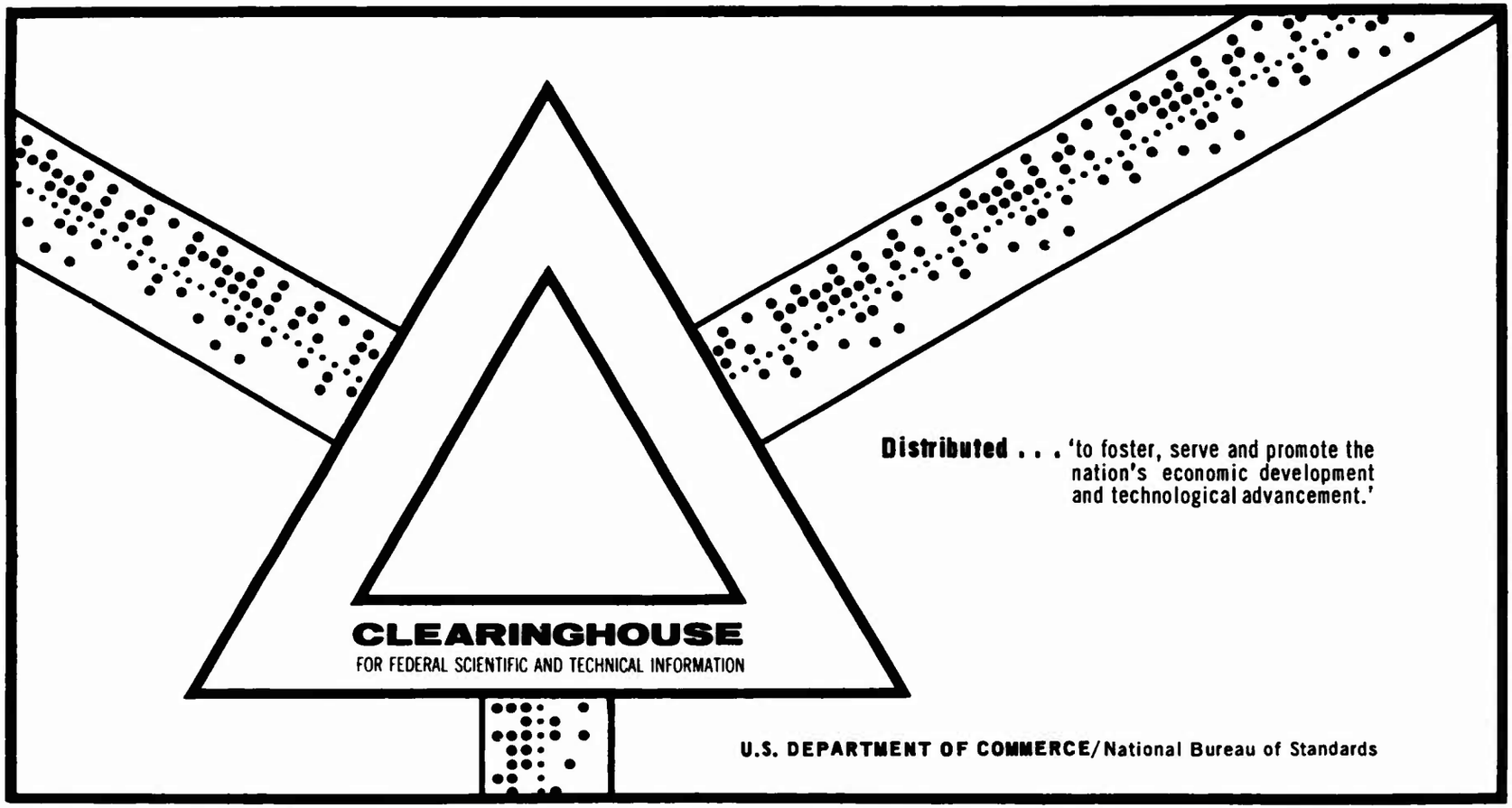
THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY

Abram G. Bayroff, et al.

February 1970



BLANK PAGES
IN THIS
DOCUMENT
WERE NOT
FILMED

Technical Research Report 1161

AD

THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY

Abram G. Bayroff and Edmund F. Fuchs

MILITARY SELECTION RESEARCH DIVISION

BESRL

U. S. Army
Behavior and Systems Research Laboratory

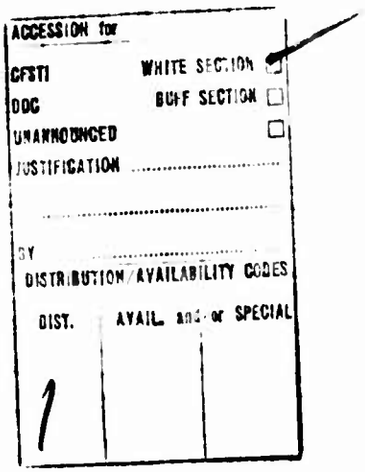
February 1970

45

BEHAVIOR AND SYSTEMS RESEARCH LABORATORY

An activity of the Chief, Research and Development

J. E. UHLANER
Director



NOTICES

DISTRIBUTION: Primary distribution of this report has been made by BESRL. Please address correspondence concerning distribution of reports to: U. S. Army Behavior and Systems Research Laboratory, Attn: CRDBSRL, Room 239, Commonwealth Building, 1320 Wilson Blvd., Arlington, Virginia 22209.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U. S. Army Behavior and Systems Research Laboratory.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Technical Research Report 1161

AD

THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY

Abram G. Bayroff and Edmund F. Fuchs

MILITARY SELECTION RESEARCH DIVISION
Edmund F. Fuchs, Chief

U. S. ARMY BEHAVIOR AND SYSTEMS RESEARCH LABORATORY

Office, Chief of Research and Development
Department of the Army

Room 239, The Commonwealth Building
1320 Wilson Boulevard, Arlington, Virginia 22209

February 1970

Army Project Number
2Q062106A722

Enlisted Manpower 00-01

BESRL 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.

THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY

BRIEF

Requirement:

To identify, among classification tests used by the several services, those which are interchangeable in terms of abilities and aptitudes measured; and from these to develop shortened forms to constitute an alternate form of a battery for service-wide use.

Procedure:

Comparability of the tests in the batteries used by the three services was determined from test intercorrelations in a consolidated sample of enlisted input (1000 each from the Army, Navy, and Air Force; 300 from the Marine Corps). The sample was stratified on AFQT to provide a mobilization distribution. Correlation coefficients were corrected first for restriction on AFQT and then for unreliability (test-retest with alternate forms). The new battery (Armed Services Vocational Aptitude Battery, ASVAB) based on tests found to be interchangeable was standardized on a 3000-man sample of Selective Service registrants again stratified on AFQT. Raw scores were converted to percentiles of the mobilization population.

Findings:

Seven sets of tests were identified as interchangeable: tests of word knowledge, arithmetic reasoning, space perception, mechanical comprehension, shop information, automotive information, and electronics information. The Army Coding Speed Test was selected as the measure of clerical aptitude, on the basis of separate validity studies. An eighth test, tool knowledge, was added to provide AFQT scores. Patterns of relationships among ASVAB tests and of ASVAB tests with AFQT were similar to those of the parent tests.

Utilization of Findings:

The ASVAB tests are currently being used to test potential recruits in the last year of high school. The tests may also be used as service classification batteries, supplemented as needed by tests unique to a given service.

FOREWORD

Maintenance and improvement of the U. S. Army system for screening potential enlisted input is a continuing requirement of BESRL's ENLISTED MANPOWER Work Unit. The unit provides differential screening batteries and related instruments, develops appropriate basic tools for development of military aptitude tests, and devises and explores the feasibility of innovative testing techniques for extracting more predictive information.

The Assistant Secretary of Defense (Manpower and Reserve Affairs) has requested research on a common aptitude battery that can be used by all the services. The Army with BESRL as its research agency has been the lead service in an accelerated program to determine to what extent the aptitude tests of the several services are interchangeable. The development of the Armed Services Vocational Aptitude Battery (ASVAB) consisting of abbreviated forms of tests found to be interchangeable is the subject of the present Technical Research Report. First use of the ASVAB is in testing potential recruits in high schools.

The entire Work Unit is responsive to RDT&E Project 2Q062106A722, "Selection and Behavioral Evaluation," FY 1970 Work Program objectives and to special requirements of the Deputy Chief of Staff for Personnel.

J. E. UHLANER, Director
U. S. Army Behavior and Systems
Research Laboratory

THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY

CONTENTS

TABLES

THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY

INTRODUCTION AND BACKGROUND

Personnel testing programs are an essential component of the screening and classification systems of all the armed services. The testing programs vary in many ways with the service which developed and uses them.

In screening for overall trainability, all the services use the Armed Forces Qualification Test (AFQT), as required by Congressional legislation.[1] The AFQT is administered to all potential enlisted input, both applicants for enlistment and Selective Service registrants. The services, however, differ in the sequence of testing and in the aptitude measures employed to supplement the AFQT. For example, the Army administers the overall test first, followed by more specific measures. The Air Force reverses the procedure, administering the more specific tests first.

In testing for more specific aptitudes as a basis for classification of enlisted men for military training and jobs, the Army, Navy, and Air Force have each developed their own batteries to meet their own needs. The Marine Corps uses the Army tests in screening and classification. Each service derives a set of composite scores from its battery. The composite scores, each based on two or more tests, are used as measures of trainability in groups of jobs which have generally similar requirements.

The batteries of the several services contain tests which appear to be similar in content, although differing in format, length, difficulty pattern, and other characteristics. For example, tests of verbal ability and of arithmetic reasoning appear in the batteries of all the services. The question has repeatedly been raised: Why not a single test to be used by all the services rather than three different tests all of which appear to measure the same aspect of trainability?

Essentially, the answer is that interchangeability of tests cannot be determined solely on the basis of similarity in content. A number of other factors must be considered such as selection standards, job requirements, and performance standards. In addition, there are test characteristics that cannot be inferred from inspection of the content--range of ability measured, pattern of difficulty of the test questions, and especially the validity of the test as a predictor of training or job performance. Only an empirical research study can produce evidence of interchangeability, and only by such a study is it possible to determine whether one battery of tests common to all the services is feasible.

---

[1] PL 759, 80th Congress, 1948; PL 51, 82d Congress, 1951; and PL 40, 90th Congress, 1967.

The need for a common battery gained increased attention recently in connection with the testing of high school seniors as part of the recruiting programs of the Army, Navy, and Air Force. For a number of years, the Air Force had been administering the Airman Qualifying Examination in a large number of high schools. Test scores were made available to school counselors for use in student guidance, as well as to Air Force recruiters. When the Army and Navy sought to test in the high schools, each with its own test battery, the additional testing time required brought considerable resistance from the schools. If testing in the high schools for recruiting purposes by all the services was to survive, the testing time required would have to be reduced. A logical solution was for all the services to use the same battery.

The Manpower Management Planning Board, of which the Assistant Secretary of Defense (Manpower and Reserve Affairs) is chairman, requested the research representatives of the services to review the technical problems involved in developing a single test battery for use of all the services. The battery was to serve the following purposes: "1) testing high school seniors, 2) establishing mental qualifications for enlistment and induction, 3) selection of enlistment applicants for particular occupational or training systems, and 4) classification and assignment." The review indicated that with an appropriate research design, the development of a common aptitude battery appeared feasible. However, there was uncertainty that one battery could be used for all the purposes desired.

In February 1966, after receiving the recommendations of the research representatives of the services, the ASD (M and RA) directed the services to begin development of a common aptitude battery that would be appropriate at least for the first stated purpose, testing high school seniors.[2] The battery was to provide common aptitude measures to be used by all the services as well as an overall measure for the Armed Forces Qualification Test. Again, it was recognized that the successful development of one all-purpose battery was uncertain.

The present study, then, was directed at 1) identifying counterpart tests of the three service classification batteries which were interchangeable, and 2) from the tests so identified, selecting items to produce standardized tests shorter than the parent tests so that total testing time would not exceed two and one-half hours. The short tests would be comparable to the longer classification tests and to the four content areas of the AFQT.

---

[2] Memorandum for the Undersecretaries of the Military Departments from the Assistant Secretary of Defense (Manpower), Subject: "Development of a common aptitude battery," dated 3 February 1966.

Allocation of Responsibility

All four services participated in the study, the Army as lead service having major responsibility.[3] The general plans were agreed to by all the services. Each service administered the batteries to samples of its in-service personnel and furnished punched cards containing the scores to the U. S. Army Behavior and Systems Research Laboratory for statistical processing. All four services participated in interpretation of the data, identification of the interchangeable tests, selection of items for the abbreviated tests, and standardization of the abbreviated tests.

IDENTIFICATION OF INTERCHANGEABLE TESTS

Administration of the Tests

The Army, Navy, and Air Force batteries were administered to 3900 enlisted men (1200 each in the Army, Navy, and Air Force; 300 in the Marine Corps) during reception processing, at the installations shown in the box below. Each enlisted man took 30 tests distributed over three days, no more than one battery a day. The study was considered of such importance that the testing was given priority over conflicting activities.

ADMINISTRATION OF
SERVICE CLASSIFICATION TEST BATTERIES

Service	N	Site	N
Army	1200	Fort Jackson, S.C.	400
		Fort Leonard Wood, Mo.	400
		Fort Dix, N. J.	400
Navy	1200	Great Lakes, Ill.	600
		San Diego, Calif.	600
Air Force	1200	Lackland AFB, Tex.	1200
Marine Corps	300	Parris Island, S. C.	150
		San Diego, Calif	150
Total	3900		3900

[3] The following research personnel had major responsibility:
Army: Edmund F. Fuchs, Abram G. Bayroff; Navy: C. Leonard Swanson; Air Force: Lonnie D. Valentine, Robert B. Stephens; Marine Corps: Edward A. Dover.
Others who made substantial contributions were: Army: Leonard C. Seeley, Robert B. Ross; Navy: Martin F. Wiskoff, Charles I. Hodges, Bernard Rimland; Air Force: Ernest C. Tupes, Bart M. Vitola; Marine Corps: Howard F. Uphoff, Joseph R. Beggun

Sampling Adjustments

The service samples differed because of differences in acceptance standards and screening systems. It was necessary to provide a standard sample that would represent the more stable mobilization population rather than sampling a particular input. Several adjustments were made, including the statistical selection of examinees for the consolidated sample in such proportions as to correspond to the expected distribution of AFQT scores in the full population of young men of military age. Such a procedure has been standard in the development of many military tests as, for example, the AFQT.

Statistical Analysis

The tests of each service battery had undergone extensive validation study. Because of differences among the services in acceptance standards, training programs, and job requirements, a test with known degree of validity in one service might have reduced validity if applied in another service. Hence, it was necessary to require high correlation among counterpart tests if they were to be considered sufficiently interchangeable to be equally effective in predicting success in training and on the job. The scores on all tests were correlated with each other and statistical adjustments made to provide stability and generalizability to the base mobilization population.

DEVELOPMENT OF THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY (ASVAB)

The Interchangeable Tests

The following sets of service tests were found to be correlated with each other sufficiently for the tests in a set to be considered interchangeable:

word knowledge
arithmetic reasoning
space perception
mechanical comprehension
shop information
automotive information
electronics information

The tests in the clerical aptitude area were not found to be interchangeable. The Army Coding Speed Test was chosen as the test of clerical aptitude to be included in the new battery because the Navy had found it to be more valid than the other clerical tests. A test of tool knowledge was added so that all four content areas were represented to provide an AFQT score.

Abbreviating the Tests

It was necessary to abbreviate the tests identified as interchangeable so that the new battery would not exceed the time limit imposed. To this end, 25 items were selected from each set of three interchangeable tests to provide one test about half the length of each parent test. Items with a wide range of difficulty, from very easy to very difficult, were selected so that each new test could measure a wide range of ability, although with not as fine discrimination as the longer parent test. After editing, the selected items were organized into a battery of nine tests, the Armed Services Vocational Aptitude Battery (ASVAB). The tests could be considered to be short alternate forms of the parent tests.

Standardizing the ASVAB

The next step was to standardize the ASVAB. The purpose of standardization is two-fold: to prepare standard instructions for administering the tests and to convert the raw scores on each test to scores that reflect the percentage of men in the mobilization population making each score. To accomplish the conversions, the ASVAB was administered to 3050 Selective Service registrants at 11 Armed Forces Examining and Entrance Stations (AFEES) throughout the country. The AFEES involved in the ASVAB standardization are shown in the box below.

AFEES INVOLVED IN STANDARDIZATION OF ASVAB

AFEES	Service Responsible	Number Tested
New York, N. Y.	Army	470
Baltimore, Md.	Army	330
Cleveland, Ohio	Army	310
Detroit, Mich.	Navy	550
Los Angeles, Calif.	Navy	225
Oakland, Calif.	Navy	225
San Antonio, Tex.	Air Force	140
Houston, Tex.	Air Force	170
Des Moines, Ia.	Air Force	220
Atlanta, Ga.	Marine Corps	200
Columbia, Ga.	Marine Corps	210
	Total	3050

As before, a standard sample of examinees was established with a distribution of AFQT scores that would be expected in the mobilization population. Another test (the "norming reference test") was administered to provide the mobilization percentile equivalents of the ASVAB raw scores, i.e., the percentage of men in the mobilization population expected to make the various ASVAB test scores. The results of the standardization study indicated that all the ASVAB tests could be used for screening and qualifying potential enlisted men. In the case of two tests, an unexpectedly large proportion made the highest scores. A similar excess of high scores occurred on the norming reference test. These findings were particularly unexpected since reports of resistance to preinduction processing suggested the possibility of an excess of low scores and are another indication of the need to study the mobilization population and to develop, if necessary, new reference standards.

DIFFERENCES IN SCREENING AND CLASSIFICATION SYSTEMS OF THE SERVICES

A description of the screening and classification systems of the several services is provided as background for the problems involved in developing a differential aptitude battery that all the services could use effectively and the research procedures devised to deal with the problems.

Prescreening and Qualification

As noted in the body of the report, all the services use the AFQT as an overall screen, but differ in the points of input flow at which the AFQT is administered. The services also differ in the prescreening and differential measures employed as supplement to the AFQT.

For overall prescreening, the Army and Marine Corps administer the Enlistment Screening Test (EST) to varying numbers of their applicants prior to administration of the AFQT. No prescreening test is administered to Selective Service registrants prior to the AFQT. The Army Qualification Battery (AQB) is administered after the AFQT to all marginal examinees (AFQT Category IV, 10th to 30th percentile, inclusive) and to all applicants for enlistment who seek commitment to a particular training program. The AQB yields aptitude area composites comparable to those derived from the Army Classification Battery and used in deciding whether the applicant qualifies for training in particular job areas.

The Navy administers its Applicant Qualification Test (AQT) for overall prescreening of most of its applicants for enlistment and the Short Basic Test Battery (SBTB) to about half the applicants for differential prescreening prior to the AFQT. It also administers the AQB for differential screening of AFQT category IV applicants.

In contrast to the other services, the Air Force employs only one instrument in addition to the AFQT. This instrument is the Airman Qualifying Examination (AQE), a differential aptitude battery administered as a prescreen prior to the AFQT.

Classification for Training and Jobs

All the services employ differential aptitude measures for classifying their personnel. The Army and the Marine Corps administer the Army Classification Battery (ACB) where needed to measure higher levels or where the AQB has not been administered during input processing (inductees in AFQT categories I, II, and III and applicants for enlistment in AFQT

categories I, II, and III who did not seek commitment to a particular program). Similarly, the Navy administers the NBTB for such purposes. The Air Force uses as its classification instrument the AQE which measures the higher levels of aptitude. The Air Force administers no other tests for classification purposes. All the services derive aptitude area composites of two or more tests from their batteries for use in classification.

The differences in personnel system and in mission of the services which, in turn, require differences in training and job requirements as well as in occupational and force structure, have heretofore been considered to stand in the way of developing one set of tests that could be used by all the services with no loss in effectiveness. Differences in the psychometric characteristics of tests which are similar in content imposed another consideration. One concerted effort resulted in the development of common core classification tests of verbal, arithmetic reasoning, spatial relationship abilities, and mechanical knowledge. However, they were not incorporated in the classification batteries of the services, although the Verbal and Arithmetic Reasoning tests were incorporated in the Army Classification Battery (June 1957).[4] Aside from these studies, no sustained effort was made until recently to determine the feasibility of a common aptitude battery, even though the desirability of a common battery was recognized.

OBJECTIVES

The present study was directed at 1) identifying the counterpart tests of the three service classification batteries which are interchangeable, and 2) from the tests so identified, selecting items to produce standardized tests shorter than the parent tests so that total testing time would not exceed two and one-half hours. The intention was that the short tests would be comparable to the longer classification tests and to the four content areas of the AFQT. As indicated earlier in the report, the immediate purpose was to develop tests which could be used by all the services in high schools by recruiting personnel and as a basis for counseling.

---

[4] Trump, James B., Richard K. White, Cecil D. Johnson, and Edmund F. Fuchs. Standardization of common core tests. BESRL Technical Research Report 1109. December 1957.
Helme, W. H., J. B. Trump, and D. J. Fitch. Validation of common core pattern analysis and mechanical knowledge tests for mechanical maintenance courses. BESRL Technical Research Note 107. July 1960.

Identification of interchangeable tests was based on intercorrelation matrixes obtained from administration of the Army Classification Battery, the Navy Basic Test Battery, and the Air Force Airman Qualifying Examination to a consolidated sample of in-service personnel in proportions corresponding to the expected distribution of AFQT scores in the full population of young men of military age. To the AQE were added easier items to make the tests comparable in range of difficulty to the ACB and NBTB. Items were arranged in subtest format.

The principal problem was the extent to which tests from different services which were identified as interchangeable could be expected to be equally good predictors in all the services. The likelihood that such interchangeable tests would be equally valid would be greater the higher the correlation coefficients among the tests. To indicate the extent to which the interchangeable tests would be equivalent to alternate forms of the tests, it was necessary to correct the correlation coefficients for attenuation because of unreliability.

Another aspect of this problem concerned the abbreviated tests. Since these tests were composed of items from the interchangeable tests, it was expected that substantial correlation would exist between the abbreviated tests and their parent tests.

Another problem arose from the fact that samples of unselected input were not available to provide the data from which the correlation matrixes were to be computed. The testing time required was far beyond the time available at Armed Forces Examining and Entrance Stations (AFEES) where input to all the services is examined prior to selection. Hence, it was necessary to test in-service samples at reception centers and to correct the correlation coefficients for restriction in range by selection on AFQT.

The restriction problem was further complicated by the fact that the in-service AFQT distributions would be biased, in part a result of the variation in AFQT acceptance standards and in part a result of the differential screening applied in addition to AFQT, sometimes before and sometimes after. Thus, men at the low end of the in-service distribution on AFQT had been preselected for specific aptitude in a variety of ways, and enlisted men in the lower end of the AFQT distribution would be expected to have higher scores on some specific aptitude measures than would unselected civilians of military age (the "mobilization population") with comparable AFQT scores. To offset the effects of this bias, limits in the AFQT score would have to be set below which cases would be excluded from the sample.

Still another problem was the effect of order of administration on scores and intercorrelations. A completely counterbalanced order of tests within a battery and of the three service batteries was clearly not possible. Instead, it was expected that in consolidating the four

service samples, adequate control of order effects would be possible if each service battery were equally often administered first, second, and third. With the tests in a battery administered in the order prescribed for operational testing, any order effects would not be considered as contaminating the test intercorrelations. A slight difficulty existed in the case of the Air Force tests, since the operational AQE is organized and administered as a single spiral omnibus test, whereas the modified AQE administered in this study was organized in subtest format.

IDENTIFICATION OF INTERCHANGEABLE TESTS

Research Design

The basic design is encompassed in a correlation matrix of 30 test variables (AFQT, 9 Army tests, 10 Navy tests, and 10 Air Force tests) computed in the consolidated sample of all the services. This matrix was replicated in each of the service samples. All correlation coefficients were corrected for restriction in range of AFQT and for attenuation due to unreliability.

Test-retest or alternate form reliability measures were considered more appropriate than internal consistency measures because correction was to be applied to correlation between separate tests and because internal consistency measures are generally underestimates with the consequence of over-correction of the correlation coefficients. Each service supplied test-retest reliability estimates of its own tests. The Army had no such data. Hence, the test-retest reliability of AFQT and the ACB variables was measured in an additional Army sample.

Testing Samples

To provide data for the intercorrelation study, the Army, Navy, and Air Force each tested enough enlisted men to provide a minimum of 1000 complete cases; the Marine Corps, a minimum of 300 complete cases. Other than to insure wide-range samples, no specific structuring was attempted. To obtain geographical representation, the tests were administered at installations in all parts of the country. To provide data for measuring the reliability of the AFQT and the ACB, the alternate forms of the tests were administered to another sample (N = 367) at two Army installations according to the following schedule:

	Tests Administered		
Installation	Day 1	Day 2	
Fort Jackson	ACB	1st half, ACB alternate	2d half, ACB alternate AFQT 7C[a]
Fort Leonard Wood	ACB	1st half, ACB alternate	2d half, ACB alternate AFQT 8C[a]

[a] Interval between test and retest with AFQT was variable, sometimes a matter of weeks.

Testing Procedures

All tests were administered at reception centers during classification processing prior to the beginning of basic combat training. Each examinee received the 30 tests of the three classification batteries, distributed over three days, with no more than one battery a day. Testing was during normal duty hours.

The order of testing was such that in the consolidated sample, each battery was administered first, second, and third to the same number of enlisted men, except for the small imbalance introduced by the Marine Corps sample. Each service administered its own battery first. All services administered the same test forms of each battery.

Analysis Samples

The correlation matrix which provided the data for identification of interchangeable tests was computed in a standard trainee-mobilization sample. The sample was constituted by consolidating the service test samples and stratifying on AFQT. Since it was expected that the AFQT distributions would be differentially biased at the low end because of differences in acceptance standards and in the differential screening applied in addition to the AFQT, it was necessary to establish AFQT scores above which the biases would be expected to be at a minimum. The differential bias in the Army and Marine Corps distributions would be the result primarily of the differential screening with the AQB applied to marginal passers (category IV) on AFQT; an AFQT percentile score of 20, the middle of category IV, as a lower limit would be expected to reduce the effects of this bias to a minimum.

In the Navy and Air Force samples, the problem was more difficult because differential prescreening was applied prior to testing with the AFQT. The differential prescreening was expected to introduce curvilinearity in the regression of the prescreening tests on AFQT. Scatterplots showing the regression of each Navy and Air Force test on AFQT were examined to estimate the AFQT score at which the regression lines departed from linearity. The AFQT distribution above this point would be considered unbiased.

The minimum acceptable AFQT scores for the Army and Marine Corps samples were expected to be lower than those for the Navy and Air Force samples. The difference resulted from the fact that the Navy and Air Force obtained all their enlisted input through enlistment and hence could establish qualifying scores which were higher than those established by law for Selective Service registrants. On the other hand, a large portion of the Army and Marine Corps input consisted of Selective Service registrants. In the consolidated stratified sample, the cases in the lowest portion of the distribution were selected exclusively from the Army and Marine Corps samples. Throughout the rest of the distribution, all services were represented.

The Consolidated Analysis Sample

As a result of the examination of the scatterplots showing the regression of Navy and Air Force tests on AFQT, a minimum AFQT score of 30 was set as reducing bias in the AFQT distribution attributable to the differential prescreening. As it turned out, both the Navy and the Air Force had relatively few cases below AFQT 30. As indicated above, the minimum acceptable AFQT score for the Army and Marine Corps sample was set at 20. Stratification of the consolidated sample on AFQT was accomplished by multiplying the frequencies in each AFQT half-decile by a factor such that each product equaled 100. This procedure avoided the necessity of discarding cases in excess of the frequencies needed for stratification, with the resulting advantage of greater reliability with larger numbers of cases.

Effect of Order of Testing

The three service batteries were administered in the following order:

	Service Batteries Administered		
Sample	Day 1	Day 2	Day 3
Army	Army	Air Force	Navy
Navy	Navy	Army	Air Force
Air Force	Air Force	Navy	Army
Marine Corps	Army	Navy	Air Force

It was possible that order of testing would affect the test intercorrelations. To test for order effect, the grand means of the three days of testing were compared to determine if they varied enough to affect the intercorrelations. In the four test samples, the means for the three successive days were 243, 251, and 241, respectively. The somewhat higher mean for the second day could be attributed to the fact that more men (Air Force and Marine Corps) took the longest (Navy) of the three service batteries on day 2. Otherwise, the means for the three days were so similar that the intercorrelations were not likely to be affected by order of testing.

Test Reliability

Estimates of reliability of the tests of the three service batteries were obtained by retesting with alternate forms, one to two days after the first testing, and computing the correlation coefficients for the respective alternate forms. The coefficients shown in Table 1 served as the basis for correcting the entries in the intercorrelation matrix (Appendix B) for unreliability. The AFQT retest was given a variable period (sometimes weeks) after the first AFQT testing, an interval

considerably longer than that for retesting with the other tests. The reliability of AFQT (r = .94, .92) as given by Bayroff and Anderson[5] in 1963 was based on immediate retesting with alternate forms.

The Interchangeable Tests

The intercorrelation matrix (Appendix B) indicated that there were seven sets of counterpart tests in which the tests could be considered interchangeable, that is, one test in the set could be expected to be as good a predictor as any other test in the set. In most instances, as indicated in Table 2, the minimum correlation of r = .90, corrected for restriction and for unreliability, was achieved. (Correlation coefficients greater than 1.00 are artifacts resulting from the double correction.) There were some additional sets of tests which achieved the minimum correlation, although differing in apparent content.

The seven sets of interchangeable tests (tests of word knowledge, arithmetic reasoning, space perception, mechanical comprehension, shop information, automotive information, and electronics information) served as a source of items for the ASVAB tests. The service tests used to measure clerical aptitude did not intercorrelate highly enough to be considered interchangeable. However, the Navy reported that the Army Coding Speed Test had shown higher validity than any of the other tests used by the services for measuring clerical aptitude. Accordingly, the test was added as a source of ASVAB items. Since the Navy study had been based on a test which was twice as long as the current Army Coding Speed Test, the longer test was used.

DEVELOPMENT OF THE ASVAB

Preparation of the Tests

The tests of the Armed Services Vocational Aptitude Battery (ASVAB) were expected to meet three requirements: 1) to be essentially alternate forms of the parent tests, 2) to be half the length of the parent tests, and 3) to be appropriate for a wide range of ability. The approach was to treat each set of interchangeable tests as a pool of items, supplemented as necessary from available pools of comparable items.

---

[5] Bayroff, A. G. and A. A. Anderson. Development of Armed Forces Qualification Test, 7 and 8. BESRL Technical Research Report 1132. May 1963.

Table 1

TEST-RETEST RELIABILITY COEFFICIENTS CORRECTED FOR SELECTION ON AFQT

Army Tests[a]			Navy Tests			Air Force Tests		
	uncorrected	corrected		uncorrected	corrected		uncorrected	corrected
VE	.90	.91	GC	.94	.97	AC	.80	.82
AR	.83	.85	AR	.84	.90	WK	.94	.95
PA	.71	.75	MK	.87	.89	AR	.89	.91
MA	.71	.74	MC	.87	.92	HF	.82	.84
CS1	.65	.66	CL	.79	.80	GM	.86	.88
CS2	.73	.74	SP	.85	.90	EI	.92	.94
SM	.68	.71	MATH	.88	.92	MP	.91	.93
AI	.86	.87	SCI	.79	.86	SP	.94	.95
ELI	.65	.68	ER	.70	.75	DI	.82	.86
AFQT[b]	.85	.87	AK	.92	.94	PC	.87	.90

[a] See Appendix C for means and standard deviations of test and retest.
[b] Computed in Army sample.

Table 2

INTERCHANGEABLE CLASSIFICATION TESTS

Content Area	Variable No.	Name		r^a and Variables	
WORD KNOWLEDGE (WK)	2	A	Verbal	1.00 (2-11)	.96 (2-22)
	11	N	General Classification		1.03 (11-22)
	22	AF	Word Knowledge		
ARITHMETIC REASONING (AR)	3	A	Arithmetic Reasoning	.93 (3-12)	.95 (3-23)
	12	N	Arithmetic		.89 (12-23)
	23	AF	Arithmetic Reasoning		
SPACE PERCEPTION (SP)	4	A	Pattern Analysis		.95 (4-30)
		N	(None)		
	30	AF	Pattern Comprehension		
MECHANICAL COMPREHENSION (MC)	5	A	Mechanical Aptitude	.94 (5-14)	.88 (5-27)
	14	N	Mechanical Comprehension		.92 (14-27)
	27	AF	Mechanical Principles		
SHOP INFORMATION (SI)	8	A	Shop Mechanics	.90 (8-16)	.95 (8-28)
	16	N	Shop Practices		.84 (16-28)
	28	AF	Shop Practices		
AUTOMOTIVE INFORMATION (AI)	9	A	Automotive Information	.97 (9-20)	1.03 (9-25)
	20	N	Automotive Knowledge		1.00 (20-25)
	25	AF	General Mechanics		
ELECTRONICS INFORMATION (EI)	10	A	Electronics Information		1.01 (10-26)
		N	(None)		
	26	AF	Electrical Information		

[a] Corrected for selection (AFQT) and unreliability.

Selection of Items. To provide a common format throughout all tests except the Clerical Aptitude Test, it was necessary in a number of instances to edit items so that all items had four alternatives. The format of the items in the parent tests for space perception was different from the common format; hence items of appropriate format were substituted from a pool of items prepared for AFQT 7 and 8, but not used in the final AFQT forms. The items of each parent test were considered highly homogeneous; consequently, selection was made primarily on the basis of available item difficulty indexes (p-values). Available item statistics for the Army tests were incomplete; hence, an item analysis was made in the test-retest reliability sample. Since the parent tests were used for classification, the p-values had been computed in samples which excluded those who had not met the qualifying scores for acceptance, approximately the lowest 20% of the mobilization population. Accordingly, the p-values were adjusted to indicate the proportion of the full population which could be expected to pass each of several desired points of discrimination.

In developing the difficulty pattern for the tests, consideration was given to the classification functions of the tests. It appeared that one standard deviation above the mean and one standard deviation below contained the score levels important for classification. Hence, it was expected that reliable measurement extending one and one-half standard deviations above and one and one-half standard deviations below the mean would provide adequate discrimination for classification, and that a test 25 items in length would discriminate adequately throughout the full range. In light of these considerations, the p-value distribution shown in Table 3 was adopted.

Table 3

DISTRIBUTION OF P-VALUES FOR SELECTION OF ITEMS

Standard Deviation	Number of Items	p-values[a]
+2.0 to +2.5	1	1
+1.5 to +2.0	2	3-7
+1.0 to +1.5	3	7-20
+0.5 to +1.0	3	20-40
0.0 to +0.5	3	40-60
-0.5 to 0.0	3	60-73
-1.0 to -0.5	3	73-84
-1.5 to -1.0	3	84-91
-2.0 to -1.5	2	91-95
-2.5 to -2.0	1	95-99
-3.0 to -2.5	1	99
Total	25	

[a] Corrected for chance success.

The Experimental Tests. The selected items were organized by subject matter into separate tests, each with its own time limit. Within each test, items were placed in ascending order of difficulty.

The Coding Speed Test was prepared in two formats. One was the format of the Army tests, in which all words to be coded have as alternatives the same ten numerical codes in the same order and the answer spaces are adjacent to the lead words. The second format followed the format of the other tests, each lead word being followed by five alternatives, which were not the same for all the items. Answers were recorded on a separate answer sheet. Since this is a speeded test and the differences in format could affect test performance, it was necessary to determine the comparability of the two formats.

Provision for AFQT Score. In addition to providing aptitude composites, the ASVAB was required to provide an AFQT score. Three of the four content areas of the AFQT were represented by interchangeable tests (word knowledge, arithmetic reasoning, space perception). It was therefore considered necessary to add a test to the ASVAB to cover the fourth AFQT content area, tool knowledge, to be used only in computing AFQT scores. One third of the tool knowledge items for the ASVAB--the easiest items--were selected from unused items prepared for earlier forms of the AFQT. The remainder were selected from similar items of the Navy Mechanical Knowledge Test, none of which were selected for the ASVAB Mechanical Comprehension Test.

TEST STANDARDIZATION

Initial application of the ASVAB was to be in testing high school seniors as part of the joint services recruiting program and as a tool for use by high school counselors for vocational guidance of the students. However, the ASVAB was also to provide AFQT scores. A sample of the mobilization population was therefore required. Also, classification scores derived from the ASVAB need to reflect the mobilization norms on which service classification is based. Accordingly, the AFQT Reference Test R-9, an editorial revision of the Army General Classification Test (AGCT) which provided the mobilization distribution of World War II personnel, was adopted as the norming reference test for the conversion of raw scores to standard scores, or their percentile equivalents, in the mobilization distribution.

General Design

The general design was essentially similar to the design employed in the standardization of recent forms of the AFQT. A full-range sample on AFQT was tested at AFEES with the ASVAB tests and the AFQT Reference Test R-9. After stratification on AFQT to produce a mobilization sample, the raw scores on the ASVAB tests were converted to percentile scores on the R-9.

Sampling

The data were collected at eleven AFEES throughout the country to provide geographic sampling. Only Selective Service registrants were tested. All others were excluded because of possible bias in test scores resulting from prescreening, as in the case of applicants for enlistment, or from previous testing as in the case of prior service personnel, reservists, and 1-Y personnel. An additional reason was that only registrants could be held over for more than one day, if it proved necessary, whereas prospective enlistees could not.

A total of 3050 examinees were tested with the ASVAB. To one half, the Coding Speed Test was administered in the original format; to the second half, the revised format was administered. From each half a sample of 1400, stratified on AFQT to produce a mobilization sample, was obtained. To provide data on the comparability of the two Coding Speed tests, an additional sample totaling 200 examinees at two AFEES was tested with both tests.

Testing Procedures

The operational AFQT was administered first at each AFEES. One-fourth of the examinees were then tested in each of the following orders:

1. ASVAB (with original format of the Coding Speed Test) followed by R-9

2. ASVAB (with revised format of the Coding Speed Test) followed by R-9

3. R-9, followed by ASVAB (with original format of the Coding Speed Test)

4. R-9, followed by ASVAB (with revised format of the Coding Speed Test)

Other operational tests such as the Army Qualification Battery were administered after completion of ASVAB and R-9 testing. The tests of the ASVAB were administered in fixed order with the exception of the Coding Speed Test. The original format of the test was administered as the fifth ASVAB test, the revised format as the first ASVAB test. A short rest period was introduced in the middle of the testing. For study of the comparability of the two formats of the Coding Speed Test, half the examinees at each of the two AFEES were tested with the original format first, and half with the revised format first.

Statistical Analysis

Mobilization samples for standardization of the ASVAB tests were established by stratifying on the operational AFQT. Two samples of 1400 examinees each were established, each tested with one of the two formats of the Coding Speed Test. No stratification was needed of the samples in which the comparability of the two Coding Speed Test formats was studied.

Raw scores on the ASVAB tests (scored R - W/3, except the Coding Speed tests which were scored rights only) were converted to percentiles in the mobilization samples. Frequency distributions of the raw scores on each test were prepared for each sample and the two samples combined. Corresponding distributions were prepared of the percentile equivalents of the raw scores (scored R - W/3) of R-9. To each ASVAB raw score was assigned the R-9 percentile score which had the same cumulative frequency in the samples as did the ASVAB score. To provide the percentile norms for the AFQT derived from the four ASVAB tests (Word Knowledge, Arithmetic Reasoning, Trade Knowledge, Space Perception), two methods were tried: 1) The percentile norms for the four tests were averaged. 2) The raw scores for the four tests were added together and then converted to percentiles. The second method was decided upon, since the cumulative frequencies were closer to the mobilization percentiles.

Correlation coefficients were computed among the ASVAB variables, between the ASVAB-AFQT and the operational AFQT, and between the two formats of the Coding Speed Test.

Development of the Norms

The raw score distributions of most of the ASVAB tests were of the expected wide range, the highest frequencies being in the middle of the distribution. The distributions on two of the tests, Word Knowledge and Arithmetic Reasoning, were peaked at the upper end, the highest scores being the most frequent. The three highest scores on the Word Knowledge Test were obtained by 42% of the sample; in the Arithmetic Reasoning Test, by 22% of the sample. A similar anomaly appeared in the distribution of the norming reference test R-9. Since the sample was stratified on AFQT, itself normed with the R-9, it had been expected that the highest R-9 decile would contain 10% of the sample. Instead, it contained 21% of the sample, this in spite of the fact that the R-9 puts a considerable emphasis on speed.

To throw some light on what happened in the two ASVAB tests, the item p-values were computed in the standardization samples and compared with the original values. In general, the easiest items became somewhat more difficult and the most difficult items became substantially easier, with the word knowledge items showing these differences much more than the arithmetic reasoning items (Table 4).

Table 4

CHANGES IN ITEM DIFFICULTY

Item No.	Word Knowledge		Arithmetic Reasoning	
	Corrected p-values[a]		Corrected p-values[a]	
	Original	Standardization	Original	Standardization
1	97	84	95	87
2	95	87	93	93
3	95	87	93	91
4	92	87	91	88
5	91	83	84	75
6	87	80	83	85
7	84	79	81	80
8	81	73	79	59
9	79	76	73	72
10	76	80	69	71
11	73	71	67	64
12	68	65	61	61
13	63	73	57	60
14	57	67	48	41
15	51	77	41	48
16	44	52	37	44
17	39	55	31	21
18	31	73	25	47
19	21	57	20	28
20	19	65	15	43
21	15	60	15	31
22	08	73	15	13
23	07	67	11	24
24	01	11	09	35
25	00	21	00	28

[a] Decimal points omitted.

The explanation of these anomalies is not immediately clear. The apparent regression of the item p-values may account for the excessive number of high scores, but does not account for the low ends of the distributions being as expected. Nor is it clear why regression effects should appear in these two tests and not the others. Perhaps most puzzling is the fact that these anomalies are in the opposite direction from what might have occurred--greater proportions of low scores resulting from the reported resistance to preinduction testing and the poor quality of schooling received by many of the examinees.

In the distributions, a disparity was noted between the cumulative frequencies of AFQT and R-9 for the respective percentiles. Through the first three deciles agreement was close (Table 5), but beyond that divergencies appeared, as was to be expected from the peaking at the high end of the R-9 distribution. For use in converting the raw scores to percentiles, the cumulative frequencies of AFQT and of R-9 were averaged. To each ASVAB raw score was assigned the percentile score of the equivalent averaged cumulative frequency. The smoothing adjustments of these equivalents were based on the same assumptions as were involved in the development of AFQT conversion tables.[6]

Table 5

DISTRIBUTIONS OF REFERENCE TEST SCORES

Decile	Cumulative Frequency		
	AFQT	R-9	Average
1	280	295	288
2	560	585	572
3	840	796	818
4	1120	922	1021
5	1400	1092	1246
6	1680	1288	1484
7	1960	1567	1764
8	2240	1785	2012
9	2520	2184	2352
10	2800	2800	2800

[6] See footnote 5 on page 15.

Coding Speed Test

The two formats of the Coding Speed Test were substantially correlated (r = .86). Correlation coefficients of the revised format with the other ASVAB tests and the two reference tests were slightly higher than the corresponding coefficients to the original format, although the pattern of correlation was essentially the same for the two formats (Table 6). Because of the somewhat greater convenience of the revised format, it was selected for inclusion in the ASVAB recommended for operational use.

The AFQT

Correlation between the ASVAB-AFQT and the operational AFQT was substantial (r = .89), almost as high as between the two alternate forms of the operational AFQT (.94 and .92). Coefficients were computed in two ways--corresponding to the two methods tried in developing percentile norms: 1) The raw score sum and the percentile score sum of the four ASVAB tests were correlated with the operational AFQT percentile scores (r = .89 and .90, respectively). 2) The percentile scores of each of the four ASVAB tests were correlated with operational AFQT percentile scores, and, through correlation of sums, the correlation coefficient of the four tests with the operational AFQT was obtained. This coefficient was the same (r = .89) as that obtained by the first method. Correlation of the ASVAB-AFQT with the operational AFQT, Forms 7 and 8, was the same (r = .90, .89) as between AFQT 7 and 8 in its standardization form with the then operational AFQT, Forms 5 and 6 (r = .90, .89).[7]

Comparability of ASVAB Tests and Parent Tests

The ASVAB tests were developed to be essentially alternate forms of the parent classification tests. In the present study, it was not feasible to determine directly the degree of correlation between the two sets of tests. Instead, as a guide, the correlation coefficients of the respective tests with the operational AFQT were examined. As Table 7 indicates, the correlation coefficients of the ASVAB tests with AFQT were in most instances similar to those of the parent tests with AFQT. One exception was the Navy's Electronics and Radio Test, presumably because it is a higher level test than the corresponding tests of the other services. The second exception was the Coding Speed Test of the ASVAB which had higher correlation (r = .65) with AFQT than did the Army's Coding Speed Test (r = .34). Whether the difference in format and administration and the fact that the ASVAB test was twice the length of the Army test account for this difference is not clear.

---

[7] See footnote 5 on page 15.

Table 6

ASVAB INTERCORRELATIONS FOR STRATIFIED SAMPLE
(N = 2800)[a]

AFQT	1.00											
R-9	.84	1.00										
WK	.77	.85	1.00									
AR	.80	.88	.79	1.00								
TK	.48	.27	.25	.28	1.00							
SP	.74	.63	.55	.63	.45	1.00						
MC	.77	.68	.65	.67	.55	.67	1.00					
SI	.71	.60	.62	.58	.67	.58	.71	1.00				
AI	.67	.54	.54	.52	.69	.50	.69	.77	1.00			
EI	.76	.69	.69	.68	.54	.63	.74	.76	.71	1.00		
CS(rev)	.65	.74	.69	.72	.19	.50	.51	.48	.39	.51	1.00	
CS(orig)	.57	.67	.63	.65	.11	.45	.41	.39	.30	.46	.86[b]	1.00
	AFQT	R-9	WK	AR	TK	SP	MC	SI	AI	EI	CS(rev)	CS(orig)
Means[c]	49.4	57.1	17.5	14.7	12.7	12.7	12.2	12.3	11.5	13.0	45.1	54.4
Standard Deviations	28.8	32.4	8.1	7.6	5.9	6.3	5.9	5.9	7.4	6.4	17.9	20.6

[a] CS (rev) and CS (orig) each administered to one half of stratified sample.
[b] Average of two non-stratified samples, N = approx. 200 each.
[c] Raw scores except AFQT.

Table 7

CORRELATION OF ASVAB AND PARENT TESTS WITH AFQT

ASVAB		Parent Classification Tests[a] Army		Navy		Air Force	
WK	.77	VE	.70	GC	.78	WK	.72
AR	.80	AR	.72	AR	.71	AR	.72
TK	.48	---		MK	.54	---	
SP	.74	PA	.77	---		PC	.71
MC	.77	MA	.67	MC	.73	MP	.78
SI	.71	SM	.74	SP	.65	SP	.64
AI	.67	AI	.61	AK	.61	GM	.69
EI	.76	ELI	.72	ER	.42	EI	.72
CS	.65 (rev)	CS2	.34	---		---	
CS	.57 (orig)	---		---		---	

[a]Corrected for selection on AFQT.

Examination of the intercorrelations of the ASVAB tests and of the service classification tests (Table 8) revealed that in half the 36 intercorrelations the ASVAB coefficients were within the range of the parent test coefficients; that is, the ASVAB coefficient differed no more from one of the classification test coefficients than one such coefficient differed from another. In two-thirds of the intercorrelations, the ASVAB coefficients were within .06 of the coefficients of the parent tests. The remaining intercorrelations, which exceeded these limits, involved the Navy Electronics and Radio Test (ER) and the Coding Speed tests, as was the case in the correlation of these tests with the AFQT. In general, then, the pattern of ASVAB intercorrelations differed no more from the parent test intercorrelations than the pattern in one classification battery differed from another. On this basis, the ASVAB tests may be considered to be alternate forms of the parent tests.

Problems encountered in the development of the ASVAB pointed to the need for re-study of the mobilization population and the development of a battery of aptitude tests to serve as reference standards for use in the development of military selection and classification tests.

Table 8

INTERCORRELATIONS WITHIN BATTERIES

ASVAB		Parent Classification Tests[a] Army		Navy		Air Force	
WK-AR	.79	VE-AR	.66	GC-AR	.75	WK-AR	.68
TK	.25	---		MK	.31	---	
SP	.55	PA	.54	---		PC	.52
MC	.65	MA	.51	MC	.61	MP	.64
SI	.62	SM	.56	SP	.54	SP	.44
AI	.54	AI	.57	AK	.45	GM	.51
EI	.69	ELI	.52	ER	.42	EI	.63
CS	.69	CS2	.36	---		---	
AR-TK	.28	AR-		AR-MK	.26	AR-	
SP	.63	PA	.65	---		PC	.59
MC	.67	MA	.54	MC	.52	MP	.67
SI	.58	SM	.56	SP	.46	SP	.44
AI	.52	AI	.39	AK	.40	GM	.48
EI	.68	ELI	.52	ER	.38	EI	.60
CS	.72	CS2	.44	---		---	
TK-SP	.45			MK-			
MC	.55			MC	.72		
SI	.67			SP	.70		
AI	.69			AK	.75		
EI	.54			ER	.35		
CS	.19			---			
SP-MC	.67	PA-MA	.63			PC-MP	.69
SI	.58	SM	.65			SP	.54
AI	.50	AI	.47			GM	.51
EI	.63	ELI	.63			EI	.57
CS	.50	CS2	.34			---	
MC-SI	.71	MA-SM	.70	MC-SP	.71	MP-SP	.69
AI	.69	AI	.60	AK	.74	GM	.72
EI	.74	ELI	.63	ER	.42	EI	.77
CS	.51	CS2	.26	---		---	
SI-AI	.77	SM-AI	.73	SP-AK	.74	SP-GM	.81
EI	.76	ELI	.69	ER	.44	EI	.70
CS	.48	CS2	.25	---		---	
AI-EI	.71	AI-ELI	.63	AK-ER	.34	GM-EI	.77
CS	.39	CS2	.12	---		---	
EI-CS	.51	ELI-CS2	.20				

[a] Corrected for selection on AFQT.

APPENDIXES

APPENDIX A IDENTIFICATION OF TESTS

The tests used in the analysis are identified by the following variable numbers and abbreviations.

1.	Armed Forces Qualification Test	AFQT
	Army Classification Battery	ACB
2.	Verbal	VE
3.	Arithmetic Reasoning	AR
4.	Pattern Analysis	PA
5.	Mechanical Aptitude	MA
	Army Clerical Speed[a]	
6.	Digit Substitution	CS1
7.	Coding Speed	CS2
8.	Shop Mechanics	SM
9.	Automotive Information	AI
10.	Electronics Information	ELI
	Navy Basic Test Battery	BTB
11.	General Classification	GC
12.	Automotive Reasoning	AR
13.	Mechanical Knowledge	MK
14.	Mechanical Comprehension	MC
15.	Clerical	CL
16.	Shop Practices	SP
	Electronic Technicians Selection Test[b]	
17.	Mathematics	MATH
18.	Science	SCI
19.	Electronics and Radio	ER
20.	Automotive Knowledge[c]	AK
	Air Force Airman Qualifying Examination[d]	AQE
21.	Arithmetic Computation	AC
22.	Word Knowledge	WK
23.	Arithmetic Reasoning	AR
24.	Hidden Figures	HF
25.	General Mechanics	GM
26.	Electrical Information	EI
27.	Mechanical Principles	MP
28.	Shop Practics	SP
29.	Data Interpretation	DI
30.	Pattern Comprehension	PC

---

[a] The Army Coding Speed Test consists of two parts but is operationally scored as one test. However, for purposes of this study, two part scores were obtained.

[b] The Navy Electronics Technicians Selection Test is operationally not a part of the Basic Test Battery.

[c] At the time of this study, the Navy Automotive Knowledge Test had not yet been incorporated in the Basic Test Battery.

[d] The Air Force Airman Qualifying Examination used in this study differed from the operational AQE in that easier items were added.

APPENDIX B

INTERCORRELATIONS OF ARMY, NAVY, AND AIR FORCE CLASSIFICATION TESTS
(Corrected for selection on AFQT and unreliability; stratified sample, N = 2000)

	ARMY										NAVY										AIR FORCE									
Variable	1 AFQT	2 VE	3 AR	4 PA	5 MA	6 CS1	7 CS2	8 SM	9 AI	10 ELI	11 GC	12 AR	13 MK	14 MC	15 CL	16 SP	17 MATH	18 SCI	19 ER	20 AK	21 AC	22 WK	23 AR	24 HF	25 GM	26 EI	27 MP	28 SP	29 DI	30 PC
1	100[a]	79	83	96	84	47	43	95	70	93	93	78	61	86	36	72	71	78	49	69	56	81	82	51	83	82	87	75	86	79
2	79	100	75	65	62	42	44	70	41	66	100	70	22	54	35	49	64	70	39	41	54	96	71	40	53	63	63	43	75	52
3	83	75	100	81	69	56	55	72	45	69	88	93	29	61	45	52	83	72	45	45	78	77	95	46	54	64	71	47	89	65
4	96	65	81	100	84	52	45	88	58	88	81	73	58	84	41	63	72	73	47	58	58	67	76	59	71	74	88	69	87	95
5	84	62	69	84	100	43	35	96	75	89	77	63	69	94	29	73	62	78	51	72	47	67	71	49	83	81	88	74	80	68
6	47	42	56	52	43	100	79	42	20	33	50	49	11	30	52	19	48	34	20	13	58	45	54	31	21	28	38	20	55	39
7	43	44	55	45	35	79	100	34	15	28	49	49	07	22	57	16	47	34	20	11	59	46	53	26	18	25	32	17	53	34
8	95	70	72	88	96	42	34	100	93	100	84	66	82	98	27	90	61	77	52	86	49	74	72	48	103	92	92	95	80	71
9	70	41	45	58	75	20	15	93	100	82	52	42	83	83	14	78	33	53	35	97	25	46	47	32	103	78	72	85	53	47
10	93	66	69	88	89	33	28	100	82	100	80	62	83	94	21	87	65	84	67	85	41	70	67	42	97	101	91	86	77	72
11	93	100	88	81	77	50	49	84	52	80	100	85	36	73	42	61	78	84	51	53	64	103	87	52	66	76	80	56	92	67
12	78	70	93	73	63	49	49	66	42	62	85	100	28	59	46	48	77	67	43	43	72	72	89	42	51	60	66	45	83	60
13	61	22	29	58	69	11	07	82	83	83	36	28	100	82	11	76	25	45	40	82	14	28	33	29	90	70	65	84	42	49
14	86	54	61	84	94	30	22	98	83	94	73	59	82	100	27	80	54	73	49	85	36	60	63	44	92	83	92	83	71	73
15	36	35	45	41	29	52	57	27	14	21	42	46	11	27	100	11	41	28	14	11	50	37	45	27	17	21	30	17	43	34
16	72	49	52	63	73	19	16	90	78	87	61	48	76	80	11	100	44	62	49	81	29	55	52	32	89	81	72	84	60	55
17	71	64	83	72	62	48	47	61	33	65	78	77	25	54	41	44	100	72	49	36	71	68	80	43	44	59	64	38	80	60
18	78	70	72	73	78	34	34	77	53	84	84	67	45	73	28	62	72	100	58	56	50	76	72	42	67	77	77	58	79	63
19	49	39	45	47	51	20	20	52	35	67	51	43	40	49	14	49	49	58	100	39	40	42	46	41	46	62	52	41	52	42
20	69	41	45	58	72	13	11	86	97	85	53	43	82	85	11	81	36	56	39	100	20	46	47	26	100	80	73	84	53	50
21	56	54	78	58	47	58	59	49	25	41	64	72	14	36	50	29	71	50	40	20	100	56	76	50	30	44	49	28	72	46
22	81	96	77	67	67	45	46	74	46	70	103	72	28	60	37	55	68	76	42	46	56	100	77	39	60	70	70	50	81	56
23	82	71	95	76	71	54	53	72	47	67	87	89	33	63	45	52	80	72	46	47	76	77	100	48	56	67	74	51	92	64
24	51	40	46	59	49	31	26	48	32	42	52	42	29	44	27	32	43	42	41	26	50	39	48	100	35	41	52	34	55	48
25	83	53	54	71	83	21	18	103	103	97	66	51	90	92	17	89	44	67	46	100	30	60	56	35	100	91	84	97	64	69
26	82	63	64	74	81	28	25	92	78	101	76	60	70	83	21	81	59	77	62	80	44	70	67	41	91	100	85	80	75	62
27	87	63	71	88	88	38	32	92	72	91	80	66	65	92	30	72	64	77	52	73	49	70	74	52	84	85	100	77	83	74
28	75	43	47	69	74	20	17	95	85	86	56	45	84	93	17	84	38	58	41	84	28	50	51	34	97	80	77	100	60	60
29	86	75	89	87	80	55	53	80	53	77	92	83	42	71	43	60	80	79	52	53	72	81	92	55	64	75	83	60	100	74
30	79	52	65	95	68	39	34	71	47	72	67	60	49	73	34	55	60	63	42	50	46	56	64	48	59	62	74	60	74	100

[a]Decimal points omitted.

APPENDIX C

MEANS AND STANDARD DEVIATIONS OF
ARMY CLASSIFICATION BATTERY - TEST AND RETEST[a]
(N = 367)

	Test		Retest	
	Mean	Standard Deviation	Mean	Standard Deviation
VE	33.0	12.2	32.7	11.9
AR	18.7	9.3	18.4	9.5
PA	25.5	12.3	24.3	13.4
MA	29.1	6.9	28.5	6.8
CS1	29.0	8.8	28.0	9.0
CS2	29.6	9.6	26.6	8.6
SM	22.6	5.5	22.3	5.7
AI	21.6	9.2	22.8	8.7
ELI	18.8	7.3	18.1	7.4
AFQT[b]	50.8	25.6	57.2	25.7

[a] See Table 1 on page 16.
[b] Computed in Army sample.

DISTRIBUTION

BESRL Distribution List

1 O/Dep ASD (Ed)
1 Dir., Army Rsch, OCRD
2 Dir., Army Rsch.--Dep Dir & Scient Dir.
2 OCRD, DA (C/Phys. Engr Sciences; C/Scient Anal Br.)
5 ODCSPER, DA--(CSD, OCP, Pers Rsch Div; Dir., Pers Studies & Rsch.; C/Distr Div.)
7 OACSFOR, DA--(Ofc Info & Data Sys; Cml & Nuc Opns; Doctr & Sys; Org, Unit Tng & Readiness; Mpr & Forces)
1 ACSI, DA, Attn: ACSI-DSRS
1 CINFO, DA--Community Rel Div.
5 OPO, DA--C/Stds Sys Ofc.
4 DDR&E
1 TAGO, DA--Sep & Ret Div.
1 ODWAC, DA
1 O/Chief of Chaplains, DA, Attn: CHPL
1 Dep Pres., USA Sec Bd.
2 CG, USCONARC--Attn: G-1, ATPER
20 CG, USCONARC--DCS Ind Tng, Attn: C/Rsch Div.
6 CG, USACSC, Ft. Belvoir
6 CG, USACDC, Cmbt Army Gp., Ft. Leavenworth
2 CG, USA Security Agcy.
1 CG, USAMC, Attn: AMCRD-E
3 CG, USAHEL, Aberdeen PG--(Tech Dir; Dir/HEL; Sptg Rsch Lab.)
1 AVCofSA (Asst Vice Chief of Staff, Army)
1 CG, USAECOM, Ft. Monmouth--(AMSEL-[illegible] GDA; AMSEL-CB-AS)
2 USACDC, AG Agcy, Ft. Benjamin Harrison
3 CO, USACDC, Fld Agcys (Ft. Benning, Ft. Bragg, Ft. Knox, Ft. Rucker, Ft. Sill.)
2 USACDC Exper Cmd, Ft. Ord--(Tech Library; Plans Off. (Psychol))
1 AVCofSA (DMIS)--Dir, Mgt Info Sys
2 CO, USAEEC, Ft. Benjamin Harrison
1 USAEEC, C/Eval & Anal Br., Ft. Benjamin Harrison
2 USAF--(AFPDCM, AFNICAD)
1 Cmdt, Marine Corps (Code AOIB)
2 Dir, HumRRO, Attn: Library
6 Dirs. of Rsch., HumRRO Fld Divs. (Ft. Benning, Ft. Bliss, Ft. Knox, Ft. Rucker, Presido Monterey)
1 USAMRL, C/Exper Psych., Ft. Knox
3 USA Natick Labs, Natick--(Tech. Library; Psych Labs.; Pioneer Rsch Div.)
1 CO, Naval Tng Dev Cen., Natick--Attn: Tech Library
2 CO, USA Garrison, Tech Ref Ofc., Ft. Huachuca
1 CO, USA Gen Equip Test Actv., Ft. Lee, Attn: C/HF Div., MED.
2 CO, USA Electronic PG, Ft. Huachuca--(Attn: STEEP-SU-I; Test Dte, STEEP-T-BI)
1 CO, U. S. Army Edgewood Arsenal, Attn: SMUEA-TS-D
1 CO, USAR&D Gp (FE)
3 CO, USAR&D Office--(Durham, Alaska, Arctic Test Cen.)
1 C/USA R&D Liaison Gp (Europe)
1 C/Intel Mtrl Dev Ofc (EWL), Ft. Holabird
1 Per Rsch Div (HRP), AFHRL, Lackland
2 Cmdt, USAF Sch Aerospace Med., Aeromed. Library, Brooks AFB.
3 ONR, Pers Tng Rsch.--Attn: Assoc. Dir., Code 458
1 OIC, USN Med NP Rsch Unit., San Diego
1 Naval Pers & Tng Rsch Lab., San Diego, Attn: Library
2 Pers Rsch Lab., WNY, Attn: Library
3 C/Naval Pers (Pers A3), Arlington Annex
2 ONR--(Dir., Physiol Psych Br (Code 454); Head, Gp Psych Br (Code 452))
1 C/BuM&S, Rsch Psych Br (Code 712), Navy
1 Hq AFLC (MCACE), Wright-Patterson AFB
2 Dir., WRAIR--(Dept Clin Soc Psych; NP Div.)
1 C/Per Rsch Staff, OP, USDA
1 USAMERDC, HF Engr Tech & Rsch Spt Ofc., Ft. Belvoir
1 USDA Forest Svc., Pacif NW., F&R Exper Sta., Seattle Unit
1 Dir., NSA, Attn: M 13, Ft. Meade
1 CIA, Attn: OCR/DD/Publ.
1 Cen for Pop Rsch., Natl Inst for Child Hlth & Hum Develop.
1 OPMG, DA--(Admin-Mgt Ofc., C/MP Plans & Opns Div.)
1 Sixth US Army Library, Presidio SF
1 USA Med R&D Cmd., Attn: MEDDH-RB
1 OTSG, DA--C/Preventive Med Div.
1 CINCPACAF (IGSAG)
1 C/Special Proj., Sta 3-12 (P-1), USCG Hq.
4 USA Stdzn. Gp.--(Canadian Forces Hq.; UK.)
1 USA Tropic Test Cen., Attn: STETC-RE, Ft. Clayton
1 US Department of Labor, Attn: MA, USTES
1 BuM&S (513), Department Navy, Attn: Head, Avn Opnl Psych Br.
1 Dir., US Naval Rsch Lab., Attn: Code 2027
1 C/Engr Dept., Indust Engr Br., USA Engr R&D Labs., Ft. Belvoir
1 C/Behavioral Sci Div., USA Rsch Inst of Environ Med.
1 USA Liaison Ofc., Tact Air Recon Cen., Shaw AFB
1 Naval Tng Dev Cen., Orlando, Attn: Head, HF Lab (Code 55)
1 Naval Pers & Tng Rsch Lab., San Diego--(Dirs., Stat Dept.; Pers Msrmt & Rsch Dept.)
1 USA Adv Mtrls Concept Agcy (AMXAM-AC)
1 Info Acq Unit, Natl Clghse Mntl Hlth Info., Natl Inst of Mntl Hlth.
1 Cmdt., C&GSC, Ft. Leavenworth, Attn: Ed Advsr.
1 Cmdt., US Army Mgt Sch., Ft. Belvoir
1 Cmdt., USA Arty Msl Sch., Ft. Sill, Attn: Dir Ed., Ed Svcs Div.
1 Cmdt., USA Msl Mun Cen Sch., Redstone Arsenal, Attn: Tech Librarian
1 Cmdt., USAWC, Carlisle Barracks, Attn: Library
2 Cmdt., USA Air Def Sch., Ft. Bliss, Attn: Tech Library
1 Cmdt., USASA Tng Cen & Sch., Ft. Devens
1 Cmdt., USA Armor Sch., Ft. Knox, Attn: AHBAAS-DMP
2 Cmdt., USAIS, Ft. Holabird, Attn: AHBHS-S-CB (ASRC)
1 Cmdt., USA SE Signal Sch., Ft. Gordon, Attn: Ed Advsr.
6 Cmdt., C&GSC, Ft. Leavenworth
2 Dir., Mil Psych & Ldrship., USMA
3 Supt., PDO 4013A, USAF Academy
1 Dir., Air Univ Library, Maxwell AFB, Attn: AUL (3T-AUL-60-113)
1 Dir., Admissions, USCG Academy
1 CG, USA Inf Cen., Ft. Benning, Attn: AJIGP
1 Dir of Rsch., O/Rsch., West Point
1 Dir., Gnd Cmbt Div., Air Gnd Opns Sch., Eglin AF Aux #9
1 CO, Naval Aerospace Med Inst., Aerospace Med Cen., Pensacola, Attn: Code 31
1 Chrmn., Test Rsch & Develop., USAFI
1 Pres., USA Avn Test Bd., Ft. Rucker, Attn: STEGB-TP-V
1 USA Inf Bd., Ft. Benning, Attn: Secy-ATBD
1 O/Dir of Instruct., USASW Sch., Ft. Bragg

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
U. S. Army Behavior and Systems Research Laboratory, OCRD, Arlington, Virginia	Unclassified
	2b. GROUP

3. REPORT TITLE

THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Abram G. Bayroff and Edmund F. Fuchs

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February 1970	42	

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. DA R&D Proj. No. 2Q062106A722	Technical Research Report 1161
c. Enlisted Manpower	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d. 00-01	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office, Chief of Research and Development, DA

13. ABSTRACT In response to the directive issued by the Assistant Secretary for Defense (Manpower and Reserve Affairs), in February 1966, to develop a common aptitude battery for use by all the services, the Army and BESRL served as the lead service. The battery was to provide common aptitude measures to be used by all the services, as well to provide an overall measure for the Armed Forces Qualification Test. The objective of the study reported was to identify among classification tests of the Army, Navy, and Air Force, those which were interchangeable in terms of abilities and aptitudes measured; and from those so identified, to develop shortened forms to constitute an alternate inter-service battery which would not require testing time in excess of two and one-half hours. Comparability of the several service tests was determined from test intercorrelations in a consolidated enlisted input sample (N = 1000 each Army, Navy, Air Force; 300 Marine Corps) which was stratified on AFQT to provide a mobilization distribution. Correlation for restriction on AFQT and for unreliability (test-retest with alternate forms) was made. The new battery derived (Armed Services Vocational Aptitude Battery, ASVAB) was standardized on a 3000-man sample of Selective Service registrants, again stratified on AFQT.

Seven sets of tests were identified as interchangeable: word knowledge, arithmetic reasoning, space perception, mechanical comprehension, shop information, automotive information, and electronics information. The Army Coding Speed Test was selected as a measure of clerical aptitude on the basis of separate validity studies. Tool Knowledge, an eighth test, was added to provide AFQT scores. Similarity of patterns of relationships was revealed among ASVAB tests and of ASVAB with AFQT to those of the parent tests. The ASVAB is currently being used to test potential recruits in the senior year of high school.

DD FORM 1 NOV 68 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Psychological measurement						
*Armed Services Vocational Aptitude Battery						
*ASVAB						
*Classification tests						
*Sampling techniques						
Interchangeable tests						
AFQT						
ACB						
Navy Basic Test Battery						
Air Force Airman Qualifying Examination						
Common test battery - Services						
Classification test battery - Services						
Screening and Classification Systems						
Psychometrics						
Military psychology						