	4. [illegible]	Yes	[illegible]	Words	Driver 1	Cognitive	Multiple choice
[illegible]ht [illegible]	[illegible] signal-start [illegible]	Yes	No[d]	Picture	Driver 1	Perceptual	Multiple choice
	[illegible] signal-back[illegible]	Yes	No[d]	Picture	Driver 1	Perceptual	Multiple choice
	[illegible] signal-left turn	Yes	No[d]	Picture	Driver 1	Perceptual	Multiple choice
	[illegible]	Yes	No[d]	Picture	Driver 1	Perceptual	Multiple choice
[illegible] down [illegible]	[illegible]	Yes	Yes	Words	Driver 1	Cognitive	Multiple choice
	1[illegible]. Maximum [illegible]	Yes	Yes	Words	Driver 1	Cognitive	Multiple choice
[illegible]d round	[illegible]. Identify ammo	Yes	Yes	Picture	Loader 1	Perceptual	Multiple choice
	[illegible]. Identify ammo	Yes	Yes	Picture	Loader 1	Perceptual	Multiple choice
	[illegible]. Safety rule-[illegible]	Yes	No[d]	Words	Loader 1	Cognitive	Error detection
	14. Safety rule-[illegible]	Yes	No[d]	Words	Loader 1	Cognitive	Error detection
	[illegible]. Safety rule-[illegible]	Yes	No[d]	Words	Loader 1	Cognitive	Error detection
	[illegible] time	Yes	No[d]	Words	Loader 1	Cognitive	Multiple choice
[illegible]ck [illegible] level [illegible] main gun	1[illegible]. [illegible]	Yes	Yes	Words	Loader 1	Perceptual	Multiple choice
[illegible]esight and [illegible]o	1[illegible]. Prepare computer	Yes	No[d]	Words	Gunner 2	Cognitive	Error detection
	1[illegible]. Aiming point for boresight	Yes	No[d]	Words	Gunner 2	Perceptual	Motor manipulation
	20. Set [illegible]	Yes	No[d]	Words	Gunner	Cognitive	Multiple choice
	21. Aiming point for shot group	Yes	No[d]	Words	Gunner [illegible]	Perceptual	Motor manipulation
	[illegible]. Align reticle	No[b]	No[d]	Words	Gunner 2	Cognitive	Multiple choice
[illegible]fire procedure	2[illegible]. Safety rule	Yes	Yes	Words	Gunner 2	Cognitive	Error detection
[illegible] gun - [illegible]tial sight [illegible]ture	24. M[illegible] sight-stationary target	Yes	Yes	Drawing	Gunner 2	Perceptual	Motor manipulation
	25. [illegible] sight-moving target	Yes	Yes	Drawing	Gunner 2	Perceptual	Motor manipulation
	26. M1[illegible] sight-stationary target	Yes	Yes	Drawing	Gunner 2	Perceptual	Motor manipulation
	27. M1[illegible] sight-moving target	Yes	No[c]	Drawing	Gunner [illegible]	Perceptual	Motor manipulation

[illegible]em omitted because tested at different point in training.

[illegible]em omitted because of poor television picture.

[illegible]em not included on any hands-on test.

[illegible]em not covered during training.

Table 1 (continued)

...k and response component		Item used on final TV test	Item or task on hands-on test	Type of paper & pencil stimulus	Crew position and skill level	TV test response component type	TV test response type
...gun - ...ust fire	28. Burst-on-target - M32 sight	Yes	Yes	Drawing	Gunner 2	Perceptual	Motor manipulation
	29. Target form - M32 sight	No^d	No^d	Drawing	Gunner 2	Perceptual	Motor manipulation
	30. Burst-on-target - M1-5 sight	Yes	Yes	Drawing	Gunner 2	Perceptual	Motor manipulation
	31. Target form - M1-5 sight	No^d	No^d	Drawing	Gunner 2	Perceptual	Motor manipulation
...ect and ...ntify ...gets	[illegible]. Identify vehicle	Yes	No^d	[illegible]	Crew 2	Perceptual	Multiple choice
	[illegible]. Identify vehicle	Yes	No^d	[illegible]	Crew 2	Perceptual	Multiple choice
...radio ...operation	[illegible]	No^d	No^d	Words	Crew [illegible]	Cognitive	Error detection
...rgen... ...cedure	[illegible]. Evaluate ... tank ...	No^d	No^d	Words	Crew [illegible]	Cognitive	Error detection
...wires ...inter...	[illegible]	No^d	No^d	Words	Crew [illegible]	Cognitive	Multiple choice
... ...cedure	[illegible]. Enter radio net	No	No^d	Words	Tank Cmdr 3	Cognitive	Error detection
...map	[illegible]. Determine location	No^d	No^d	Drawing	Tank Cmdr 3	Perceptual	Multiple choice
	[illegible]. Measure distance	No^d	No^d	Drawing	Tank Cmdr 3	Perceptual	Error detection
...pare ...ge card	40. Read instruments and plot	No^d	No^d	Picture	Tank Cmdr 3	Cognitive	Multiple choice
	41. Record data	No^d	No^d	Picture	Tank Cmdr 3	Cognitive	Multiple choice
...vice ...hine gun	42. Assemble a part	No^d	No^d	Words	Tank Cmdr 3	Cognitive	Multiple choice
...age enemy	43. Fire command - tank	No^d	No^d	Words	Tank Cmdr 3	Cognitive	Error detection
	44. Fire command - truck	No^d	No^d	Words	Tank Cmdr 3	Cognitive	Error detection
	45. Fire command - coax	No^d	No^d	Words	Tank Cmdr	Cognitive	Error detection
	46. Fire command - troops	No^d	No^d	Words	Tank Cmdr	Cognitive	Error detection
	47. Choose priority target	No^d	No^d	Words	Tank Cmdr 3	Perceptual	Multiple choice

...em omitted because tested at different point in training.

...em omitted because of poor television picture.

...em not included on any hands-on test.

...em not covered during training.

response.

Table 1 provides a description of the final television tape. The categorizing of response components into perceptual, cognitive, or motor types was somewhat intuitive. The intent was to show the predominant element of each response component and not to imply that other elements were not present.

Of the 47 items shown in Table 1, only 37 were administered to the examinees in the experiment and only 30 were scored. Most skill level 3 items were eliminated before the experiment upon the recommendation of the military staff at the Armor Center. These items were considered too advanced for the examinees. After the start of the experiment, several military advisers recommended the elimination of six more items, and one item was eliminated due to a poor television picture. These seven items were administered but not scored. The footnotes in Table 1 give the reason for the elimination of any item and also explain why certain items were not included on the hands-on test.

A more specific description of each response type shown in Table 1 follows:

(1) Multiple choice. The examinee was required to select one answer from a list of three, four, or five alternatives. These alternatives were sometimes the same as those in the usual paper-and-pencil test--namely, words on the screen--and sometimes consisted of images on the screen.

(2) Error detection. The examinee was required to watch a procedure being performed on the screen and to indicate the time and location of an error, if one occurred, at the time it occurred. The examinee was shown the procedure twice and responded on the second showing.

(3) Motor manipulation. The examinee placed a plastic gun reticle (those reticles used with the main gun in the M60A1 tank) on various stationary and moving targets as if preparing to fire the main gun. The reticles were also used to simulate the adjustment of fire that would be made if the first round missed the target. The motor-manipulation response was supposed to be a crude simulation of the actual response in aiming the main gun. However, the movements required were so far down on the scale of response fidelity that the motor component appeared not to be measured at all. Perhaps the reticle response was primarily perceptual and cognitive.

Paper-and-Pencil Test

The paper-and-pencil test paralleled the television test on an item-by-item basis. The stimulus input on this test was primarily printed words, but some pictures and drawings were used on perceptual items. Table 1 shows the stimulus input for each item.

As with the television test, only 37 of the paper-and-pencil test items were administered and only 30 were scored. The items scored were the same as those scored for the television test.

The paper-and-pencil items and the television items differed greatly in the amount of time allotted to respond to each item. The total time limit was the same for both tests; however, examinees could allocate the response time any way they chose on the paper-and-pencil test but were restricted to 10 seconds per item on the television test.

On the paper-and-pencil test, examinees could change their answers, skip items and answer later, and review their answers; on the television test, none of this flexibility was permitted.

These differences between the two tests were retained because each medium lends itself most readily to the type of procedure used. Any other procedures or a common procedure for both tests would have required much more control and thereby reduced administration feasibility.

Hands-On Performance Test

The hands-on test was one routinely administered to tank crewmen trainees as a final examination for the Advanced Individual Training Armor course. This test was given in the form of a county fair with stations and 30 performance measures. Examinees were graded on a "go/no-go" basis for each performance measure. For each no-go, examinees were required to seek out remedial training and report back later for a retest. If the retest was a no-go the examinee had to report back the next day, after further remedial training, for a second and final test. For the purposes of the present experiment, the score recorded for each examinee was the number of first-round no-go's. This was not a particularly good criterion because the number of no-go's was very small.

Response Equipment

A secondary objective of the study was to check out the television response equipment (Telestrator). This equipment consists of a clear plastic electronic tablet and associated recording and programing components. The electronic tablet covers the television screen (the tablet is approximately ½ inch away from the screen at the center of the screen and approximately 1 inch away at the edges of the screen).

The examinee looks through the tablet to view the test items. Responses are made by touching the face of the tablet with an electronic stylus or an electronic gun reticle at a particular time and location. Before the test, the correct answers (time and location) are programed on the television tape. During the test, examinees are credited with a correct answer if they touch the screen at the correct preprogramed time and location. Any other response by an examinee is counted as incorrect. Only one answer is permitted for each time, and the first answer--correct or incorrect--made during the 10-second response period is counted.

The response equipment was in prototype form and because of operational difficulties could not be used for the experiment. However, it proved possible to test the operating concept of the equipment by placing a human grader behind each examinee and having this observer record on a sheet of paper whether the examinee touched the correct location at the correct time. This grading task was quite simple, and during a pilot run with eight examinees there were no difficulties in grading.

The television monitors were black and white and measured 15 inches diagonally. The examinees sat approximately 2 feet from the sets at self-regulated distances so that they could manipulate the response implements comfortably. Prior to the start of the experiment, it was decided to remove the electronic tablets from in front of the screens because of parallax problems. After the tablets were removed, the accuracy of the responding and scoring improved to a very precise level.

The response implements consisted of a stylus used for all multiple-choice and error-detection items, and two plastic gun reticles used for motor-manipulation items. The stylus was simulated by using the eraser [illegible] an ordinary lead pencil. The two plastic gun reticles, the same design [illegible] the M32 and M105D main gun reticles in the M60A1 tank, were manipulated by small wooden knobs glued to the plastic reticles.

Examinees

The examinees were tank crewmen who had just completed the Advanced Individual Training/Armor course. Altogether, 134 examinees assigned from three different companies were tested. Examinees were drawn from the companies by a selection process best described as haphazard rather than random; however, there is no reason to believe that selective bias was present. As each group of examinees arrived for the experiment for each session, the group was randomly assigned to the television or paper-and-pencil test. Originally, 144 examinees were scheduled for the experiment, but 2 were lost due to scheduling problems and 8 were lost due to scoring problems.

Procedure

Testing was conducted over a 5-day period in three morning and five afternoon sessions. The actual schedule and distribution of examinees are given in Table 2.

Table 2

Schedule and Distribution of Examinees

Test and time of day	Days 1	2	3	4	5	Totals
Television	16	16	16	11	11	70
Morning	8	8	8	--	--	24
Afternoon	8	8	8	11	11	46
Paper-and-pencil	16	11	18	11	8	64
Morning	8	4	10	--	--	22
Afternoon	8	7	8	11	8	42

Each group of subjects reported at 0800 or 1300 and was given an orientation session explaining the purpose of the experiment. All of the paper-and-pencil group was administered the paper-and-pencil test right after an orientation. The television test was administered to four examinees at a time; the rest of the television group was assigned to a waiting room. Both the television and the paper-and-pencil tests required approximately 1 hour to complete.

Approximately 10 minutes of training were required to teach the examinees the methods for responding to the television items. Most of this training was concentrated on the use of the plastic reticle. The examinees were trained by having them respond to the four practice test items. If any examinee had difficulty with the reticles, such as choosing the incorrect reticle or holding reticles incorrectly, the tape was stopped and the four practice items presented again. In no case was it necessary to present the practice items more than twice.

RESULTS

Feasibility of Using Television in Testing

The examinees did not appear to have any difficulty in understanding the items. All of the content had been covered in the Advanced Individual Training course, and the examinees had been tested on similar items several times. All of the items were also performance based and posed questions that occur normally in everyday operations.

The responding proceeded smoothly for most items. The examinees responded very quickly on the easy items (approximately 1-2 seconds with the stylus, 3-4 seconds with the reticles). On difficult items, the amount of response time allotted (10 seconds) still appeared ample, although there usually would be a lot of hesitating over the answers. On only a few items the examinees failed to respond. When queried after the completion of the test about the amount of response time, most examinees indicated that for the most part the response time was adequate. A few examinees said that more response time should have been allotted to some items.

The administration of the television test was more time consuming than that of the paper-and-pencil test because of the need to provide preliminary training in the correct way to respond and the limit of four examinees per session. Administration could be made more feasible by increasing the number of television monitors, but it would still be advisable to have one test administrator for each four examinees because of the examinees' unfamiliarity with the response method. Compared to the administration time for hands-on testing, however, television testing is much less costly.

Acceptance of Television Testing

The reaction of the examinees to the television test appeared to be quite favorable. Postexamination interviews indicated that most examinees actually preferred the television test to the hands-on test and all examinees thought the television test was fair. Even when queried about the test's being used as a basis for promotion or extra pay, the examinees still thought it was fair. Some examinees preferred the hands-on mode of testing, but no one preferred the paper-and-pencil mode.

Some reasons mentioned for preferring the television mode follow:

1. Scoring is fairer and not dependent upon the whims of the test administrator.

2. Testing is faster and not so drawn out.

3. In television testing no one is shouting at you and ordering you around.

Some of the reasons for preferring the hands-on mode follow:

1. There is more time to think and to respond.

2. Testing is more spread out and doesn't come so fast.

3. Television hurts the eyes.

4. There is a chance to walk around between items.

The examinees also indicated that television testing would be better than paper-and-pencil testing because the questions would be more understandable and require much less reading.

Comparison of Television, Paper-and-Pencil, and Hands-On Tests

The comparison between the mean percent error made on the television test and that made on the paper-and-pencil test is shown in Table 3. The means for the television and paper-and-pencil tests do not differ to any great degree, indicating that the difficulty levels of the two tests are fairly equal.

Table 3

Mean Percent Error Made on the Television and Paper-and-Pencil Tests

Test and time of day	Days 1	2	3	4	5	Mean
Television						
Morning	19.63	15.00	14.75	-----	-----	16.46
Afternoon	27.88	29.38	28.00	26.09	27.09	27.54
				Unweighted mean		22.00
Paper-and-pencil						
Morning	28.38	20.75	22.3	-----	-----	24.23
Afternoon	27.13	24.71	1[illegible].5	29.90	27.38	26.05
				Unweighted mean		25.14

One interesting facet of the data is that afternoon television examinees made many more errors than the morning groups. These results are convincing because they are consistent across the first 3 days of the experiment and because the afternoon means for Days 4 and 5 are consistent with the other afternoon means. There does not appear to be any morning-afternoon effect for the paper-and-pencil test.

The analysis of variance using the unweighted means analysis for unequal cell frequencies (Winer, 1962) is shown in Table 4. This analysis shows no difference between the television test and the paper-and-pencil test in terms of item difficulty. There was a significant morning-afternoon effect, but the more meaningful result is the significant mean square (MS) interaction. Analysis of this MS interaction reveals that the morning-afternoon effect is concentrated on the television test and not on the paper-and-pencil test.

In order to check on whether the afternoon examinees may have been less qualified than the morning examinees, the first round no-go's from the hands-on test were analyzed. These results are shown in Table 5. Inspection of the means indicates little difference between the television and paper-and-pencil groups, or between the morning and afternoon groups. If anything, the afternoon group performed slightly better than the morning group. An analysis of variance of these results showed no significant difference for any of the variables.

Table 4

Analysis of Variance for Television and Paper-and-Pencil Tests

Source	df	MS	F	p
TV vs. P & P (method)	1	9.50	1.04	ns
Morning vs. afternoon (session)	1	124.57	13.60	< .01
Method x session	1	58.66	6.40	< .0[illegible]
Within cell	130	9.16		

Although overall scores on the television and paper-and-pencil tests did not differ, there might be differences among the various items. Accordingly, the items were grouped by response type (multiple choice, error detection, and motor manipulation) and log linear Chi-square tests (Shaffer, 1973) were computed for each item. Table 6 shows that there was a wide variation of difficulty among the items ranging from 10 to 81% error. For the multiple choice-items, there was little difference between the television and paper-and-pencil versions. Only 1 of 13 items showed a significant difference. For the error-detection items there was a substantial difference, with six out of nine items showing a significant difference.

Table 5

Mean Percent Error Made by the Television and Paper-and-Pencil Groups on the Hands-On Test

Test and time of day	Days 1	2	3	4	5	Mean
Television						
Morning	9.00	6.00	9.52	----	----	8.16
Afternoon	9.00	9.00	6.52	6.56	2.96	6.52
				Unweighted mean		7.34
Paper-and-pencil						
Morning	12.00	6.00	8.00	----	----	9.08
Afternoon	7.52	9.72	5.52	12.00	6.52	8.48
				Unweighted mean		8.78

It is interesting to note that errors of commission are more difficult to detect on television; whereas errors of omission and no-error items are more difficult to detect on paper-and-pencil. Three of the eight motor-manipulation items show some significant difference, and all three of these items show more difficulty for the television test. The net result of this item difficulty analysis shows five items more difficult on television tests and five items more difficult on paper-and-pencil tests. This canceling effect is reflected in the overall nonsignificant difference between the television test and the paper-and-pencil test.

The last analysis, in Table 7, shows the correlations of the hands-on test with the paper-and-pencil test and the television test. Those correlations are also broken down for the morning and afternoon groups. There is a low positive correlation between the television and hands-on tests and also between the paper-and-pencil and hands-on tests. The paper-and-pencil correlation is significantly different from zero; however, there is no significant difference between the television versus the hands-on and the paper-and-pencil versus hands-on correlations. The breakdown for morning and afternoon groups shows a somewhat higher positive correlation for the afternoon group and very little correlation for the morning group. Once again, there is no significant difference between the television and paper-and-pencil correlations with the hands-on test.

Table 6

Comparison Between Television and Paper-and-Pencil Items Percent Error Arranged by Response Types

Response type and item number	TV	% error P & P	X^2
Multiple choice			
1	1	0	ns
2	0	2	ns
4	43	33	1.43
5	21	14	.82
6	38	72	13.83*
7	47	52	ns
8	26	20	ns
9	16	12	ns
10	17	16	ns
11	3	5	ns
12	0	0	ns
17	1	3	ns
32	81	77	ns
Error detection (commission)			
3	51	28	7.47*
13	11	11	ns
14	40	12	11.09*
37.	46	56	1.42
Error detection (omission)			
16	19	5[illegible]	14.88*
18	4	47	18.90*
23	13	45	11.26*
Error detection (no error)			
15	16	19	ns
20	17	66	30.36*
Motor manipulation (reticles)			
19	46	16	12.24*
21	51	17	16.17*
24	4	5	ns
25	27	11	[illegible].75**
26	11	2	ns
27	14	11	ns
28	23	34	1.45
30	23	28	ns

*$p < .01$.
**$p < .10$.

Table 7

Correlations Between the Television and Paper-and-Pencil Scores and the Hands-On Performance Scores

Time of day and tests correlated	r	p	
Overall			
Television vs. hands-on	.24	ns	difference is nonsignificant
Paper-and-pencil vs. hands-on	.33	<.01	
Morning			
Television vs. hands-on	-.09	ns	difference is nonsignificant
Paper-and-pencil vs. hands-on	.16	ns	
Afternoon			
Television vs. hands-on	.47	<.01	difference is nonsignificant
Paper-and-pencil vs. hands-on	.40	<.01	

CONCLUSIONS AND DISCUSSION

The results from this research indicate that it is possible to produce a synthetic test using television as the stimulus input. The examinees can understand the problems, make proper responses, and accept the test as "fair" for career evaluation.

The experience gleaned from the production and administration of this prototype test indicates that television testing is more costly than paper-and-pencil testing but far less costly than hands-on testing. The production of the tape, from conception to final editing, required several months and used the services of a substantial number of professional people. Television tests are also somewhat inflexible, not only in the difficulty in effecting changes in the test, but also in the timing decisions--the amounts of time to allot for posing each question and for each response--that have to be made before the production of the test.

Television testing will have a much more promising future if a presentation and response device can be designed which will permit the examinee to advance to the next item as soon as the present one is answered, to see the same item twice, to change answers to an item, and to review the entire test. Such a capability would permit the flexibility of presenting multipart items, such as in troubleshooting and would permit the presentation of multimedia items, such as using both television and technical manuals in the same item.

The present experiment provides evidence that television testing is highly acceptable to the examinees. Their predominant attitude was that the test was little different from the hands-on tests in the Advanced Individual Training course, except that television was quicker and less subject to scoring error. All of the scenes were quite familiar to the examinees, and the items were ones that they had been studying for 8-13 weeks.

Television used in the multiple-choice format appears to offer an advantage over slide or paper-and-pencil formats. Before the experiment, it was felt that television would offer an advantage for those items in which motion was an integral part of the stimulus. For example, Spangenburg (1973) has shown that watching a television display of a procedure involving motion leads to more learning than watching a sequence of still shots. However, this advantage of motion proved to be true for one motion item in the present research (item 6, Table 6) but not true for two items (5 and 7). Perhaps if more motion-type items had been included in the multiple-choice category an advantage might have been shown.

In the error-detection category there did appear to be a clear-cut difference between television and paper-and-pencil items. Here the fidelity of the stimulus did seem to play a role, and the enriched stimulus of the television picture may have presented cues to the examinees. The two error-detection items that proved to be more difficult for television examinees (items 3 and 14, Tables 1 and 6) were two of the first error-detection items to be presented. Since error detection was an unfamiliar response for the examinees, this unfamiliarity may have caused some difficulty. This same phenomenon can be seen in the motor-manipulation items which involve an even more unfamiliar response. Here the television examinees had more difficulty with the first few items than did the paper-and-pencil examinees.

The correlations between the synthetic and the hands-on tests are too low to warrant recommending the substitution of synthetic for hands-on tests. However, the correlations for the afternoon groups are high enough to encourage further research. The hands-on criterion test used in the present experiment was somewhat unsatisfactory because of the large number of perfect scores.

The drop in the scores on the television test for the afternoon group as compared to the morning group was interesting but unexpected. One possible explanation for the drop may be that the examinees were required to stare continuously at a fairly large television screen from a very close distance for approximately 1 hour. A human being may be able to tolerate this strain in the morning but by the afternoon accumulated fatigue plus a heavy Army lunch may have combined with the strain to produce a letdown.

The results of this study favored the feasibility of television testing. The tests can be produced at a reasonable cost, can be administered in a reasonable manner, are understandable by the examinees, and have high acceptability with the examinees.

The validity of the results was inconclusive. The criterion scores for the hands-on test were unsatisfactory in that all hands-on examinees made a perfect score. The correlation between the television and hands-on tests was low positive but nonsignificant. Comparison between parallel television and paper-and-pencil tests also showed no difference on an overall basis, although there were significant differences between many items.

Evidence from this study is insufficient to conclude that synthetic performance tests with television inputs can be used to replace hands-on performance tests.

REFERENCES

Cockrell, J. T. [illegible] Evaluation of Minicourse [illegible] Role Playing in the Classroom. Far West Laboratory for Educational Research and Development, 197[illegible].

Department of the Army. Report of the Board of Inquiry [illegible] Army [illegible] System ([illegible] Board). Washington, D.C.: [illegible]

[illegible], J. [illegible] An approach to [illegible] Human Performance Assessment (HumRRO Prof. Paper [illegible]). Alexandria, Va.: Human Resources Research Organization, October 197[illegible].

[illegible], R. [illegible] Development of a Work Sample [illegible] (HumRRO Tech. Rep. [illegible]). Alexandria, Va.: Human Resources Research Organization, July 197[illegible].

Engel, J. D., & Rehder, R. J. A Comparison of Correlated Job and Work Sample Measures for General Vehicle Repairmen (HumRRO Tech. Rep. 70-16). Alexandria, Va.: Human Resources Research Organization, October 1970.

Hayes, J., & Pulliam, R. Development and Evaluation of Video Systems for Performance Testing and Student Monitoring (AFHRL TR-74-67). Brooks Air Force Base, Tex.: Air Force Human Resources Laboratory, July 1974.

[illegible], Jr. [illegible] Study: The Use of Qualification Tests as Performance [illegible] for Enlisted Combat MOS Evaluation (Research Study No. [illegible]). Fort Benjamin Harrison, Indianapolis, Ind.: U.S. Army Enlisted Evaluation Center, Jan. 197[illegible].

Osborn, W. C. An Approach to the Development of Synthetic Performance Tests for Enlisted Combat MOS Evaluation. (Paper for 12th Annual Military Testing Association Conference, French Lick, Ind. HumRRO Prof. Paper 30-70). Alexandria, Va.: Human Resources Research Organization, December 1970.

Osborn, W. C., & Ford, J. P. Research on Methods of Synthetic Performance Testing (HumRRO-FR-CD(L)-76-1). Alexandria, Va.: Human Resources Research Organization, April 1976.

Shaffer, J. P. Defining and Testing Hypotheses in Multidimensional Contingency Tables. Psychological Bulletin, 1972, 79, 127-141.

[illegible], E. C. Preliminary Validity Report of the MOS Evaluation Test for Medical Specialist (Research Study No. 31). Fort Benjamin Harrison, Indianapolis, Ind.: U.S. Army Enlisted Evaluation Center, 1965.

[illegible], [illegible], Hayes, J. F., & Hufhand, W. R. [illegible] Maintenance Performance: A Video Approach to Symbolic Testing of Electronics Maintenance Tasks (AFHRL-TR-74-57 (IV)). Brooks Air Force Base, Texas: Air Force Human Resources Laboratory, July 1974.

[illegible], [illegible] W. The Motion Variable in Procedural Learning. AV Communication Review, Winter [illegible], [illegible], [illegible].

[illegible], [illegible], [illegible], R. C., & [illegible], W. A. A Preliminary Validity Report on the MOS Evaluation Test for Personnel Specialist MOS Code 71[illegible] (Research Study No. [illegible]). Fort Benjamin Harrison, Indianapolis, Ind.: U.S. Army Enlisted Evaluation Center, [illegible].

[illegible], B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, [illegible].

APPENDIX

[illegible] FOR THE DEVELOPMENT OF TELEVISION TESTS

Very little evidence is available as to the best way to present test items on television. The only published research for military tests is Shriver's (Shriver et al., 1974) and as noted before, the conclusions from this research were very negative. Most decisions made for the present test were based on paper-and-pencil-test, development, expert opinion, and experience. Very few hard-and-fast guidelines can be offered because so many decisions depend upon the format chosen, the type of questions, and the amount of time available.

An important limiting factor in the development of television tests is the amount of running time available for each test item and for the complete test. The maximum desirable time for a television test such as the present one is 45-50 minutes for a number of reasons, including eyestrain, general fatigue, and administrative cost. One advantage of television testing over hands-on testing is the low administrative cost per examinee. The longer the television test, the less the advantage.

Although experience with television testing is too limited to offer much in the way of guidelines, it may be useful to describe the developmental stages and some of the difficulties encountered.

Prior to the development of the test it was decided to aim for a 50-minute running time, to cover the MOS of tank crewmen at skill levels 1, 2, and 3 and the job positions of driver, loader, gunner, and tank commander. The test was to be a group test with individual TV monitors and the examinees were to respond to the items by touching the face of the television screen with a stylus or reticle.

The first selection step was to ask various military training departments (gunnery, automotive, and such) to submit a list of critical tasks which should be tested. These departments submitted a total of 75 tasks. Because only a limited number of tasks could be used on the final tape, the list had to be pared down considerably. Many tasks were eliminated in order to balance the test among skill levels and crew positions. For example, 40 of the tasks received from the departments were for skill level 3 and only 5 of these tasks were on the final tape. Most remaining excess tasks were eliminated simply by deciding to limit the test to tasks associated with the actual operation of the tank. Critical areas such as drug abuse, first aid, leadership, and tactical decisions, and complex tasks, such as sketching an area map and tasks that required excessively long television running times, were eliminated.

The first step in developing specific test items for each task was to list all response components making up a task and decide whether each response component was primarily cognitive, perceptual, or motor. Each cognitive or perceptual response component was then examined for criticality and suitability for television testing. Depending upon the number of critical and testable response components making up a task, a decision was made to include one or more test items for the task (the number of test items per task ranged from one to six as shown in Table 1). Primarily cognitive response components were tested in error detection and multiple choice test items. Primarily perceptual components were tested in error detection [illegible] and multiple choice items.

Some of the critical and general observations that can be made [illegible] [illegible] as follows:

1. Multiple choice. Items of this sort are very simple to construct and develop and require very little test time (about 10-15 seconds) per item. The choices are presented simultaneously on screen on a four-way split screen or [illegible] words on the screen. Presenting the choices serially creates difficulties not only in terms of overall increased running time but also because the examinees often forget the first choice by the time the last is [illegible]. Either the items have to be presented twice (to [illegible] the second presentation) or the examinee must respond "yes" or "no" [illegible] as it appears. Neither method is very satisfactory.

Additionally, presenting multiple choice items on television does not offer too much advantage over a paper-and-pencil format except in terms of reducing the need for reading and perhaps presenting a more easily understood item. For example, the motion and sounds associated with [illegible] may be helpful in understanding the item.

2. Error detection. This response type has been criticized harshly by Shriver (see Page 2), and there are other difficulties as well. One major difficulty is in producing the item (televising the procedure accurately). If a no error item is desired, it is necessary to find an actor who can carry out the procedure without error. All too often, expert advisers cannot agree on the correct procedure. Many repetitions of each scene have to be made before the experts and actors can reach some sort of compromise, and even then there remain logical and inherent difficulties which cannot be resolved. For example, in televising items for the load round into main gun task it was necessary to choose between showing the action at normal speed or in slow motion. When the action was shown at normal speed, no examinee could discriminate the crucial element (hand position) and the item had no meaning. When the action was shown in slow motion (so the crucial element could be seen), examinees criticized the slowness itself as an error.

Another major difficulty with error-detection items concerns very slight deviations from prescribed procedure which often escape the scrutiny of expert advisers. Exceptionally well-skilled examinees may be lured into pointing to the slight deviations as errors, while the less skilled may not notice the slight deviations and point to the major intended error. This was particularly true for the response format used in the present study, where the first response made by the examinee was scored and all subsequent responses to the same item were ignored.

The error-detection response type was included to test the examinees' knowledge of incorrect as well as correct actions. Many incorrect actions occur very infrequently but can be very serious when they do occur. Performing the correct action in a hands-on test does not necessarily indicate awareness of danger points, and examinees need to be tested directly as to awareness of incorrect actions. However, as Shriver points out, watching someone else perform is very different from doing it yourself. Failing to notice an error may indicate lack of knowledge or it may indicate inability to notice error in others.

The overall conclusion is that error detection is a doubtful response type and more thought and research are needed prior to its acceptance as a useful procedure.

3. Motor manipulation. This response type was rather specific to this particular study and the response equipment being evaluated. In fact, one strong selling point of the response equipment was its provision for testing the motor-manipulation items. All test items under this particular response type pertained to where the examinee should place the reticle on the television screen when simulating firing the main gun under various conditions. However, analysis reveals that this response type is not really a test of motor ability, but rather a test of a combination of perceptual and cognitive abilities. The cognitive element was knowing the correct lead and elevation for each target and the perceptual element was being able to discriminate the correct lead and elevation. There is no evidence to indicate that the ability to manipulate a plastic reticle on a television screen has any correlation with the motor element involved in aiming an actual gun. On the plus side, this response type is more of a recall item than a multiple-choice question and therefore should provide a more exact measure of recall. On the minus side is the requirement to learn a new response quickly (manipulating plastic reticles). Incorrect responses may be caused by lack of knowledge or perceptual ability, or merely by failure to master the new response of manipulating plastic reticles.

This response type, like the error-detection response type, needs much more thought and research prior to acceptance as a useful procedure.

Some comments on a few miscellaneous topics may also be useful:

1. Use of a time period to indicate error. One item (3, Tables 1 and 6) in the television test used the passage of time as the cue for the examinee to note an error. That is, the actor waited only 5 seconds before continuing a procedure, where the prescribed procedure in the technical manual calls for a 120-180-second wait. Some criticism has indicated that this time-passage technique might confuse examinees because Americans have been conditioned through exposure to motion pictures and television to accept any length time period shown on a screen as the appropriate time. Item 3 did seem to confuse many examinees.

2. Long items. Several items on the television tape had relatively long running times (approximately 3 minutes). Some critics claim that including such long items may be unwise because coverage of the total subject matter is restricted at best, given a 50-60 minute time limit for the test. Although a long item may not necessarily confuse the examinees, it is noteworthy that 3 out of 4 long items retained in the test did prove very difficult for the television examinees and all 4 of the long items omitted from the study appeared to be confusing during pilot runs. Another reason for omitting long items is the difficulty in getting an actor to perform a long sequence in letter-perfect fashion. One very long item on the tape (evacuate injured crewman) was never completely satisfactory. The final take was accepted because the director became concerned with the safety of the actor playing the role of injured crewman.

3. Resolution. Unlike the human eye, television cannot capture both a wide-angle view and good resolution at the same time. For scenes that require good resolution it is a good idea to zoom in on a scene and remain there. To attempt to show more than one closeup in any one sequence tends to confuse the viewer.

4. Restricted view. Even with a wide-angle lens, television gives a very restricted view and care must be taken to provide setting shots. Precise judgments as to the placement of controls are difficult to make from a television picture.

5. Poor depth perception. Much depth perception is lost in a television picture. Items that depend upon judgment of depth should be omitted.

6. Closeup and motion. Any kind of motion in a closeup shot is confusing. Necessary movements must be very slow and precise. However, it should be noted that slowness is perceived as an error by many people.

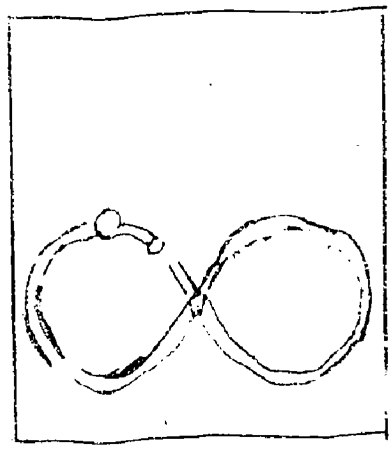
More research needs to be accumulated before a more precise set of guidelines can be produced for television testing. Particularly needed is development of a more adequate stimulus presentation and response-recording device. Also needed are researchers well grounded in the capabilities and limitations of television and the use of television cameras, lighting, and editing equipment. Television testing offers much potential but before this potential can be reached, much preparatory work remains to be done.

APPENDIX B

EXAMPLES OF PAPER-AND-PENCIL ITEMS

A. Multiple choice.

12. You are the driver of an M60A1 tank. What response do you make to the following ground guide signal given at night with a flashlight?



a. Move backward.

b. Start engine.

c. Stop engine.

d. Turn left.

B. Error Detection

6. You are the loader in an M60A1 tank. The tank commander gives the following fire command:

"GUNNER, BEEHIVE TIME, TROOPS, ONE SIX HUNDRED"

The firing switch has been checked and the breech is open. You do the six steps in order:

(1) Select a BEEHIVE round.
(2) Insert the round 2/3 of the way into the chamber.
(3) Push the round into the chamber with the heel of [illegible]ight hand.
(4) Clear the path of the recoil.
(5) Turn the firing switch to FIRE.
(6) Announce "UP."

Did you do anything wrong?

a. Step (2) is wrong.

b. Step (3) is wrong.

c. A step is missing.

d. All of the steps are correct.

C. Motor Manipulation (Reticle Manipulation)

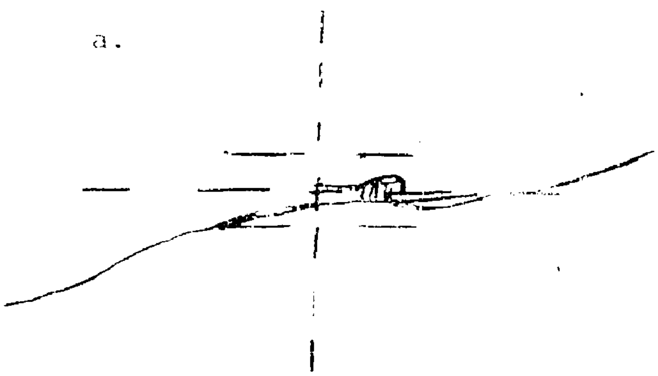
FOR THE NEXT FOUR PROBLEMS ENGAGE ALL STATIONARY TARGETS AT CENTER OF MASS AND ASSUME ALL MOVING TARGETS ARE TRAVELING AT 15 MPH.

23. You are the gunner on an M60A1 tank. The tank commander gives the fire command:

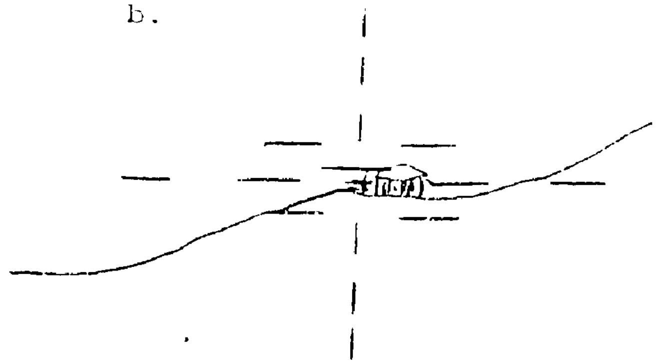
"GUNNER, SABOT, TANK"

Which of the following sight pictures would you take up using the periscope reticle?

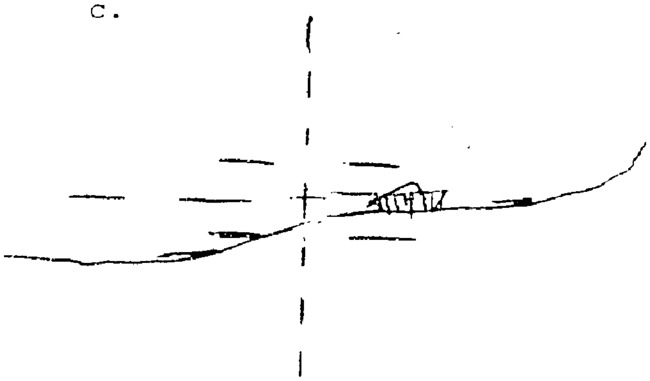
a.



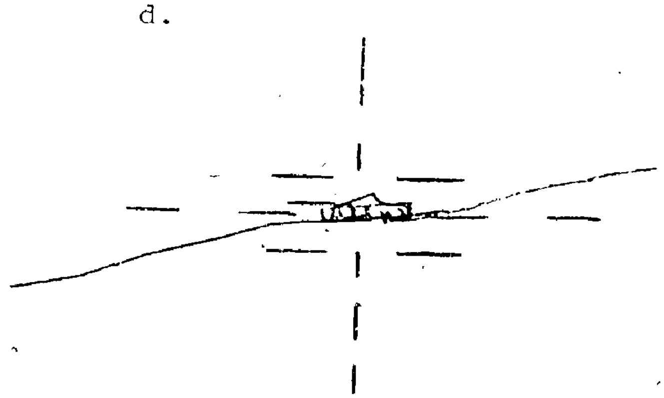
b.



c.



d.



APPENDIX C

EVALUATION OF THE TELESTRATOR EQUIPMENT[4]

One of the reasons the television test was designed and produced was to evaluate the Telestrator equipment (also known as the Telestar equipment). The novel component of this equipment is an electronic tablet which can be fitted over the face of a television screen. The tablet will record the horizontal and vertical (XY) location when it is touched with an electronic contact point embedded in a stylus or similar indicator (such as a gunsight reticle). By the proper use of auxiliary recording equipment it is possible to record the place and time the screen is touched. The recording equipment includes a counter which keeps a running total of the number of items, number of answers attempted, and number of correct answers.

The complete system includes the electronic tablet, a programing unit, and several student units. The programming unit is used by the test developer to place electronically on the television tape the XY coordinates for the correct answer for each test item and the time period during which the equipment will accept this answer. The student unit compares electronically the programed answer and the examinee's answer and records the results.

The student unit provides three types of feedback to the examinee for each test item. Immediate feedback is provided by a high-pitched tone and a small red light that comes on for a correct answer versus a low-pitched tone and no light for an incorrect answer. Slightly delayed feedback comes from a counter which shows new totals of items and correct answers at the end of the programed time period for answering each problem.

As to whether the Telestrator equipment has any merit or not, it is necessary to examine both:

1. The equipment itself, as designed and produced, and
2. The overall testing strategy which includes (a) Individual responding, (b) Television stimulus, (c) Immediate feedback, and (d) Time limit on each response.

A. The Equipment

As with most newly designed equipment, the Telestrator contained many bugs and never worked properly. However, it was possible to test some aspects of the equipment by using human graders to record right or wrong answers by the examinees according to the time and place the screen was touched. Several pilot tests were run with the following results.

[4]This summary of Telestrator operation was submitted previously to the Training Support Division, TRADOC.

(1) Accuracy. There is a fundamental flaw in the Telestrator design insofar as precise responding is concerned. The equipment was claimed to be accurate to 1/4 inch. However, due to parallax the actual accuracy was more on the order of 1 or 2 inches. This grossness effectively eliminated the use of the reticle test items because with any reasonable size reticle no discrimination was possible for leads or ranges. The grossness also eliminated many test items in which the examinee was required to discriminate among several tank controls. The spacing between these controls as shown on the screen was not great enough to permit exact programing of the answers, and one answer box would overlap another. The parallax results from mounting the electronic tablet at some distance from the actual television screen (due to curvature of the television screen the parallax increases as one approaches the edge of the screen).

In order to continue the experiment and test the idea of responding to television, the electronic tablets were removed from the television monitors and the examinees were required to touch the face of the actual television screens. This completely eliminated all parallax and permitted the use of reticle items and other precision responding.

(2) Video presentation. The electronic tablet is constructed in such a manner that it blocks a 1-inch-wide area around the outer edge of the television screen. This is a serious limitation because it is necessary to use a small-size monitor for such closeup work and this outer 1 inch covers a substantial part of the available screen area.

B. <u>Testing Strategy</u>

Because of the device's failure to work properly and the poor design of the electronic tablet, it was not possible to evaluate the testing strategy completely. However, by eliminating the parallax (removing the electronic tablets) and using human graders to record responses, it was possible to make a limited test of the strategy.

(1) Responding to television. The examinees seemed to have a little trouble understanding the test items, and responded very precisely. Three types of test items were used; Multiple choice, Error detection, and Reticle manipulation.

No training other than instructions was required for learning to respond to the multiple-choice and error-detection items. Approximately 10 minutes were required for training on the reticle-manipulation items.

(2) Time to respond. Ten seconds were allowed for responding to each item. The time limit was generous and most examinees responded to most items in less than 5 seconds.

(3) Perception of test. The examinees perceived the test as being "fair" and most actually preferred the television test over the hands-on test.

(4) Comparison with other tests. The television test was compared to parallel paper-and-pencil and hands-on tests.

(a) Paper-and-pencil test. Overall there was little difference between the mean scores on the television and the paper-and-pencil tests. However, on an item-by-item basis there was considerable difference for some items. On error detection, the television scores were much better; on vehicle manipulation, the television scores were worse.

(b) Hands-on test. There was a low positive correlation between the television test and the hands-on test. This correlation was less than desirable.

The experiment was too limited to permit any conclusions at this time with reference to the reliability and validity of the above results.

(5) Feedback. Because the examinees were tested four at a time and because the Telestrater equipment was not working, it was not possible to provide immediate feedback after each item.

(6) Eye fatigue. The television test and the responding mode required the examinees to stare continually at the television monitor. There were many complaints of eyestrain and there is some evidence that the afternoon television examinees performed more poorly than the morning television examinees.

Conclusions and Recommendations

(1) The Telestrator equipment as presently designed should be rejected because of the parallax problem.

(2) The television method appears to offer enough promise to warrant the testing of other response devices which do not have the parallax problem.

(3) There are many unknowns in television testing and the overall testing strategy, and the research effort needs to be greatly expanded such as:

(a) A more definitive comparison with hands-on tests.

(b) Research on the "immediate feedback" idea.

(c) Using alternative response devices.

(d) Comparison with slide-tape devices.

(e) Further research on eye fatigue.

ARI Distribution List

4 OASD (M&RA)
2 HQDA (DAMI-CSZ)
1 HQDA (DAPE-PBR
1 HQDA (DAMA-AR)
1 HQDA (DAPE-HRE-PO)
1 HQDA (SGRD-ID)
1 HQDA (DAMI-DOT-C)
1 HQDA (DAPC-PMZ-A)
1 HQDA (DACH-PPZ-A)
1 HQDA (DAPE-HRE)
1 HQDA (DAPE-MPO-C)
1 HQDA (DAPE-DW)
1 HQDA (DAPE-HRL)
1 HQDA (DAPE-CPS)
1 HQDA (DAFD-MFA)
1 HQDA (DARD-ARS-P)
1 HQDA (DAPC-PAS-A)
1 HQDA (DUSA-OR)
1 HQDA (DAMO-RQR)
1 HQDA (DASG)
1 HQDA (DA10-PI)
1 Chief, Consult Div (DA-OTSG), Adelphi, MD
1 Mil Asst. Hum Res, ODDR&E, OAD (E&LS)
1 HQ USARAL, APO Seattle, ATTN: ARAGP-R
1 HQ First Army, ATTN: AFKA-OI-TI
2 HQ Fifth Army, Ft Sam Houston
1 Dir, Army Stf Studies Ofc, ATTN: OA[illegible]A (DSP)
1 Ofc Chief of Stf, Studies Ofc
1 DCSPER, ATTN: CPS/OCP
1 The Army Lib, Pentagon, ATTN: RSB Chief
1 The Army Lib, Pentagon, ATTN: ANRAL
1 Ofc, Asst Sect of the Army (R&D)
1 Tech Support Ofc, OJCS
1 USASA, Arlington, ATTN: IARD-T
1 USA Rsch Ofc, Durham, ATTN: Life Sciences Dir
2 USARIEM, Natick, ATTN: SGRD-UE-CA
1 USATTC, Ft Clayton, ATTN: STETC-MO-A
1 USAIMA, Ft Bragg, ATTN: ATSU-CTD-OM
1 USAIMA, Ft Bragg, ATTN: Marquat Lib
1 US WAC Ctr & Sch, Ft McClellan, ATTN: Lib
1 US WAC Ctr & Sch, Ft McClellan, ATTN: Tng Dir
1 USA Quartermaster Sch, Ft Lee, ATTN: ATSM-TE
1 Intelligence Material Dev Ofc, EWL, Ft Holabird
1 USA SE Signal Sch, Ft Gordon, ATTN: ATSO-EA
1 USA Chaplain Ctr & Sch, Ft Hamilton, ATTN: ATSC-TE-RD
1 USATSCH, Ft Eustis, ATTN: Educ Advisor
1 USA War College, Carlisle Barracks, ATTN: Lib
2 WRAIR, Neuropsychiatry Div
1 DLI, SDA, Monterey
1 USA Concept Anal Agcy, Bethesda, ATTN: MOCA-WGC
1 USA Concept Anal Agcy, Bethesda, ATTN: MOCA-MR
1 USA Concept Anal Agcy, Bethesda, ATTN: MOCA-JF
1 USA Artic Test Ctr, APO Seattle, ATTN: STEAC-MO-ASL
1 USA Artic Test Ctr, APO Seattle, ATTN: AMSTE-PL-TS
1 USA Armament Cmd, Redstone Arsenal, ATTN: ATSK-TEM
1 USA Armament Cmd, Rock Island, ATTN: AMSAR-TDC
1 FAA-NAFEC, Atlantic City, ATTN: Library
1 FAA-NAFEC, Atlantic City, ATTN: Hum Engr Br
1 FAA Aeronautical Ctr, Oklahoma City, ATTN: AAC-44D
2 USA Fld Arty Sch, Ft Sill, ATTN: Library
1 USA Armor Sch, Ft Knox, ATTN: Library
1 USA Armor Sch, Ft Knox, ATTN: ATSB-DI-E
1 USA Armor Sch, Ft Knox, ATTN: ATSB-DT-TP
1 USA Armor Sch, Ft Knox, ATTN: ATSB-CD-AD
2 HQUSACDEC, Ft Ord, ATTN: Library
1 HQUSACDEC, Ft Ord, ATTN: ATEC–EX–E–Hum Factors
2 USAEEC, Ft Benjamin Harrison, ATTN: Library
1 USAPACDC, Ft Benjamin Harrison, ATTN: ATCP–HR
1 USA Comm–Elect Sch, Ft Monmouth, ATTN: ATS[illegible]–EA
1 USAEC, Ft Monmouth, ATTN: AMSEL–CT–HDP
1 USAEC, Ft Monmouth, ATTN: AMSEL–P[illegible]
1 USAEC, Ft Monmouth, ATTN: AMSEL–SI[illegible]B
1 USAEC, Ft Monmouth, ATTN: C, Facl Dev Br
1 USA Materials Sys Anal Agcy, Aberdeen, ATTN: AMXSY–P
1 Edgewood Arsenal, Aberdeen, ATTN: SAREA–BL–H
1 USA Ord Ctr & Sch, Aberdeen, ATTN: ATSL–TEM–C
2 USA Hum Engr Lab, Aberdeen, ATTN: Library/Dir
1 USA Combat Arms Tng Bd, Ft Benning, ATTN: Ad Supervisor
1 USA Infantry Hum Rsch Unit, Ft Benning, ATTN: Chief
1 USA Infantry Bd, Ft Benning, ATTN: STEBC–TE–T
1 USASMA, Ft Bliss, ATTN: ATSS–LRC
1 USA Air Def Sch, Ft Bliss, ATTN: ATSA–CTD–ME
1 USA Air Def Sch, Ft Bliss, ATTN: Tech Lib
1 USA Air Def Bd, Ft Bliss, ATTN: FILES
1 USA Air Def Bd, Ft Bliss, ATTN: STEBD–PO
1 USA Cmd & General Stf College, Ft Leavenworth, ATTN: Lib
1 USA Cmd & General Stf College, Ft Leavenworth, ATTN: ATSW–SE–L
1 USA Cmd & General Stf College, Ft Leavenworth, ATTN: Ed Advisor
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: DepCdr
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: CCS
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: ATCASA
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: ATCACO–E
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: ATCACC–CI
1 USAECOM, Night Vision Lab, Ft Belvoir, ATTN: AMSEL–NV–SD
3 USA Computer Sys Cmd, Ft Belvoir, ATTN: Tech Library
1 USA[illegible]RDC, Ft Belvoir, ATTN: STSFB–DQ
1 USA E[illegible]Sch, Ft Belvoir, ATTN: Library
1 USA Topographic Lab, Ft Belvoir, ATTN: ETL–TD–S
1 USA Topographic Lab, Ft Belvoir, ATTN: STINFO Center
1 USA Topographic Lab, Ft Belvoir, ATTN: ETL–GSL
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: CTD–MS
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATS–CTD–MS
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI–TE
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI–TEX–GS
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI–CTS–OR
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI–CTD–DT
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI–CTD–CS
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: DAS/SRD
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI–TEM
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: Library
1 CDR, HQ Ft Huachuca, ATTN: Tech Ref Div
2 CDR, USA Electronic Prvg Grd, ATTN: STEEP–MT–S
1 CDR, Project MASSTER, ATTN: Tech Info Center
1 Hq MASSTER, USATRADOC, LNO
1 Research Institute, HQ MASSTER, Ft Hood
1 USA Recruiting Cmd, Ft Sherdian, ATTN: USARCPM–P
1 Senior Army Adv., USAFAGOD/TAC, Elgin AF Aux Fld No. 9
1 HQ USARPAC, DCSPER, APO SF 96558, ATTN: GPPE–SE
1 Stimson Lib, Academy of Health Sciences, Ft Sam Houston
1 Marine Corps Inst., ATTN: Dean–MCI
1 HQUSMC, Commandant, ATTN: Code MTMT 51
1 HQUSMC, Commandant, ATTN: Code MPI–20
2 USCG Academy, New London, ATTN: Admission
2 USCG Academy, New London, ATTN: Library
1 USCG Training Ctr, NY, ATTN: CO
1 USCG Training Ctr, NY, ATTN: Educ Svc Ofc
1 USCG, Psychol Res Br, DC, ATTN: GP 1/62
1 HQ Mid–Range Br, MC Det, Quantico, ATTN: P&S Div

6 USATRADOC, Ft Monroe, ATTN: ATPR–AD
1 USATRADOC, Ft Monroe, ATTN: ATTS–EA
1 USA Forces Cmd, Ft McPherson, ATTN: Library
2 USA Aviation Test Bd, Ft Rucker, ATTN: STEBG–PO
1 USA Agcy for Aviation Safety, Ft Rucker, ATTN: Library
1 USA Agcy for Aviation Safety, Ft Rucker, ATTN: Educ Advisor
1 USA Aviation Sch, Ft Rucker, ATTN: PO Drawer O
1 HQUSA Aviation Sys Cmd, St Louis, ATTN: AMSAV–ZDR
2 USA Aviation Sys Test Act, Edwards AFB, ATTN: SAVTE–T
1 USA Air Def Sch, Ft Bliss, ATTN: ATSA TEM
1 USA Air Mobility Rsch & Dev Lab, Moffett Fld, ATTN: SAVDL–AS
1 USA Aviation Sch, Res Tng Mgt, Ft Rucker, ATTN: ATST–T–RTM
1 USA Aviation Sch, CO, Ft Rucker, ATTN: ATST–D–A
1 HQ, DARCOM, Alexandria, ATTN: AMXCD–TL
1 HQ, DARCOM, Alexandria, ATTN: CDR
1 US Military Academy, West Point, ATTN: Serials Unit
1 US Military Academy, West Point, ATTN: Ofc of Milt Ldrshp
1 US Military Academy, West Point, ATTN: MAOR
1 USA Standardization Gp, UK, FPO NY, ATTN: MASE–GC
1 Ofc of Naval Rsch, Arlington, ATTN: Code 452
3 Ofc of Naval Rsch, Arlington, ATTN: Code 458
1 Ofc of Naval Rsch, Arlington, ATTN: Code 450
1 Ofc of Naval Rsch, Arlington, ATTN: Code 441
1 Naval Aerospc Med Res Lab, Pensacola, ATTN: Acous Sch Div
1 Naval Aerospc Med Res Lab, Pensacola, ATTN: Code L51
1 Naval Aerospc Med Res Lab, Pensacola, ATTN: Code L5
1 Chief of NavPers, ATTN: Pers–OR
1 NAVAIRSTA, Norfolk, ATTN: Safety Ctr
1 Nav Oceanographic, DC, ATTN: Code 6251, Charts & Tech
1 Center of Naval Anal, ATTN: Doc Ctr
1 NavAirSysCom, ATTN: AIR–5313C
1 Nav BuMed, ATTN: 713
1 NavHelicopterSubSqua 2, FPO SF 96601
1 AFHRL (FT) William AFB
1 AFHRL (TT) Lowry AFB
1 AFHRL (AS) WPAFB, OH
2 AFHRL (DOJZ) Brooks AFB
1 AFHRL (DOJN) Lackland AFB
1 HQUSAF (INYSD)
1 HQUSAF (DPXXA)
1 AFVTG (RD) Randolph AFB
3 AMRL (HE) WPAFB, OH
2 AF Inst of Tech, WPAFB, OH, ATTN: ENE/SL
1 ATC (XPTD) Randolph AFB
1 USAF AeroMed Lib, Brooks AFB (SUL–4), ATTN: DOC SEC
1 AFOSR (NL), Arlington
1 AF Log Cmd, McClellan AFB, ATTN: ALC/DPCRB
1 Air Force Academy, CO, ATTN: Dept of Bel Scn
5 NavPers & Dev Ctr, San Diego
2 Navy Med Neuropsychiatric Rsch Unit, San Diego
1 Nav Electronic Lab, San Diego, ATTN: Res Lab
1 Nav TrngCen, San Diego, ATTN: Code 9000–Lib
1 NavPostGraSch, Monterey, ATTN: Code 55Aa
1 NavPostGraSch, Monterey, ATTN: Code 2124
1 NavTrngEquipCtr, Orlando, ATTN: Tech Lib
1 US Dept of Labor, DC, ATTN: Manpower Admin
1 US Dept of Justice, DC, ATTN: Drug Enforce Admin
1 Nat Bur of Standards, DC, ATTN: Computer Info Section
1 Nat Clearing House for MH–Info, Rockville
1 Denver Federal Ctr, Lakewood, ATTN: BLM
12 Defense Documentation Center
4 Dir Psych, Army Hq, Russell Ofcs, Canberra
1 Scientific Advsr, Mil Bd, Army Hq, Russell Ofcs, Canberra
1 Mil and Air Attache, Austrian Embassy
1 Centre de Recherche Des Facteurs, Humaine de la Defense Nationale, Brussels
2 Canadian Joint Staff Washington
1 C/Air Staff, Royal Canadian AF, ATTN: Pers Std Anal Br
3 Chief, Canadian Def Rsch Staff, ATTN: C/CRDS(W)
4 British Def Staff, British Embassy, Washington
1 Militaerpsykologisk Tjeneste, Copehagen
1 Military Attache, French Embassy, ATTN: Doc Sec
1 Medecin Chef, C.E.R.P.A.–Arsenal, Toulon/Naval France
1 Prin Scientific Off, Appl Hum Engr Rsch Div, Ministry of Defense, New Delhi
1 Pers Rsch Ofc Library, AKA, Israel Defense Forces
1 Ministeris van Defensie, DOOP/KL Afd Sociaal Psychologische Zaken, The Hague, Netherlands