(5) 81 640

N-63-4-3

(1)

Technical Research Note 129

EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST

Scale-3

U. S. ARMY
PERSONNEL
RESEARCH
OFFICE

$1.60

U. S. ARMY PERSONNEL RESEARCH OFFICE

An activity of the Chief, Research and Development

Charles S. Gersoni
Colonel, GS
Commanding

J. E. Uhlaner
Director, Research Laboratories

Hubert E. Brogden
Chief Scientist

AD Div 23/1, 28/4

U. S. Army Personnel Research Office, OCRD, DA
EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST by Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson. December 1962. Rept on Input Quality c-13 Proj.--18 p. incl. tables, 22 Ref. (USAPRO Technical Research Note No. 129).
(DA Project 0J95-60-001) Unclassified Report

An exploratory study of the sequential item test technique applied to Army screening test content was made to gain some knowledge of its potential as a useful method of reducing Army testing time. The SIT was compared in field tryout with other short test techniques and with operational Army screening tests as to feasibility of administration and screening effectiveness. The SIT was moderately satisfactory in terms of its correlation with conventional tests. However, it revealed a number of drawbacks (increased time and cost in construction, administration, and scoring, and greater difficulty for low level personnel to understand) which led to the conclusion that further experimentation with the test in its present hand-scored paper-and-pencil format is not justified. The technique appears adaptable, however, to administration by testing machine, a procedure which would eliminate most of the difficulties encountered.

UNCLASSIFIED
Human Resources Research
--Personnel Utilization

AD Div 23/1, 28/4

U. S. Army Personnel Research Office, OCRD, DA
EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST by Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson. December 1962. Rept on Input Quality c-13 Proj.--18 p. incl. tables, 22 Ref. (USAPRO Technical Research Note No. 129).
(DA Project 0J95-60-001) Unclassified Report

An exploratory study of the sequential item test technique applied to Army screening test content was made to gain some knowledge of its potential as a useful method of reducing Army testing time. The SIT was compared in field tryout with other short test techniques and with operational Army screening tests as to feasibility of administration and screening effectiveness. The SIT was moderately satisfactory in terms of its correlation with conventional tests. However, it revealed a number of drawbacks (increased time and cost in construction, administration, and scoring, and greater difficulty for low level personnel to understand) which led to the conclusion that further experimentation with the test in its present hand-scored paper-and-pencil format is not justified. The technique appears adaptable, however, to administration by testing machine, a procedure which would eliminate most of the difficulties encountered.

UNCLASSIFIED
Human Resources Research
--Personnel Utilization

AD Div 23/1, 28/4

U. S. Army Personnel Research Office, OCRD, DA
EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST by Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson. December 1962. Rept on Input Quality c-13 Proj.--18 p. incl. tables, 22 Ref. (USAPRO Technical Research Note No. 129).
(DA Project 0J95-60-001) Unclassified Report

An exploratory study of the sequential item test technique applied to Army screening test content was made to gain some knowledge of its potential as a useful method of reducing Army testing time. The SIT was compared in field tryout with other short test techniques and with operational Army screening tests as to feasibility of administration and screening effectiveness. The SIT was moderately satisfactory in terms of its correlation with conventional tests. However, it revealed a number of drawbacks (increased time and cost in construction, administration, and scoring, and greater difficulty for low level personnel to understand) which led to the conclusion that further experimentation with the test in its present hand-scored paper-and-pencil format is not justified. The technique appears adaptable, however, to administration by testing machine, a procedure which would eliminate most of the difficulties encountered.

UNCLASSIFIED
Human Resources Research
--Personnel Utilization

AD Div 23/1, 28/4

U. S. Army Personnel Research Office, OCRD, DA
EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST by Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson. December 1962. Rept on Input Quality c-13 Proj.--18 p. incl. tables, 22 Ref. (USAPRO Technical Research Note No. 129).
(DA Project 0J95-60-001) Unclassified Report

An exploratory study of the sequential item test technique applied to Army screening test content was made to gain some knowledge of its potential as a useful method of reducing Army testing time. The SIT was compared in field tryout with other short test techniques and with operational Army screening tests as to feasibility of administration and screening effectiveness. The SIT was moderately satisfactory in terms of its correlation with conventional tests. However, it revealed a number of drawbacks (increased time and cost in construction, administration, and scoring, and greater difficulty for low level personnel to understand) which led to the conclusion that further experimentation with the test in its present hand-scored paper-and-pencil format is not justified. The technique appears adaptable, however, to administration by testing machine, a procedure which would eliminate most of the difficulties encountered.

UNCLASSIFIED
Human Resources Research
--Personnel Utilization

(4) # 1.60 (5) 81 640
(18) APRO
(19) Technical Research Note 129.

(6) EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST

(7) NA
(8) NA
(9) NA

(10)

Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson

~~A. G. Bayroff, Task Leader~~

(20) 41.
(21) NA

(11) Dec 62,
(12) 11p.
(13) NA
(14) NA
(15) NA
(16) Proj 0>95 ± 60 001 ±

Submitted by Edmund F. Fuchs
Chief, Military Selection Research Laboratory

Approved by

J. E. Uhlaner
Director, Research Laboratories

Hubert E. Brogden
Chief Scientist

(17) NA

Charles S. Gersoni
Colonel, GS
Commanding

Army Project Number
OJ95-60-001

Input Quality c-13

December 1962

USAPRO Technical Research Reports and Technical Research Notes 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.

PREFACE

The present publication reports on a portion of Subtask c, "Methods for Improving Testing," of the INPUT QUALITY Task, FY 1963 Work Program. The entire research Task is responsive to special requirements of the Department of Defense AFES Policy Board and the Deputy Chief of Staff for Personnel.

Research on screening and induction techniques is a continuing effort which must reflect developing military policy and organization involving all the armed services. Changes in the screening system, change in the input population, new developments in techniques of measurement, of administering and scoring tests, all are taken into account. Current research embraces the following activities: (1) devising methods to increase effectiveness of overall screening through new tests and test content; (2) improving effectiveness of short tests for the differential measurement of aptitude areas for the middle ability level; (3) exploring the feasibility of very short, limited-range tests; and (4) devising new approaches to the detection of deliberate failures. Methodological research is conducted with a view to the development of future operational measures. Applications of computer technology and automation are explored which may permit more efficient screening and in addition provide a speedily available basis for quality manpower control.

EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST

BRIEF

Requirement:

Sequential item techniques, intended to reduce the length of tests needed to provide reliable scores, required evaluation as a potentially useful method of reducing Army testing time.

Procedure:

The Army Sequential Item Test (SIT), adapted to Army screening test content, was administered to samples of selective service registrants and inductees at Armed Forces Examining Stations and at Training Centers. The test was compared with operational Army screening tests as to feasibility of administration and effectiveness as a means of assessing military trainability.

Findings:

From a technical standpoint, the Sequential Item Test was moderately satisfactory. Relationships of SIT subtest scores to total Armed Forces Qualification Test scores and to tests of the Army Classification Battery corresponding in content to the SIT subtests were reasonably high. However, the test as constructed yielded a disproportionate number of maximum scores.

The test represented no net saving in testing time, since the complicated instructions took up more time than is required to administer a conventional test. More important, examinee failure to follow the sequential routing as instructed resulted in a large proportion of answer sheets that could not be scored.

Utilization of Findings:

Further experimentation with the Sequential Item Test in its present hand-scored paper-and-pencil format was not considered justified. However, the technique appears adaptable to administration by testing machine, a procedure which would eliminate most of the drawbacks encountered in the present study.

EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST

CONTENTS

TABLES

FIGURES

EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST

OBJECTIVE

One mission of the Input Quality Task is to explore new approaches to the assessment of general military trainability. A sequential item test was one such approach chosen for exploratory study, principally because of the saving in testing time it might afford without loss of effectiveness. The present study was undertaken to gain some knowledge of the technical promise of this novel technique, particularly in comparison with other short test techniques being studied.

Work done by Krathwohl and Huyser (1956) of Michigan State University served as a point of departure for this study. Their sequential item technique provided for a range of multiple-choice items presented at each of a limited number of "stages". In a given stage, an examinee answered only that item to which he had been directed--based on the p-value of his response to the item he did in the previous stage. Thus, he did not have to work on many items which were entirely inappropriate to his ability level, a feature which had the possibility of reducing both testing time and error in the scores. Examinee's choice of alternative in one item guides him to the item he is to work on next--the "self-routing" feature of the sequential item technique. For Krathwohl and Huyser's 6-item version of a 60-item conventional type test, a product-moment correlation coefficient of .76 with the 60-item parent test was obtained on a sample of 100 college freshmen.

An adaptation of Krathwohl and Huyser's methods was used in the construction of four subtests of an Army Sequential Item Test (SIT) covering the content areas and difficulty levels of the Armed Forces Qualification Test (AFQT) (Bayroff, Thomas, and Anderson, 1960). An exploratory field tryout was conducted on two of the subtests, Verbal (SIT-VE) and Arithmetic Reasoning (SIT-AR), primarily to determine the feasibility of the technique and to obtain some indication of the significance of the scores. The purpose of the present report is to present results of this field tryout and to evaluate the technique as a basis for further research.

CONSTRUCTION OF EXPERIMENTAL ARMY SEQUENTIAL ITEM TEST

The model used in designing each SIT subtest was a 6-stage model, a modified form of Krathwohl and Huyser's 8-stage model (one of several they studied). An examinee was required to take a total of six items, each item being at a different level of difficulty. The SIT has four independent subtests, each of homogeneous item content--Verbal, Arithmetic Reasoning, Tool Functions, and Spatial Relations. To allow for all possible routes in the 6-stage test, 26 items were required for each subtest. However, 37 items were added to each subtest to conceal clues to correct alternatives, resulting in a total of 63 items for each of the four subtests.

For each SIT subtest, all examinees take the same item in the first stage. The examinee's routing to each subsequent stage is determined by the alternative he selects. The p-value of the first item is approximately the average (.70) of the highest and lowest p-values of the test. If an examinee marks this item correctly, he is routed to a more difficult item at stage 2. Of the three incorrect alternatives to this item, two are about the same in drawing power. If the examinee marks either of these, he is routed to a stage 2 item of similar p-value. The other incorrect alternative to the first item has very poor pulling power. Examinees who mark this alternative are routed to an item of higher p-value. This procedure was adopted in an attempt to minimize initial errors of classification. Differentiation among incorrect alternatives was followed only in routing from first to second stage items; beginning with Stage 2, the examinee who passes one item is routed to a more difficult item, and the examinee who marks any of the incorrect alternatives goes to an easier item. The procedure is presented graphically in Figure 1.

Each alternative in all stages--except the last--is numbered. Alternatives of items in the final stage are lettered, indicating the end of the test. The process of routing to a more or a less difficult item permits correction of previous errors of classification, enabling an examinee finally to approach his "true level of ability" by the sixth stage. Since each subtest consists of homogeneous items scaled in difficulty, an examinee's score is dependent both upon which item he attempts in stage 6, and upon his performance on that item. If he marks this item correctly, his score is the scale value of the item. If he marks it incorrectly, his score is one point less. Each subtest score ranges from 0 to 9. This system is a modification of the one preferred by Krathwohl and Huyser over several others involving weighting procedures. A specially designed answer sheet and directions for administration were prepared for an experimental tryout of SIT-VE and SIT-AR.

FIELD TRYOUT OF THE TEST

The complexities of the routing instructions in the SIT made it desirable to determine whether examinees could follow them and if so with what results. Also of concern were the relationships among SIT-VE, SIT-AR, their ACB counterparts, and the AFQT. The SIT-VE and SIT-AR subtests were administered to available groups at certain of the Armed Forces Examining Stations--Denver, Colorado and Columbia, South Carolina--and Training Divisions--Ft. Carson, Colorado and Ft. Jackson, South Carolina. Because of the exploratory nature of this feasibility tryout, no attempt at stratification on AFQT was made in selecting examinees. The cases obtained did, however, show representation in all AFQT categories. Recorded operational AFQT 5 or 6 scores were obtained for all examinees. Recorded operational Army Classification Battery Verbal and Arithmetic Reasoning test scores (ACB-VE and ACB-AR) were obtained for roughly half the SIT examinees to whom ACB was given.

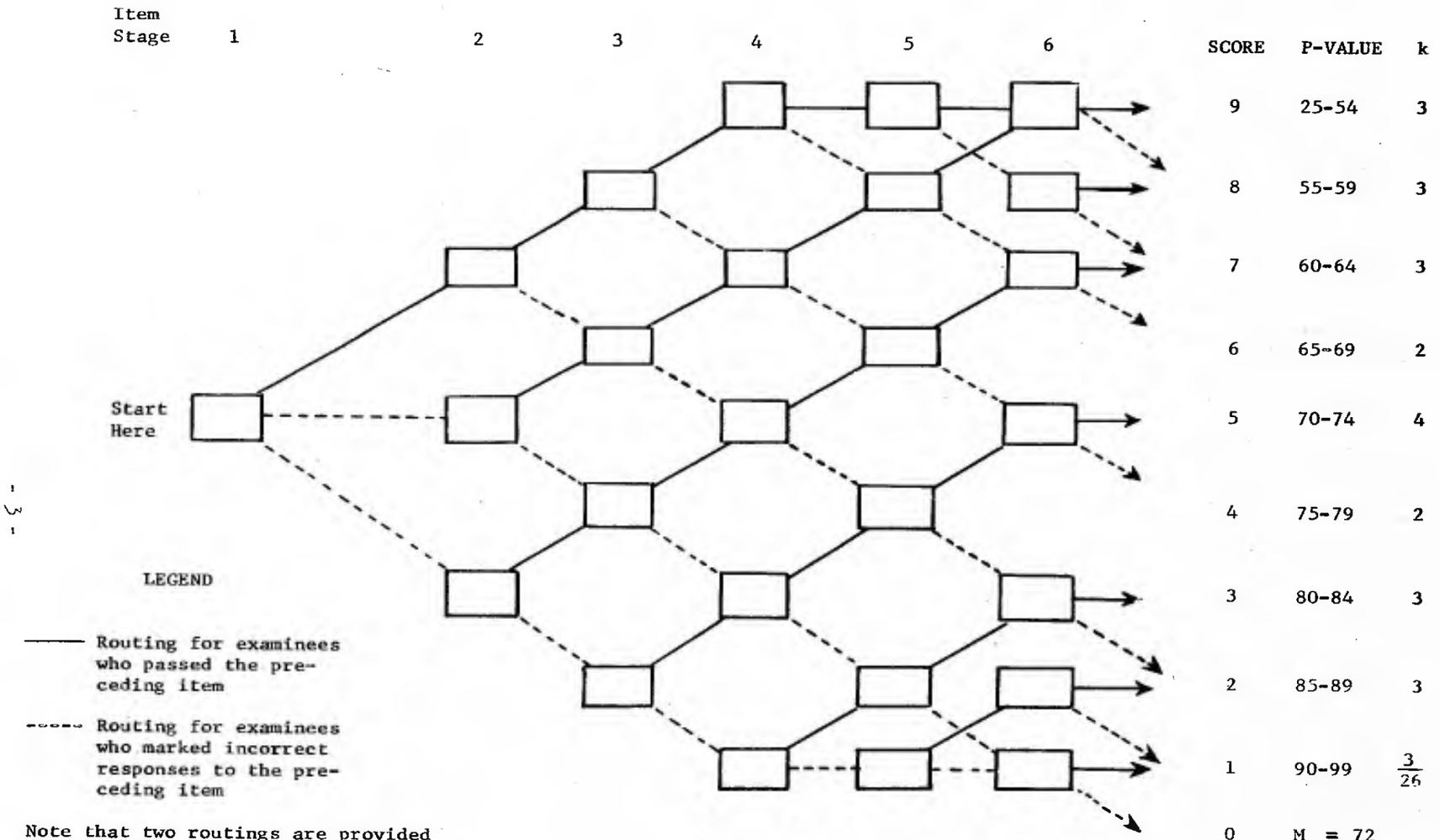

Figure 1. General schema for routing items for each AFQT-SIT subtest.

Inspection of the cases collected revealed a sizable number of SIT papers which were unscorable and hence could not be included in the statistical analysis. Scorable cases were divided into three overlapping samples: Sample A consisted of 125 cases for which the ACB-VE score was available for each SIT-VE answer sheet. Sample B consisted of 116 cases for which an ACB-AR score was available for each SIT-AR answer sheet. Sample C consisted of 247 cases for which both SIT-VE and SIT-AR answer sheets were available, regardless of whether ACB scores were obtained. The first two samples contained no AFQT Category V cases, since Category V men are not given the ACB. Sample C contained 9 Category V cases.

RESULTS OF STATISTICAL ANALYSIS

Table 1 contains a summary of the results of the statistical analysis. In all three samples, mean raw scores on the SIT-VE and SIT-AR were high (7.17 to 7.31 for a score range of 0 to 9). All SIT score distributions showed high degrees of skewness, with marked bunching of scores at the maximum score point. On SIT-VE, the maximum score of 9 was attained by 43 percent of Sample A and 44 percent of the larger Sample C; on SIT-AR, by 38 percent of Sample B and 37 percent of Sample C. In Sample C, 22 percent scored 9 on both SIT-VE and SIT-AR. Despite the high degree of skewness, the standard deviations obtained (2.02 to 2.19) can be considered appropriate for a test of 10 score points.

In comparison with full mobilization samples used in earlier studies, Samples A and B--in which there were no Category V men--scored higher both on the ACB tests and on total AFQT. On ACB-VE, the mean was 108.6 and on ACB-AR, the mean was 108.2. Both standard deviations were 21.4. The AFQT mean was a percentile score of 58, the equivalent of a raw score of 73. In recent mobilization samples, the following means and standard deviations have been obtained:

Test	Mean	S. D.
ACB-VE (standard score)	100.95	26.94[1/]
AFQT 5-6 (total raw score)	66.25	20.22[2/]
ACB-AR (standard score)	99.61	23.77[1/]
ACB-AR (standard score)	99.25	19.77[2/]

1/ Bayroff, Seeley, and Anderson, 1959

2/ Bayroff, Seeley, and Kehr, 1958

Table 1

MEANS, STANDARD DEVIATIONS, AND COEFFICIENTS OF CORRELATION FOR SIT-VE, SIT-AR, AND BACKGROUND VARIABLES

Sample	N	Variable	No. of Items	Mean	S.D.	r^a
A	125	SIT-VE (raw)	6	7.24	2.19	(SIT-VE)
		ACB-VE (s.s.)	50	108.64	21.46	.63 (ACB-VE)
		Total AFQT-5-6 (%-ile)	100	58.12	26.54	.49 .68
B	116	SIT-AR (raw)	6	7.17	2.08	(SIT-AR)
		ACB-AR (s.s.)	40	108.20	21.43	.63 (ACB-AR)
		Total AFQT-5-6 (%-ile)	100	58.85	27.01	.55 .74
C	247	SIT-VE (raw)	6	7.31	2.15	(SIT-VE)
		SIT-AR (raw)	6	7.17	2.02	.50

[a]Uncorrected for restriction in range.

Although this higher ability level in the present sample would have contributed to the high mean scores on the SIT, there is no available evidence that the SIT subtests, as now constructed, would provide a sufficiently high ceiling for the normal flow of preinductees. The limited ceiling might in part have been a result of the scoring system employed or merely the consequence of making the test too easy for the group tested.

Relationship of SIT Subtests to ACB Tests

The 6-item SIT subtests showed favorable correlation with the longer ACB tests. For both SIT-VE and SIT-AR, the coefficient of correlation with the corresponding ACB test was .63. In comparison, the first operational Verbal test of the Army Qualification Battery (each form was composed of the 25 vocabulary items of AFQT 5 or 6) yielded a corrected r of .90 with ACB-VE; the 25-item Arithmetic Reasoning Test of the AQB correlated r = .85 with ACB-AR (Bayroff, Seeley, and Anderson, 1959). Unlike the SIT, the longer AQB tests were in the same conventional format and required the same test-taking procedure as the ACB tests with which they were correlated.

SIT Subtests as Predictors of Total AFQT Score

Uncorrected coefficients of correlation between SIT subtests and total AFQT 5-6 were promising for a novel 6-item test of one content area vs a conventional 100-item test of four different content areas--.49 for SIT-VE and .55 for SIT-AR. However, for the same examinees, the longer conventional-type ACB tests were better predictors of total AFQT scores. Uncorrected correlation coefficients were .68 for ACB-VE and .74 for ACB-AR.

Relationship of SIT-VE and SIT-AR

For Sample C examinees, an r of .50 between SIT-VE and SIT-AR was obtained. As an indication of the degree of independence of the abilities measured by the two tests, this coefficient could be interpreted as a desirable feature of the SIT. In the APRO standard matrix of ACB intercorrelations, the r between the longer ACB VE and AR is .72. However, it is not possible to determine how much of the relatively low r is a result of the low ceiling in the SIT subtests.

Reliability Indications

Reliability was inferred from the uncorrected correlation coefficient of .63 between the 6-item SIT subtests and their 40 and 50-item ACB counterparts, which suggested a considerable degree of stability of measurement for the SIT. No direct estimates of reliability were made since there were no alternate forms. No conventional method--other than correlation between alternate forms--was considered appropriate.

SUMMARY OF RESULTS OF STATISTICAL ANALYSIS

The statistical results obtained are summarized and evaluated below. Some of the results are encouraging as findings in a study of the feasibility of a new technique.

1. The distributions of SIT scores were markedly skewed with large numbers obtaining the maximum score. This low ceiling could be expected to limit the magnitude of the coefficients obtained. It is possible that the low ceiling was a function of the simple scoring system used, or that the test was just too easy.

2. The correlation coefficient of .63 between SIT-VE and SIT-AR with their ACB counterparts was promising, considering that a 6-item test of unconventional format was being correlated with a longer test of conventional format. Since the SIT format probably introduced measurement factors not present in the ACB counterpart, this correlation coefficient is not taken as indicating as low a reliability for the SIT as would be implied if the correlation between the two tests reflected only the function of the respective reliabilities of the ACB and SIT counterparts.

3. Correlation coefficients of SIT's with total AFQT were reasonably high--.49 for SIT-VE and .55 for SIT-AR. For the same SIT examinees, however, ACB tests were better predictors of total AFQT than the SIT's--r's were .68 for ACB-VE, and .74 for ACB-AR.

4. The correlation coefficient for SIT-VE vs SIT-AR was .50.

NONSTATISTICAL APPRAISAL OF THE SEQUENTIAL ITEM TEST

Quite apart from considerations of means and correlation coefficients, a number of nonstatistical considerations are significant in an appraisal of the Sequential Item Test as a useful tool for screening enlisted personnel. In the course of constructing the four subtests of the SIT and administering two of the subtests to a sample of enlisted men, it became apparent that the SIT possessed some characteristics not entirely advantageous in terms of intended Army use.

1. It was much more costly and time-consuming to construct a 6-step SIT subtest than a 25-item subtest of conventional format such as the AFQT. In terms of the number of items alone, the difference was substantial. Each 6-step subtest required 63 items of appropriate prescribed difficulty levels instead of the 25 required in the conventional AFQT type measure.

2. Arranging the items in the test and numbering each alternative of each item to yield a fool-proof routing pattern was a tedious and time-consuming procedure. Despite repeated checking and cross-checking, the SIT subtests administered in the field showed a number of construction oversights which would require correction before further use could be made of the tests.

3. The SIT was more time-consuming to administer than is apparent at first glance. Any 6-item test would seem on the surface to offer substantial time advantage over a 25-item test of the same subject matter content and difficulty level. With the SIT, however, this was not the case. Because of the unusual and complex nature of the examinee's task, instructions must of necessity be intricate and much more lengthy than for the ordinary AFQT type of test. The time required to give the instructions for the VE and AR subtests and to work the practice problems ranged from 10 to 15 minutes. The time allowed to do the items in the two subtests was 15 minutes. If four subtests were used, it is estimated total administration time would be approximately 50 minutes--not much less than the time required for the 100-item AFQT.

4. The SIT presented scoring problems. It could not readily be hand-scored by merely noting the score of the item marked in Stage 6. Such a procedure would have overlooked the possibility that the examinee had strayed from a proper routing or indeed marked items randomly without regard for the test instructions. In scoring the cases analyzed in the present study, the scorer had to follow step by step each examinee's route on the answer sheet, with the result that several types of error

were discovered. These included doing consecutive items as in conventional tests without regard to the indicated routing until a lettered (stage 6) item was encountered; marking items apparently randomly; stopping before reaching a lettered stage 6 item; following a legitimate route except for one or two departures apparently through clerical or reading error, as in going to 52 instead of 32; in the AR subtest, using the two-digit numerical answer to find the next item instead of the number preceding the answer; crossing out the number that directed the examinee to the next item so heavily that the examinee apparently could not read the number and hence went to the wrong next item. All these errors required the test scorer to spend far more time than he would have in scoring a conventional test considerably longer than 6 items.

5. Because of the necessary complexity of its instructions, the SIT was more difficult for examinees to understand than are conventional AFQT type tests. Further, the lower the AFQT level of the examinee, the more likely he was to misunderstand, or to understand only partially. Krathwohl and Huyser stated that even their college student subjects had trouble with the instructions. Because of the frequent occurrence of errors, many test papers gathered during the data collection for this project could not be scored, and hence could not be included in the analysis described above. Table 2 presents a comparison of numbers of scorable and unscorable papers obtained from examinees in AFQT mental categories I through V. A striking relationship between AFQT category and number of scorable SIT papers was noted. A substantial number of Category V men produced unscorable papers. The higher the AFQT category, the fewer the percentage of unscorable papers, with none among Category I men. This result is understandable in view of the complexity of the task involved in the SIT. Unfortunately, much of the Army's interest in the SIT as a potential input screening device centers around its usefulness for measuring relatively low level personnel (Category IV and V). If scorable answer sheets cannot be obtained from most low level personnel, then the usefulness of the SIT is severely limited. Further, there were considerably larger percentages of unscorable SIT-AR's than of SIT-VE's. The fact that the two subtests were constructed from the same pattern and of items with the same p-values would obviate explanation on the grounds that SIT-AR is the more difficult test. Part of the trouble with SIT-AR answer sheets was due, as noted above, to the error made by some examinees in using the numerical answer to an item to identify the next item to be attempted.

Of perhaps even greater importance was the fact that the AR items constituted the second portion of the test (given with one time limit for both subtests); many of the unscorable AR's were unscorable because they were incomplete--the examinee never reached Stage 6. Either a counterbalanced order of administering the two subtests or allowing more time might have remedied this deficiency. However, a substantial increase in time would further cast doubts on the desirability of the tests as an operational instrument.

Table 2

DISTRIBUTION, BY AFQT MENTAL CATEGORY, OF TOTAL SIT EXAMINEES AND OF NUMBERS OF SCORABLE SIT-VE AND SIT-AR TESTS

AFQT		Total No. SIT-VE, -AR Examinees[a]	No. of SIT-VE Tests			No. of SIT-AR Tests		
				Not Scorable			Not Scorable	
Mental Category	Percentile Score Interval		Scorable	No.	% of total examined	Scorable	No.	% of total examined
I	93-100	31	31	0	0.0	31	0	0.0
II	65-92	86	84	2	2.3	77	9	10.5
III	31-64	123	117	6	4.9	106	17	13.8
IV	10-30	39	33	6	15.4	34	5	12.9
Total Cat. I-IV		279	265[b]	14	5.0	248[c]	31	11.1
V	1-9	48	33	15	31.2	10	38	79.2
Total I-V		327	298	29	8.9	258	69	21.1

[a] All SIT examinees were given both SIT-VE and SIT-AR.
[b] Of these, recorded ACB-VE scores were obtained on 125 (47.2%).
[c] Of these, recorded ACB-AR scores were obtained on 116 (46.8%).

CONCLUSIONS

In summary, on the basis of the tryout on the sample described here, the Sequential Item Test has some satisfying characteristics in terms of its correlation with conventional tests. On the other hand, the drawbacks noted--increased time and expense required to construct, administer, and score the test, and particularly the fact that the test was difficult for low level personnel to understand--indicate that there are a number of difficulties that would have to be removed before the technique could be used in input screening.

These difficulties have led to the conclusion that further experimentation with the instrument in its present format is not worthwhile. However, the basic concept may have considerable utility if applied in a testing machine in which items cued to the correctness of the prior response could be programmed for automatic presentation. Since the machine would make the decision as to the item to be presented at each stage, many of the difficulties encountered in the present format would be removed. In case such adaptation is developed and tried out, results of the present experiment will have been useful in indicating the need for a larger number of difficult items in the basic item pool. The need to assemble a pool of pretested items and to develop normative values for given routes would remain. Research to determine optimal scoring systems, time limits, and prediction of external criteria such as school performance would also be needed.

REFERENCES

Bayroff, A. G., Seeley, L. C., and Anderson, A. A. Development of the Army Qualification Battery, AQB-1. Personnel Research Branch Technical Research Report 1117. October 1959.

Bayroff, A. G., Seeley, L. C., and Kehr, Carol J. Relationship between the Armed Forces Qualification Test and aptitude areas. Personnel Research Branch Research Memorandum 58-20. December 1958.

Bayroff, A. G., Thomas, J. A., and Anderson, A. A. Construction of an experimental sequential item test. Personnel Research Branch Research Memorandum 60-1. January 1960.

Krathwohl, David R., and Huyser, Robert J. The Sequential Item Test. (Revision of paper presented at annual convention of the American Psychological Association, 1956.)

Additional Selected Publications of the U. S. Army Personnel Research Office Dealing with INPUT QUALITY Research

Bayroff, A. G., and Anderson, A. A. Background and development of Armed Forces Qualification Test, AFQT 7 and 8. (In press)

Bayroff, A. G., Anderson, A. A., and Heermann, E. F. Development of failure keys for AFQT 7 and 8. (In press)

Bayroff, A. G., Fuchs, E. F., and Seeley, L. C. New instruments for screening Army input. American Psychological Association meeting, 7 September 1960. (Abstr.) The American Psychologist, vol 15, pp. 496-97. 1960.

Bayroff, A. G., Mundy, J. P., and Uhlaner, J. E. Development of new forms of the Army Qualification Test, AFQT 5 and AFQT 6. American Psychological Association meeting, 4 September 1956. (Abstr.) The American Psychologist, vol 11, p. 427. 1956.

Bayroff, A. G., Seeley, L. C., and Anderson, A. A. Relationship of AFQT to rated basic training performance. Technical Research Note 106. February, 1960.

Bayroff, A. G., Thomas, J. A., and Kehr, Carol J. Evaluation of EST for predicting AFQT performance. Technical Research Report 1114. February 1959.

Bolanovich, D. J., Harper, Bertha, Birnbaum, A. H., Lovelace, N. R., and Uhlaner, J. E. Development and standardization of short forms of the Armed Forces Qualification Test. Technical Research Report 934. April 1952.

Bolanovich, D. J., Jones, W., Frankfeldt, E., Lovelace, N. R., and Uhlaner, J. E. Development of instruments for selecting men for military service under reduced mental and literacy standards for induction. Technical Research Report 927. February 1952.

Bolanovich, D. J., Jones, W., Lovelace, N. R., and Fuchs, E. F. Development of a scoring key for identifying failures due to negative motivation on the Armed Forces Qualification Test. Technical Research Report 960. June 1952.

Bolanovich, D. J., Jones, W., Lovelace, N. R., and Uhlaner, J. E. Follow-up of the standardization of the Armed Forces Qualification Test. Technical Research Report 956. June 1952.

Bolanovich, D. J., Mundy, J. P., Jones, W., Klieger, W. A., Houston, T. J., and Marks, M. R. Test performance of administrative inductees. Technical Research Report 1080. October 1953.

Morton, Mary A., Houston, T. J., Mundy, J. P., and Bayroff, A. G. Mental screening tests for women in the armed forces. Technical Research Report 1103. May 1957.

Mundy, J. P., Meyer, L. A., Leedy, H. B., Sternberg, J. J., Goldstein, L. G., and Bayroff, A. G. Development of instruments for screening AFQT 5 and 6 failures. Technical Research Report 1104. May 1957.

Mundy, J. P., Schenkel, K. F., and Burke, Laverne, K. Interrelationships of Armed Forces Qualification Test (AFQT 3 and 4) subtests. Research Memorandum 55-2. May 1955.

Schenkel, K. F., Burke, Laverne, K., and Marks, M. R. Construction of the Examen Calificacion de Fuerzas Armadas. Technical Research Report 1090. December 1954.

Schenkel, K. F., and Fuchs, E. F. Development of an aptitude test for a foreign population. American Psychological Association meeting. 4 September 1956. (Abstr.) The American Psycholgist, vol 11, p. 427. 1956.

Schenkel, K. F., Meyer, L. A., Rosenberg, N., and Bayroff, A. G. Evaluation of the Puerto Rican screening test (ECFA) against success on the job. Technical Research Report 1106. June 1957.

Uhlaner, J. E., and Bolanovich, D. J. Development of the Armed Forces Qualification Test and predecessor Army screening tests, 1946-1950. Technical Research Report 976. November 1952.

AD Div 23/1, 28/4

UNCLASSIFIED
Human Resources Research
--Personnel Utilization

U. S. Army Personnel Research Office, OCRD, DA
EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST by Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson. December 1962. Rept on Input Quality c-13 Proj.--18 p. incl. tables, 22 Ref. (USAPRO Technical Research Note No. 129).
(DA Project 0J95-60-001) Unclassified Report

An exploratory study of the sequential item test technique applied to Army screening test content was made to gain some knowledge of its potential as a useful method of reducing Army testing time. The SIT was compared in field tryout with other short test techniques and with operational Army screening tests as to feasibility of administration and screening effectiveness. The SIT was moderately satisfactory in terms of its correlation with conventional tests. However, it revealed a number of drawbacks (increased time and cost in construction, administration, and scoring, and greater difficulty for low level personnel to understand) which led to the conclusion that further experimentation with the test in its present hand-scored paper-and-pencil format is not justified. The technique appears adaptable, however, to administration by testing machine, a procedure which would eliminate most of the difficulties encountered.

AD Div 23/1, 28/4

UNCLASSIFIED
Human Resources Research
--Personnel Utilization

U. S. Army Personnel Research Office, OCRD, DA
EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST by Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson. December 1962. Rept on Input Quality c-13 Proj.--18 p. incl. tables, 22 Ref. (USAPRO Technical Research Note No. 129).
(DA Project 0J95-60-001) Unclassified Report

An exploratory study of the sequential item test technique applied to Army screening test content was made to gain some knowledge of its potential as a useful method of reducing Army testing time. The SIT was compared in field tryout with other short test techniques and with operational Army screening tests as to feasibility of administration and screening effectiveness. The SIT was moderately satisfactory in terms of its correlation with conventional tests. However, it revealed a number of drawbacks (increased time and cost in construction, administration, and scoring, and greater difficulty for low level personnel to understand) which led to the conclusion that further experimentation with the test in its present hand-scored paper-and-pencil format is not justified. The technique appears adaptable, however, to administration by testing machine, a procedure which would eliminate most of the difficulties encountered.

AD Div 23/1, 28/4

UNCLASSIFIED
Human Resources Research
--Personnel Utilization

U. S. Army Personnel Research Office, OCRD, DA
EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST by Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson. December 1962. Rept on Input Quality c-13 Proj.--18 p. incl. tables, 22 Ref. (USAPRO Technical Research Note No. 129).
(DA Project 0J95-60-001) Unclassified Report

An exploratory study of the sequential item test technique applied to Army screening test content was made to gain some knowledge of its potential as a useful method of reducing Army testing time. The SIT was compared in field tryout with other short test techniques and with operational Army screening tests as to feasibility of administration and screening effectiveness. The SIT was moderately satisfactory in terms of its correlation with conventional tests. However, it revealed a number of drawbacks (increased time and cost in construction, administration, and scoring, and greater difficulty for low level personnel to understand) which led to the conclusion that further experimentation with the test in its present hand-scored paper-and-pencil format is not justified. The technique appears adaptable, however, to administration by testing machine, a procedure which would eliminate most of the difficulties encountered.

AD Div 23/1, 28/4

UNCLASSIFIED
Human Resources Research
--Personnel Utilization

U. S. Army Personnel Research Office, OCRD, DA
EXPLORATORY STUDY OF A SEQUENTIAL ITEM TEST by Leonard C. Seeley, Mary A. Morton, and Alan A. Anderson. December 1962. Rept on Input Quality c-13 Proj.--18 p. incl. tables, 22 Ref. (USAPRO Technical Research Note No. 129).
(DA Project 0J95-60-001) Unclassified Report

An exploratory study of the sequential item test technique applied to Army screening test content was made to gain some knowledge of its potential as a useful method of reducing Army testing time. The SIT was compared in field tryout with other short test techniques and with operational Army screening tests as to feasibility of administration and screening effectiveness. The SIT was moderately satisfactory in terms of its correlation with conventional tests. However, it revealed a number of drawbacks (increased time and cost in construction, administration, and scoring, and greater difficulty for low level personnel to understand) which led to the conclusion that further experimentation with the test in its present hand-scored paper-and-pencil format is not justified. The technique appears adaptable, however, to administration by testing machine, a procedure which would eliminate most of the difficulties encountered.

UNCLASSIFIED