NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.
Methods For Improving Enlisted Input--
Status Report, 30 June 1962
Initial distribution of this report has been made by the US Army Personnel Research Office. Additional copies may be obtained as follows. Department of Defense agencies and their contractors should address requests for additional copies to: Commander, Armed Services Technical Information Agency, ATTN: TIPDR, Arlington Hall Station, Arlington 12, Virginia. Other agencies and persons may obtain information on cost of microfilm or photostatic copies from Office of Technical Services, Department of Commerce, Washington 25, D. C., ATTN: Technical Reports Section. Copies may also be obtained on loan from local depository libraries. A list of these libraries may be obtained from Documents Expediting Project, Library of Congress, Washington 25, D. C. The mailing address of USAPRO is: US Army Personnel Research Office, Washington 25, D. C.
METHODS FOR IMPROVING ENLISTED INPUT --
STATUS REPORT, 30 JUNE 1962

A. G. Bayroff

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

Approved by
Julius E. Uhlaner
Director, Research Laboratories

Hubert E. Brogden
Chief Scientist

June 1962

USAPRO Technical Research Reports and Technical 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. Status reports in the blue cover depart from the usual format in that they do not necessarily contain separate management and technical supplements.
BRIEF
METHODS FOR IMPROVING ENLISTED INPUT--STATUS REPORT, 30 JUNE 1962

Requirement:

Congressional legislation has set a requirement for procedures to screen selective service registrants so that individuals who do not have sufficient trainability to succeed in the services may be rejected (PL 51, 82d Congress). Successive forms of the Armed Forces Qualification Test are developed to meet the screening requirement for both potential inductees and enlistees. Supplementary screening authorized by PL 564, 85th Congress has taken the form of differential aptitude measures as directed by DCSPER.

Procedure:

During FY 1962, research was conducted on the following phases of input screening:

- Evaluation of the effectiveness of the interim form of a differential ability test battery for input screening, the Army Qualification Battery AQB-I.
- Standardization of new tests constructed for the Army Qualification Battery.
- Construction and standardization of Women's Army Classification Battery; Standardization of Armed Forces Women's Selection Test (jointly with Air Force).
- Exploring feasibility of new testing techniques.
- Devising a feasible method of estimating the mental abilities in the civilian manpower pool eligible for military service (initial planning phase).
- Sampling current AFQT data.

Findings:

Differential aptitude measures provided by AQB-I were found to be satisfactorily related to their aptitude area counterparts based on tests of the Army Classification Battery (ACB).

Experimental AQB tests were substantially correlated with ACB counterpart tests. Standard scores on individual tests yielded the desired distribution of scores in the middle range.

Of experimental test techniques, the most promising was a method for developing very short (8-item) tests to predict pass-fail AFQT within a narrow ability range.

Utilization of Findings:

The Army Qualification Battery (AQB-I) was introduced at Armed Forces Examining Stations in September 1961 as an interim measure for differential aptitude screening of AFQT Category IV men and for determining eligibility of applicants for enlistment for a specific training program. New forms of the AQB were prepared for implementation 1 July 1962.

Short, limited-range tests may be developed by USAPRO scientists as substitutes for the Enlistment Screening Test for quick estimation of eligibility under various enlistment standards.

AFWST 5 and 6 and WEST 3 and 4 were implemented by the Army 1 December 1961 for screening applicants for the Women's Army Corps. The WACB was implemented 11 September 1961.
METHODS FOR IMPROVING ENLISTED INPUT--
STATUS REPORT, 30 JUNE 1962

CONTENTS

OPERATIONAL REQUIREMENTS FOR INPUT SCREENING........ 1
RESEARCH EMPHASIS--FY 1962................................. 2
DIFFERENTIAL ABILITY TESTS FOR INDUCTION................. 3
Follow-Up of Interim Tests, AQB-1.......................... 3
STANDARDIZATION OF NEW TESTS OF THE ARMY QUALIFICATION BATTERY (AQB)..... 3
TESTS FOR WOMEN........................................... 4
METHODS OF IMPROVING TESTING.............................. 5
Experimental Short Tests for the Enlistment Screening Test 5
Experimental Disguised Aptitude Test.......................... 6
Self-Instruction Technique Applied to an Experimental Nonlanguage Qualification Test 7

AFQT SAMPLE, 1961-1962.................................... 7

PROTOTYPE STUDY OF THE MOBILIZATION BASE................ 8

REFERENCES................................................. 9
METHODS FOR IMPROVING ENLISTED INPUT--
STATUS REPORT, 30 JUNE 1962

OPERATIONAL REQUIREMENTS FOR INPUT SCREENING

The screening of enlisted input for mental quality has been a continuing need of the Army and the other Services. Prior to World War II, no systematic testing with psychometric devices was applied to reject those whose mental ability was so low as to make it inadvisable to attempt to train them for even minimal utility. In 1946, a modified form of the Army General Classification Test (AGCT-la) which was used for classification purposes during the war, was introduced as a screening test for applicants for enlistment in the Army.

In 1950, as a result of Congressional legislation and directives from the Secretary of Defense, the Armed Forces Qualification Test was introduced to screen both applicants for enlistment and Selective Service registrants coming into all the services. The use of one test by the Army, the Navy, the Air Force, and the Marine Corps made it possible to attempt an equitable distribution of mental ability. As a result of Congressional legislation in 1958, the Army was authorized to apply supplementary screening for more specific aptitudes to marginally acceptable personnel—those scoring in the 10th to 30th percentiles on AFQT (Category IV). In September 1961, the Army Qualification Battery (AQB-I) was introduced as a means of screening Category IV personnel to meet differential aptitude requirements. AQB-I, a battery of short middle-range tests corresponding to tests of the Army Classification Battery, is also used to determine the eligibility for enlistment of applicants who desire a commitment for a specific training program.

In addition to the needs met by AFQT and AQB, other screening needs continue. One is the need to provide recruiters with a test they can administer to prospective applicants so that only those who are likely to qualify on the AFQT will be forwarded to the AFES for processing. The Enlistment Screening Test (EST) has been developed in a number of successive forms for this purpose. Another need is to provide devices which will aid in identifying deliberate failures on AFQT. For this purpose, failure keys and the Individual Picture Recall Test have been developed. To screen women applicants, a number of forms of the Armed Forces Women's Selection Test (AFWST) and of the Women's Enlistment Screening Test (WEST) have been developed, with the Air Force bearing the major responsibility. The relatively short tests of the AQB, with the substitution of AFWST subtests for AFQT subtests, were appropriate to the classification needs of women. Accordingly, the Women's Army Classification Battery (WACB) was developed and standardized for the purpose.

Other needs have existed in the past. One of these, the screening of Insular Puerto Ricans, has been complicated by the fact that most of them lacked sufficient familiarity with the English language to take the AFQT. Hence, to meet this need, a Spanish-language equivalent to AFQT,
Examen Calificación de Fuerzas Armadas (ECFA), standardized on Puerto Rican samples, was developed. Whether new forms of ECFA will be needed in the future is uncertain. On the one hand, it is likely that eventually the public education program in English will increase English literacy sufficiently to eliminate the need for a separate Spanish-language equivalent of AFQT. On the other hand, it is uncertain that differences in Puerto Rican and continental U.S. culture will decrease sufficiently to permit use of AFQT without restandardization.

One other need has existed in the past, namely, the classification of AFQT failures according to their likelihood of benefiting from English literacy training programs. This need arose from the possibility that in event of full mobilization the acceptance of those failures might be required. For this purpose, verbal-arithmetic subtest keys and the Non-language Qualification Test (NQT 1) were developed to measure level of English literacy and level of ability. However, in view of the rising standards of acceptance, the standby mobilization categorization of those who failed at the lower standard (10th percentile) is considered unprofitable. It is expected that procedures for categorization of AFQT failures will have been suspended by the end of FY 1962.

RESEARCH EMPHASIS--FY 1962

With the implementation on 1 July 1960 of new forms of Armed Forces Qualification Test (AFQT 7 and 8) and the accompanying failure keys and verbal-arithmetic subtest keys, research effort of the INPUT QUALITY Task was shifted to meet other requirements. Major emphasis was placed on continuing the research involved in the development of new tests for the Army Qualification Battery (AQB). Increased effort was devoted to new methods of developing tests and to the tryout of new testing techniques which could lead to improvement in conventional operational measures: effective short tests for the differential measurement of aptitudes at the middle ability level; short, limited-range tests for use as new forms of the Enlistment Screening Test; new types of test and new test techniques; and new approaches to the detection of deliberate failures.

Planning was initiated for studying potential resources for investigating the mental abilities of the civilian manpower pool available for military service. Tentative selection was made of the Selective Service System as the resource for the overall study. Discussions were held with Selective Service headquarters personnel, and visits were made to nearby Selective Service local boards where pertinent records and procedures were examined.

Steps were taken to study distributions of AFQT scores and other variables in a 10-percent sample of Army input.
DIFFERENTIAL ABILITY TESTS FOR INDUCTION

FOLLOW-UP OF INTERIM TESTS, AQB-1

After completion of research to develop the first form of the Army Qualification Battery, AQB-1, a follow-up study was undertaken at the request of DCSPER. The operational effectiveness of the battery was assessed in a cross-standardization study which provided information on the stability of the norms, the correlation of individual AQB tests with their Army Classification Battery counterparts, and acceptance rates resulting under the established standards. Tests of AQB-1 and of the Army Classification Battery (ACB) were administered to 1050 Selective Service registrants who had obtained AFQT percentile scores 10-30. Particular attention was paid to obtaining a rectangular distribution within this interval, corresponding to the expected distribution of a mobilization distribution of AFQT scores.

Statistical analysis has not yet been completed. Results now available indicate that the original norms were verified, although some minor adjustments in the conversion tables may be desirable. The correlation pattern of AQB-1 tests and their ACB counterparts was verified.

Study of acceptance rates was based on AQB composites of tests corresponding to aptitude areas derived from the ACB tests used in enlisted classification. Individuals are considered acceptable only if they achieve a standard score of 90 or higher on two or more aptitude areas. The acceptance rate, that is, the percentage meeting the differential aptitude requirements, was higher than had been found in previous studies. The increase was noted also when aptitude area scores were computed on the basis of ACB tests, suggesting that the increased rate is not attributable solely to characteristics of AQB-1. In part, the increased rate may be the result of as yet unidentified operational factors. What may be more important is the likelihood that the increase reflects instability of the passing rate. Such instability will occur because of sampling fluctuations and because the nature of the sample of preinductees changes from one time period to another.

STANDARDIZATION OF NEW TESTS OF THE ARMY QUALIFICATION BATTERY (AQB)

The initial operational form of the Army Qualification Battery, AQB-1, developed and standardized for operational use in a relatively short time, was considered an interim measure. Development of alternative forms of the eight tests comprising AQB-1 and of two additional tests had been undertaken concurrently with research to develop AQB-1. Major activity during FY 1962 was the standardization of new tests for the final forms of the battery. Experimental tests previously constructed were administered and the statistical analysis required for item selection was completed.
Four of the tests were derived from the current Armed Forces Qualification Test: Verbal Ability (VE), Arithmetic Reasoning (AR), Shop Mechanics (SM), and Pattern Analysis (PA). The six additional tests were: Army Clerical Speed (ACS), Automotive Information (AI), Mechanical Aptitude (MA), Electronics Information (EI), General Information (GI), and the Classification Inventory (CI). Two forms of each of the ten tests had been prepared. In general, the tests prepared for the AQB paralleled comparable tests of the Army Classification Battery (ACB), differing only in the reduced length and range of the AQB tests. No counterpart to the Army Radio Code Aptitude Test of the ACB was prepared, inasmuch as the test is used primarily to meet a specialized and restricted personnel requirement.

To obtain standardization data, the new AQB tests were administered at ten Armed Forces Examining Stations representing the six Army Areas, thus providing geographic coverage. Each of the two sets of forms was administered to 1200 examinees (Selective Service registrants and applicants for enlistment) who had not had prior service and who had passed AFQT at or above the 10th percentile. Each examinee was also administered the Army Classification Battery tests after the experimental AQB tests.

For the statistical analysis, 900 or more cases for each of the two sets of forms were selected so that each of the upper nine AFQT deciles contained an equal number of cases distributed rectangularly within each decile. The raw scores were converted to Army Standard Scores using the operational AFQT and the counterpart ACB tests as reference variables.

The Standard Score ranges in the resulting conversion tables were narrower than the ranges for the counterpart ACB tests, a distribution in line with the discrimination desired for the AQB. Correlation between AQB tests and their ACB counterparts was generally substantial, although not quite as high as with AQB-1.

Plans have been made for implementation of the new AQB forms 1 July 1962.

TESTS FOR WOMEN

New forms of the Armed Forces Women's Selection Test (AFWST 5 and 6) and of the Women's Enlistment Screening Test (WEST 3 and 4) were prepared and standardized in collaboration with the Air Force, which bore major responsibility. Samples of enlisted women in basic training at the WAC Center, Fort McClellan, as well as Air Force enlisted women, were tested.

In addition, a differential aptitude battery, the Women's Army Classification Battery (WACB), incorporating tests from the Army Qualification Battery, was developed and standardized on trainee samples.
at the WAC Center. The WACB is intended to serve a measurement function comparable to that of the AQB in the case of men. The WACB provides test scores on Verbal Ability (VE) and Arithmetic Reasoning (AR), two tests derived from the AFVST, and on additional short tests: Army Clerical Speed (ACS), Automotive Information (AI), Mechanical Aptitude (MA), Electronics Information (ELI), Pattern Analysis (PA), Shop Mechanics (SM), and Army Radio Code (ARC). Combinations of these tests provide aptitude area scores for Electronics (EL), General Maintenance (GM), Motor Maintenance (MM), Clerical (CL), General Technical (GT), and Radio Code (RC). Since women are not assigned in occupational areas involving combat activities, no combat aptitude area scores were included.

AFVST 5 and 6 and WEST 3 and 4 were implemented by the Army 1 December 1961. WACB was implemented 11 September 1961. Subsequently, the first phase of a restandardization of AFVST 5 and 6 was completed and plans made for a longer range follow-up.

METHODS OF IMPROVING TESTING

EXPERIMENTAL SHORT TESTS FOR THE ENLISTMENT SCREENING TEST

In developing aptitude tests for selection purposes, the objective is usually maximum reliability of the measure coupled with adequate validity over a full range of the aptitude in question. To attain this objective, relatively long tests with items finely graduated in difficulty are required. However, there are occasions when the selection problem is of a "go-no go" nature, when all examinees below a predicted criterion level are rejected and all those above that level are selected, as typified by the Enlistment Screening Test (EST). In such instances, the selection test must predict criterion performance at a particular level rather than at all levels. Recent developments in test theory have indicated that for such "go-no go" selection, long wide-range tests are unnecessary. Further, where a relatively low ability level is to be measured, the longer test may prove to have lower validity than a restricted-range test appropriate for lower ability examinees. Experience with the comparatively short tests of the AQB supported these research developments.

In anticipation of a requirement for new forms of the Enlistment Screening Test, the foregoing principles were applied in exploring the feasibility of short, limited-range tests to predict pass-fail at the 31st percentile, current standard for enlistment in the Army, and for other operational AFQT cutting scores which define the mental standards for various input programs.

Preliminary analysis indicated that as few as eight items with the necessary psychometric characteristics would be adequate for the purpose. Two methods were applied to the selection of items from AFQT 7 and 8 for the experimental short forms. One method required that the mean difficulty of the eight items be appropriate to the criterion level to be predicted (for example, a mean p-value of .70 to predict pass-fail at
AFQT percentile 31), high internal consistency indexes, and high validity (correlation with operational AFQT 5 and 6 score). In addition, it was assumed that all item intercorrelations were equal. The second method was essentially an application of test selection methods to item selection as a means of identifying the eight items whose linear combination gave the best prediction of the criterion level. Availability of computer facilities made it possible to compute the necessarily large number of item intercorrelations required as well as item validity coefficients for the several criterion levels.

Validity estimates were computed in the samples in which the items were selected (back-validity) and in independent samples (cross-validity). The prediction of pass-fail at the 31st percentile was satisfactory and about the same for tests constructed by the two methods \((r = .70, approximately)\). Prediction of other AFQT levels was somewhat less satisfactory, presumably because there were fewer items from which to select at those levels. (The difficulty distribution of AFQT items is not rectangular---there are more items appropriate in difficulty to the 31st AFQT decile than to any other decile.) With adequate numbers of items with the required psychometric characteristics to select from, eight-item tests could in all likelihood be developed for other AFQT levels that would be as valid as the short tests predicting the 31st percentile.

**EXPERIMENTAL DISGUISED APTITUDE TEST**

The Army has for some time employed terminal screening procedures to identify individuals deliberately failing AFQT. Research and operational experience have indicated that these procedures are, in general, accomplishing their purpose. However, there is room for improvement in research method and operational effectiveness. One approach which appears promising is derived from the literature on projective tests. With appropriate scoring, substantial correlation has been found between some projective tests and conventional cognitive aptitude tests. Hence an attempt is being made to determine whether the serious limitations of projective tests can be overcome and the technique modified to provide a useful measure of cognitive aptitude.

The first step was the construction of an experimental test designed to give the impression of measuring personality characteristics. Fifty unstructured visual stimuli (inkblots) have been prepared; and a second series of more structured stimuli is in preparation. Present plans are to administer the stimuli to input samples to obtain free responses stating what the stimuli look like to the examinee. From the response pools, four alternatives will be selected for each stimulus. All four selected responses will have the same frequency (popularity) in the entire sample; yet one response will be associated with a higher mean AFQT score than will the other three responses. If a final form of the test adequately correlated with AFQT can be attained, an alternative approach to detection of deliberate failures will be available.
The Nonlanguage Qualification Test used to measure level of ability among AFQT failures who do not understand English is difficult to administer. An experimental test which does not require the understanding of English has been devised. Content of the experimental test is unconventional in that examinees are first required to learn pictorial codes and then to integrate the codes so as to solve the problems presented. Means of self-instruction was introduced to minimize the amount of instruction to be given by the examiner. The self-instruction is in the form of immediate knowledge of results and the requirement that the examinee continue with a problem until he finds that he has selected the correct answer.

Two major problems were encountered. Since examinees had to learn the codes while taking the test, it was not possible to determine empirically which problems were serving as training problems and which were functioning as test problems. Further, the fact that examinees were instructed to make more than one response to each problem, if necessary to obtain the correct answer, made the test difficult to score. Scores resulting from several scoring methods were analyzed in relation to scores on the AFQT and on the operational Nonlanguage Qualification Test. When no consistent or significant relationships were found, plans for further research on the experimental test were abandoned. The effectiveness of immediate feedback as a test administration technique will be further explored by means of more suitable testing methods.

AFQT SAMPLE, 1961-1962

Information is frequently required concerning the relation between AFQT scores and variables such as level of education and geographical area, as well as on AFQT score distributions of special categories of personnel. Comprehensive data on the distribution of AFQT scores and related variables are basic to research planning, both for well-defined technical requirements such as sampling design for standardization and validation studies and for identifying aspects of operational testing programs in need of research study. Army personnel management often has need for information which can be obtained only by analyzing the test scores and other variables in specific segments of the Army population. Official statistical reports do not provide data on test scores and other pertinent variables in a form that can be used to answer many important research and user questions.

Arrangements were made for all Armed Forces Examining Stations to provide the necessary data on a 10 percent sample of Army input for the period 15 April 1961 through 15 March 1962. Planning for Statistical analysis has been initiated.
The objective of the study is to develop a prototype method for estimating the usable ability resources available in the civilian manpower pool eligible for military service. The method is directed at developing aptitude test and personal background data on an unbiased sample of Selective Service registrants, including those in deferred classes. Because of the unavailability of personnel, no activity occurred in this study during this period.
REFERENCES

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)


U.S. Army Personnel Research Office, ODRC, DA

METHODS FOR IMPROVING ENLISTED INPUT—STATUS REPORT, 30 June 1962
15 p., 18 ref. (USAFRO Technical Research Report No. 1125)
(DA Project 095-60-001) Unclassified Report

To meet the continuing requirement under Congressional legislation and
DDO policies for procedures to screen Armed Services input, major
research emphasis during FY 1962 was placed on: 1) evaluation of
an internal differential aptitude measure—Army Qualification Battery,
AQB; 2) standardization of new tests for the Army Qualification
Battery; 3) construction and standardization of a differential apti-
tude battery for WAC, Women's Army Classification Battery (WACB); 4)
standardization of new forms of the Armed Forces Women's Selection
Test (APFT 5-6) and Women's Enlistment Screening Test (WEST 3-6) in
collaboration with the Air Force; 5) exploring feasibility of short,
limited-range tests to predict pass-fail at the 31st percentile on
APQT, current standard for enlistment in the Army; 6) devising a re-
liable method of estimating mental abilities in the civilian manpower
pool eligible for military service; and 7) planning for sampling
current APQT data. New forms of the AQI were prepared for imple-
mentation on 1 July 1962. Research findings led to operational use of
the women's tests, APFT 5-6, WEST 3-4, and WACB in late 1961.

U.S. Army Personnel Research Office, ODRC, DA

METHODS FOR IMPROVING ENLISTED INPUT—STATUS REPORT, 30 June 1962
15 p., 18 ref. (USAFRO Technical Research Report No. 1125)
(DA Project 095-60-001) Unclassified Report

To meet the continuing requirement under Congressional legislation and
DDO policies for procedures to screen Armed Services input, major
research emphasis during FY 1962 was placed on: 1) evaluation of
an internal differential aptitude measure—Army Qualification Battery,
AQB; 2) standardization of new tests for the Army Qualification
Battery; 3) construction and standardization of a differential apti-
tude battery for WAC, Women's Army Classification Battery (WACB); 4)
standardization of new forms of the Armed Forces Women's Selection
Test (APFT 5-6) and Women's Enlistment Screening Test (WEST 3-6) in
collaboration with the Air Force; 5) exploring feasibility of short,
limited-range tests to predict pass-fail at the 31st percentile on
APQT, current standard for enlistment in the Army; 6) devising a re-
liable method of estimating mental abilities in the civilian manpower
pool eligible for military service; and 7) planning for sampling
current APQT data. New forms of the AQI were prepared for imple-
mentation on 1 July 1962. Research findings led to operational use of
the women's tests, APFT 5-6, WEST 3-4, and WACB in late 1961.