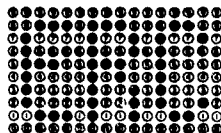


407460



EDUCATIONAL RESEARCH CORPORATION

TEN CRAIGIE STREET • CAMBRIDGE 38 • MASSACHUSETTS

BOSTON TELEPHONE CONNECTING ALL DEPARTMENTS UNIVERSITY 4-2400 CABLE ADDRESS EDRESCORP BOSTON

ESD-TR-61-25-3

THE DEVELOPMENT OF A PROFICIENCY TEST FOR SAGE IDENTIFICATION BRANCH PERSONNEL

Robert J. Curran

Phillip J. Rulon, Principal Investigator
ERC Project 53

FINAL REPORT
Contract AF 19(604)-5195

June 1961

Operational Applications Office
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE

Bedford, Massachusetts

AD 1
DDC FILE COPY

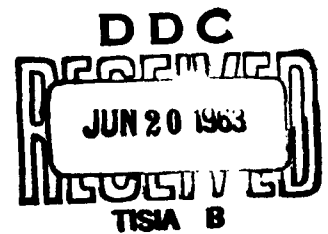
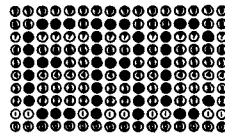
407460

5.60

(5) 283 600

63-4-1





Reproduction in whole or in part is permitted
for any purpose of the United States Government

(4) 260

(5) 283 600

(7)-(8) NA

(18)

(19)

ESD-TR-61-25-3

(6)

THE DEVELOPMENT OF A PROFICIENCY TEST FOR
SAGE IDENTIFICATION BRANCH PERSONNEL,

(19) by

Robert J. Curran,

Phillip J. Rulon, Principal Investigator

ERC Project 53

(9)

FINAL ^{rept.} REPORT

(15)

Contract AF 19604-5195

(16)

Project No. 1975

(17)

Task Order 76892

(11)

June 1961,

(12)

16 p.

(13)

NA

(14)

NA

Operational Applications Office
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE

Bedford, Massachusetts

(20) 22.

(21) NA

Ydc

ABSTRACT

↙ An earlier part of this project was to develop a prototype, job-oriented performance test for the SAGE positions of Identification Officer and Technician. The primary goals of the present project were to up-date the test and modify it so that (a) it could be administered and scored by Air Force personnel and (b) further forms of the test applicable to other SAGE Direction Centers could be constructed by in-service personnel.

In order to replicate the job as closely as possible, photographs were made of Situation and Digital Information Displays under actual job conditions. Even though the photographs were made at the Boston Air Defense Sector, the situations were found to be applicable to other sectors.

The test was administered at ~~three SAGE Direction Centers~~ and at the SAGE Training School. Altogether, seventy-one (71) subjects were tested.

↘ The techniques employed in the construction and modification of the test seem to be useful for the development of similar tests for local needs. The tests were administered and scored with little difficulty by a number of different people, some of these being Air Force personnel. Within the limits of sample size and of the criterion information available, the form of test as modified appeared to be reliable and valid. ↙

TABLE OF CONTENTS

	<u>Page</u>
Foreword	ii
Introduction.	1
Description of the Test	4
Administration of the Test.	6
Validity	13
Reliability	15
Conclusions.	16

FOREWORD

This is one of three terminal reports under Contract AF 19(604)-5195, Development of Proficiency Measures for Operator Positions in the Data Processing Sub-system of SAGE. The other two reports deal with the Track Initiation and Monitoring positions and with the Manual Data Inputs positions.

The tests reported on herein were developed from an earlier prototype generated under Contract AF 41(657)-95, under which a number of workers contributed to the development and techniques employed. Among those contributors P. J. Rulon and R. F. Schweiker were particularly active. Ruth W. Bredon had the direct responsibility for constructing the prototype test.

Major H. R. Burkhart and Captain J. E. Sullivan of the BOADS Direction Center Training Office supported the development and administration efforts of this testing. Members of the BOADS Identification served both as critics and subjects.

Captain Raigh Mason, Captain Robert Leestamper, Mr. Ira Goldstein, and Mr. Arthur Marcus, all of the Operator Performance Branch, Operational Applications Laboratory, Air Force Cambridge Research Center, assisted in administering the test and inscuring background information.

INTRODUCTION

This report describes the development of a test to measure the proficiency of SAGE Identification personnel. The test is a refinement of two earlier prototypes which were used with the Experimental SAGE Sector personnel. The prototype tests were developed under Contract No. AF 41(657)-95, Operational Applications Laboratory, Air Force Cambridge Research Center. The report on this prototype development is entitled "Development of Prototype Proficiency Tests for SAGE Operators: Identification Personnel" (AFCRC-TN-58-62).

In up-dating and refining the prototype test, several factors were taken into consideration. These were: (1) In active Direction Centers, all Identification personnel were required to be familiar with the IDO, IDT 1, and IDT 2 consoles. This requirement meant that separate tests for IDO's and IDT's were unnecessary. (2) In the original tests, situational and digital displays were depicted by drafted presentations. In order to attain a higher degree of verisimilitude, actual console displays were photographed. (3) The primary considerations were to construct a test which was easy to administer and score by untrained personnel. (4) Where operational procedures and regulations conflicted, it was decided that operational procedures were to be deemed correct. (5) Since one of the goals of the study was to administer the tests at a variety of direction centers, it was decided to utilize questions which had wide application. At the same time, it was decided to use displays based on the geographical boundaries of one direction center; this necessitated the inclusion of a

GEOREF map in direction centers other than the one depicted in the displays. Needless to say, the acceptance of these factors ruled out other possible approaches. For instance, it was found difficult to construct a test which might measure proficiency and at the same time be diagnostic under the above circumstances.

In addition to the considerations listed above that somewhat limited the scope of the test, other limitations also developed. The time limit of the test was originally based on the fact that Identification personnel were required to identify a Pending track within one minute of its original presentation. By the time the test was first administered, this requirement no longer held, which to some extent invalidated the time limits of the test. All Identification personnel were required to perform certain administrative duties such as the preparation of reports; no attempt was made to include such activities in the test.

One of the serious defects of the test, as it was developed, was the failure to include displays of emergency situations. To a large extent, photographs of such situations can only be obtained during training missions which occur sporadically. After an extensive period of waiting, such photographs were obtained but were subsequently made useless by poor photographic development. Because of time limitations, it was impossible to obtain more photographs. In the development of any future test, it is strongly recommended that such emergency situations be tested.

At each direction center, various job supports were employed in order to accomplish successfully certain tasks. These job supports were not standard from one direction center to another; consequently it was

impossible to include all of them in a test that was to be used at different direction centers. In time, these job supports will probably be standardized. At that time, they should be incorporated into any proficiency measures.

The test was originally constructed to measure proficiency on Program Model No. 3. At the time the test was administered, not all the provisions of Program Model No. 3 were actually being carried out. As a result, five questions were eliminated in the final scoring of the test.

This document reports on the administration of the test at three direction centers in the 26th Air Division and at the SAGE Training School at Richards-Gebaur Air Force Base. The active direction centers were at different stages of training and, in fact, were using different program models at the times the tests were administered. While the test is based on actual operational procedures, the personnel at the SAGE Training School were familiar only with regulations since they had no opportunity to observe operating procedures. Low scores by trainees should not be interpreted as indicating inadequate training.

DESCRIPTION OF THE TEST

When revisions of the prototype were first being considered, it was decided that areas of performance should be tested in direct proportion to the importance of the area relative to other aspects of the job. After observing a sample of operators, it was found that this aim would be difficult to achieve. For example, while some operators utilize all the capabilities of the console, other operators restrict their choices to very few procedures. Some operators may rarely use the correlation box, but do just as well, or better, than those who frequently use it. Consequently, it was impossible to allocate a certain percentage of questions to such areas as the special expanded displays or the correlation boxes.

While most of the average operator's day was taken up with the normal air traffic in his vicinity, he should have definite knowledge of actions to perform in emergency situations. Photographs of such emergency situations were obtained but were later destroyed in the developing with the result that they were not included in the test.

The selection of the situations to be photographed was made in conference with experts in the field of Identification. The photographs were taken during a ten-day period at the Boston Air Defense Sector while operations were going on. Photographs were taken of situational and digital displays. The final selection of items was also made in cooperation with these same experts and supposedly represented a good sample of an operator's job..

The test in its final form consisted of fifty-four items based on situation displays and twenty-two items based on digital displays. With

the exception of "Hostile", "Big Photo", and "Keystone" identifications, all possible areas of choice and decision-making were included in the test. The regulations governing identification of aircraft by speed and altitude were changed after the test items concerning digital displays were prepared. These changes made the questions based on digital displays somewhat easier than they were originally intended to be.

In the scoring of the test, the first item was used as a sample questions and was not scored. After the original administration of the test at the Boston Air Defense Sector, five other questions were not scored. Three were liminated because the photographs were difficult to interpret and two were eliminated because there was disagreement about the possible right answers.

ADMINISTRATION OF THE TEST

The test was originally administered at the Boston Air Defense Sector, Stewart Air Force Base, New York, in June of 1959. Subsequently, it was administered to the personnel of the New York Air Defense Sector, McGuire Air Force Base, New Jersey. Then, it was administered to training personnel at the SAGE Training School, Richards-Gebaur Air Force Base, Missouri, in December of 1959. Finally, it was administered to SAGE Identification personnel at the Syracuse Air Defense Sector, Syracuse Air Force Station, New York, in January of 1960.

After the administration of the test at the Boston Air Defense Sector, a map of that area was included with each test to be used as a reference by individuals not assigned to that base. While this aid probably helped Identification personnel from other sectors, it probably did not fully compensate them for having to take a test designed for another sector. Even though the operational procedures are supposedly standard for each sector, there are situations which are unique for each sector, including such factors as air routes and checkpoints. These differences probably allowed the personnel from the Boston Air Defense Sector some advantages in taking the test.

Originally a time limit for completion of the test was set at seventy-five minutes, allowing one minute per identification, much the same as the rule that was in operation in actual practice. It was found that this time period was much too generous and at the last two test sites, the time limit for the test was reduced to fifty minutes. While all the individuals at the Syracuse Air Defense Sector completed the test within these new time limits,

four of the personnel at Richards-Gebaur did not finish the test.

It should be noted that the personnel at the BOADS and The NYADS operate in coastal areas which require a great deal of vigilance in their air surveillance efforts. The Syracuse Air Defense Sector became operational as an inland sector and much later than the BOADS and the NYADS. Identification at the Syracuse sector is generally under control of the computer, and performed automatically. The Richards-Gebaur SAGE Training School personnel were being taught procedures based on a new program model and were also being taught "book" procedures rather than actual operational procedures. As can be seen, these groups are by no means comparable.

Altogether, twenty-two Identification personnel were tested at the Boston Air Defense Sector, fourteen were tested at the New York Air Defense Sector, twelve were tested at the Syracuse Air Defense Sector, and twenty-four were tested at the SAGE Training School. The BOADS and NYADS samples include some cross-trained personnel.

At the Boston Air Defense Sector, each of the examinees was interviewed after completing the test concerning his attitudes toward the items and the test in general. At the other bases, the operators were asked to submit written criticisms, if any, after they completed the test. Other than the five questions which were eventually eliminated, there was very little criticism of the items. On the other hand, there seemed to be a general feeling that the test was an adequate representation of the job.

Table I presents the percentages of correct responses for each item in terms of the upper half and lower half of the distributions and for the

total sample. To obtain these percentages, the entire sample was divided into an upper and a lower half according to total score. Four individuals fell at the median. Two of these individuals were randomly assigned to the upper half, and two were assigned to the lower half. For all practical purposes, the upper half consists almost completely of cases obtained at the Boston and New York Air Defense Sectors. The lower half, on the other hand, consists almost completely of cases obtained from the testing at the Syracuse Air Defense Sector and at the SAGE Training School at Richards-Gebaur Air Force Base. Two individuals from the Boston Air Defense Sector obtained scores that placed them in the lower half of the distribution. One individual from the Syracuse Air Defense Sector and one individual from the SAGE Training School achieved scores that placed them in the upper half of the distribution.

TABLE I
Percentages of Correct Responses in the Upper Half,
Lower Half, and Total Sample

<u>Item Number</u>	<u>Upper Half of Sample</u>	<u>Lower Half of Sample</u>	<u>Total Sample</u>
2	77.	52.	64.5
3	80.	26.	53.0
5	63.	46.	54.5
6	34.	34.	34.0
7	80.	31.	55.5
8	74.	46.	60.0
9	86.	86.	86.0
10	77.	63.	70.0
11	74.	49.	61.5
12	100.	71.	85.5

Table I (continued):

<u>Item Number</u>	<u>Upper Half of Sample</u>	<u>Lower Half of Sample</u>	<u>Total Sample</u>
13	74.	51.	62.5
14	46.	29.	37.5
15	80.	91.	85.5
16	94.	71.	82.5
17	51.	17.	34.0
18	63.	23.	43.0
19	57.	43.	50.0
20	49.	23.	36.0
21	43.	49.	46.0
22	83.	49.	66.0
25	60.	09.	34.5
26	11.	71.	41.0
27	89.	43.	66.0
28	86.	74.	80.0
29	60.	14.	37.0
30	60.	09.	34.5
31	77.	69.	73.0
33	37.	37.	37.0
34	37.	31.	34.0
35	66.	51.	58.0
36	100.	51.	75.5
37	74.	60.	67.0
38	74.	66.	70.0
39	69.	57.	63.0
40	49.	23.	36.0
41	69.	69.	69.0
42	69.	37.	53.0
43	74.	17.	45.5
44	54.	26.	40.0
45	77.	60.	68.5

Table I (continued):

<u>Item Number</u>	<u>Upper Half of Sample</u>	<u>Lower Half of Sample</u>	<u>Total Sample</u>
46	74.	71.	72.5
47	91.	34.	62.5
48	71.	49.	60.0
49	43.	23.	33.0
50	89.	66.	77.5
51	86.	54.	70.0
52	51.	29.	40.0
53	06.	09.	07.5
54	94.	80.	87.0
55	37.	34.	35.5
56	89.	80.	84.5
57	40.	31.	35.5
58	57.	43.	50.0
59	94.	68.	81.0
60	71.	43.	57.0
61	54.	34.	44.0
62	57.	11.	34.0
63	34.	11.	22.5
64	51.	34.	42.5
65	60.	51.	55.5
66	57.	37.	47.0
67	80.	83.	81.5
68	80.	31.	55.5
69	77.	51.	64.0
71	54.	46.	50.0
72	74.	29.	51.5
73	71.	51.	61.0
74	77.	43.	60.0
75	94.	71.	82.5
76	83.	60.	71.5

Thirteen items seemed to provide no or little discrimination between the upper and lower groups. These items are numbers, 6, 9, 33, 34, 38, 39,

41, 46, 55, 56, 57, 65, and 71. Five items proved to be negatively discriminating. These items are 15, 21, 26, 53, and 67. On these negatively discriminating items, however, the differences between the upper and lower groups were slight with one exception. Where such negative discrimination did occur, the items may require revision. Table II presents a tabulation of the number of items on the test of specified difficulty. The average difficulty for all items was approximately 56%. A revision of the test should eliminate the very difficult items. Easy items are useful from a motivational standpoint, especially if they are placed at the beginning of a test.

TABLE II
Number of Items on ID Proficiency Test of Specified Difficulty

<u>Percent of Subjects Choosing Correct Response</u>	<u>Number of Items</u>
00 - 10	1
11 - 20	0
21 - 30	2
31 - 40	16
41 - 50	10
51 - 60	12
61 - 70	15
71 - 80	6
81 - 90	9
91 and over	0

Table III provides a summary of the mean scores of Identification personnel at the different bases. These mean scores fall within a pattern that one might expect in view of the differences in length of experience of the various groups. The relatively high scores of the cross-trained personnel at the BOADS and the NYADS may seem to be out of line with others scores, but they may be explained in terms of the efficacy of the cross-training program at each of these direction centers. Most of these cross-trained personnel also had considerable experience in a SAGE direction center, which may be a contributing factor.

TABLE III
Summary of Mean Scores of Several Groups

<u>Group</u>	<u>Section I Situation Displays</u>	<u>Section II Digital Displays</u>	<u>Total Scores</u>
All ID Personnel (BOADS)* (N = 22)	33.0	13.9	49.6
All ID Personnel (NYADS)* (N = 14)	31.0	15.1	46.1
All ID Personnel (R-G AFB)* (N = 24)	21.3	11.6	32.9
All ID Personnel (SYADS)* (N = 12)	21.7	09.4	31.1
BOADS IDO's (N = 4)	31.0	15.8	46.8
NYADS IDO's (N = 6)	37.0	15.0	52.0
R-G AFB IDO's (N = 13)	21.6	12.5	34.2
SYADS IDO (N = 1)	24.0	02.0	26.0
BOADS IDT's (N = 8)	33.1	12.5	45.6
NYADS IDT's (N = 6)	31.5	15.2	46.7
R-G AFB IDT's (N = 11)	21.0	10.5	31.5
SYADS IDT's (N = 11)	21.5	10.1	31.5
BOADS CT ID Personnel (N = 10)*	31.7	13.5	45.2
NYADS CT ID Personnel (N = 2)	29.5	12.5	42.0
Total Sample (Average)	26.5	12.6	39.3

*BOADS = Boston Air Defense Sector

*NYADS = New York Air Defense Sector

*R-G = Richards-Gebaur AFB, SAGE Training School

*SYADS = Syracuse Air Defense Sector

*CT = Cross-trainees

VALIDITY

If one accepts the fact that the photographs are reasonable simulations of the displays encountered on the job, and the fact that the photographs were taken during actual operational activities, validity is established by these facts alone. On the other hand, if one insists that the test scores be related to criterion data, it is difficult to establish validity. There are no other measures of job proficiency that simulate the conditions of the job as closely as this test does. In addition to the small number of cases involved, there was a general lack of adequate criterion information which might be employed to establish validity.

Nevertheless, attempts were made to relate the test scores to other data in order to estimate validity. All of these attempts were inconclusive or without statistical significance. Twelve personnel, consisting of six Identification officers and six Identification technicians, were rated on an evaluation scale at the New York Air Defense Sector. The rating scale had a total maximum of 45 points. Four of the officers were rated at the total maximum. On the other hand, all of the Identification technicians were rated considerably lower. Their test scores were correlated with their scores on the rating scale and a correlation of $-.467$ was obtained. For a sample size of 12, this correlation is not significantly different from zero. At the Boston Air Defense Sector, twenty-two cases consisting of six Identification officers and sixteen Identification technicians had their experience evaluated first, in number of

days that they had served in the Identification section, and secondly in terms of numbers of months that they had worked in the direction center. This experience was correlated with their test scores. A correlation coefficient of .056 was obtained between length of SAGE experience and test scores, and a correlation of .59 was obtained between number of days on an Identification job and test scores. No other pertinent data were available for even these small numbers of cases.

Therefore, if one insists on a statistically significant basis for the establishment of validity, the validity of this test is not proven. On the other hand, if one accepts as valid a test that requires decisions as they are made on the job, and which has been submitted to subject-matter experts and approved by them as representing the job, then the test is valid.

RELIABILITY

Two methods were used to estimate the reliability of the test. The split-half technique and the Kuder-Richardson Formula 20 were both employed. The Spearman-Brown formula was applied to both coefficients. Using the Kuder-Richards Formula 20, a reliability coefficient of .973 was obtained; using the split-half technique, a reliability coefficient of .0807 was obtained.

CONCLUSIONS

The feasibility of using photographs in the construction of proficiency tests for SAGE Identification Branch personnel has been demonstrated. In view of the changes in program models and the many revisions in operational procedures, photography would seem to be an ideal method of keeping current the display aspects of SAGE proficiency tests.

Although there are no statistical indications as to the validity of the SAGE Identification personnel proficiency test, its continued use is recommended. The approach used in test construction is felt to be adequate to justify validity by definition.

Changes in program models and operational procedures will undoubtedly render some of the items obsolete. However, up-dating of the test is feasible, since it is not felt that the types of decisions that the Identification Branch personnel make will be changed, but the way in which these decisions are implemented.

The test, in its present form, proved easy to administer and score, even by untrained personnel.