DEPARTMENT OF DEFENSE

Hem
0309

ARMED SERVICES VOCATIONAL APTITUDE BATTERY

(ASVAB)

TEST MANUAL

FOR ASVAB FORMS 8, 9, 10, 11, 12, 13, and 14

DOD 1304. 12AA

1 JUL 84

Test Manual for the

Armed Services Vocational Aptitude Battery

United States Military Entrance Processing Command

2500 Green Bay Road

North Chicago, Illinois 60064

DoD 1304.12AA July 1984

PREFACE

This manual documents the procedures used in development of the Armed Services Vocational Aptitude Battery (ASVAB), reports evidence of its technical merit, and reports the extent of its compliance with ethical principles for the development, validation, and use of personnel selection procedures. It is written for technical personnel in the test development and analysis field.

This document is based upon technical publications of the personnel research activities of the Army, Navy, Marine Corps and Air Force, and upon formal and informal memoranda from the Department of Defense offices involved in policy oversight of the Armed Services vocational aptitude testing programs. The volume of research generated in the development and operation of the testing program is so great that it is not practical to cite all publications.

The authors wish to acknowledge gratefully the contributions of Dr. W. S. Sellman, Office of the Assistant Secretary of Defense; Dr. Malcolm J. Ree, Major John R. Welsh, Mr. James M. Wilbourn, and Mr. John J. Mathews, Air Force Human Resources Laboratory; Dr. William H. Sims, Center for Naval Analyses; Dr. Clessen J. Martin, Army Research Institute for the Behavioral and Social Sciences; and Colonel Edward E. Gerding, Military Entrance Processing Command. A special acknowledgement is made to Dr. Leland D. Brokaw, who compiled and organized much of the information in this manual from a wide variety of sources. The authors of this manual are Dr. C. Wayne Shore and Dr. Benjamin A. Fairbank, Jr., both of McFann-Gray & Associates, Inc.

Questions as to policy and management of the program may be addressed to:

Office of the Assistant Secretary of Defense (MI&L)
Director for Accession Policy
The Pentagon
Washington, D. C. 20301

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

Test Manual for the Armed Services Vocational Aptitude Battery

Chapter 1

Development of the Various Forms

A. Background

The use of selection and classification tests has become an accepted procedure for assigning persons to occupational specialties. Since World War I, the Armed Services have understood that the more accurate the match between the capabilities of recruits and the requirements of military occupations, the more effective the use of personnel resources. Mass testing procedures were used to test millions of entering military personnel in World War II to provide measures of potential for training and to screen for a few selected career fields.

In recent years the Armed Services Vocational Aptitude Battery (ASVAB) has been the enlisted military personnel selection and classification test and has been used for recruitment activities in high schools. The ASVAB is a Department of Defense (DoD) instrument developed jointly by the Armed Services.

The content of the ASVAB is based upon research programs conducted by the manpower and behavioral science laboratories of the various Services, programs which began immediately after World War II. Materials selected for inclusion in the ASVAB have demonstrated ability to predict performance in technical training.

The differential measurement of abilities needed in various occupations across all vocational areas became the focus of military classification research in the late 1940s. By the 1950s the Army, the Navy, and the Air Force were using classification batteries developed separately by each Service for enlisted personnel.

Competition between the Services for highly able recruits led to the Selective Service Act of 1948, which addressed the appropriate distribution of manpower. As a device to promote equitable distribution of both higher and lower ability personnel, the Armed Forces Qualification Test (AFQT) was developed as a joint-Service project with the Army providing the lead laboratory. The AFQT was standardized against the Army General Classification Test (AGCT) score distribution of all men under arms as of December 31, 1944. That population is hereinafter referred to as the 1944 reference population. It was scored in a percentile metric, but provided the same qualitative categories as

yielded by the AGCT. The percentile limits of the categories are shown in Table 1 (Uhlaner & Bolanovich, 1952).

Table 1

Percentile Limits of Mental Category Scores

Category	Percentile Limits
I	93 - 99
II	65 - 92
III	31 - 64
IV	10 - 30
V	1 - 9

Initially the AFQT was used to assign established proportions of high ability (Category I and II) personnel to each Service as well as fair shares of the lesser ability (Category IV) personnel, and to serve as a screen for denial of enlistment to the least qualified applicants. Since the implementation of the All-Volunteer Force, the AFQT categories have remained as indices of ability for comparison of the distribution of recruit ability in the various Services.

In 1958, the Air Force first introduced a military aptitude test battery into the nation's high schools. The Airman Qualifying Examination, a short version of the Airman Classification Battery, was provided without charge to high schools for use in their vocational counseling programs. This practice also provided Air Force recruiters with test results which were useful for recruiting purposes. Shortly thereafter the Army and the Navy instituted similar programs.

In 1966, the DoD directed the Services to explore jointly the development of a testing instrument to be used for recruiting purposes by all the Services, replacing the short tests used in the high schools and the longer enlisted classification batteries. The new test battery was expected to determine mental qualification for selection of applicants, and for classification and assignment of recruits. It was also to provide a measure similar to the mental category scores provided by the AFQT (Bayroff & Fuchs, 1970).

The resulting ASVAB is the sole instrument used for military enlistment and classification testing. The ASVAB program is directed by the Manpower Accession Policy Steering Committee, composed of high-ranking officers from

the personnel division of each Service headquarters, the Commander of the Military Entrance Processing Command (MEPCOM), and chaired by the Director for Accession Policy from the Office of the Assistant Secretary of Defense (OASD) (Manpower, Installations, and Logistics) (MI&L).[1] Planning, research, and development are accomplished by the Joint-Service Selection and Classification Working Group, composed of testing policy staff officers from each Service, research scientists from each Service's personnel research laboratory, and representatives from MEPCOM. The efforts of the steering committee and the working group are reviewed by the Defense Advisory Committee on Military Personnel Testing, composed of eminent personnel measurement experts from the civilian community (OASD/MRA&L, 1980).

The Air Force Human Resources Laboratory is the lead laboratory for research and development in support of the ASVAB program. MEPCOM implements the Army's responsibility for accomplishment of operational ASVAB testing and score processing.

B. Chronology of ASVAB Form Development

Form 1 of ASVAB, used in high schools for school years 1968-69 through 1973-74, was developed from tests with counterparts in each of the Service batteries. Items were selected to produce tests shorter than the parent tests, so that total testing time would not exceed two and one-half hours.

Forms 2 and 3 of ASVAB were alternate forms, similar to Form 1. Form 2 was used in the high school program during school years 1973-74 through 1975-76. Form 3 was used for Air Force selection and classification from 1973 through 1976, and for Marine Corps selection and classification starting in 1975.

Form 4 was developed as a back-up to Form 2, but was never implemented as an operational test. The demonstrated effectiveness of ASVAB Forms 2 and 3 prompted DoD to direct the Military Services jointly to develop and employ a single battery for use in both high school testing and in the Military Entrance Processing Stations (MEPSs). The new battery would screen for enlistment and provide aptitude data for initial classification and assignment decisions.

---

[1] Until 1984, OASD(MI&L) was designated as the Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs and Logistics) (OASD/MRA&L).

Forms 5, 6, and 7 replaced the Service-unique classification batteries administered before entry into basic training, and nearly all mental testing for selection and classification was conducted either by the MEPSs or in the high school testing program. Forms 6 and 7 were implemented in the MEPSs in January, 1976, and Form 5 was introduced into the high school program in July of the same year.

Forms 8, 9, and 10 replaced Forms 6 and 7 as military selection and classification measures in October, 1980. Forms 8, 9, and 10 were designed to be more accurate at lower levels of ability than were the predecessor tests. They also provide a broader measure of verbal skill than did the earlier forms.

Forms 11, 12, and 13 have been developed as forms parallel to Forms 8, 9, and 10 (Prestwood, Vale, & Massey, in press). Form 14 is also parallel to Forms 8, 9, and 10 (J. M. Wilbourn, personnel communication, February 28, 1984). During 1984, Forms 11, 12, and 13 are scheduled to replace Forms 8, 9, and 10, and Form 14 is scheduled to replace Form 5. Information contained in this report concerning content, length, and administration times of Forms 8, 9, and 10 also applies to Forms 11, 12, 13, and 14. The most recent, comprehensive information involves Forms 8, 9, and 10. Therefore, this manual contains more information on those versions than on other versions.

C. Application and Content

The ASVAB continues the series of military vocational selection and classification instruments based upon continuing programs of research in each of the Services. The Army, the Navy, and the Air Force independently had developed in-house batteries to meet their specific needs. Service-unique differences were present, but the general approach and the validation strategies were common to all Services. The general approach consisted of identifying the criteria of interest, then assessing the potential of various tests to predict those criteria. The validation strategy consisted of continuously validating the selected subtests against the same criteria which were to be predicted (Thomas, 1970; Uhlaner, 1968; Uhlaner & Bolanovich, 1952; Weeks, Mullins, & Vitola, 1975).

During the period from 1945 through 1980, most military selection and classification tests were evaluated in terms of their ability to predict success in specific technical training courses. These technical training courses have been

based upon requirements for trained personnel in military occupations. Measures of job performance were not readily available for many of the occupations. Since 1970, military technical training has been objectively based upon job requirements defined through computer-based task analyses of vocational specialties (Maier & Fuchs, 1972; Maier & Grafton, 1981; McCormick, 1979; Morsh & Archer, 1967; Swanson, 1979; Vitola & Alley, 1968; Yellen & Foley, 1978).

The content of the ASVAB reflects those subject areas which have shown validity through prediction of training criteria in each of the Services. Forms 8, 9, and 10 evolved from previous ASVAB forms and from Service classification batteries which had been found valid for use in personnel classification programs. In the first instance, the content of the classification battery was selected to provide measurement of the different types of skills and knowledge areas found necessary in military jobs by occupational analysts. Further modification of ASVAB content was accomplished through factor analytic methods. Factor analyses have been performed of the combined classification batteries of all the Services (Zachert, 1952) and of various forms of the ASVAB (Bock & Moore, 1984; Fischl, Ross, & McBride, 1979; McBride, 1981; Ree, Mullins, Mathews, & Massey, 1982; Sims & Mifflin, 1978).

D. Subtest Selection

The test content of Forms 8, 9, and 10 was approved by the Joint Service Selection and Classification Working Group on the basis of research studies accomplished by the various Services. These forms contain three subtests not on the previous forms: Coding Speed, Paragraph Comprehension, and Auto and Shop Information. The General Information, Space Perception, and Attention to Detail subtests, and the Classification Inventory of the prior batteries were deleted from the revised forms because they made little unique contribution to the validity of the composites in which they appeared. The subtests contained in the various forms are listed in Table A-1 in Appendix A.

Coding Speed had demonstrated useful validity in the prediction of some Army criteria. In previous forms, Automotive Information and Shop Information, each with 20 questions, were highly intercorrelated and had shown similar validity patterns, so the two topical areas were combined as a single subtest, Auto and Shop Information, with 25 questions.

A need for a measure of reading ability, as well as a need for better measurement of verbal ability, brought about the introduction of Paragraph Comprehension.

The AFQT composite score is used by all Services as an indicator of general trainability. The test content providing the AFQT score from Forms 8, 9, and 10 includes Word Knowledge, Arithmetic Reasoning, Paragraph Comprehension, and Numerical Operations. The composite score contains measures of numeric, verbal, and reasoning factors, as well as a measure of reading ability. The AFQT score derived from subtests in Forms 8, 9, and 10 involves more subtests and more items than previous AFQT composites, and therefore is expected to be more resistant to compromise.

Subtests comprising Forms 8, 9, and 10 are listed in Table 2. The content of the current generation of subtests can be compared with the content of earlier forms in Appendix A, which presents descriptions of tests previously used in the ASVAB program and the content of each generation of tests.

Because the AFQT score establishes applicants' qualifications for enlistment, both recruiters and applicants have a strong interest in the applicants' passing the test. Most test compromise has been in the AFQT portions of the ASVAB. Therefore, six versions of the AFQT subtests were prepared for use in Forms 8, 9, and 10, and three versions of the non-AFQT subtests were constructed. Thus, Forms 8, 9, and 10 consist of six forms. These six forms are designated 8a, 8b, 9a, 9b, 10a, and 10b. The like-numbered forms (e.g., 8a and 8b), however, differ only in the AFQT subtests. Thus, Form 8a contains one of the six versions of the AFQT and the same set of non-AFQT subtests as Form 8b, but the positions of the items in the non-AFQT subtests in Forms 8a and 8b differ from each other, so that the scoring keys are different. No two forms of the ASVAB contain the same AFQT items.

In summary, the AFQT subtests, prepared in six versions, include Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, and Numerical Operations. The non-AFQT subtests, prepared in three versions, include General Science, Coding Speed, Auto and Shop Information, Mathematics Knowledge, Mechanical Comprehension, and Electronics Information.

Table 2

Subtest Composition of Forms 8, 9, and 10

Content Area or Subtest	Abbreviation	Number of Questions	Testing Time (minutes)
General Science	GS	25	11
Arithmetic Reasoning[a]	AR	30	36
Word Knowledge[a,b]	WK	35	11
Paragraph Comprehension[a,b]	PC	15	13
Numerical Operations[a]	NO	50	3
Coding Speed	CS	84	7
Auto and Shop Information	AS	25	11
Mathematics Knowledge	MK	25	24
Mechanical Comprehension	MC	25	19
Electronics Information	EI	20	9
TOTAL		334	144

Note. From Normalization of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 8, 9, and 10 using a sample of Service Recruits (CRC 438) by W. H. Sims and A. R. Truss, 1980, Alexandria, VA: Center for Naval Analyses. Adapted by permission.

[a] Armed Forces Qualification Test score:
AFQT = AR+WK+PC+½NO (Raw Scores)

[b] Verbal score: VE = WK+PC (Raw Scores)

Chapter 2
Calibration, Equating and Score Development

A. Calibration Methods[1]

Raw test scores do not have meaningful units, in part because they vary with the difficulty of the items which make up the test. Test scores are needed which are meaningful even when test difficulty changes. It has thus become standard to express AFQT scores as percentiles. Changes in the ability of the enlisted population, however, can lead to changes in the percentile score corresponding to a given ability. The DoD therefore has referred AFQT percentile scores not to the norms for a current version of the AFQT, but to the abilities of the 1944 reference population. Through test equating and calibration, it has been possible to report AFQT scores on all AFQT tests up to and including those which use subtests of Forms 8, 9, and 10 in terms of the 1944 reference population. The remainder of this section describes the methods used to calibrate Forms 8, 9, and 10 to the 1944 reference population.

Three independent studies were designed by the Joint-Service Selection and Classification Working Group to calibrate Forms 8, 9, and 10. The design specified that only one reference test, AFQT-7a, would be used. (Note: AFQT-7a is a pre-existing, standardized form of the AFQT, originally introduced in 1960 and used operationally through 1972, not a subsection of Form 7 of the ASVAB.)

The AFQT-7a and Forms 8, 9, and 10 were administered in counterbalanced order to more than 8,000 examinees. Each examinee took AFQT-7a and one form of the ASVAB. To ensure that calibrations would apply to all relevant populations, three samples were specified: applicants for enlistment, new recruits from all Services, and high school students in grades 11 and 12. Since the 1944 reference population contained only males, the calibration samples were also restricted to males. The conventional equipercentile equating technique was used in all the studies.

---

[1] Much of the material in this section was taken from Maier (1981b, pp. 11-22) and was analyzed by Educational Testing Service, Inc. (Boldt, 1980a; Sims & Truss, 1980).

The analysis of each sample was carried out independently. The sample of applicants for enlistment was analyzed through the combined efforts of the OASD (MRA&L) and the Army Research Institute, with Dr. Milton H. Maier as the principal investigator (Maier, 1981b). The sample of Service recruits was analyzed by the Center for Naval Analyses (CNA), with Dr. William H. Sims as the principal investigator. The sample of high school students was analyzed by Educational Testing Service, Inc. (ETS) with Dr. Robert Boldt as the principal investigator (Boldt, 1980c; Sims & Truss, 1980).

The score scale for Forms 8, 9, and 10 was based on the combined sample of applicants and recruits. The calibration was to express the scores of the calibrated version on the same metric as that used for the 1944 reference population. Such comparability would make possible a comparison of the relative ability of the 1944 reference population and the ability of those tested by Forms 8, 9, and 10.

1. Enlistment Applicants

A nationally representative sample of applicants for enlistment was tested at the MEPSs. The reference test (AFQT-7a) and Form-8aX (preoperational version of Form 8a) were administered to all enlistment applicants in the sample. The data collection began in January, 1980, and was completed in February, 1980. Each MEPS was briefed on the study by a representative of the Joint-Service Selection and Classification Working Group. Each representative reported that personnel at the MEPSs followed good testing practices in the sessions observed. Of equal importance was the cooperation of the recruiters in forwarding applicants for testing. On past occasions recruiters may have selectively withheld applicants to avoid experimental testing or sent them to the mobile testing stations (alternate testing sites when testing at a MEPS was not feasible) where no extra testing occurred. This potential problem was avoided in this study as enlistment applicants tested at the mobile testing stations were included in the study. The sample was designed to be representative of the applicants processed by the MEPSs at that time.

All experimental tests were administered before the operational tests. Fatigue, therefore, should not have affected the test scores, and because of the counterbalanced administration of AFQT-7a and Form-8aX, motivation should have been equal for both the reference and new tests.

As a check on the quality of the test data, regression analyses were used to identify deviant test scores. One analysis was to predict the Form-8aX AFQT

score, called AFQT-8aX, from AFQT-7a, and another analysis was to predict the Numerical Operations score from the Arithmetic Reasoning score. Persons whose scores deviated by more than two standard errors of estimate were deleted from the sample.

The original sample consisted of 2,620 male applicants. Of this number, five percent had deviant AFQT-8aX or AFQT-7a scores. An additional four percent had deviant Numerical Operations or Arithmetic Reasoning scores. The final sample of MEPS applicants consisted of 2,375 cases of which about 33 percent were black and about 10 percent were Hispanic.

2. Service Recruits

A sample of recruits drawn from the current population of new enlisted accessions was used for this analysis. Each Service provided its proportional share of the sample (Army — 43 percent; Navy — 23 percent; Air Force — 20 percent; Marine Corps — 13 percent (percents were rounded)). Form-8aX and AFQT-7a were administered to 3,799 male recruits from all Services. The tests were administered at special sessions conducted by personnel from the recruit reception centers. Each reception center was briefed on the study by a representative of the Joint-Service Selection and Classification Working Group who observed at least one testing session. The CNA also applied regression-based editing to the data to remove cases with deviant test scores.

The editing methodology differed from that which was used for the applicant sample. The intent was to remove both deviant test sessions and deviant individuals.

The first step was to compute AFQT-8aX and AFQT-7a means for each testing session. There were 44 test sessions. A regression analysis was used to identify deviant testing sessions. Sessions that deviated more than 2.5 standard errors of estimate from the regression line were deleted. Nine of the 44 sessions were deviant, and all cases from these sessions were deleted.

The second step was to identify individuals with deviant scores. The average regression between AFQT-7a and AFQT-8aX was computed, and cases found to be more than 2.5 standard errors of estimate from the average regression line were deleted. Of the original 3,799 cases, 13 percent were deleted because of faulty testing sessions and another three percent were deleted because of deviant AFQT scores. Finally, another five percent were deleted because their operational test scores were not available. The final recruit sample was 3,001 cases.

An additional factor that may affect the calibration is the racial/ethnic mix of the sample. The final recruit sample was weighted to represent the assumed mix in 1959, when AFQT-7a was calibrated. The assumed mix was 82 percent white, 12 percent black, and 6 percent other.

3. High School Students

Schools throughout the country that had participated in the ASVAB High School Testing Program were requested by ETS to administer the experimental tests. Of the 180 schools contacted, 40 agreed to participate. In their editing of the data, ETS deleted nine percent of the cases because the examinees attempted very few items on one or more tests. Another one percent of the examinees were deleted because their answer sheets were lost, or mutilated, or because of a testing irregularity. The scores of all female students were deleted, which left 1,745 usable male cases.

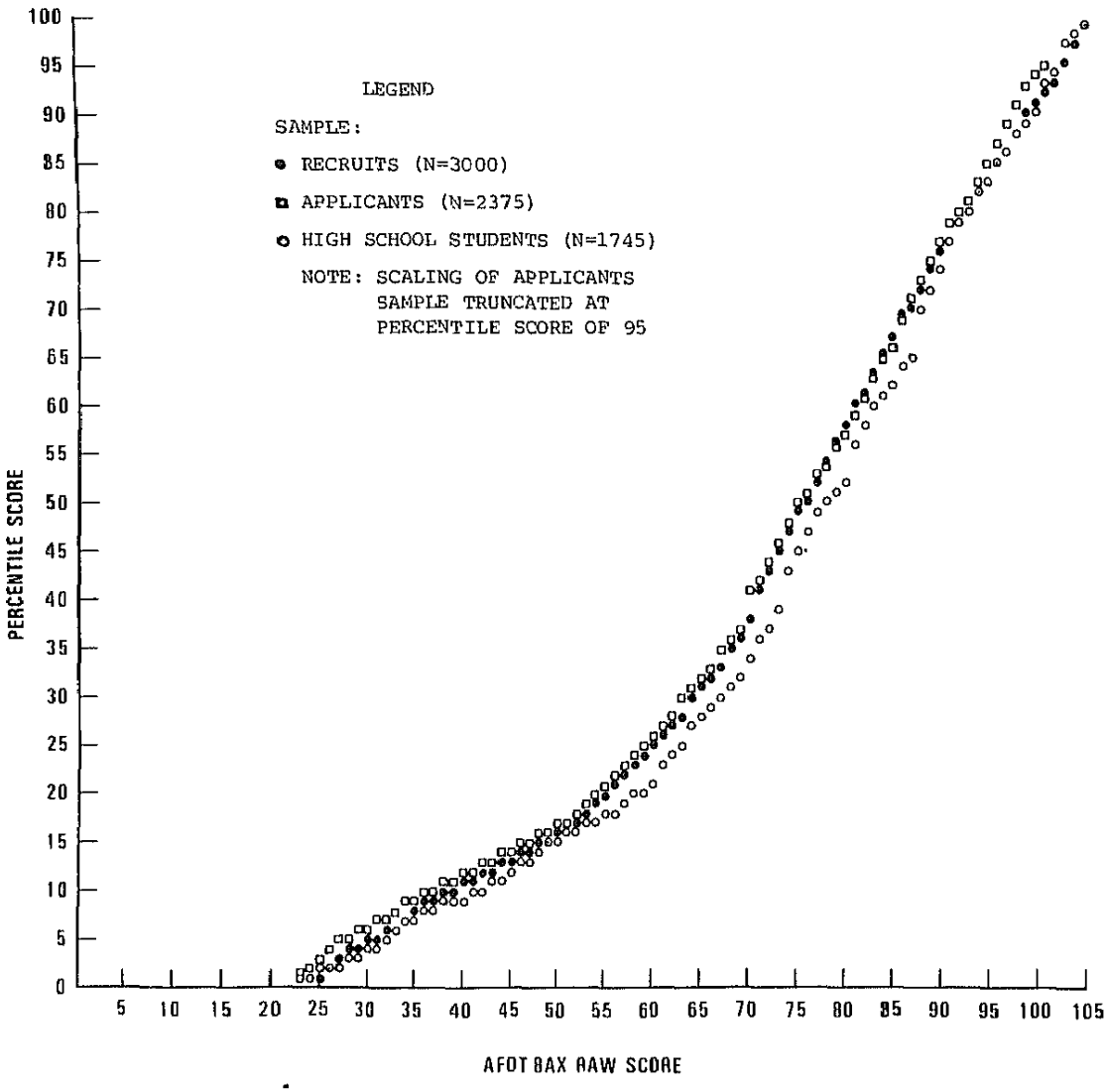
B. Calibration Results

The conversions from AFQT-8aX raw score to percentile score in the three studies are shown in Figure 1. The conversion lines were similar in the bottom end of the scale. There was a tendency for the high school sample to fall to the right of the two military samples. This means that a higher AFQT-8aX raw score is required in the high school sample to convert to a given percentile score. The high school sample starts deviating markedly at about the 20th percentile score, and then becomes similar to the military samples again at about the 75th percentile score. The applicant and recruit samples were similar throughout the scale.

In all three studies, the editing of the data had little effect on the score scale. Similarly, the weighting of the recruit sample to obtain the desired racial-ethnic mix had little effect on the scale. Furthermore, using recruits does not result in calibrations which differ significantly from those obtained from applicants. The only consistent difference was that conversions based on high school students result in somewhat lower scaled scores than those based on military samples. A reasonable explanation, advanced by the Defense Advisory Committee on Military Personnel Testing, is that high school students are more literate than school dropouts, but are relatively less superior on nonverbal tests. Since AFQT-8aX has a large literacy component, high school students scored higher on AFQT-8aX than on AFQT-7a whereas military samples, which

Figure 1

Calibration of ASVAB AFQT in Three Independent Samples



contained larger percentages of school dropouts, tended to score relatively higher on AFQT-7a.

Based on the similarity of the results for the applicants and recruits, the two samples were combined to construct the final Forms 8, 9, and 10 score scale. The cumulative frequency distributions of the AFQT-7a percentile scores and AFQT-8aX raw scores are shown in Figure 2. The combined sample of 5,375 cases contained more cases at both extremes than either one alone, and therefore should result in more reliable conversions in Categories I and V.

The final conversion adopted for operational use is shown in Figure 3. The conversion shows the following properties:

- o Differentiation between individuals with small differences in aptitude, and who are in Categories IV and V, is reflected in the test scores; one or two raw scores correspond to each percentile score,
- o Differentiation throughout the score range appears to be adequate, and
- o The progression in percentile scores is relatively smooth.

This conversion was applied to all six forms of the AFQT in Forms 8, 9, and 10. It is presented in tabular form in Appendix B.

C. Verification of the Calibration

When Forms 8, 9, and 10 became operational, their performance was closely monitored for the purposes of carrying out an Initial Operational Test and Evaluation (IOT&E). The data collected during that period were used to establish a calibration to verify the accuracy of the operational equating tables. The following material is quoted from the report of that IOT&E (Ree, Mathews, Mullins, & Massey, 1982):

> A sample of applicants for military enlistment was administered (one) form of ASVAB and the AFQT-7a in counterbalanced order. From this target sample of 22,400, a "males only" sample of 15,115 was developed through data editing techniques designed to exclude females and cases with incomplete or unusable data. For analytic purposes, this edited sample was separated into six samples based on the six forms of ASVAB administered. Data were collected at 20 geographically dispersed (MEPS) on the six forms of ASVAB and the AFQT-7a. Each of the six males only samples was edited and scored, and descriptive statistics were computed. Percentiles for both the ASVAB and the AFQT-7a were equated and smoothed by a polynomial regression procedure. Each sample was split in half, and the equating and smoothing were repeated on each half sample.

Figure 2

Cumulative Frequency Distribution of AFQT-7a and ASVAB Scores In Combined Samples of Recruits and Applicants

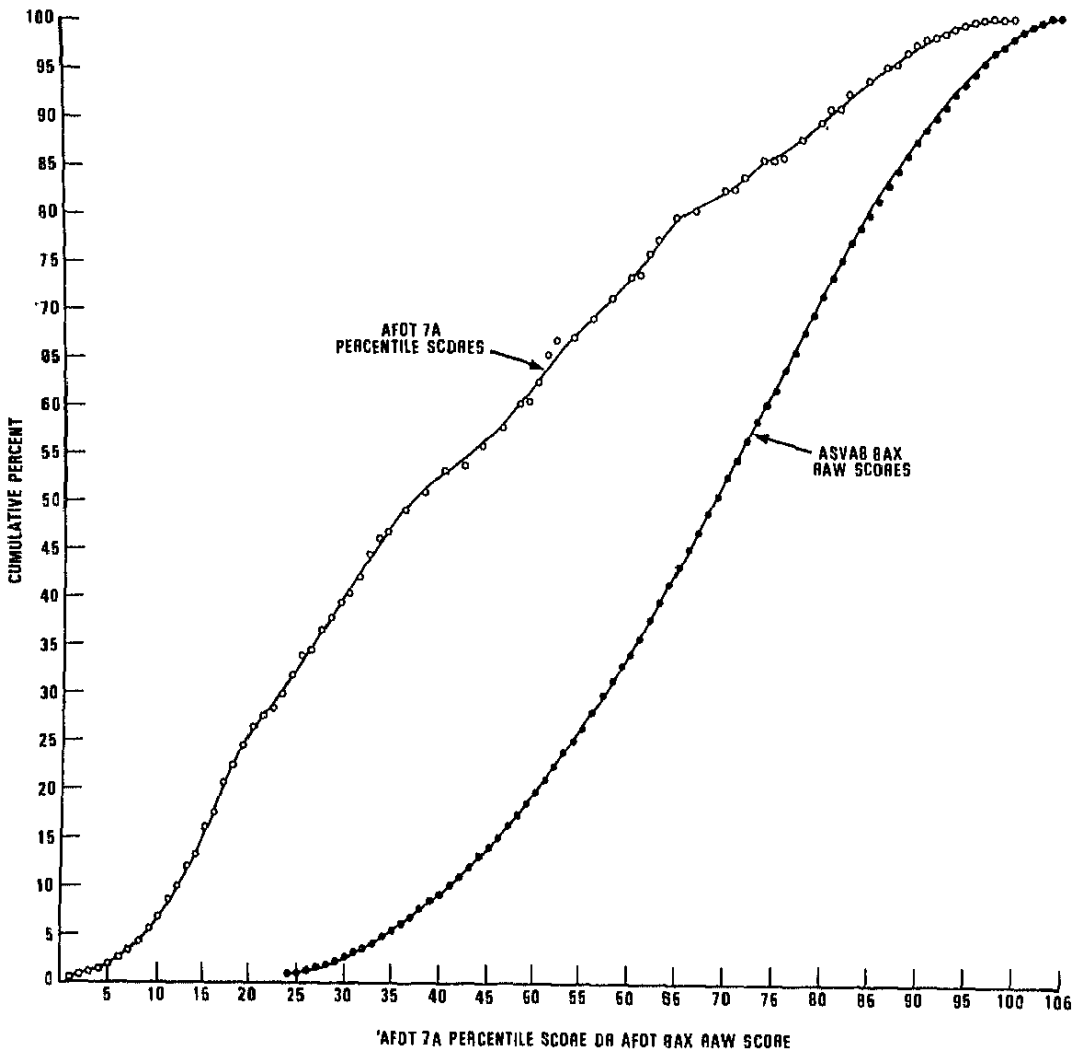
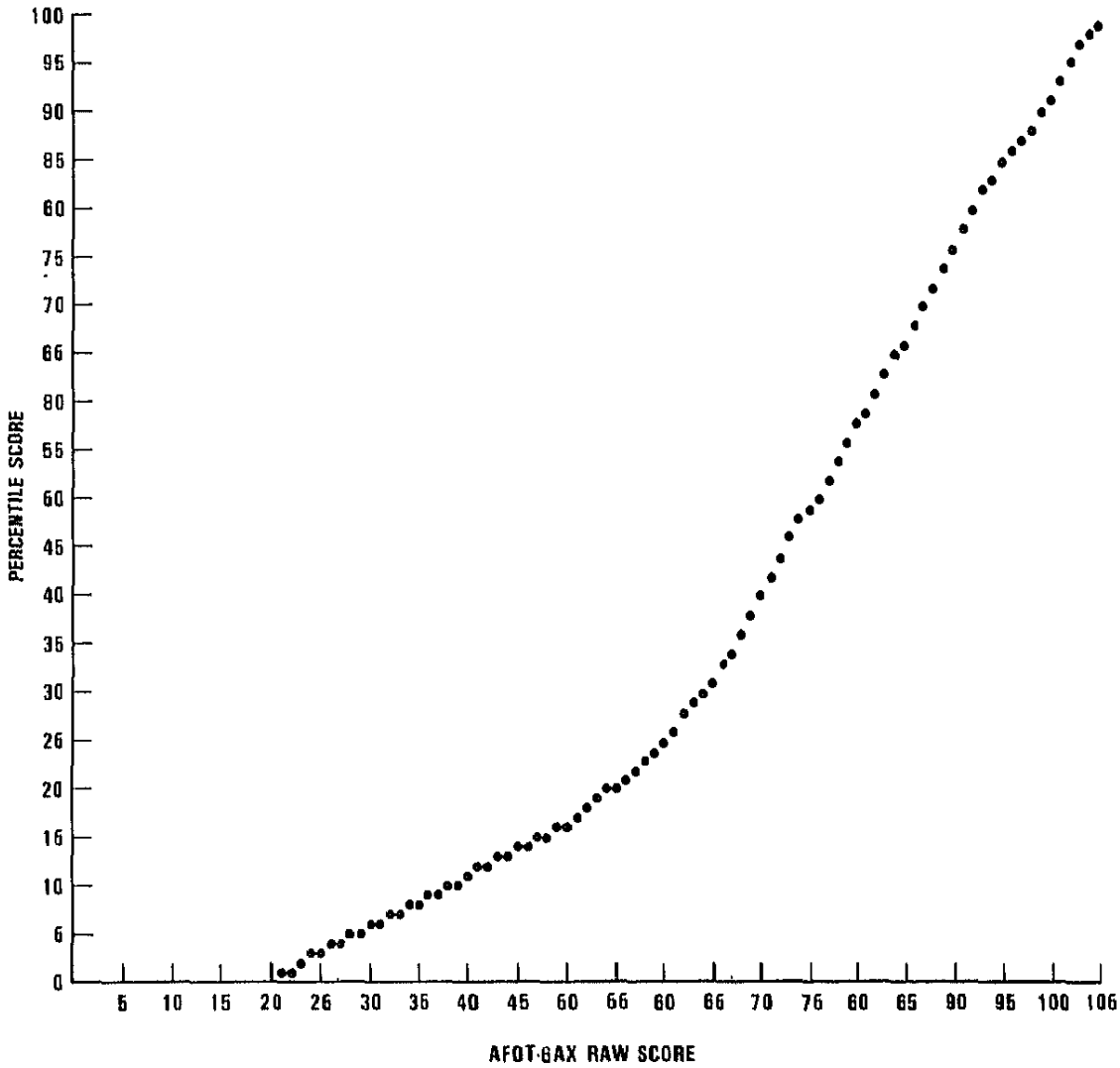


Figure 3

Final Calibration of ASVAB Based on Combined Samples of Recruits and Applicants



Since results were consistent among the large sample and the two half samples, they were accepted. In order to investigate the similiarity of the equated scores across the forms, root-mean-square (RMS) and average absolute deviation (AAD) measures were computed between the various equating tables. A comparison of the forms found them to be equivalent when they were equated to AFQT-7a. The RMS and AAD measures showed only small differences among the operational table and tables developed during this study. Forms 8, 9, and 10 of ASVAB were found to be parallel when equated to AFQT-7a, and a single conversion table was deemed appropriate for operational enlistment processing.

A presentation of the agreement between the six individual tables created in this verification study and the operational conversion table is shown in Table 3. This table presents the results obtained from the six tables together with the results obtained from the operational table in terms of assignment of cases to AFQT mental categories. The applicant group is low in Category I personnel, and somewhat restricted in Category II. The operational table is seen to place somewhat more cases in Categories I, II, and V, and slightly fewer cases in Categories III and IV. The differences, however, are not great. The use of one table rather than six would have the operational advantages of convenience and fewer opportunities for errors brought about by the use of an inappropriate table.

The conversion tables developed for each of the six forms of ASVAB, a conversion table prepared by averaging across those six forms, and the operational conversion table are presented in Appendix B. Further details of calibration analyses can be found in Maier (1981b); Ree, Mathews, Mullins and Massey (1982) and Sims and Truss (1980).

D. Subtest Standard Scores

A scoring feature introduced with Forms 8, 9, and 10 was the conversion of subtest raw scores to standard scores prior to computing aptitude composite scores. Aptitude composite scores, as discussed below, are formed by adding scores of certain subtests. In earlier versions of ASVAB, subtest raw scores (number of items correct) were summed and converted to aptitude composite scale scores (standard scores for the Army and the Marine Corps and percentile scores for the Air Force). Raw scores are computed for each subtest of ASVAB by counting the number of correct responses. After the raw scores have been derived, the raw score composite is computed for the AFQT. The raw score composite for AFQT is standardized into a percentile metric calibrated to the 1944 reference population.

Table 3

Classification by Mental Category Based on Operational versus Six Tables

Category by Six Tables	Category by Operational Table					Total	Percentage by Six Tables
	I	II	III	IV	V		
I	244	-	-	-	-	244	1.6
II	156	3045	-	-	-	3201	21.2
III	-	224	5199	121	-	5544	36.7
IV	-	-	-	5015	177	5192	34.3
V	-	-	-	-	934	934	6.2
Total	400	3269	5199	5136	1111	15115	
Percentage by Operational Table	2.6	21.6	34.4	34.0	7.4		

Note. From Calibration of Armed Services Vocational Aptitude Battery Forms 8, 9 and 10 (AFHRL-TR-81-49) by M. J. Ree, J. J. Mathews, C. J. Mullins and R. H. Massey, 1982, Brooks AFB, TX: Air Force Human Resources Laboratory. Adapted by permission.

Before the composites used for vocational classification purposes by the Services are computed, the subtest scores are standardized by application of the formula appearing below:

$$\text{ASVAB Subtest Standard Score} = 50 + \frac{10(X_i - \overline{X})}{SD_X}$$

where

X_i = the subject's raw score on subtest X,

$\overline{X}$ = the mean raw score of subtest X, in the reference population and

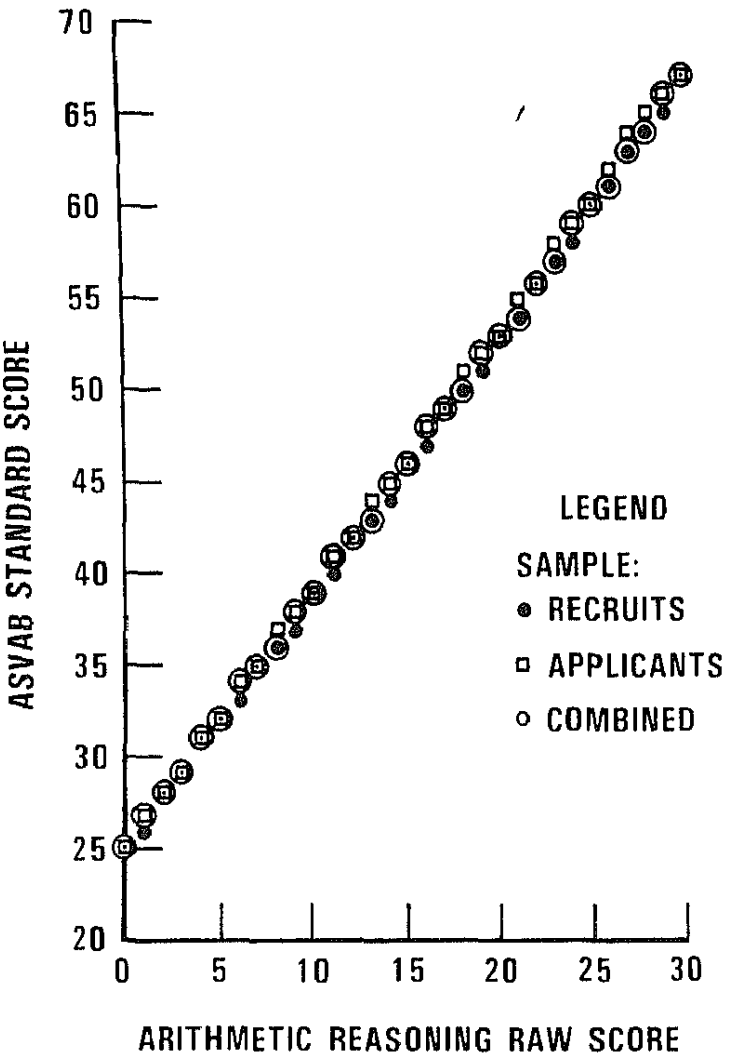
SD_X = the standard deviation of subtest X, in the tested population.

The conversion from Arithmetic Reasoning raw score to subtest standard score is shown in Figure 4 for the sample of recruits, for the sample of applicants, and for the combined sample. The three conversions are almost identical. For operational purposes, subtest standard scores are summed and then converted to standard scores for the Army or percentiles for the Air Force.

A separate conversion was computed for each subtest in the sample of recruits, the sample of applicants, and in the combined sample. As with Arithmetic Reasoning, the conversions in the three samples are similar.

Figure 4

Converting Arithmetic Reasoning (AR) Raw to Standard Scores



E. Aptitude Composite Scores

Each Service has developed its own set of aptitude composites to classify enlistees for job training programs. The Services also use composites to supplement the AFQT for determining qualification for enlistment (Atwater & Abrahams, 1980; Maier & Fuchs, 1972; Maier & Grafton, 1981; Sims & Hiatt, 1981; Thomas, 1970; Vitola & Alley, 1968).

Based on classification battery experience, the Services have differing configurations of selector composites. Within the Form 8, 9, and 10 programs, the Army has used 10 composites, the Navy has used 12 composites, the Marine Corps has used 6 composites, and the Air Force has used 4. These composites, by name and abbreviation, appear in Table 4. Eleven composites are listed for the Navy; the twelfth is VE, or the sum of the scores on Word Knowledge and Paragraph Comprehension. VE is considered a selector composite for the purposes of the Navy, but more generally it is used as a measure of verbal ability. The different services require various different scores on the composites to qualify applicants for entrance into particular occupations. The Army and the Marine Corps use a standard score conversion of the composite, the Air Force uses a percentile metric, and the Navy applies different raw score minimums for assignment to various ratings.

In addition to the AFQT score, the Services each use three identical composites, although they call them by different names. For example, the Army, Navy, and Marine Corps use a "General Technical" composite identical in composition to the Air Force "General" aptitude index. In other cases, identical composite names across Services include different subtests. The Army and the Marine Corps each use a "Mechanical Maintenance" composite, but the two composites differ in test composition. Table 5 shows the test composition of each Service composite.

When the composites, consisting of summed standard scores, had been computed, they were equated to the distribution of ability in a wartime mobilization population. The Navy does not standardize the composite score after it is computed.

The scores on the composites are the raw data that the Services use for determining eligibility for the different specialties. The Services from time to time establish or modify the scores on the composites which are required to qualify applicants for various career fields. Based on composite scores, on

Table 4

Titles and Abbreviations of Selector Composites
by Service for Forms 8, 9, and 10

Title	Abbreviation	Title	Abbreviation
Army		Navy	
Electronics	(EL)	General Technical	(GT)
Operators/Foods	(OF)	Mechanical	(MECH)
Surveillance/Communications	(SC)	Electronics	(ELEC)
Mechanical Maintenance	(MM)	Clerical	(CLER)
Clerical	(CL)	Aviation Structural Mechanic	(AM)
Skilled Technical	(ST)	Basic Electricity/Electronics	(BE/E)
Combat	(CO)	Boiler Technician/Engineman/	
Field Artillery	(FA)	Machinists Mate	(BT/EN/MM)
General Technical	(GT)	Machinery Repairman	(MR)
General Maintenance	(GM)	Submarine	(SUB)
		Communications Technician	(CT)
		Hospitalman	(HM)
Marine Corps		Air Force	
Combat	(CO)	Mechanical	(M)
Field Artillery	(FA)	Administrative	(A)
Clerical	(CL)	General	(G)
Electronics Repair	(EL)	Electronics	(E)
Mechanical Maintenance	(MM)		
General Technical	(GT)		

Table 5

Forms 8, 9, and 10 Test Composition of Selector Composites by Service

Army	Navy	Marine Corps	Air Force	ASVAB Tests
AFQT	AFQT	AFQT	AFQT	AR + .5NO + VE
GT	GT	GT	G	AR + VE
EL	ELEC	EL	E	GS + AR + MK + EI
CL	CLER	CL	A	NO + CS + VE
MM	—	—	—	NO + AS + MC + EI
—	MECH	—	—	AS + MC + VE
—	—	MM	—	AR + AS + MC + EI
—	—	—	M	GS + 2AS + MC
CO	—	—	—	AR + CS + AS + MC
—	—	CO	—	NO + AS + VE
FA	—	—	—	AR + CS + MK + MC
—	—	FA	—	AR + AS + VE
OF	—	—	—	NO + AS + MC + VE
—	AM	—	—	MC + VE
SC	—	—	—	NO + CS + AS + VE
—	BE/E	—	—	GS + AR + 2MK
ST	—	—	—	GS + MK + MC + VE
—	BT/EN/MM	—	—	AS + MK
GM	—	—	—	GS + AS + MK + EI
—	MR	—	—	AR + AS + MC
—	SUB	—	—	AR + MC + VE
—	CT	—	—	AR + NO + CS + VE
—	HM	—	—	GS + MK + VE

Service needs at the time of application, and on the applicant's preference, an applicant may be given a choice of career fields within three or four different occupational specialties. Such policies and procedures extend beyond issues relevant to technical aspects of the ASVAB and so are not discussed here in detail.

F. Norms

The population base to be used for normative studies has been an issue for a number of years. Although the 1944 reference population provided a completely representative sample of the male population of the United States which was eligible for service, passage of time led to a concern that the norms based upon that population might no longer represent the distribution of ability in current populations.

Specifically, there was concern that the chaining of test forms and the changes in test content had somewhat diminished the precision with which more recent versions of ASVAB could be related back to the 1944 reference population. That population was tested with instruments whose technical merit reflected the state of the science in the 1940s, whereas more recent versions of ASVAB have incorporated many advances in psychometric knowledge and technique.

To develop a new reference population against which ASVAB scores could be interpreted, DoD sponsored a study called the 1980 Profile of American Youth (OASD/MRA&L, 1982b). Another objective was to assess the vocational aptitudes of individuals, ages 16 to 23. This study was unique in that it was the first time that a vocational aptitude battery had been administered to a nationally representative sample. Specifically, the ASVAB was administered during 1980 to about 12,000 men and women, ages 16 to 23. The sample contained individuals both from urban and rural areas, from all major regions, and nearly equal proportions of males and females. To provide more precise subgroup analyses, certain small subgroups (e.g., blacks and Hispanics) were oversampled. An independent panel of sampling experts concluded that the sample design was appropriate, and all of the statistical procedures used in the development of sample case weights and sampling statistics were proper.

Form 8a was administered to the examinees. The test was evaluated by authorities on educational and psychological testing to determine its suitability

for measuring vocational aptitudes and its equity for minorities and females (Bock & Mislevy, 1981). They reported:

> Data from responses of the Profile of American Youth sample to the ASVAB are free from major defects such as high levels of guessing or carelessness, inappropriate levels of difficulty, cultural test-question bias, and inconsistencies in test administration procedures. They provide a sound basis for the estimation of population attributes such as means, medians and percentile points, for the youth population as a whole and for subpopulations defined by age, sex, and race/ethnicity.

The Profile of American Youth represents a major research effort which produced the 1980 reference population designed to establish new national norms for the ASVAB. Henceforth, it will be possible to refer future test calibrations to the abilities of the 1980 sample. Such norms will allow continued meaningful comparisons of the abilities of future potential or real enlisted military populations to the abilities of the 1980 sample.

Chapter 3
Reliability and Validity

A. Subtest Reliability

Aptitude scores used for selection and classification purposes must be reliable, i.e., be stable in their measurement and consistent in the manner in which they rank persons. Reliability coefficients have not been published for the composite scores used by the various Services, but, as a general rule, the reliability of a composite is equal to or greater than the average reliability of the tests which are included (Guilford, 1950, p. 524).

The power subtests included in Forms 8, 9, and 10 are uniformly reliable as determined by measures of internal consistency. The average of the Kuder-Richardson Formula 20 reliabilities is .86, and the range of subtest reliabilities is from .80 to .93. A summary of the reliabilities, as computed by Ree, Mullins, Mathews and Massey (1982) appears in Table 6.

Direct evidence of the reliability of the two speeded subtests in the battery is not available. It can be observed that the correlation between Numerical Operations and Coding Speed varies between .70 in the 1980 reference population (Table 7) and .53 in a sample of Navy recruits (Table 8). The generally lower intercorrelations in the Navy matrix may be ascribed to the restriction of range accompanying selection of the cases into the Navy. Intercorrelations in the range of .5 to .7 suggest test reliabilities no lower than .7, so that it may reasonably be concluded that the subtests are of satisfactory reliability. Further data for all subtests in Forms 8-13, concerning subtest intercorrelations, item distribution statistics, reliabilities, and item statistics are consistent with satisfactory reliabilities (see Appendix C).

Parallel (alternate) form reliabilities for high school composites were computed following the administration of Forms 8a and 14 to a sample 11th and 12th grade high school students, two-year college students, and others ages from 18 through 23 (DoD, 1984). These reliabilities, which range from .84 to .99, are shown in Table 9.

B. Composite Score Validity

1. Restriction of Range

The validation of an operational test is complicated by the fact that applicants falling below a cut-off score cannot appear in a validation sample.

Table 6

Kuder-Richardson Formula 20 Reliabilities for Forms 8, 9, and 10 Power Subtests

Subtest	ASVAB 8a	8b	9a	9b	10a	10b	Avg.
General Science (GS)[a]	.84	.85	.88	.87	.86	.86	.86
Arithmetic Reasoning (AR)	.90	.91	.91	.91	.90	.91	.91
Word Knowledge (WK)	.92	.92	.92	.92	.93	.92	.92
Paragraph Comprehension (PC)	.80	.80	.81	.80	.84	.80	.81
Auto-Shop Information (AS)[a]	.88	.88	.89	.81	.87	.88	.87
Mathematics Knowledge (MK)[a]	.87	.87	.87	.88	.86	.87	.87
Mechanical Comprehension (MC)[a]	.86	.86	.85	.84	.86	.85	.85
Electronics Information (EI)[a]	.83	.83	.82	.81	.81	.80	.81
Average	.86	.87	.87	.86	.87	.86	.86

Note. From Armed Services Vocational Aptitude Battery: Item and Factor Analysis of Forms 8, 9, and 10 (AFHRL-TR-81-55) by M. J. Ree, C. J. Mullins, J. J. Mathews and R. H. Massey, 1982, Brooks AFB, TX: Air Force Human Resources Laboratory. Adapted by permission. Number of cases ranges from 2420 to 2620.

[a] Identical items appear within a and b versions of each numbered form, differently ordered so scoring keys are not identical.

Table 7

Intercorrelations of Form 8 Subtests for
Males and Females in the 1980 Reference Population

	Subtest										Test Raw Score	
Subtest[a]	GS	AR	WK	PC	NO	CS	AS	MK	MC	EI	Mean	SD
GS	-	72	80	69	52	45	64	69	70	76	16.0	5.01
AR	72	-	71	67	63	51	53	83	69	66	18.0	7.37
WK	80	71	-	80	60	55	53	67	60	68	26.3	7.71
PC	69	67	80	-	60	56	42	64	52	57	11.0	3.36
NO	52	63	60	60	-	70	30	62	40	41	34.5	10.99
CS	45	51	55	56	70	-	22	52	34	34	46.3	16.25
AS	64	53	53	42	30	22	-	41	74	75	14.3	5.55
MK	69	83	67	64	62	52	41	-	60	59	13.6	6.39
MC	70	69	60	52	40	34	74	60	-	74	14.2	5.35
EI	76	66	68	57	41	34	75	59	74	-	11.6	4.24

Note. From Validity of ASVAB Forms 8, 9, and 10 for Marine Corps Training Courses: Subtests and Current Composites (Memorandum No. 83-3107) by M.H. Maier and A.R. Truss, 1983, Alexandria, VA: Center for Naval Analyses. Reprinted by permission. Decimals are omitted from intercorrelation values.

[a] ASVAB Subtests:

GS - General Science
AR - Arithmetic Reasoning
WK - Word Knowledge
PC - Paragraph Comprehension
NO - Numerical Operations
CS - Coding Speed
AS - Auto/Shop Information
MK - Mathematics Knowledge
MC - Mechanical Comprehension
EI - Electronics Information

Table 8

Means, Standard Deviations and Intercorrelations
Among Forms 8, 9, and 10 Subtests for a Full-Range Recruit Sample (N=66,459)
(Navy)

	Subtest											Test Standard Score	
Subtest[a]	GS	AR	WK	PC	NO	CS	AS	MK	MC	EI	VE	Mean	SD
GS	100	50	68	53	07	11	51	50	56	60	69	51.91	7.90
AR	50	100	46	46	32	27	37	70	52	43	50	53.62	7.86
WK	68	46	100	61	08	16	38	45	43	49	96	52.54	6.73
PC	53	46	61	100	14	20	32	43	39	39	80	53.31	6.49
NO	07	32	08	14	100	53	-04	35	05	01	11	53.28	7.31
CS	11	27	16	20	53	100	01	31	10	06	19	53.09	7.90
AS	51	37	38	32	-04	01	100	29	63	64	40	51.25	8.67
MK	50	70	45	43	35	31	29	100	49	41	48	52.54	8.80
MC	56	52	43	39	05	10	63	49	100	61	45	51.24	8.27
EI	60	43	49	39	01	06	64	41	61	100	50	51.45	8.05
VE	69	50	96	80	11	19	40	48	45	50	100	52.88	6.37

Note. From Predictive Validation of Armed Services Vocational Aptitude Battery Forms 8, 9, and 10 Against Performance at 47 Navy Schools (Draft Report) by S. Booth-Kewley, 1983, San Diego, CA: Navy Personnel Research and Development Center. Reprinted by permission. Decimals are omitted from intercorrelation values.

[a] ASVAB Subtests:
GS - General Science
AR - Arithmetic Reasoning
WK - Word Knowledge
PC - Paragraph Comprehension
NO - Numerical Operations
CS - Coding Speed
AS - Auto/Shop Information
MK - Mathematics Knowledge
MC - Mechanical Comprehension
EI - Electronics Information
VE - Verbal Test (WK+PC)

Table 9

Parallel Forms Reliabilities for High School Composites for Forms 8a and 14

	18-23 Year Olds	11th Grade		12th Grade		Two-year College	
		Male	Female	Male	Female	Male	Female
Composite							
Academic							
Academic Ability (WK+PC+AR)	.94	.94	.92	.93	.93	.88	.88
Verbal (GS+WK+PC)	.94	.94	.93	.93	.93	.89	.89
Math (AR+MK)	.94	.93	.91	.93	.91	.92	.90
Occupational							
Mechanical & Crafts (AR+AS+MC+EI)	.93	.92	.84	.92	.86	.91	.88
Business & Clerical (WK+PC+MK+CS)	.94	.94	.93	.93	.92	.90	.90
Electronics & Electrical (GS+AR+MK+EI)	.94	.94	.91	.93	.92	.92	.90
Health, Social & Technology (WK+PC+AR+MC)	.95	.95	.92	.94	.92	.92	.90

Note. From Technical supplement to the counselor's manual for ASVAB-14 by the Department of Defense, 1984, North Chicago, IL: U.S. Military Entrance Processing Command. Reprinted by permission.

The validity coefficient of interest is that which would be found in an unselected sample. The accompanying restriction in the range of abilities in the selected sample results in lowered correlation coefficients when those coefficients are computed with either the selector scores or with other scores which are correlated with the selection scores. Formulae to adjust obtained values to the magnitude to be expected in an unselected sample have been given by Thorndike (1949, pp. 173-174). These formulae are fully appropriate if only those cases below a given cut-off point are missing from a sample taken from a normal population. They provide for both direct (selector score) restriction and indirect (a correlated measure) restriction.

The assumptions required for the use of the Thorndike formulae are not entirely met by the ASVAB data. For example, the restriction within a given validation sample may come both from the non-acceptance of individuals falling below a certain score and from the absence of persons falling above a higher score because of their prior selection for a specialty with a higher minimum aptitude cut-off. Consequently, truncation occurs on both ends of the score range for certain specialties.

Assumptions about the normality of the population from which the samples were selected may be considered with respect to the standard scores used by the Army and the Marine Corps, and with respect to the raw score composite used by the Navy. Both of these scoring systems maintain the original shape of the distribution of test scores. The Air Force data indicate acceptable validity (as discussed later in this chapter), but the Air Force's use of a percentile metric results in a flat, rather than normal, distribution, and so the use of formulae derived for normally distributed data may underestimate the validities. In most occupational specialties, the effect of restriction from selection is to provide lower correlations for the selection measure than for other available selection indices. Correction for restriction of range to provide more sensitive information as to the relative performance of various measures is indicated.

Correction for restriction of range also permits meaningful comparisons of validity coefficients between groups which differ in the amount of restriction to which their measures are subject.

2. Criterion Identification

The preferred criterion for validation of a selection measure for civilian occupations has been identified as job performance ("Uniform Guidelines," 1978).

However, there are no uniformly available, common measures of such performance across the Military Services. As a result, training school performance is commonly used for validating selector composites.

The Services establish the content of training courses based upon objective occupational analyses for each of their specialties. These analyses, as part of the Instructional Systems Development process followed by the Services, help ensure that the content of technical training courses reflects the content of jobs in the field. Therefore, to the extent that objective measures of performance in training are available, training grades are useful criteria for evaluating the performance of selection measures.

The appropriateness of training-school scores as a criterion is further established by two considerations. First, attrition from training schools represents ineffective manpower utilization. Individuals who are not trained cannot do the job. Prediction of training success is therefore valuable. Second, variables associated with individual assignments introduce extraneous variance into job performance. Such extraneous variance does not correlate with aptitude test scores and so obscures true validity relationships.

The Services each use a computer-based task inventory system for objectively monitoring the content of their occupational specialties. Tailoring of course content to the observed requirements of each specialty ensures that training is in content areas relevant to the work to be done in the field (McCormick, 1979; Morsh & Archer, 1967; Pass, 1980; Yellen & Foley, 1978).

Modern training technology has affected the usefulness of training grades. Many courses are no longer graded along a numerical continuum, but are graded as simply pass or fail. Often students who have difficulty retake difficult phases until they achieve a passing grade. Some courses are self-paced, and the measure of performance is the time required to complete the course.

3. Validation Sample Collection and Report Organization

Each of the Services has accomplished preliminary validation of Forms 8, 9, and 10 against performance in technical schools.

These forms of ASVAB were implemented in October, 1980. Recruits, tested at MEPSs, were sent to training units for basic training lasting for several weeks. At the end of basic training, and in some instances after a leave, they reported to the technical training facilities.

Technical training classes enter weekly, bi-weekly, or monthly. Each class may contain as few as eight or as many as several hundred recruits. Course

lengths vary from a few weeks to many months.

From the foregoing, it is apparent that accumulation of validation samples of sufficient size for statistical stability across the spectrum of occupations found in each Service is a time-consuming project of significant complexity.

Variations in the ability of those who enter a given course may occur as extended time is required to build a sample of adequate size for statistical stability. Course content may vary during the time required. Conversely, in some large courses, full samples were developed quickly. Such samples, representative of the students entering during only a few weeks, may differ from samples developed over a longer period of time (Lecznar, 1962).

Each validation study has included data screening activities to ensure that the cases treated in the sample are reasonably homogeneous in terms of data availability and meaning. The Navy has included validation analyses involving time required for completion as a criterion measure. For those ratings, the Navy has reported the validity correlations as negative, reflecting the association of higher selection scores with shorter completion times.

The Services have performed extensive analyses to determine individual ASVAB test validities, the validity of operational composites, and the identification of potentially more powerful "new" composites.

In this manual, summaries are presented of the validity of the AFQT composite score used by the Services, and of the current selector composites used by each Service for various military occupations.

Material presented is grouped within the categories of the DoD Occupational Grouping System. This system categorizes enlisted specialties in the four Services into nine occupational areas, each of which is subdivided into highly related groups of occupations, within which homogeneous subgroups are identified. It is a three-digit system; the left-most digit identifies the area, the center digit identifies the group, and the right-most digit identifies the subgroup (OASD/MRA&L, 1982a). Validation data are reported in DoD areas in which two or more Services presented specialties.

Although a given occupation in one Service may differ from an occupation of similar content in another Service, the DoD Occupational Grouping System does provide a framework for general comparisons and for various kinds of manpower studies.

The Army included 11 specialties in its report of the validation of Forms 8, 9, and 10 (Rossmeissl, Martin & Wing, 1983), the Navy included 47 ratings (i.e.,

specialties) (Booth-Kewley, 1983), the Marine Corps included more than 50 specialties (Maier & Truss, 1983), and the Air Force included about 70 specialties (Wilbourn, Valentine & Ree, in press).

C. Notes about the Data

Each of the Services completed validation analyses of Forms 8, 9, and 10 in the first half of 1983. Data from preliminary reports were made available and data from them are cited in this chapter.

The Army and Navy provided both restricted correlational values and correlations corrected for restriction of range. The Marine Corps furnished only the corrected data, and the Air Force only the uncorrected data. The Army and the Marine Corps corrected the restriction on the basis of correlational and distributional data from the 1980 reference population, as presented in Table 7. The Navy based their corrections upon a sample of Navy recruits, with correlational and distributional data as presented in Table 8.

In the validation tables which follow, all Army data were provided by Rossmeissl et al. (1983), Navy data by Booth-Kewley (1983), Marine Corps data by Maier and Truss (1983), and Air Force data by Wilbourn et al. (in press).

Occupational specialties are identified in the following tables by the Service Occupational Code (SOC) reported by the authors of the validation reports. The Army's Military Occupational Specialty (MOS) consists of five characters, the first three of which are usually sufficient to distinguish one specialty from another. The three characters are two numbers and one letter, collectively identifying the specialty without regard to skill level. Navy ratings are identified by a two or three letter designation. Like the MOS, the rating designation indicates a general field of expertise. The Marine Corps uses a four-digit MOS. The first two digits designate an occupational field, the third identifies the promotional channel, and the fourth identifies the specialty within the occupational field. The Air Force uses a five-digit Air Force Specialty Code (AFSC). The first two digits indicate a career field, while the third and fifth digits indicate further specialization within that field. The fourth digit indicates skill level.

In some cases terminology in the DoD Occupational Conversion Manual (OASD/MRA&L, 1982a) differs from that appearing in the Service report. In those cases the Service terminology has been used.

As indicated in Chapter 2, the Services use a variety of selector composites based upon different combinations of tests taken from the ASVAB. The names of the selector composites used by each Service appear in Table 4. The subtests composing each selector composite are indicated in Table 5.

D. DoD Area 0 Validation: Infantry, Gun Crews, and Seamanship Specialists

The Army, Marine Corps, and Navy provided a widely varied set of specialties for this occupational area. The corrected validities of the AFQT composite ranged from a low of .30 for an Army specialty to a high of .69 for a Navy rating.

The selector composite validities were slightly less variable, ranging from .36 for the Army specialty to a high of .65 for the Navy rating. These values, typical for enlisted selection indices, appear in Table 10.

By title, the selector composites applied are: Army, Operator/Foods and Mechanical Maintenance; Navy, General Technical; and Marine Corps, Field Artillery. Apparent differences are minimized when the tests included in each composite are reviewed. The composites include quantitative measures, clerical speed measures, mechanical aptitude, and verbal measures.

In a report on the validation of Forms 5, 6, and 7, Swanson (1979) reported the selector composite for the Quartermaster (QM) rating to have a corrected validity of .73 as compared to .54 in the current sample.

E. DoD Area 1 Validation: Electronic Equipment Repairmen

This is one of the larger occupational areas reported, with 20 specialties reported by the Services. The Air Force reported more specialties in this area than any other Service.

Within the group of Radio/Radar Repairmen reported in Table 11, the corrected AFQT validities reported by the Army, Navy, and Marine Corps are moderate to strong, ranging from .32 to .84. The negative correlations for the Navy AX rating represent the negative relationship between training times and selector scores.

The selector composite validities reported by the Services tend to be uniform, although the comparisons must take into account the Air Force use of uncorrected validity coefficients. The Air Force values tend to fall within the range of the uncorrected selector composite validities reported by the Army and the Navy. The corrected validity coefficients range from .31 to .87.

Table 10

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Area 0: Infantry, Gun Crews, and Seamanship Specialists

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	Selector Composite (Abbrev.)
04 Artillery, Gunnery, Rockets and Missiles					
Army Specialty and DoD Subgroup					
16P	101	Short Range Missile Crewman (043)	.15/.30	.21/.36	(OF)
16S	514	Man Portable Air Defense Crewman (043)	.17/.40	.23/.44	(OF)
Marine Corps Specialty & DoD Subgroup					
0844	208	Field Artillery Fire Control Crewman (041)	—/.66	—/.63	(FA)
06 Seamanship					
Army Specialty and DoD Subgroup					
61B	92	Watercraft Operator (062)	.49/.69	.45/.65	(MM)
Navy Rating and DoD Subgroup					
QM	473	Quartermaster (061)	.47/.53	.47/.54	(GT)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 11

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Group 10: Electronic Equipment Repairmen—Radio/Radar

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	Composite (Abbrev.)
101 Communications Radio					
Army Specialty					
32D	120	Station Technical Controller	.44/.67	.43/.67	(EL)
Air Force Specialty					
30430	219	Wideband Communications Equipment Specialist	—/—	.55/—	(E)
30434	366	Ground Radio Communications Specialist	—/—	.49/—	(E)
30730	180	Telecommunications Systems Control Specialist	—/—	.37/—	(E)
32830	351	Avionics Communications Specialist	—/—	.56/—	(E)
102 Navigation, Communication, and Countermeasure, etc.					
Army Specialty					
33S	103	Electronic Warfare Intercept Systems Repairer	.46/.84	.56/.87	(ST)
Navy Rating					
AX	288	Aviation Antisubmarine Warfare Technician	-.34/-.49	-.28/-.45	(ELEC)
Air Force Specialty					
32232	244	Avionics Sensor System Specialist	—/—	.49/—	(E)
32530	245	Automatic Flight Control System Specialist	—/—	.41/—	(E)
32831	297	Avionic Navigation System Specialist	—/—	.45/—	(E)
32833	244	Electronic Warfare System Specialist	—/—	.53/—	(E)
32834	218	Avionic Inertial and Radar Navigation System Specialist	—/—	.44/—	(E)
104 Surveillance/Target Acquisition and Tracking Radar					
Marine Corps Specialty					
7222	107	Hawk Missile System Operator	—/.32	—/.31	(GT)
Air Force Specialty					
30333	113	Automatic Tracking Radar Specialist	—/—	.38/—	(E)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.

[a] Service Occupational Code

Comparisons within three subgroups in the electronic repair area are given in Tables 12, 13, and 14. For both AFQT and the selector composites, the validities fall in the intermediate range of values usually found in enlisted selection measure validation studies. An exception is the Navy Fire Control Technician (FTM) in which validities of .71 for AFQT and .80 for the selector composite (ELEC) are found.

In reports dealing with the validation of Forms 5, 6, and 7 Swanson (1979) and Valentine (1977) cite data comparable to the validities reported for Forms 8, 9, and 10. Swanson cites corrected validities of .82 for the Aviation Antisubmarine Warfare Technician (AX) and .77 for the Aviation Fire Control Technician (AQ). Valentine reports uncorrected validities of .44 for Communications-Electronics Systems (AFSC 30X3X) and .33 for Avionics Systems (AFSC 32X3X).

Swanson (1979) cites a corrected validity of .81 for the selector composite for the Fire Control Technician (FTM), and of .67 for the Data Systems Technician (DS).

Direct comparison of data from validation studies separated in time is complicated by possible changes in input populations, criterion composition, and differences in the base for correction for restriction in range. Nevertheless, the data suggest that Forms 8, 9, and 10 are comparable in predictive efficiency to the preceeding forms.

F. DoD Area 2 Validation: Communications and Intelligence Specialists

Validity of selector composites used for the selection of Radio Operators seems relatively consistent across the Services (Table 15). The Air Force restricted validity coefficient of .16 for the Administrative aptitude index is atypically low, but the amount of restriction of range involved is unknown. The Navy uses a variety of selectors, each of which shows validity at intermediate levels in the corrected coefficients.

The selector composite validities reported for the Signal Intelligence and Electronic Warfare specialists are greater than those reported for Radio Operators (Table 16). The Army reports the highest validity with a corrected coefficient of .81 for AFQT and .79 for their Surveillance/Communications composite. The Air Force uncorrected Administrative aptitude index falls well within the range of the Navy's uncorrected values.

Table 12

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Subgroup 112: Airborne Fire Control

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
Navy Rating					
AQ	475	Aviation Fire Control Technician	-.33/-.47	-.26/-.43	(ELEC)
Air Force Specialty					
32132	288	Weapon Control Systems Mechanic	—/—	.49/—	(E)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 13

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Subgroup 121: Missile Guidance and Control

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
Marine Corps Specialty					
7212	112	Redeye Gunner	—/.44	—/.62	(FA)
Navy Rating					
FTM	172	Fire Control Technician	.40/.71	.52/.80	(ELEC)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 14

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Subgroup 150: ADP Computers, General

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
Navy Rating					
DS	118	Data Systems Technician	.26/.52	.32/.57	(ELEC)
Air Force Specialty					
30534	237	Electronic Computer & Switching Systems Specialist	—/—	.45/—	(E)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 15

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Group 20: Communications and Intelligence Specialists-Radio and Radio Code

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
Marine Corps Specialty & DoD Subgroup					
2531	903	Field Radio Operator (202)	—/.43	—/.47	(EL)
Navy Rating and DoD Subgroup					
SM	377	Signalman (203)	.39/.54	.32/.50	(GT)
RM	302	Radioman (201)	-.34/-.54	—/—	
			—/—	-.27/-.52	(ELEC)
			—/—	-.35/-.49	(CLER)
			—/—	-.19/-.47	(SUB)
			—/—	-.09/-.36	(VE)
Air Force Specialty and DoD Subgroup					
29333	132	General Radio Operator (201)	—/—	.16/—	(A)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 16

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Group 23: Communications and Intelligence Specialists—Signal Intelligence & Electronic Warfare

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
Army Specialty & DoD Subgroup					
05G	91	Signal/Security Specialists (231)	.55/.81	.48/.79	(SC)
Navy Rating and DoD Subgroup					
EW	400	Electronic Warfare Technician (230)	-.30/-.45	-.20/-.39	(ELEC)
EW	408	Electronic Warfare Technician (230)	-.26/-.43	-.21/-.41	(ELEC)
CTR	140	Cryptologic Technician (231)	.51/.59	.50/.59	(GT)
CTT	63	Cryptologic Technician (231)	.39/.50	.44/.53	(GT)
CTT	259	Cryptologic Technician (231)	.56/.65	.60/.68	(GT)
Air Force Specialty & DoD Subgroup					
20731	138	Morse System Operator (231)	—/—	.33/—	(A)
20230	135	Radio Communications Analysis Specialist (232)	—/—	.45/—	(G)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

In the general area of Communications Center Operations, the Marine Corps and the Air Force report median validities for enlisted specialties as shown in Table 17.

With reference to selection indices derived from Forms 5, 6, and 7, Sims and Hiatt (1981) report corrected validities of .49 for Marine Corps Field Radio Operator (2531) and .51 for Communication Center Operations (2542). Swanson (1979) reports .51 for Signalman (SM) and .17 for Radioman (RM). Valentine (1977) found an uncorrected correlation of .25 with Radio Operator training (29130) for the Air Force Administrative aptitude index.

G. DoD Area 4 Validation: Technical Specialists, etc.

The Marine Corps and the Air Force reported validities in the occupational subgroup of Firefighting and Damage Control as reported in Table 18. The Marine Corps AFQT and selector composite validities are typical for Marine

specialties. The validity of selection for the Air Force Protection Specialist occupation is supported by the restricted correlation reported by the Air Force which exceeds the corrected values reported by the Marine Corps, and falls relatively high in the ranking of uncorrected validation correlation coefficients.

Table 17

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Subgroup 260: Communications Center Operations, General

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
Marine Corps Specialty & DoD Subgroup					
2542	334	Communications Center Operator	—/.49	—/.49	(CL)
Air Force Specialty					
29130	348	Telecommunications Operations Specialist	—/—	.32/—	(G)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 18

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Subgroup 495: Firefighting and Damage Control

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
Marine Corps Specialty					
7051	158	Aircraft Firefighting and Rescue Specialist	—/.29	—/.42	(MM)
Air Force Specialty					
57130	817	Fire Protection Specialist	—/—	.44/—	(G)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

H. DoD Area 5 Validation: Functional Support and Administration

Validities in the Administration career subgroup are in intermediate ranges with AFQT corrected validation correlations of .37 to .65. Selector composites ranged from .15 to .38 for uncorrected correlations and from .37 to .64 for corrected correlations. Forms 5, 6, and 7 showed validity in the same ranges with the Marine Corps reporting corrected values of .51 for AFQT and .53 for the CL Composite (Sims & Hiatt, 1981). Swanson (1979) reports a corrected correlation of -.25 for the Navy Cryptologic Technician (CTA) using completion time as a criterion, and Valentine (1977) found uncorrected values of .32 for AFQT and .20 for the Air Force Administrative (A) composite. Table 19 includes the validation data for Forms 8, 9, and 10.

The performance of Data Processing Operators is less well predicted than that of administrative personnel, according to the data in Table 20. The Navy found identical figures for AFQT and their General Technical (GT) composite—uncorrected correlations of .23, correcting to .39. The Air Force found an uncorrected validity correlation for their General (G) aptitude index of .43. For Forms 5, 6, and 7, Swanson (1979) reports an uncorrected correlation of .48 with a corrected value of .77 for the Data Processing Technical (DP), while Valentine (1977) reports uncorrected values of .32 for AFQT and .26 for the General (G) aptitude index.

In the field of Supply Administration, the Army and the Marine Corps report consistently high validities for Forms 8, 9, and 10 as shown in Table 21. Corrected validities for AFQT range from .59 to .75; values for the corrected selector composite (Clerical) range from .60 to .73. These values compare favorably with Marine Corps validation data for Forms 5, 6, and 7 reported by Sims and Hiatt (1981). They found corrected selector composite values of .46 for Basic Stock Clerk (MOS 3043) and .51 for Aviation Supply Clerk (MOS 3072).

I. DoD Area 6 Validation: Electrical/Mechanical Equipment Repairmen

As might be expected in modern military organizations, this occupational area included more specialties than any other for which validation data were reported. Prediction levels were uniformly high, with values in the group associated with aircraft repair ranging from .50 to .83 for corrected selector composite validities and .47 to .76 for the AFQT (Table 22). Prediction was similarly high for Forms 5, 6, and 7, with Swanson (1979) reporting corrected

Table 19

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Group 51: Administration

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	Selector Composite (Abbrev.)
Army Specialty & DoD Subgroup					
71D	96	Legal Clerk (512)	.38/.65	.27/.64	(CL)
Air Force Specialty & DoD Subgroup					
90630	240	Medical Administrative Specialist (513)	—/—	.38/—	(G)
510 Administration, General					
Navy Rating					
CTA	107	Cryptologic Technician	.25/.27	.23/.37	(CLER)
Marine Corps Specialty					
0151	640	Administrative Clerk	—/.58	—/.59	(CL)
0151	640	Administrative Clerk	—/.47	—/.47	(CL)
Air Force Specialty					
70230	1841	Administration Specialist	—/—	.15/—	(A)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 20

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Subgroup 531: Data Processing Operators, Analysts

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	Selector Composite (Abbrev.)
Navy Rating					
DP	373	Data Processing Technician	.23/.39	.23/.39	(GT)
Air Force Specialty					
51130	192	Computer Operator	—/—	.43/—	(G)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 21

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Subgroup 551: Supply Administration

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
Army Specialty					
76P	613	Material Control and Accounting Specialist	.40/.68	.26/.60	(CL)
Marine Corps Specialty					
3043	665	Basic Supply Stock Clerk	—/.75	—/.73	(CL)
3072	381	Aviation Supply Clerk	—/.59	—/.60	(CL)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Aviation Boatswains Mate validities of .76 (ABE), .91 (ABF), and .85 (ABH). Valentine (1977) reports uncorrected validities for three Aircraft Mechanics specialties of .45 (AFSC 43130), .34 (43131), and .40 (43132).

In the field of Automotive Repair, the Marine Corps and the Air Force selector composites demonstrate relatively higher validity than do those used by the Navy (Table 23). Specialties in the Armament and Munitions group are better predicted than specialties in Wire Communications, with Army, Navy, and Marine Corps all reporting corrected validity in the low .70s for the Armament and Munitions area. Forms 8, 9, and 10 showed higher validities in the Armament and Munitions area than Forms 5, 6, and 7, which yielded validities of .47 for Aviation Ordnance (Marine Corps MOS 05XX) and .51 for Ammunition Technician (Marine Corps MOS 2311) (Sims & Hiatt, 1981).

For Forms 8, 9, and 10 the Air Force reported uncorrected validities for the Mechanical (M) aptitude index for the Special Vehicle Mechanic (47231) and the General Purpose Vehicle Mechanic (47232) of .52 and .47, respectively. For Forms 5, 8, and 7 the corresponding values were .39 and .29 (Valentine, 1977).

Table 22

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities within DoD Occupational Group 60: Electrical/Mechanical Equipment Repairman— Aircraft and Aircraft Related

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
600 Aircraft, General					
Army Specialty					
67Y	137	Attack Helicopter Repairer	.29/.66	.39/.75	(MM)
Marine Corps Specialty					
6011	521	Aviation Mechanic	—/.56	—/.63	(MM)
Air Force Specialty					
43130	155	Helicopter Mechanic	—/--	.46/—	(M)
43131	2179	Tactical Aircraft Maintenance Specialist	--/—	.47/—	(M)
43132	2216	Airlift/Bombardment Aircraft Maintenance Specialist	--/—	.49/—	(M)
601 Aircraft Engines					
Navy Rating					
AD	880	Aviation Machinists Mate	-.32/-.47	-.35/-.50	(ELEC)
Air Force Specialty					
42632	1238	Jet Engine Mechanic	—/—	.46/—	(M)
42633	165	Turboprop Propulsion Mechanic	—/—	.43/—	(M)
602 Aircraft Accessories					
Marine Corps Specialty					
6077	105	Aviation Maintenance Ground Support Equipment Electrician	—/.76	—/.83	(MM)
Air Force Specialty					
42331	361	Aircraft Environmental Systems Mechanic	—/—	.33/—	(M)
42333	431	Aircraft Fuel Systems Mechanic	—/—	.41/—	(M)
42330	561	Aircraft Electrical System Specialist	—/—	.55/—	(E)
604 Aircraft Launch Equipment					
Navy Rating					
ABE	72	Aviation Boatswains Mate	.32/.51	.41/.56	(GT)
ABF	96	Aviation Boatswains Mate	.38/.50	.38/.50	(GT)
ABH	69	Aviation Boatswains Mate	.39/.52	.42/.54	(GT)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.

[a] Service Occupational Code

Table 23

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities within DoD Occupational Area 6: Electrical/Mechanical Equipment Repairmen

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	Selector Composite (Abbrev.)
61 Automotive					
Navy Rating and DoD Subgroup					
CM	79	Construction Mechanic (612)	.10/.25	.23/.37	(MECH)
EO	181	Equipment Operator (612)	.22/.34	.22/.36	(MECH)
Marine Corps Specialty & DoD Subgroup					
1341	169	Engineer Equipment Mechanic (612)	—/.57	—/.70	(MECH)
3521	459	Organizational Automotive Mechanic (610)	—/.50	—/.72	(MM)
2145	144	Tracked Vehicle Repairer, Tank (611)	—/.50	—/.54	(MM)
6072	130	Aviation Maintenance Ground Support Equipment Mechanic (Hydraulics)(610)	—/.59	—/.74	(MM)
Air Force Specialty and DoD Subgroup					
47231	134	Special Vehicle Mechanic (610)	—/—	.52/—	(M)
47232	135	General Purpose Vehicle Mechanic (610)	—/—	.47/—	(M)
62 Wire Communications					
Navy Rating and DoD Subgroup					
IC	658	Interior Communications Electrician (623)	-.37/-.50	-.32/-.47	(BE/E)
Air Force Specialty & DoD Subgroup					
36130	127	Cable & Antenna Systems Installation/Maintenance Specialist (621)	—/—	.37/—	(M)
64 Armament and Munitions Army Specialty and DoD Subgroup					
68J	128	Attack Fire Control Repairer (646)	.28/.62	.44/.73	(EL)
Navy Rating & DoD Subgroup					
GMT	99	Gunner's Mate Technician (644)	.48/.66	.48/.71	(MECH)
Marine Corps Specialty & DoD Subgroup					
65XX	381	Basic Aviation Ordnance (646)	—/.68	—/.73	(GT)
2311	164	Ammunition Technician (645)	—/.82	—/.70	(GT)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.

[a] Service Occupational Code

As in previous tables, the negative correlations reported for some Navy ratings reflect the use of time to complete training as a criterion measure. Those who score higher on the selector composites tend to complete training in less time.

The last block of specialties in DoD Occupational Area 6—Electrical and Mechanical Equipment Repairmen-Shipboard Propulsion is shown in Table 24. These specialties, from the Army and the Navy, are characterized by relatively high predictability, even though several selector scores are used. Except for the Boiler Technician (Navy, BT) and the Engineman (Navy, EN), prediction falls between corrected values of .63 and .75, with the median coefficients above .70.

Table 24

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities in DoD Occupational Group 65: Electrical/Mechanical Equipment Repairmen—Shipboard Propulsion

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	Selector Composite (Abbrev.)
Army Specialty & DoD Subgroup					
61C	150	Watercraft Engineer (652)	.45/.73	.45/.75	(OF)
Navy Rating & DoD Subgroup					
BT	2085	Boiler Technician (651)	-.38/-.43	-.32/-.39	(BT/EN/MM)
EN	1258	Engineman (651)	-.28/-.38	-.28/-.39	(BT/EN/MM)
GSE/	117	Gas Turbine Fundamental	—/—	—/—	(BT/EN/MM)
GSM		Electrical (652)	.25/.63	.47/.74	(BT/EN/MM)
GSM	84	Gas Turbine Technician (652)	.23/.53	.35/.62	(BT/EN/MM)
GSM/	117	Gas Turbine Fundamental			
GSE		Electrical (652)	—/—	.46/.75	(ELEC)
GSM	84	Gas Turbine Technician (652)	—/—	.36/.63	(ELEC)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

J. DoD Area 7 Validation: Craftsmen

This occupational area includes very different specialties, ranging from the Air Force Pavement Maintenance Specialists to the Navy Steelworker and the Marine Corps Engineer Equipment Operator, as shown in Table 25. The observed validities are moderate, with AFQT corrected values ranging from .17 to .58

Table 25

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities within DoD Occupational Area 7: Craftsmen

SOC[a]	N	Specialty Title	AFQT Validities	Selector Validities	Composite (Abbrev.)
70 Metalworking					
Navy Specialty & DoD Subgroup					
MR	194	Machinery Repairman (702)	.16/.41	.48/.67	(MR)
Air Force Specialty & DoD Subgroup					
42735	550	Airframe Repair Specialist (700)	—/—	.27/—	(M)
42731	322	Corrosion Control Specialist (701)	—/—	.16/—	(M)
71 Construction					
Navy Rating & DoD Subgroup					
BU	203	Builder (710)	.32/.58	.43/.67	(MECH)
SW	85	Steelworker (711)	.05/.17	.20/.31	(MECH)
Marine Corps Specialty & DoD Subgroup					
1345	452	Engineer Equipment Operator (713)	—/.49	—/.57	(MM)
Air Force Specialty & DoD Subgroup					
55130	151	Pavement Maintenance Specialist (710)	—/—	.36/—	(M)
55230	100	Carpentry Specialist (710)	—/—	.29/—	(M)
55232	115	Material Fabrication Specialist (710)	—/—	.47/—	(M)
720 Utilities, General					
Navy Rating					
UT	77	Utilitiesman	.30/.35	.15/.23	(MECH)
Air Force Specialty					
56631	172	Environmental Support Specialist	—/—	.41/—	(M)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.

[a] Service Occupational Code

with a median value of .41. Corrected selector composite values range from .23 to .67 with a median of .57. Uncorrected values submitted by the Air Force fall in the same range as the Navy's uncorrected correlations.

Comparisons with the validity of Forms 5, 6, and 7 are possible for Airframe Repair (AFSC 42X3X), for which Valentine (1977) reports an uncorrected validity for the selector composite of .40; for Pavement Maintenance Specialist (AFSC 55X3X) an uncorrected validity of .36, and for Environmental Support Specialist (AFSC 56330) an uncorrected validity of .45. These data suggest that Forms 8, 9, and 10 are of the same order of validity as the prior forms.

K. DoD Area 8 Validation: Service and Supply Handlers

Composite scores used for the selection of cooks or food service personnel, as shown in Table 26, are effective. Corrected values for AFQT validities are shown as .56 and .62 for the Navy and Marine Corps, respectively. Corrected selector composite values are .57 and .65, with the Air Force reporting an uncorrected selector composite validity of .38, which is slightly below the Navy value. Among Material Handlers, prediction is slightly less accurate with the Navy reporting corrected values for AFQT and the selector composite of .33 and .32, respectively. In two Supply specialties the Air Force reported uncorrected selector composite validities of .35 and .37.

For Forms 5, 6, and 7, Sims and Hiatt (1981) report validities for the Marine Corps' Basic Food Service of .43 for AFQT and .43 for their General Technical composite. These corrected values compare with uncorrected values reported by Valentine (1977) for the Air Force of .34 for Supply (AFSC 64530) and .37 for Medical Material Specialist (AFSC 91X3X) within the General aptitude index.

L. Validation within Black/White and Male/Female Samples

The Army and the Air Force have reported validation data for black and white samples and between sex groups for Forms 8, 9, and 10. Within the time period for data collection, more samples accumulated permitting black/white comparisons among male samples than were available for comparisons between sex groups. Proscriptions against females in combat specialties eliminated some specialties from consideration in terms of sex variables.

Table 26

Cross-Service Comparison of AFQT/ASVAB 8, 9, and 10 Selector Composite Validities within DoD Occupational Area 8: Service and Supply Handlers

SOC[a]	N	Specialty Title	AFQT Validities	Selector Composite Validities	(Abbrev.)
800 Food Service, General					
Navy Specialty					
MS	1581	Mess Management Specialist	.45/.56	.47/.57	(GT)
Marine Corps Specialty					
3371	504	Cook Specialist	—/.62	—/.65	(GT)
Air Force Specialty					
62230	488	Food Service Specialist	—/—	.38/—	(G)
82 Material Receipt, Storage, and Issue					
Navy Rating and DoD Subgroup					
SH	595	Ships Serviceman (823)	.20/.33	.19/.32	(GT)
Air Force Specialty & DoD Subgroup					
64531	538	Material Facilities Specialist (822)	—/—	.35/—	(G)
91530	105	Medical Material Specialist (822)	—/—	.37/—	(G)

Note. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.
[a] Service Occupational Code

Table 27 presents data concerning the relative validity in black and white male samples of the AFQT and the Operator/Foods (O/F) composite for the Army specialty of Man Portable Air Defense Systems. Rossmeissl et al. (1983) report a corrected validity of .47 for blacks and .68 for whites for the AFQT score. The corresponding selector composite validity was shown as .53 for blacks and .51 for whites. An additional 4.3 percent of the sample was neither black nor white. In the total sample, the corrected validity was reported as .40 for the AFQT and .44 for the O/F Composite.

In Table 28 data are presented comparing black and white males and females within an Army clerical specialty. It is noted that the blacks and whites, together, comprise about 93 per cent of the total sample. Of the total sample, 63 per cent is black. Validity coefficients within the white samples are slightly higher than in the black samples, but both the AFQT and the selector composite show satisfactory levels of prediction.

Table 27

Validity of AFQT and the Army Operator/Foods (O/F) Composite for MOS 16S Man Portable Air Defense System Crewman Black and White Males

Race	N	AFQT[a]	O/F
Black	159	.03/.47	.16/.53
White	333	.21/.68	.28/.51
Total	514	.17/.40	.23/.44

Note. From Validity of ASVAB 8, 9, and 10 as Predictors of Training Success (Selection and Classification Working Paper 83-3) by P. G. Rossmeissl, C. J. Martin and H. Wing, 1983, Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences. Reprinted by permission. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.

[a] Service Occupational Code

Table 28

Validity of AFQT and the Army Clerical Composite for Material Control and Accounting Specialist (MOS 76P) by Race and Sex

Race and Sex	N	AFQT[a]	CL
Black Males	273	.28/.69	.12/.57
Black Females	116	.26/.62	-.02/.46
White Males	143	.60/.73	.47/.65
White Females	38	.51/.77	.41/.69
Total	613	.40/.68	.26/.60

Note. From Validity of ASVAB 8, 9, and 10 as Predictors of Training Success (Selection and Classification Working Paper 83-3) by P. G. Rossmeissl, C. J. Martin and H. Wing, 1983, Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences. Reprinted by permission. Validities to the left of the slash are uncorrected; to the right, they are corrected for restriction of range.

[a] Service Occupational Code

Similar data for six Air Force specialties are presented in Table 29. Prediction is least effective among Administrative Specialist (AFSC 70230) and more effective for Law Enforcement Specialist (AFSC 81132) and Aircraft Electrical Systems Specialist (AFSC 42330). There are no major differences between black and white or between male and female predictions for those specialties in which adequate samples appear, and in which validities reach useful levels. Note that the values in Table 29 have not been corrected for restriction in range.

Inasmuch as blacks were well represented in the total sample, considerations of adverse impact are minimized. There is nothing in the data to suggest that the tests are discriminatory with respect to minority members. The conclusion that there is no test bias against minority members is consistent with results from earlier studies (Bock & Moore, 1984; Boldt et al., 1977; Guinn, Tupes & Alley, 1970a, 1970b; Shore & Marion, 1972).

M. Validation for Form 14

The academic composites (see Table 9) for Form 14 measure potential for academic training. Those composites were validated as predictors of grades in civilian academic and vocational courses. The average validity for high school and two-year college courses was about .4 (DoD, 1984). In another study of 1000 high school students (Streicher & Friedman, 1983), the academic ability composite correlated highly with similar tests (e.g., .90 with the California Achievement Test and .85 with the Differential Aptitude Test).

The occupational composites were validated on more than 50 military occupational training courses, with the corrected validity coefficients averaging about .6 (Maier & Truss, 1984).

N. Summary

Forms 8, 9, 10, and 14 are found to be of satisfactory reliability with reference both to the individual subtests composing the battery and to the composite scores developed from those tests.

The Army, Navy, Marine Corps, and Air Force have completed initial validation of Forms 8, 9, and 10. In this manual, specialties from the four Services have been grouped as specified by the DoD Occupational Conversion Manual (OASD/MRA&L, 1982a). This manual includes data from 11 Army

Table 29

Validity of the Air Force Selector Aptitude Index for Training in Selected Specialties by Race and Sex

	Black		White		Total	
	N	r[a]	N	r	N	r
AFSC 70230 Administration Specialist						
Administrative						
Male	467	.01	754	.26	1280	.20
Female	163	.09	381	.06	561	.02
Total	630	.01	1138	.19	1841	.15
AFSC 62230 Food Service Specialist						
General						
Male	86	.37	202	.37	307	.37
Female	31	.44	107	.33	141	.40
Total	117	.38	309	.35	488	.38
AFSC 42330 Arcrft Electrical Sys Spec						
Electronics						
Male	66	.51	403	.52	488	.53
Female[b]	—	—	51	.49	73	.38
Total	85	.49	454	.55	561	.55

	Black		White		Total	
	N	r	N	r	N	r
AFSC 73230 Personnel Specialist						
Administrative						
Male	115	.19	351	.35	485	.33
Female	59	.07	132	.32	194	.25
Total	174	.17	483	.34	679	.31
AFSC 81132 Law Enforcement Specialist						
General						
Male	357	.38	1234	.47	1617	.49
Female	43	.23	188	.48	238	.48
Total	400	.37	1422	.48	1855	.49
AFSC 42331 Arcrft Environmental Sys Spec						
Mechanical						
Male	77	.17	218	.34	308	.31
Female[b]	--	—	49	.13	53	.20
Total	81	.21	267	.37	361	.33

Note. From Aptitude Index Validation of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 8, 9, and 10 (AFHRL-TP-84-08) by J. M. Wilbourn, L. D. Valentine, Jr. and M. J. Ree, in press, Brooks AFB, TX: Air Force Human Resources Laboratory. Reprinted by permission.

[a] All correlations are uncorrected for restriction of range.
[b] Groups of less than 25 were not considered.

occupations, 25 Marine Corps specialties, 30 Navy ratings, and 40 Air Force specialties.

Although there are examples of marginal predictive efficiency for some specialties and dramatically high prediction in others, on the average, validities computed for Forms 8, 9, and 10 are equivalent to validities computed for earlier forms.

Criteria employed in all the validation studies are training performance measures, either in terms of training grades or of time spent to achieve a given standard of performance. It is recognized that such criteria do not equate to job performance, but it is also noted that all Service training course content is now controlled by objective task analyses of work as done in the field. In this context, and pending development of objective, common job performance measures across all Services, the school performance measure is the best available criterion for assessing the value of the selection measures.

Within the limits of available data, the Services have reported the comparative validation of their selection measures for blacks, whites, and for males and females. Although blacks and women in some cases show lower validation correlations than do whites or males, there is nothing in the data to suggest that these validities are insufficient or that adverse impact results from use of the tests.

The studies summarized in this chapter together make a convincing case for the widely applicable use of the ASVAB selector composites as valid predictors of training success. There is, however, no single statement or number which can sum up the implications of the coefficients. In order to estimate the consequences of the validity coefficients, this paragraph presents some simplifying assumptions, and then refers to a well known method for assessing the expected effects of various levels of validity. Taylor and Russell (1939) developed a set of tables which collectively express the relationships between four quantities. The quantities are the validity of a test, the proportions of examinees who would be successful if all examinees were accepted into training (or if the selection were made at random among the examinees), the proportion who would be successful if the highest scoring examinees were selected, and the proportion of examinees who are selected. The proportion of applicants who would be successful if applicants were assigned at random to technical training schools is not known. The first simplifying assumption is thus that the proportion

is equal to 0.6. Any technical training schools will also be assumed to accept students who place in the top 30% of the examinees on the composite. Given those assumptions, it is possible to report the proportion of entrants who would successfully complete technical training as a function of test validity.

If the validity of the test were .3, then 73% of those accepted would be expected to complete successfully. If the validity were .5, then 82%, and if the validity were .7, then 91% would be expected to complete training successfully. The range of validities of .3 to .7 is representative of operational values.

A number of factors make these percentages higher than they might be operationally, most notably the fact that only the most demanding of the schools will be able to restrict its students to those scoring among the top ten percent. Nevertheless, the figures show that even modest selector composite validities of .3 and .4 allow a marked increase in the proportion of students who would successfully complete training, given the assumed base proportion of .6.

Thus the validities reported across all job families by all Services are sufficiently strong to provide effective predictors of training success, and thus to reduce training failure rates, decrease training time, and promote advantageous employment of enlisted personnel.

Chapter 4

Administration, Materials Control, and Service Implementation

A. Testing Personnel

Personnel technicians of the various Services who have been trained in the proper administration, proctoring, and scoring of psychological tests are assigned to the ASVAB program. These persons have been given extensive training on the ethics of testing, personal privacy, and the proper methods of test administration. Test administrators usually serve as test proctors for an extended period before assignment to test administration duties. During that period of experience they become sensitive to signs of examinee distress or confusion and are familiar with techniques for handling problems in the testing room without creating turmoil. They learn that examinee questions reflecting confusion about how to respond to a given test are to be answered by reiteration of appropriate sections of the administrative directions and not by ad lib response.

Test security and the confidentiality of test results are emphasized both in the training for the administration and in the management of the testing program.

B. Manual for Administration, Armed Services Vocational Aptitude Battery

The following material is quoted from DoD 1304.12A (DoD, 1983, pp. 1-5) titled as Section B, above:

Section 1

PREPARATION FOR TESTING

1. Introduction.

This manual prescribes the procedures and instructions for administration of the Armed Services Vocational Aptitude Battery (ASVAB).

2. Testing Conditions and Standards.

Directives and regulations of each of the services and Office of Personnel Management (OPM) describe acceptable testing standards and conditions. Test administrators are responsible for being familiar with the testing standards of their service/agency and assuring compliance with the

standards established by their service/agency. Such standards are established for the benefit of both the examinee and the test administrator; inadequate testing facilities complicate proper exercise of good test control practices. Sound judgment must be exercised in accepting testing facilities to ensure both that there is equitable opportunity for examinees and that proper testing and test control practices are possible. This is of vital concern to all the services.

The value and accuracy of test scores can be affected by the procedures and conditions of test administration. Individuals tested under poor conditions may feel that their test performance has been adversely affected to the extent that their scores cannot be compared equitably with the scores of those tested under favorable circumstances. Public acceptance of testing is a function of confidence that tests do provide a true picture of the potential, knowledge, and abilities of examinees. For this reason tests should be administered under standard conditions following procedures which give all persons the opportunity to do their best.

The procedures for administering tests should be those which elicit the best performance of which the person is capable. Particular attention should be given to ensuring that the examinees:

a. are reasonably free from distracting influences in the surrounding environment,
b. consider the test worthwhile,
c. are not distressed by substantial physical discomfort including fatigue.

While ideal testing conditions cannot always be achieved with the limited facilities available in field locations, close attention to the following features will provide conditions that are adequate:

d. The testing room must be reasonably quiet. Frequent shouting outside the windows, bells, trucks unloading, and other such noises may interfere with the test performance of the examinees. Tests will not be given to an examinee in a location where ordinary business is being conducted. The distraction of conversation, machinery, and other noises is detrimental to prolonged concentration on the part of the examinee.

e. Testing instructions must be clearly audible; the examiner's voice should be heard clearly by all persons being tested. If loudspeakers are used, care should be exercised in placing the loudspeakers and in locating the microphones. The level of amplification should be carefully controlled.

f. Lighting must be adequate. The testing room should be well lighted and the working surfaces should have sufficient uniform light. Deep shadows and strong glare on the working surface caused by poor arrangement of light fixtures should be avoided. The lighting should be adequate for comfortable reading without eye strain.

g. Ventilation, temperature, and humidity sometimes are difficult to control, but all practicable steps should be taken to provide for the examinee's comfort. Testing should not be conducted when temperatures and/or humidity conditions are so extreme as to interfere significantly with concentration.

h. The testing room should be arranged so that the test examiner can be seen by everyone while reading from the test administration manual. The desks or tables should be arranged to leave aisles for the proctors to use in distributing and collecting test materials, and in circulating about the room during the test. If possible, there also should be enough space between rows to allow passage. Examinees should be seated far enough apart to prevent an examinee from taking information from another's answer sheet. An overall space of 15 square feet per examinee is proper. This includes the space for the control aisle and aisles for proctoring.

i. Large tables may be used for testing but partitions of adequate height should be used to separate each examinee (to eliminate the possibility of one examinee looking on another's answer sheet).

j. The working surface should be flat, smooth, and free from cracks. The space allotted to each person should be large enough to accommodate an open test booklet and a separate answer sheet without overlapping.

3. Test Examinee and Test Examiner Deportment.

While examiners must demand discipline of all examinees, the examinees are also due reasonable and courteous treatment. Mental state should be such that the examinee considers it worthwhile to perform optimally and is capable of doing so.

To ensure that the examinee is in a good physical state, tests will be scheduled when the examinee is not fatigued or ill. Testing should not be scheduled after extended or strenuous periods of hard labor or at the end of a day's work. In all instances, persons in charge of testing should be alert to signs of genuine distress and the affected persons should be excused until a more appropriate time.

The test examiner should be selected for unquestionable integrity, maturity, ability to maintain test security, quality of speaking voice, and ability to handle groups of examinees effectively and in a friendly manner. Generally, a test examiner should be selected who does not have a marked regional, foreign, or other accent which may be difficult for some examinees to understand. The test examiner will generally be placed in charge of the group testing room.

The test examiner should be continuously alert and vigilantly maintain test security at all times. The [test administrator] should always be alert for signs of applicant cheating such as use of crib sheets, unauthorized testing aids, etc. Every effort should be made to discourage the use of these aids. Applicants caught using unauthorized aids will be dealt with IAW Chapter 3 MEPCOM Regulations 611-1 or appropriate OPM regulations.

The examiner should make a careful study of this manual and the directive prescribing the use of the test. The examiner should be completely familiar with the purpose of the test, the materials needed to administer it, the directions to be read, and any problems that are likely to arise. The examiner should rehearse the directions which are to be read aloud until they can be read slowly and distinctly without stumbling over words or losing the place.

Familiarity with test content itself is also valuable. Before giving the test, the examiner should make sure that enough test booklets, answer sheets, special pencils, and scratch pads are available. Scoring keys should not be brought into the testing rooms during testing, with the exception of the hand scoring keys needed by MEPCOM activities to compute the unverified raw AFQT scores at MET sites. The hand scoring keys will be maintained and safeguarded by the test administrator to preclude any access by unauthorized personnel.

4. Order of Test Administration and Time Required.

Table 1 specifies the order in which the Armed Services Vocational Aptitude Battery tests are to be administered, as well as the time limits for each test. This is the same as their order in the test booklet. The importance of adhering to the time limits cannot be over emphasized. The tests are separately timed to assure equal opportunity for all subjects on all tests. Moreover, score norms are based on these standard times.

5. List of Testing Materials (Omitted)

6. Preparation Prior to First Test Session.

It is important that test examiners and proctors become familiar with the test prior to administering it. Generally, the administration will be smoother and the proctoring more effective if both the examiner and proctors are familiar with the directions and items. It is recommended that they study this manual and familiarize themselves with the entire test and associated materials prior to their first administration of the battery. It has also been found that a "trial" test session prior to first administration of a new battery helps in preparation of examiners and proctors. It is suggested that for such a session the examiner administer the battery to the proctors; this provides practice for the examiner and helps familiarize proctors with content and structure of the battery.

C. Secure Handling of Test Data

All test material is treated as sensitive and confidential and is not released to unauthorized persons. Test score data are transmitted to the centralized recruiting facility for each Service where potential assignments are determined. Communication between the assignment facility, the recruiter, and the applicant results in determination of the specific occupational specialty for which the applicant is to been listed. Return of this information to the assignment facility

Table 30[1]

Order of Administration and Time Limits for Armed Services Vocational Aptitude Battery Tests

Order of Administration	Test	Time Limits (in minutes)
1	General Science	11
2	Arithmetic Reasoning	36
3	Word Knowledge	11
4	Paragraph Comprehension	13
5	Numerical Operations	3
6	Coding Speed	7
7	Auto & Shop Information	11
8	Mathematics Knowledge	24
9	Mechanical Comprehension	19
10	Electronics Information	9
Total Time		144

initiates the preparation of the personnel records jacket covering the applicant's enlistment.

Testing booklets, scoring stencils, completed answer sheets, and testing data are kept under lock and key when not in use. When being used they are protected from inspection by unauthorized personnel.

The specific procedures followed will be modified as the process becomes increasingly computerized.

D. Implementation of the ASVAB in the Services

The formal authorization for use of scores derived from the ASVAB lies in military regulations issued by each of the Services. These regulations specify relationships between ASVAB composite scores common to all the Services (AFQT) and specific to each Service and the qualification for entry into the Service and into specific occupational fields. The composite scores used by

[1] In the original document from which this section is excerpted, this was Table 1. It is here renumbered in order to conform to the numbering of tables in this manual.

each service have been described in Chapter 2, and typical validation data have been presented in Chapter 3. Regulations pertinent to each Service are identified in Table 31. They are not listed among the references because they are under continuous review and modification, without change in title or identifying regulation number.

Table 31

Identification by Service of Enlisted Classification Regulations

Service	Regulation
Army	AR 611-201, Personnel Selection and Classification; Enlisted Career Management Fields and Military Occupational Specialties
Navy	NAVPERS 18068D, Manual of Navy Enlisted Manpower and Personnel Classifications and Occupational Standards, Section I and II
Marine Corps	MCO P1200.7D, Military Occupational Specialties Manual (MOS Manual)
Air Force	AFR 39-1 Enlisted Personnel, Airman Classification Regulation

Chapter 5
Compliance with APA Standards for Test Development

A. Introduction

Fairness and ethical conduct in testing has long been an issue of concern to test developers, both civilian and military. Fairness issues tend to center around the effects of testing and the use of results for minority groups. Ethical concerns have addressed the potential harm that could result to examinees or to the community as a whole as a result of improper development, documentation, or use of tests. Since the 1960s, such issues have received particularly close attention from observers outside the testing community. Cases brought to court regarding civilian tests involved the presentation of evidence of the abuse of tests and their misuse in the selection of persons for employment. The professional community, acting through the American Psychological Association (APA), reacted by developing standards for the preparation and use of tests intended to ensure that neither deliberate nor inadvertent misuse of tests would occur (APA, 1974, 1980).

The development of standards for the guidance of test developers and users was paralleled by federal legislation dealing with employee selection procedures ("Uniform Guidelines," 1978).

Forms 8, 9, and 10 were under development during the period 1975-1980. This chapter examines the extent to which the Standards for Educational and Psychological Tests (APA, 1974), hereinafter referred to as "Standards," were met in the development of these forms. The standards relating to validity and reliability are fully addressed by the continuing professional review and evaluation of technical publications of the Service's laboratories. Therefore, they are not cited in this manual.

Each relevant standard listed in the APA publication will be cited, and comment will be offered as to its relevance to the ASVAB and the extent to which compliance can be documented.

The Standards were written to apply to commercially or academically developed tests measuring academic achievement, assessing personality or vocational interest, and evaluating aptitude for employment. Review of the Standards indicates that they are concerned with ensuring professional approaches to test development, standardization, and use, and inhibiting the exploitation of improperly developed instruments.

The Standards are offered in four substantive content areas: Section A, "Dissemination of Information;" Section B, "Aids to Interpretation;" Section C, "Directions for Administration and Scoring;" and Section D, "Norms and Scales."

"Dissemination of Information" deals with the test developer's responsibility to provide full information on the strengths and weaknesses of his instrument and his responsibility to include factual, objective data in publications which are accessible to potential users and examinees.

"Aids to Interpretation" must be complete and fully understandable to potential users and examinees. Technical psychometric terms and relationships must be rendered clearly, and, where appropriate, using charts and graphs which convey objective test data in terms of practical significance.

"Directions for Administration" deals with control of the testing situation to ensure that operational testing is done under the same conditions as the developmental testing during which the standardization data were collected.

The material on "Norms and Scales" is intended to ensure that derived data will be meaningful in practical terms and that the publisher will provide data equally comprehensible to examinees and to professional personnel.

There are major differences between a commercial vocational counseling or aptitude battery and the ASVAB as used by the Military Services. The commercial test is used by organizations independent of the test developer, while ASVAB is administered and applied by agencies under the same management structure as the laboratories which develop the battery. The content and format of the commercial test are controlled by competition in the marketplace and critical review in the professional literature. The content and format of ASVAB are controlled by policy boards of senior executives with the Department of Defense and the Military Services, by the Defense Advisory Committee on Military Personnel Testing, and by the Joint-Service Selection and Classification Working Group.

The APA Standards do not address the concept of a vocational testing battery unique to a single large organization, under continuing review and development by established personnel research organizations, with sequential test batteries evolving under the influences of research findings and changing administrative requirements.

In the context of a single test, the Standards call for information and data to appear in a manual which serves both to advise a potential user of the

characteristics of the test and to provide materials for administering, scoring, reporting test results, and interpreting the results to an examinee. No single document can completely meet these requirements for ASVAB, although this manual assembles material responsive to them. Each Service uses ASVAB results for the selection and classification of enlisted personnel, but no two Services use the data in exactly the same way. Each Service has its own set of rules for the application of test results (see Chapters 2 and 4).

B. Compliance with APA Standards

In this section each APA standard will be introduced by quotation of the standard. (The following lettering, A through D6.1, is that used in the Standards.)

A. Dissemination of Information

A1

When a test is published or otherwise made available for operational use, it should be accompanied by a manual (or other published or readily available information) that makes every reasonable effort to follow the recommendations of these standards and, in particular, to provide the information required to substantiate any claims that have been made for its use. (Essential)

No single publication exists presenting all available information relevant to the ASVAB as called for by this APA Standard. This manual summarizes information responsive to the requirement and provides references permitting a reader to go to any desired level of detail in any topic.

The research and development program supporting the ASVAB is conducted by the personnel research laboratories in each of the Services. Technical reports based on that research are disseminated from each laboratory to all the other laboratories and to personnel policy offices in each Service headquarters and the DoD.

The research and development programs are coordinated through Service personnel policy staff agencies as developed by the Joint-Service Selection and Classification Working Group (composed of testing professionals from the Services) and reviewed by a policy board of senior executives and by the Defense Advisory Committee on Military Personnel Testing.

A1.1

If information needed to support interpretations suggested in the manual cannot be presented at the time the manual is published, the manual should satisfy the intent of standard A1 by pointing out the absence and importance of this information. (Essential)

Technical reports describing ASVAB characteristics as determined through each Service's research and development programs are reviewed through the supervisory channels of the publishing laboratory. The reports must meet professional criteria for the scientific quality of research design, adequacy of controls, appropriateness of statistical procedures, and completeness of reporting of results.

Data which imply limited use or inadequate coverage of newly discovered problems are highlighted. For example, the standardization of Forms 8, 9, and 10 (Boldt, 1980c; Maier, 1981a; Ree, Mathews, Mullins & Massey, 1982; Sims & Truss, 1980) was accomplished on a males-only sample. This was done because the reference normative base, the 1944 reference population, contained only males. That fact was reported even though a single normative table was offered for males and females. In the interim, a new normative base, the 1980 reference population, has been developed which has been statistically adjusted to represent both sexes and the largest minority group (Sellman & Hagan, 1981).

A1.2

Where the information is too extensive to be fully reported in the manual, the essential information should be summarized and accompanied by references to other sources of information....(Very desirable)

This report presents a reference list on aspects of the ASVAB research and development program.

A1.2.1

When information about a test is provided in a separate publication, that publication should meet the same standards of accuracy and freedom from misleading impressions that apply to the manual. (Essentlal)

A1.2.2

Promotional material for a test should be accurate and not give the reader false impressions. (Essential)

A1.2.3

Informational material distributed within a using organization should be accurate, complete for the purposes of the reader's need, and written in language that will not give the reader a false impression. (Essential)

All publications dealing with the military applications of the ASVAB are subject to refereeing and review prior to publication; once published they meet the critical review of testing professionals in all the Services, and it is likely that error would be challenged.

Promotional material, in the commercial sense, does not exist for the military applications of the ASVAB. Descriptive material in the form of technical publications of the research and development programs, and in the form of administrative directives, is made available to the Service agencies responsible for procurement of testing materials, their administration, data processing and reporting, and utilization of results in personnel management decisions.

A2.

A test manual should describe fully the development of the test: the rationale, specifications followed in writing items or selecting observations, and procedures and results of item analysis or other research. (Essential)

Chapter 1, Appendix A and Appendix C present information and data responsive to this standard. Because of the evolution of current forms from experience with prior forms, a more complete understanding of the basis for the current form content might depend upon review of these selected references: Bayroff (1963); Bayroff and Fuchs (1970); Brown, Kincaid and McMorrow (1981); Frankfeldt (1970); Jensen, Massey and Valentine (1976); Maier and Fuchs (1972); Sims and Hiatt (1981); Swanson (1978, 1979); Thomas (1970); Valentine (1977); Valentine and Massey (1976); Vitola and Alley (1968); Weeks et al. (1975) and Zachert (1952).

A2.1

Data gathered during the process of developing a test before it is in final form should be clearly distinguished from data pertaining to the test in final form. (Essential)

Material in Chapter 1 dealing with preliminary materials for Forms 8, 9, and 10 is identified as being nonoperational in the text and through the designation of "X" forms.

A2.2

A test manual should specify the need for maintaining necessary test security. (Very Desirable)

Information supporting compliance with this standard is reported in Chapter 4.

A2.3

A test manual or supplementary document should provide representative sample items and a statement of the intended purpose of the test in a form that can be made available to those concerned about the nature and quality of a testing program. (Very Desirable)

Information supporting compliance with this standard is reported in Chapter 1, Chapter 4, and Appendix A.

A2.4

The identity and professional qualifications of item writers and editors should be described in instances where they are relevant; for example, when adequacy of coverage of a subject matter achievement test cannot appropriately or practically be measured against any external criterion. (Desirable)

Test outlines and test format are developed by the Joint-Service Selection and Classification Working Group, composed of professional personnel from the Air Force Human Resources Laboratory, the Army Research Institute for the Behavioral and Social Sciences, and the Navy Personnel Research and Development Center and the Center for Naval Analyses. ASVAB item selection and test assembly are done under the direction of psychologists on the staff of the Air Force Human Resources Laboratory. Review of final forms is by the Defense Advisory Committee on Military Personnel Testing.

For additional information, see Chapter 3 and these references: Bayroff and Fuchs (1970); Jensen et al. (1976); Maier and Fuchs (1972); Maier and Grafton (1981); Sims and Hiatt (1981); Vitola and Alley (1968); and Wiesen and Siegal (1976).

A3.

The test and its manual should be revised at appropriate intervals. The time for revision has arrived whenever changing conditions of use or new research data make any statements in the manual incorrect or misleading. (Very Desirable)

Historically, revisions of ASVAB forms have occurred as a consequence of compromise of current operational forms, as a result of research findings dictating the value of new kinds or new applications of test content, as a result of technological change as reflected in test validation criteria, or because of administrative constraints, such as a need to reduce testing time. Current policy dictates that the ASVAB be revised every three years.

A3.1

Competent studies of the test following its publication, whether the results are favorable or unfavorable to the test, should be taken into account in revised editions of the manual or its supplementary reports. Pertinent studies by investigators other than the test authors and publishers should be included. (Very Desirable)

Management and direction of the research and development programs are provided by the Joint-Service Selection and Classification Working Group and reviewed by the Manpower Accession Policy Steering Committee plus the Defense Advisory Committee on Military Personnel Testing. Under such oversight the research undertaken by the independent Service laboratories is comprehensive. Published reports are reviewed through technical management channels within the originating laboratory and furnished to the other laboratories and policy personnel in each Service and the DoD. Each successive revision of the ASVAB and its associated materials reflects individual Service efforts to ensure optimization of test content and format to meet Service needs.

A3.2

When the test is revised or a new form is issued, the manual should be suitably revised to take those changes into account. In addition, the nature and extent of the revision and the comparability of data from the old test and the revised test should be explicitly stated. (Essential)

Information supporting compliance with this standard is reported in Chapter 1 and Appendix A.

A3.2.1

If a short form of a test is prepared by reducing the number of items or organizing a portion of the test into a separate form, new evidence should be obtained and reported for that shorter test. (Essential)

A3.2.2

When a short form is prepared from an established test, the manual should present evidence that the items in the short form represent the items in the long form or measure the same characteristics as the items in the long form. (Very Desirable)

No short form of the ASVAB has been developed, and thus the standard is not relevant.

B. Aids to Interpretation

B1

The test, the manual, the record forms and other accompanying material should help users make correct interpretations of the test results and should warn against common misuses. (Essential)

Interpretation of test results to diagnose academic weakness, establish personality structure, or measure vocational preference is not done as a formal part of the ASVAB program. The use of test results is nonjudgmental subsequent to each Service's determining the minimum required scores for enlistment and subsequent entry into an occupational area. Strong interest of the examinee in an occupational area for which achieved test scores are below specified minimum levels (specified in Service classification manuals) may result in a request for waiver of the required minimum score. Such requests are reviewed and approved by higher echelon personnel who have the benefit of training and the advice of qualified professional personnel. For further discussion of this issue, see Chapter 4 on test implementation.

B1.1

Names given to published tests, and to parts within tests should be chosen to minimize the risk of misinterpretation by test purchasers and subjects. (Essential)

The subtests in the ASVAB carry descriptive names (see Chapter 1 and Appendix A). The ASVAB is not purchased by users.

B1.1.1

Devices for identifying interests and personality traits through self-report should be entitled "inventories," "questionnaires," or "check-lists," rather than "tests." (Very Desirable)

Forms 8, 9, and 10 contain neither vocational interest nor personality assessment items. The standard is therefore not relevant.

B1.2

The manual should draw the user's attention to data that especially need to be taken into account in the interpretation of test scores. (Very Desirable)

See discussion of standard B1.

B1.3

The manual should call attention to marked influences on test scores known to be associated with region, socioeconomic status, race, creed, color, national origin, or sex. (Essential)

Information supporting compliance with this standard is reported in Chapter 3 and Boldt et al. (1977).

B1.4

The manual should draw attention to, and warn against, any serious error of interpretation that is known to be frequent. (Essential)

See discussion of standard B1.

B2

The test manual should state explicitly the purposes and applications for which the test is recommended. (Essential)

Information supporting compliance with this standard is reported in Chapters 1, 2, and 4.

B2.1

If a test is intended for research use only and is not distributed for operational use, that fact should be prominently stated in the accompanying materials. (Essential)

This standard is not applicable to the ASVAB. Operational forms of the ASVAB on occasion are used as a basis for research and development data for psychometric and validation studies.

B3

The test manual should describe clearly the psychological, educational, or other reasoning underlying the test and nature of the characteristic it is intended to measure. (Essential)

Information supporting compliance with this standard is reported in Chapters 1, and 2, and Appendix A.

B3.1

In the case of tests developed for content-referenced interpretation, special attention should be given to defining the content domain in operational terms....(Essential)

See discussion of standard B1.

B4

The test manual should identify any special qualifications required to administer the test and to interpret it properly. (Essential)

Information supporting compliance with this standard is reported in Chapter 4.

B4.1

The test manual should not imply that a test is "self-interpreting." It should specify information to be given about test results to persons who lack the training required to interpret them. (Essential)

B4.2

Where a test is recommended for a variety of purposes or types of inference, the manual should indicate the amount of training required for each use. (Essential)

B4.3

The manual should draw the user's attention to references with which he should become familiar before attempting to interpret the test results. (Very Desirable)

Information supporting compliance with this standard is reported in discussion of standard B1.

B5.

Evidence of validity and reliability, along with other relevant research data, should be presented in support of any claims being made. (Essential)

See Chapter 3 and references Atwater and Abrahams (1980); Bayroff and Fuchs (1970); Booth-Kewley (1983); Frankfeldt (1970); Jensen and Valentine (1976); Kettner (1976); Maier and Fuchs (1972); Maier and Grafton (1981); Maier and Truss (1983); Mathews, Valentine, and Sellman (1978); Ree, Mullins, Mathews and Massey (1982); Rossmeissl et al. (1983); Sims and Hiatt (1981); Swanson (1978, 1979); Thomas (1970); Valentine (1977); Valentine and Massey (1976); Vitola and Alley (1968); Weeks et al. (1975); Wiesen and Siegel (1976) and Wilbourn et al. (in press).

B5.1

Statements in the manual reporting relationships are by implication quantitative and should be stated as precisely as the data permit. If data to support such statements have not been collected, that fact should be made clear. (Essential)

B5.2

Statistical procedures that are well known and readlly interpreted should be preferred for reporting any quantitative information. Any uncommon statistical techniques should be explained and references to descriptions of them should be given. (Essential)

B5.3

When the statistical significance of a relationship is reported, the statistical report should be in a form that makes clear the sensitivity or power of the significance test. (Essential)

Research and development studies reported by the Service laboratories are characterized by sophisticated statistical treatments involving numerous well known tests of significance. Expectancy tables are usually produced based upon derived validity information and citation of the proportion of predictable variance accounted for by predictors of interest.

Studies based upon Service populations often address samples of sizes unknown in academic research and rarely approached in industry. Validation figures too small for practical use within schools or in selecting a few hundred workers a year in an industry become significant in terms of the tens of thousands of persons processed into military service each year. See Chapter 3, Appendix C, and references Booth-Kewley (1983); Maier and Truss (1983); Rossmeissl et al. (1983); Thorndike (1949) and Wilbourn et al. (in press).

B5.4

The manual should differentiate between an interpretation that is applicable only to average tendencies of a group and one that is applicable to an individual within the group. (Very Desirable)

The basic use of ASVAB results is their application to individual personnel decisions, hence such differentiation is not required and the standard is not relevant.

B5.5

The manual should state clearly what interpretations are intended for each subscore as well as for the total test. (Essential)

All use of ASVAB subscores is through their inclusion in composite scores.

B6

Test developers or others offering computer services for test interpretation should provide a manual reporting the rationale and evidence in support of computer-based interpretation of test scores. (Essential)

This standard is not applicable to the ASVAB program.

C. Directions for Administration and Scoring

C1.

The directions for administration should be presented in the test manual with sufficient clarity and emphasis so that the test user can duplicate, and will be encouraged to duplicate, the administrative conditions under which the norms and the data on reliability and validity were obtained. (Essential)

C1.1

The directions published in the test manual should be complete enough that persons tested will understand the task as the author intended. (Essential)

C1.1.1

The directions should clearly point out such critical matters as instructions on guessing, time limits, and procedures for marking answer sheets. (Essential)

C1.1.2

The directions to the test administrator should include guidance for dealing with questions from examinees. (Very Desirable)

sed in the standards listed above are addressed

.. of comprehensibility, clarity, emphasis and

motivation to maintain testing conditions identical to the conditions of the standardization testing are difficult to document, other than by pointing out that the administrative directions for Forms 8, 9, and 10 are the latest version of instructions which have an extended history of successful use in the testing environment. Although these forms of ASVAB are new, there are no subtests of types which have not been used in earlier military batteries.

With reference to dealing with questions from examinees, the test administrator is not allowed to respond ad lib; instead, the administrator must cite appropriate portions of the standardized instructions.

C1.2
If expansion or elaboration of instructions described in the test manual is permitted, the conditions under which this may be done should be clearly stated either in the form of general rules, or in terms of giving numerous examples, or both. (Essential)

Information supporting compliance with this standard is reported in Chapter 4. Expansion or elaboration of test administration instructions is not permitted. General guidance on handling testing room problems such as ill examinees, distracting external noise, or individual refusal to respond to the tests, is given in Service-specific manuals and operating instructions, but is generally uniform. A common administration manual is used for all examinees.

C2.
Instructions should prepare the examinee for examination: Sample material, practice use of answer sheets or punch cards, sample questions, etc., should be provided. (Desirable)

Information supporting compliance with this standard is reported in Chapter 4.

C3.
The procedures for scoring the test should be presented in the test manual with a maximum of detail and clarity to reduce the likelihood of scoring error. (Essential)

C3.1
The test manual should furnish scoring instructions that maximize the accuracy of scoring an objective test by outlining a procedure for checking the obtained scores for computational or clerical errors. (Very Desirable)

Although the score for the AFQT composite may be derived locally under some conditions, ASVAB test processing is normally done in centralized facilities, with scores reported back to the recruiting offices of the Services. The centralized processing involves complete checks to ensure that equipment is operating accurately, audits of samples of processed materials, and verification of the test form taken for individuals who fail to qualify for Service entry.

C3.2

Where subjective processes enter into the scoring of a test, evidence on the degree of agreement between independent scoring under operational conditions should be presented in the test manual....(Very Desirable)

C3.2.1

The basis for scoring and the procedures for training scorers should be presented in the test manual in sufficient detail to permit other scorers to reach the level of agreement reported in studies of scorer agreement given in the manual. (Very Desirable)

C3.2.2

If persons having various degrees of supervised training are expected to score the test, studies of the interscorer agreement at each skill level should be presented in the test manual. (Desirable)

All scoring for Forms 8, 9, and 10 is objective. There are no subjective materials in the battery and thus these three standards are not relevant.

C3.3

If the test is designed to use more than one method for the examinee's recording of his responses, such as hand-scored answer sheets, or entering of responses in the test booklet, the test manual should report data to the degree to which results from these methods are interchangeable. (Essential)

All examinee responses are collected on custom designed, optically scannable answer sheets. This general topic has been explored by Valentine and Cowan (1974).

C3.4

If an unusual or complicated scoring system is used, the test manual should indicate the approximate amount of time required to score the test. (Desirable)

See the discussion of standard C3.2

C3.5
"Correction for guessing" formulas should be used with multiple-choice and true-false items when the test is speeded. (Desirable)

All ASVAB subtests are scored "rights only."

D1
Norms should be published in the test manual at the time of release of the test for operational use. (Essential)

D1.1
Norms should be established even for a test developed for local use or only for predictive purposes. (Desirable)

D1.2
Even though a test is expected to be used primarily with local norms, the test manual should nevertheless provide normative data to aid the interpreter who lacks local norms. (Very Desirable)

D2
Norms presented in the test manual should refer to defined and clearly described populations. These populations should be the groups with whom users of the test will ordinarily wish to compare the persons tested. (Essential)

D2.1
Care should be taken to avoid misleading impressions about the generality of normative data. (Essential)

The Services have been required to maintain a standardized measure of mental ability for incoming personnel under the provisions of the Selective Service Act of 1948 and subsequent revisions of that federal statute. Further, entering personnel must be comparable to prior-entering persons in terms of their qualification for various occupational fields. For this reason the AFQT scores derived from selection and classification tests used by all the Services have been referenced to the performance of the 1944 reference population, as described in Chapter 2. This normative base has been used through Forms 8, 9, and 10. Information supporting compliance with this standard is reported in Chapter 2 and references: Bayroff (1963); Boldt (1980b); Jensen et al. (1976); Lecznar (1963); Maier (1981a, 1981b); OASD(MRA&L) (1980); Ree, Mathews, Mullins and Massey (1982); Sims and Truss (1980); Uhlaner and Bolanovich (1952) and Vitola and Alley (1968).

Because of currently increased use of women in military service, and other changes which have accumulated with the passage of time, the 1944 reference population has become increasingly undesirable for use. The DoD, in concert with the Department of Labor, sponsored the development of a new normative base which included women and minorities, known as the 1980 reference population. This population is nationally representative of all young men and women, ages 18 to 23, living in the United States as of the summer of 1980. Forms 11-14, to be introduced in 1984, are standardized against this 1980 reference population so that the norms may be interpreted to address questions relevant to possible differences due to ethnicity or gender. Details of the 1980 reference population and procedures relevant to its implementation can be found in Maier and Sims (1982); OASD(MRA&L) (1982b); Ree, Valentine and Earles (in press) and Sellman and Hagan (1981).

D2.1.1

The test manual should report the method of sampling from the population of examinees and should discuss any probable bias in this sampling procedure. (Essential)

D2.1.2

Norms reported in any test manual should be based on well planned samplings rather than on data collected primarily because it is readily available. Any deviations from the plan should be reported along with descriptions of actions taken or not taken with respect to them. (Essential)

D2.1.3

In addition to reporting the numbers of individuals in a set of normative data, the manual should also report the number of sampling units from which those individuals were drawn along with the numbers of individuals in each unit. (Essential)

Information supporting compliance with this standard is reported in Chapter 2 and the references cited for standards D1 - D2.1.

D2.2

The description of the norms group in the test manual should be complete enough so that the user can judge its appropriateness to his use. The description should include number of cases, classified by one or more of such relevant variables as ethnic mix, socioeconomic level, age, sex, locale, and educational status. If cluster sampling is employed the description of the norms group should state the number of separate groups tested. (Essential)

The "user" must apply the norms established in the metric appropriate to each Service. The Army and Marine Corps use a standardized score, the Air Force uses a percentile score, and the Navy uses raw composites based upon the combinations of various tests whose scores have been standardized to a mean of 50 and a standard deviation of 10 in Service applicant populations. Descriptions of the equating methodology appears in Chapter 2.

D2.2.1
The populations upon which the psychometric properties of a test were determined and for which normative data are available should be clearly and prominently described in the manual....(Essential)

Information supporting compliance with this standard is reported in Chapters 1 and 2.

D2.3
If the sample on which norms are based is small or otherwise undependable, the user should be cautioned explicitly in the test manual regarding the possible magnitude of errors arising in interpretation of the scores. (Very Desirable)

The standardization samples used for Forms 8, 9, and 10 were large (see Chapter 2).

D2.4
Norms on subtests or groups of test items should be reported in the test manual only if the validity and reliability of such subtests or groups of items are indicated. (Essential)

Information supporting compliance with this standard is reported in Chapter 3.

D2.5
The significant aspects of conditions under which normative data were obtained should be reported in the test manual. (Essential)

Information supporting compliance with this standard is reported in Chapters 2 and 4.

D3
In reporting norms, test manuals should use percentiles for one or more appropriate reference groups or standard scores for which the basis is

clearly set forth; any exceptional type of score or unit should be explained and justified. Measures of central tendency and variability should be reported. (Essential)

Information supporting compliance with this standard is reported in Chapter 2 and Appendix C.

D3.1

In the case of tests used for prediction, expectancy tables or experience tables translating obtained scores into probabilities of success, or into proficiency levels should be included whenever possible. (Desirable)

Information supporting compliance with this standard is reported in Chapter 3 and references Booth-Kewley (1983); Maier and Grafton (1981); Maier and Truss (1983); Rossmeissl et al. (1983) and Wilbourn et al. (in press).

D4

Local norms are more important for many users of tests than are published norms. A test manual should suggest using local norms in such situations. (Very Desirable)

This standard is not applicable to the military ASVAB program.

D5

Derived scales used for reporting scores should be carefully described in the test manual to increase the likelihood of accurate interpretation of scores by both the test interpreter and the examinee. (Essential)

Information supporting compliance with this standard is reported in Chapter 2. "Interpretation" of test scores is rarely required of the user — use requires only the comparison of achieved scores with stated minimum levels.

D5.1

Derivation of any scale from normative data should be clearly and unambiguously described in terms likely to prevent misinterpretations or overgeneralization. (Essential)

This standard is not applicable to the ASVAB.

D5.2

When standard scores are used, the system should be consistent with the purposes for which the test is intended and should be described in detail

in the test manual. The reasons for choosing one scale in preference to another should also be made clear in the manual. (Very Desirable)

AFQT scores are not reported as standard scores, but as percentiles of a reference population. Standard scores are used for making up the selector composites. The metrics used by the various Services were chosen before common use of the ASVAB was directed. Broad usage of selection/classification scores within the personnel systems of each Service dictates the economy of maintaining the original metric. See Bayroff (1963), Uhlaner and Bolanovich (1952) and Wecks et al. (1975).

D5.2.1
The manual should specify whether standard scores are linear transformations of raw scores or are normalized. (Essential)

The scores developed for the tests are standardized. The composite scores are equated to the normative base through an equipercentile system (see Chapter 2).

D5.2.2
The choice of a standard scale should be based upon either the standard error of measurement of the raw scores, or on some other basis that is clearly defined. (Desirable)

See the discussion of standard D5.2.

D5.2.3
Interpretive scores that lend themselves to gross misinterpretaion such as mental age or grade equivalent scores should be abandoned or their use discouraged. (Very Desirable)

ASVAB scores are presented in forms which were devised for administrative convenience and usefulness. They are not interpretive in the sense of the Standards and, therefore, are not subject to misinterpretation.

D5.3
When it is suggested in the manual that percentile ranks are to be plotted on a profile sheet, the profile sheet should be based upon the normal probability scale or some other appropriate non-linear transformation. (Very Desirable)

As used by the Military Services, percentile ranks are not plotted. The high school version includes the plotting of percentiles on a normal probability scale.

D5.4
Normative data should be presented in a form that emphasizes the fallibility of an obtained score. (Very Desirable)

This standard is not applicable to the ASVAB as used in the Military Services. The "user" does not see the normative data, as such, but is given tables that specify minimum qualifying levels for entry, either into Service or into an occupation within that Service. The standard error of measurement is considered at the time the minimum levels are established (see Chapter 2).

D6
If scales are revised, new forms added, or other changes made, the revised test manual should provide tables of equivalence between the new and the old forms. This provision is particularly important in cases where data are recorded on cumulative records. (Desirable)

The content of a composite score of a given name in a Service tends to be consistent across sequential forms of the ASVAB. If the content is changed, usually the name is also changed. Such consistency is imperative if personnel management decisions are to be made across groups of persons tested on different forms of the battery. Tables are provided for conversion of AFQT raw scores to percentiles of the normative base (see Chapter 2 and Appendix B). As discussed in Chapter 2, the use of norms derived from the 1944 reference population and, later from the 1980 reference population, make the AFQT scores comparable across years, test forms, and Services.

D6.1
When a new form is equated with an older form of a test, the revised manual should describe the content of both the old and new forms and the nature of the norms group for each form. (Essential)

ompliance with this standard is reported in

D6.2

The manual should describe the method used to establish equivalent or comparable scores and should include an assessment of the accuracy of the equating procedure. (Very Desirable)

Information supporting compliance with this standard is reported in Chapter 2.

D7

Where it is expected that a test will be used to assess groups rather than individuals (i.e., for schools or programs) normative data based on group summary statistics should be provided. (Essential)

The primary use of the ASVAB is to select and classify individuals for military service, hence the standard is not relevant.

C. Summary

The ASVAB, under continuous review and development in a professional research environment, is in virtually full compliance with the relevant APA Standards. Of 79 specifically listed standards (APA, 1974) 65 were seen to be relevant to the ASVAB as used by the Military Services.

The 14 standards which were seen not to be relevant dealt mostly with topics associated with subjective measures, such as are found in personality scales or vocational interest measures. Other non-relevant standards focused on the development of local norms, which are mostly appropriate for measures of educational achievement or of proficiency on specific tasks. The matters of group performance as opposed to individual performance were also addressed in several standards which were deemed not relevant, in the light of the ASVAB's use for selection and classification of individuals within each of the Military Services.

The matters found non-relevant recurred in each of the topical areas of the standards. For example, the matter of subjective scores, as derived for personality scales, appeared in Section B, "Aids to Interpretation," Section C, "Directions for Administration and Scoring," and in Section D, "Norms and Scales." A summary of the numbers of standards by topical area appears in Table 32.

The one relevant standard which was not met deals with the use of formula scores involving a penalty for wrong answers. The APA standards call for such

scoring procedures for speeded tests. In the ASVAB program all tests are scored "rights only." Of the three designations "essential," "very desirable," and "desirable," that standard is designated "desirable."

Table 32

Compliance with American Psychological Association Standards for Test Development as Used in Military Testing Programs

Section	Number of APA Standards	Number Not Relevant	Number Met	Number Not Met
A[a]	16	1	15	0
B	21	5	16	0
C	14	4	9	1
D	28	4	24	0
Total	79	14	64	1

[a] Sections of APA Standards:
A. Dissemination of Information
B. Aids to Interpretation
C. Directions for Administration and Scoring
D. Norms and Scales

References

REFERENCES

American Psychological Association. (1974). Standards for educational and psychological tests. Washington, DC: Author.

American Psychological Association, Division of Industrial-Organizational Psychology. (1980). Principles for the validation and use of personnel selection procedures (2nd ed.). Berkeley, CA: University of California Printing Department.

Atwater, D. C., & Abrahams, N. M. (1980). Evaluation of alternate ASVAB composites for selected Navy technical schools (Tech. Rep. NPRDC-TR-80-15). San Diego, CA: Navy Personnel Research and Development Center.

Bayroff, A. G. (1963). The mobilization base for AFQT norms (Research Memorandum No. 63-8). Washington, DC: U. S. Army Personnel Office.

Bayroff, A. G., & Fuchs, E. F. (1970). The Armed Services Vocational Aptitude Battery (Tech. Research Rep. No. 1161). Arlington, VA: U. S. Army Behavioral and Systems Research Laboratory.

Bock, R. D., & Mislevy, R. J. (1981). Data quality analysis of the Armed Services Vocational Aptitude Battery. Chicago, IL: National Opinion Research Center.

Bock, R. D., & Moore, E. G. J. (1984). Profile of American youth: Demographic influences on ASVAB test performance. Washington, DC: Office of the Assistant Secretary of Defense (Manpower, Installations, and Logistics).

Boldt, R. F. (1980a). Item analysis for the Armed Services Vocational Aptitude Battery forms 8, 9, and 10. Unpublished manuscript.

>f the Armed Services Vocational Aptitude
l Classification Test to the Armed Forces
Memorandum No. 80-2). Washington, DC:
e, Directorate for Accession Policy.

Boldt, R. F. (1980c). Scaling of the AFQT composite of the Armed Services Vocational Aptitude Battery form 8 in a high school population (Tech. Memorandum No. 80-3). Washington, DC:, Office of the Secretary of Defense, Directorate for Accession Policy.

Boldt, R. F., Levin, M. K., Powers, D. E., Griffin, M., Troike, R. C., Wolfram, W., & Ratliff, F. R. (1977). Sociolinguistic and measurement considerations for construction of armed services selection batteries (AFHRL-TR-77-65). Brooks AFB, TX: Air Force Human Resources Laboratory.

Booth-Kewley, S. (1983). Predictive validation of Armed Services Vocational Aptitude Battery forms 8, 9, and 10 against performance at 47 Navy schools [Draft report]. San Diego, CA: Navy Personnel Research and Development Center.

Brown, C. J., Kincaid, J. P., & McMorrow, H. (1981). Assessment of numerical skills of Navy enlisted personnel (Tech. Rep. TAEG-TM-81-4). Orlando, FL: U. S. Navy, Training Analysis and Evaluation Group.

Department of Defense. (1983). Manual for administration, Armed Services Vocational Aptitude Battery (DoD 1304.12A). Washington, DC: Author.

Department of Defense. (1984). Technical supplement to the counselor's manual for ASVAB-14. North Chicago, IL: U. S. Military Entrance Processing Command.

Fischl, M. A., Ross, R. M., & McBride, J. R. (1979). Development of factorially based ASVAB high school composites (Tech. Paper No. 360). Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Frankfeldt, E. (1970). Comparison of ACB and ASVAB clerical tests for use in the Army Qualification Battery (Tech. Rep. BESRL-RM-70-6). Arlington, VA: Army Behavioral and Systems Research Laboratory.

Guilford, J. P. (1950). Fundamental statistics in psychology and education. New York: McGraw-Hill.

Guinn, N., Tupes, E. C., & Alley, W. E. (1970a). Demographic differences in aptitude test performance (AFHRL-TR-70-15). Brooks AFB, TX: Air Force Human Resources Laboratory.

Guinn, N., Tupes, E. C., & Alley, W. E. (1970b). Cultural subgroup differences in the relationships between Air Force aptitude composites and training criteria (AFHRL-TR-70-35). Brooks AFB, TX: Air Force Human Resources Laboratory.

Jensen, H. E., Massey, I. H., & Valentine, L. D., Jr. (1976). Armed Services Vocational Aptitude Battery development (ASVAB forms 5, 6, and 7) (AFHRL-TR-76-87). Brooks AFB, TX: Air Force Human Resources Laboratory.

Jensen, H. E., & Valentine, L. D., Jr. (1976). Validation of ASVAB-2 against civilian vocational-technical high school criteria (AFHRL-TR-76-16). Brooks AFB, TX: Air Force Human Resources Laboratory.

Kettner, N. (1976). Armed Services Vocational Aptitude Battery (ASVAB form 5): Comparison with GATB and DAT tests (AFHRL-TR-76-78). Brooks AFB, TX: Air Force Human Resources Laboratory.

Kuder, G. F., & Richardson, M. W. (1937). Theory of the estimation of test reliability. Psychometrika, 2, 151-160.

Lecznar, W. B. (1962). Some aptitude data on Air Force enlisted accessions (PRL-TDR-62-10). Lackland AFB, TX: Aerospace Medical Division, Personnel Research Laboratory.

Lecznar, W. B. (1963). Comparison of Air Force aptitude indexes with corresponding talent test composites (PRL-TDR-63-18). Lackland AFB, TX: Aerospace Medical Division, 6570th Personnel Research Laboratory.

Maier, M. H. (1981a). Calibrating the ASVAB (Memorandum No. 81-3098). Alexandria, VA: Center for Naval Analyses.

Maier, M. H. (1981b). Development and calibration of the Armed Services Vocational Aptitude Battery (ASVAB) forms 8, 9, and 10 [In minutes of the 29-30 June, 1981 meeting of the Defense Advisory Committee on Military Personnel Testing]. Washington, DC: Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics).

Maier, M. H., & Fuchs, E. F. (1972). An improved differential Army classification system (BESRL-TR-1177). Arlington, VA: Army Behavioral and Systems Research Laboratory.

Maier, M. H., & Grafton, F. C. (1981). Aptitude composites for ASVAB 8, 9, and 10 (Research Report No. 1308). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Maier, M. H., & Sims, W. H. (1982). Constructing an ASVAB score scale in the 1980 reference population (Memorandum No. 82-31). Alexandria, VA: Center for Naval Analyses.

Maier, M. H., & Truss, A. R. (1983). Validity of ASVAB forms 8, 9, and 10 for Marine Corps training courses: Subtests and current composites (Memorandum No. 83-3107). Alexandria, VA: Center for Naval Analyses.

Maier, M. H., & Truss, A. R. (1984). Validity of the occupational and academic composites for the Armed Services Vocational Aptitude Battery, form 14, in Marine Corps training courses (Memorandum No. 84-3043/2). Alexandria, VA: Center for Naval Analyses.

Mathews, J. J., Valentine, L. D., Jr., & Sellman, W. S. (1978). Prediction of reading grade levels of service applicants from Armed Services Vocational Aptitude Battery (ASVAB) (AFHRL-TR-78-82). Brooks AFB, TX: Air Force Human Resources Laboratory.

McBride, J. R. (1981). Factor analysis of ASVAB Forms 8, 9, and 10 on Marine recruits. Unpublished manuscript.

McCormick, E. J. (1979). Job analysis: Methods and applications. New York: AMACOM, 135 W. 50th Street, New York, NY 10020.

Morsh, J. E., & Archer, W. E. (1967). Procedural guide to conducting occupational surveys in the U. S. Air Force (PRL-TR-67-11). Lackland AFB, TX: Aerospace Medical Division, Personnel Research Laboratory.

Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics). (1980, December 31). Implementation of new Armed Services Vocational Aptitude Battery and actions to improve the enlistment standards process [A report to the House and Senate Committees on Armed Services]. Washington, DC: Author.

Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics). (1982a). Occupational conversion manual, enlisted/officer/civilian. Washington, DC: Defense Manpower Data Center.

Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics). (1982b). Profile of American youth — 1980 nationwide administration of the Armed Services Vocational Aptitude Battery. Washington, DC: Author.

Pass, J. J. (1980). Methods to evaluate scales and sample size for stable task inventory information (NPRDC-TR-80-28). San Diego, CA: Navy Personnel Research and Development Center.

Prestwood, J. S., Vale, C. D., & Massey, R. H. (in press). Development of ASVAB forms 11, 12, and 13. Brooks AFB, TX: Air Force Human Resources Laboratory.

Ree, M. J., Mathews, J. J., Mullins, C. J., & Massey, R. H. (1982). Calibration of Armed Services Vocational Aptitude Battery forms 8, 9, and 10 (AFHRL-TR-81-49). Brooks AFB,TX: Air Force Human Resources Laboratory.

Ree, M. J., Mullins, C. J., Mathews, J. J., & Massey, R. H. (1982). Armed Services Vocational Aptitude Battery: Item and factor analysis of forms 8, 9, and 10 (AFHRL-TR-81-55). Brooks AFB, TX: Air Force Human Resources Laboratory.

Ree, M. J., Valentine, L. D., Jr., & Earles, J. A. (in press). The 1980 youth population: A verification report. Brooks AFB, TX: Air Force Human Resources Laboratory.

Rossmeissl, P. G., Martin, C. J., & Wing, H. (1983). Validity of ASVAB 8, 9, and 10 as predictors of training success (Selection and Classification Working Paper No. 83-3). Alexandria, VA: Army Research Institute for the Behavorial and Social Sciences.

Sellman, W. S., & Hagan, H. T. (1981). The profile of American youth: Data audit (Tech. Memorandum No. 81-1). Washington, DC: Office of the Secretary of Defense, Directorate for Accession Policy.

Shore, C. W., & Marion, R. (1972). Suitability of using common selection test standards for Negro and white airmen (AFHRL-TR-72-53). Brooks AFB, TX: Air Force Human Resources Laboratory.

Sims, W. H., & Hiatt, C. M. (1981). Validation of the Armed Services Vocational Aptitude Battery (ASVAB) forms 6 and 7 with applications to ASVAB forms 8, 9, and 10 (Study No. 1160). Alexandria, VA: Center for Naval Analyses.

Sims, W. H., & Mifflin, T. L. (1978). A factor analysis of the Armed Services Vocational Aptitude Battery (ASVAB) forms 6 and 7 (CNA-78-3092). Arlington, VA: Center for Naval Analyses.

Sims, W. H., & Truss, A. R. (1980). Normalization of the Armed Services Vocational Aptitude Battery (ASVAB) forms 8, 9, and 10 using a sample of service recruits (CRC 438). Alexandria, VA: Center for Naval Analyses.

Streicher, A. H., & Friedman, D. (1983). Armed Services Vocational Aptitude Battery (ASVAB form 14): Comparison with CAT, DAT and FIT/FACT tests [Preliminary report] . Rockville, MD: Research Applications.

Swanson, L. (1978). Armed Services Vocational Aptitude Battery, forms 6 and 7: Validation against school performance: Interim report (NPRDC-TR-78-24). San Diego, CA: Navy Personnel Research and Development Center.

Swanson, L. (1979). Armed Services Vocational Aptitude Battery, forms 6 and 7: Validation against school performance in Navy enlisted schools (July 1976-February 1978) (NPRDC-TR-80-1). San Diego, CA: Navy Personnel Research and Development Center.

Taylor, H. C., & Russell, J. T. (1939). Relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. Journal of Applied Psychology, 23, 565-578.

Thomas, P. J. (1970). A comparison between the Armed Services Vocational Aptitude Battery (form 1) and the Navy Basic Test Battery in predicting Navy school performance (Tech. Bulletin STB 70-4). San Diego, CA: Navy Personnel Training Research Laboratory.

Thorndike, R. L. (1949). Personnel selection: Test and measurement techniques. New York: Wiley.

Uhlaner, J. E. (1968). The research psychologist in the Army 1917 to 1967 (Tech. Research Rep. No. 1955). Washington, DC: U. S. Army Behavioral Science Research Laboratory.

Uhlaner, J. E., & Bolanovich, D. J. (1952). Development of the Armed Forces Qualification Test and predecessor Army screening tests, 1946-1950 (Report No. 976). Washington, DC: Department of the Army, The Adjutant General's Office, Personnel Research Branch.

Uniform guidelines on employee selection procedures. (1978, August 25). Federal Register, 43, 166:38290.

Valentine, L. D., Jr. (1977). Prediction of Air Force technical training success from ASVAB and educational background (AFHRL-TR-77-18). Brooks AFB, TX: Air Force Human Resources Laboratory.

Valentine, L. D., Jr., & Cowan, D. K. (1974). Comparability study of Armed Services Vocational Aptitude Battery scores from answer sheet and answer card administration (AFHRL-TR-73-55). Brooks AFB, TX: Air Force Human Resources Laboratory.

Valentine, L. D., Jr., & Massey, I. H. (1976). Comparison of ASVAB test-retest results of male and female enlistees (AFHRL-TR-76-43). Brooks AFB, TX: Air Force Human Resources Laboratory.

Vitola, B. M., & Alley, W. E. (1968). Development and standardization of Air Force composites for the Armed Services Vocational Aptitude Battery (AFHRL-TR-68-110). Lackland AFB, TX: Air Force Human Resources Laboratory.

Weeks, J. L., Mullins, C. J., & Vitola, B. M. (1975). Airman Classification Batteries from 1948 to 1975: A review and evaluation (AFHRL-TR-75-78). Brooks AFB, TX: Air Force Human Resources Laboratory.

Wiesen, J. P., & Siegel, A. I. (1976). Psychometric characteristics of the Armed Services Vocational Aptitude Battery (ASVAB): An annotated bibliography. Wayne, PA: Applied Psychological Services.

Wilbourn, J. M., Valentine, L. D., Jr., & Ree, M. J. (in press). Aptitude index validation of the Armed Services Vocational Aptitude Battery (ASVAB) forms 8, 9, and 10 (AFHRL-TP-84-08). Brooks AFB, TX: Air Force Human Resources Laboratory.

Yellen, T. M. I., & Foley, P. P. (1978). Navy vocational information system (NPRDC-TR-78-22). San Diego, CA: Navy Personnel Research and Development Center.

Zachert, V. (1952). Factor analysis of the Army, Navy, and Air Force classification batteries (Research Bulletin No. 52-12). Lackland AFB, TX: Air Training Command, Human Resources Research Center.

Appendix A

ASVAB Content by Form

Table A-1

Subtest Content by Form

Subtest	First Generation 1968-1975 Forms 1-4	Second Generation 1976-1980 Forms 5-7	Third Generation 1980-present Forms 8-14
Word Knowledge (WK)	25	30	35
Arithmetic Reasoning (AR)	25	20	30
Mechanical Comprehension (MC)	25	20	25
Electronics Information (EI)	25	30	20
Space Perception (SP)	25	20	
Coding Speed (CS)	100		84
Shop Information (SI)	25	20	
Automotive Information(AI)	25	20	
Auto & Shop Information (AS)			25
Tool Knowledge (TK)	25		
Numerical Operations (NO)		50	50
Mathematics Knowledge (MK)		20	25
General Science (GS)		20	25
Classification Inventory (CI)		87	
Attention to Detail (AD)		30	
Paragraph Comprehension (PC)			15
General Information (GI)		15	
Total number of items	300	382	334

Descriptions of Subtests and Sample Items

Word Knowledge (WK): requires the examinee to select an alternative word whose meaning is most nearly the same as the meaning of a word underlined in a phrase.

Sample Question: It was a small table.

A. sturdy

B. round

C. little

D. cheap

The correct answer is "little," therefore C is the right answer.

Sample Question: Similar most nearly means

A. simmer.

B. alike.

C. compliment.

D. incomparable.

The correct answer is "alike," therefore B is the correct answer.

Arithmetic Reasoning (AR): arithmetic word problems.

Sample Question: A student bought a sandwich for 80 cents, milk for 20 cents, and pie for 30 cents. How much did the meal cost?

A. $1.00

B. $1.20

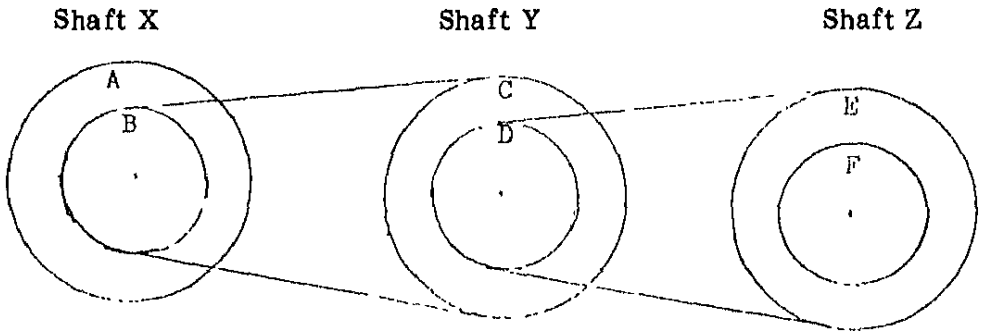
C. $1.30

D. $1.40

The total cost is $1.30, therefore C is the right answer.

Mechanical Comprehension (MC): requires answers to questions illustrating basic mechanical principles.

Sample Question:



Pulleys A and B turn with Shaft X; Pulleys C and D turn with Shaft Y; Pulleys E and F turn with Shaft Z.

Sample Question: When the system is running, which pulley makes more revolutions per minute than Pulley C?

A. Pulley A

B. Pulley D

C. Pulley E

D. Pulley F

The correct answer is A.

Electronics Information (EI): requires answers to electronic and electrical information questions.

Sample Question: What does the abbreviation AC stand for?

A. additional charge

B. alternating coil

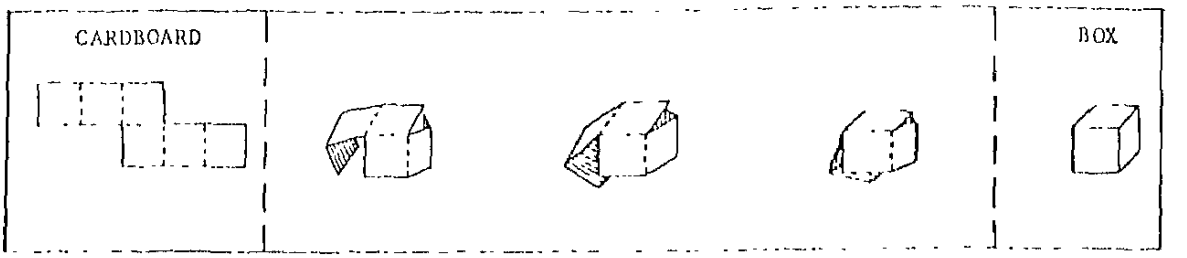
C. alternating current

D. ampere current

The correct answer is alternating current, so C is the correct response.

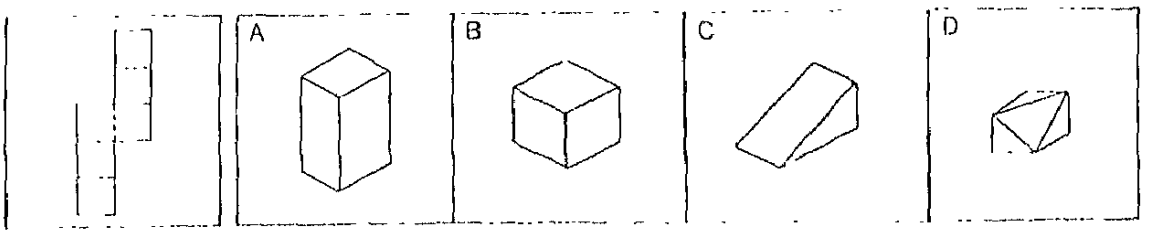
Space Perception (SP): involves visualizing the folding of flat patterns into three dimensional objects.

This test has questions about folding cardboard patterns into boxes. The first row of pictures below shows what this means. The dotted lines show where folds are to be made. The last picture shows the box that has been made by folding.



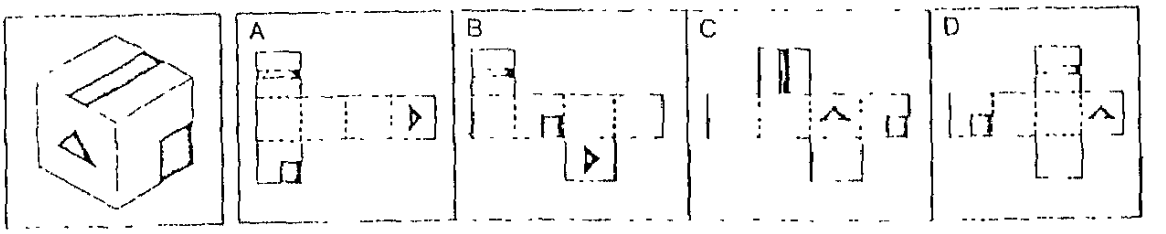
In this test, the first picture in each row shows a cardboard pattern that is to be folded. There are also four boxes labeled A, B, C and D. Your job is to find which box could be made by folding the pattern.

Look at the sample question below. Which box could this pattern make?



The B answer is correct.

Here is another type of question. Which of the four patterns could be made by unfolding the box?



The D answer is correct.

Coding Speed (CS): a reference list of 100 words matched with four-digit code numbers is used to select the correct code number for each of 84 words administered under speeded conditions.

Sample Questions:

Each question in the test is a word taken from the key at the top of that page. From among the possible answers listed for each question, you are to find the one which is the correct code number for that word, then, you blacken the space for that answer on your separate answer form.

Look at the practice key and the five sample questions below.

KEY

GREEN...2715	MAN.....3451	SALT...4586
HAT.....1413	ROOM...2864	TREE...5972

SAMPLE QUESTIONS	ANSWERS				
	A	B	C	D	E
S1. ROOM	1413	2715	2864	3451	4586
S2. GREEN	2715	2864	3451	4586	5972
S3. TREE	1413	2715	3451	4586	5972
S4. HAT	1413	2715	3451	4586	5972
S5. SALT	1413	2864	3451	4586	5972

Notice that each of the questions is one of the words in the key table. To the right of each question are possible answers listed under the letters A, B, C, D, and E. The word in Question S1 is "ROOM." By looking in the key you see that the code number for room is 2864. Among five possible answers for Question S1, 2864 is listed under choice C; so C is the correct answer. The word for question number S2 is "GREEN." By looking in the key you see that the code number for green is 2715. Among the possible answers, 2715 is listed under choice A, so A is the correct answer.

Shop Information (SI): determines the examinee's previous knowledge about shop practices and the use of specific tools.

Automotive Information (Al): determines specific knowledge about automobiles and automobile engines.

Auto and Shop Information (AS): requires responses to questions about automobiles, shop practices, and the use of tools.

Sample Questions:

The fuel used most commonly for automobile engines is

A. kerosene.

B. benzine.

C. crude oil.

D. gasoline.

Gasoline is the most commonly used fuel, so D is the correct response.

Sample Questions:

Thin sheet metal should be cut with

A. ordinary scissors.

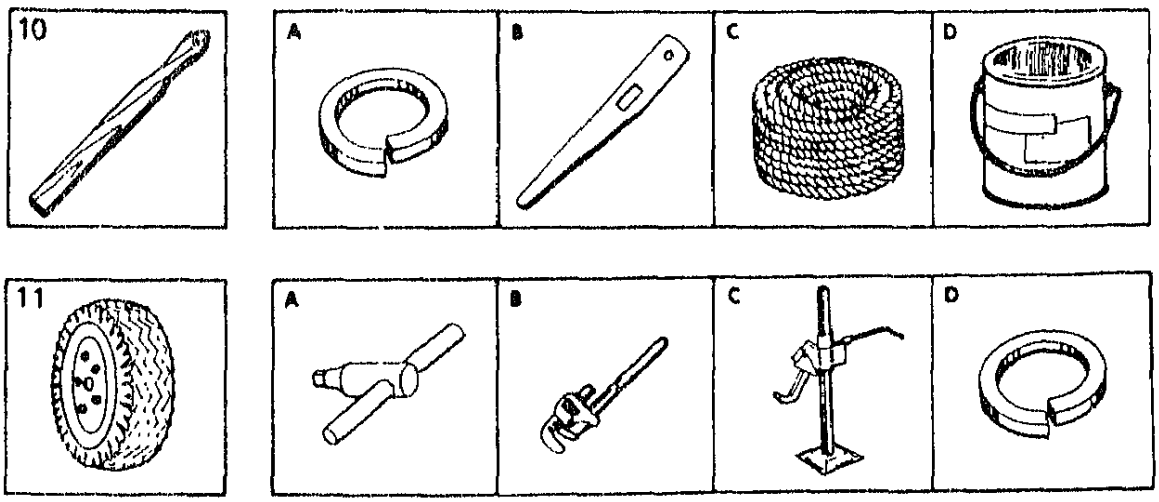
B. a hack saw.

C. tin shears.

D. a jig saw.

Tin shears are used to cut thin metal, so C is the correct answer.

Tool Knowledge (TK): is a pictorial test which requires the examinee to identify pictured tools and determine related items with which they are used.



Numerical Operations (NO): a speeded test requiring the working of simple arithmetic problems.

Sample Problem:

3 x 3

A. 1

B. 6

C. 9

D. 12

The answer is 9, so answer C is correct.

Mathematics Knowledge (MK): a test of ability to solve general mathematical problems.

Sample problem: The area of a rectangle 2 feet by 3 feet is equal to

A. 2 square feet.

B. 4 square feet.

C. 6 square feet.

D. 8 square feet.

The correct answer is 6 square feet, so C is the correct response.

General Science (GS): basic questions about biological and physical sciences.

Sample Question: A rose is a kind of

A. animal.

B. bird.

C. flower.

fish.

ɔre C is the right answer.

Classification Inventory (CI): A vocational interest measure based upon experience in, and preference for activities related to mechanical, electronic, clerical-administrative, and masculine/outdoor pursuits.

Attention to Detail (AD): A speeded test to count the number of c's embedded in a series of o's.

Instructions: This is the other speed test on the ASVAB that you will not finish, but you should work as quickly and as accurately as you can. This one is a test of your ability to find an important detail.

Look at this sample problem, S1:

c o o o o o o o o c o o o o o c o o o o o o c o o o c o o o o o c o o o o o o o
o c o o o o o o o c o c o o o o o o o o o c o o o o o o o c o o o o o o c o o o c o o

The two lines have a mixture of o's and c's.

You are to count the total number of c's in both lines of the problem.

Do this now, and you will find that there are 13 c's in the two lines. So 13 is the correct answer. After the Number S1 below are five numbers 11, 12, 13, 14, 15. The space under the 13 has been marked the way you would mark your answer sheet.

In the sample test below, count the number of c's in both lines of each problem. There may be 11, 12, 13, 14 or 15 c's. Blacken the space on your answer sheet that shows your choice as the correct answer for each problem.

Remember, on the ASVAB you should do this test as fast as you can without making mistakes.

Paragraph Comprehension (PC): requires the examinee to read a paragraph and answer questions about it.

Sample Question: The duty of the lighthouse keeper is to keep the light burning no matter what happens, so that ships will be warned of the presence of dangerous rocks. If a shipwreck should occur near the lighthouse, even though he would like to aid in the rescue of its crew and passengers, the lighthouse keeper must

A. stay at his light.

B. rush to their aid.

C. turn out the light.

D. quickly sound the siren.

The correct choice is A.

Sample Question: In certain areas water is so scarce that every attempt is made to conserve it. For instance, on one oasis in the Sahara Desert the amount of water to be given each date palm tree has been carefully determined.

How much water is each tree given?

A. no water at all

B. exactly the amount required

C. water only if it is healthy

D. water on alternate days

The correct choice is B.

Appendix B

Comparison of Conversion Tables Developed During Verification of Forms 8, 9, and 10 Calibration and the Operational Conversion Table

Table B - 1

Comparison of Conversion Tables Derived from Calibration Verification with the ASVAB-AFQT Operational Conversion for all Versions of Forms 8, 9, and 10

Operational Conversion Table		Calibration Verification Raw Scores						
Percentile	Raw Score	Average	8a	8b	9a	9b	10a	10b
99	105	-	-	-	-	-	-	-
98	104	-	-	-	-	-	-	-
97	103	-	-	-	105	-	-	-
96	-	105	105	105	104	105	-	-
95	102	104	104	104	103	104	105	105
94	-	103	103	103	102	103	104	104
93	101	102	102	102	101	102	103	103
92	-	101	101	-	100	101	102	102
91	100	100	100	101	-	100	101	101
90	99	-	99	100	99	99	100	100
89	-	99	-	99	98	98	99	99
88	98	98	98	97-98	97	97	98	98
87	97	97	97	-	96	96	97	97
86	96	96	-	-	-	-	-	-
85	95	-	96	96	95	-	96	96
84	-	-	-	-	-	-	-	-
83	94	95	95	95	94	95	95	95
82	93	94	94	94	93	94	94	94
81	-	93	93	93	92	93	93	93
80	92	92	92	92	91	92	92	92
79	-	-	-	-	-	-	-	-
78	91	91	91	91	90	91	91	91
77	-	-	-	-	-	-	-	-
76	90	90	90	90	89	90	90	90
75	-	-	-	-	-	-	-	-
74	89	89	89	89	88	89	89	89

Table B - 1 (Cont.)

Operational Conversion Table		Calibration Verification Raw Scores						
Percentile	Raw Score	Average	8a	8b	9a	9b	10a	10b
73	-	-	-	-	-	-	-	-
72	88	88	88	88	87	88	88	88
71	-	-	-	-	-	-	-	-
70	87	87	87	87	86	87	87	87
69	-	-	-	-	-	-	-	-
68	86	-	-	-	-	-	-	-
67	-	86	-	-	85	86	-	86
66	85	-	-	-	-	-	-	-
65	84	-	86	88	84	85	86	85
64	-	85	-	-	-	-	-	-
63	83	84	85	85	83	84	85	84
62	-	83	84	84	82	83	84	83
61	82	82	83	83	-	82	83	81-82
60	-	-	82	82	81	81	82	-
59	81	81	-	-	-	-	-	-
58	80	-	81	81	80	80	81	80
57	-	80	-	-	-	-	-	-
56	79	-	80	80	79	-	80	79
55	-	79	-	-	-	-	-	-
54	78	-	78	79	78	79	79	78
53	-	78	-	-	-	-	-	-
52	77	-	78	78	77	78	78	77
51	-	76-77	77	76-77	78	77	77	76
50	78	-	76	-	75	78	78	75
49	75	75	75	75	74	75	75	74
48	74	74	74	74	73	74	74	73

Table B - 1 (Cont.)

Operational Conversion Table		Calibration Verification Raw Scores						
Percentile	Raw Score	Average	8a	8b	9a	9b	10a	10b
47	-	-	-	-	-	-	-	-
46	73	73		73	72	73	73	72
45	-	-	-	-	-	-	-	-
44	72	72	73	72	71	72	72	71
43	-	-	-	-	-	-	-	-
42	71	71	72	71	70	71	71	70
41	-	-	-	-	-	-	-	-
40	70	70	71	70	69	70	70	69
39	-	-	-	-	-	-	-	-
38	69	69	70	-	68	69	69	-
37	-	-	-	-	-	-	-	-
36	68	68	69	69	67	68	68	68
35	-	-	-	-	-	-	-	-
34	67	67	68	68	66	67	67	67
33	66	66	67	67	65	66	66	66
32	-	65	68	66	64	65	65	65
31	65	64	65	65	-	64	64	64
30	64	63	64	64	63	63	63	63
29	63	62	63	62	62	-	62	62
28	62	-	62	62	61	62	-	61
27	-	61	-	61	60	61	61	60
26	61	60	61	60	59	60	60	59
25	60	59	60	59	58	59	59	58
24	59	58	59	58	57	58	58	57
23	58	57	58	57	56	57	57	56
22	57	56	57	56	55	56	56	55

Table B - 1 (Cont.)

Operational Conversion Table		Calibration Verification Raw Scores						
Percentile	Raw Seore	Average	8a	8b	9a	9b	10a	10b
21	56	55	56	55	54	55	55	-
20	54-55	54	55	54	53	54	54	54
19	53	53	54	53	52	53	53	53
18	52	51-52	52-53	52	50-51	51-52	51-52	51-52
17	51	49-50	50-51	50-51	48-49	49-50	49-50	49-50
16	49-50	47-48	49	48-49	47	47-48	47-48	47-48
15	47-48	45-46	47-48	46-47	45-46	45-46	46	45-46
14	45-46	43-44	45-46	44-45	43-44	43-44	44-45	43-44
13	43-44	42	43-44	42-43	41-42	41-42	42-43	42
12	41-42	40-41	41-42	40-41	39-40	39-40	40-41	40-41
11	40	38-39	39-40	38-39	37-38	38-38	38-39	38-39
10	38-39	36-37	38	36-37	35-36	-	36-37	36-37
9	36-37	34-35	36-37	35	33-34	34-35	34-35	34-34
8	34-35	32-33	34-35	33-34	31-32	32-33	32-33	33
7	32-33	31	32-33	31-32	29-30	30-31	31	31-32
6	30-31	29-30	30-31	29-30	28	28-29	29-30	29-30
5	28-29	27-28	28-29	27-28	26-27	26-27	27-28	27-28
4	26-27	25-26	26-27	25-26	24-25	24-25	25-26	25-26
3	24-25	21-24	24-25	22-24	21-23	22-23	21-24	23-24
2	23	19-20	21-23	19-21	18-20	18-21	18-20	20-22
1	22&below	18&below	20&below	18&below	17&below	18&below	17&below	19&below

Note. From Calibration of Armed Services Vocational Aptitude Battery Forms 8, 9 and 10 (AFHRL-TR-81-49) by M. J. Ree, J. J. Mathews, C. J. Mullins and R. H. Massey, 1982, Brooks AFB, TX: Air Force Human Resources Laboratory. Adapted by permission.

Appendix C

Statistical Data Pertinent to ASVAB Forms 8, 9, 10, and 11a

Intercorrelations

Distribution Statistics

Reliabilities

Item Statistics

Table C-1

Subtest Analysis of Form 8

Subtest	Number of Items	Mean	Standard Daviation	Skew	Kurtosis	KR 20 Reliability
	Form 8a: N = 2,620 Service applicants					
General Science (GS)	25	16.10	5.05	-.30	-.66	.84
Arithmetic Reasoning (AR)	30	17.82	7.13	.05	-1.08	.90
Word Knowledge (WK)	35	25.72	7.80	-.80	-.31	.82
Paragraph Comprehension (PC)	15	10.52	3.40	-.81	-.23	.80
Numerical Operations (NO)	50	35.35	10.28	-.45	-.38	-[a]
Coding Speed (CS)	84	42.64	15.15	-.16	-.02	-[a]
Auto-Shop Information (AS)	25	18.20	5.86	-.48	-.81	.88
Mathematics Knowledge (MK)	25	12.38	5.95	.41	-.75	.87
Mechanical Comprehension (MC)	25	15.50	5.81	-.32	-.82	.86
Electronics Information (El)	20	12.28	4.42	-.41	-.72	.83
	Form 8b: N = 2,510 Service applicants					
General Science (GS)	25	15.92	5.12	-.31	-.61	.85
Arithmetic Reasoning (AR)	30	18.52	7.41	.11	-1.10	.81
Word Knowledge (WK)	35	24.60	7.74	-.88	-.41	.92
Paragraph Comprehension (PC)	15	10.33	3.38	-.65	-.41	.80
Numerical Operations (NO)	50	35.77	10.14	-.63	-.01	-[a]
Coding Speed (CS)	84	43.04	15.41	-.19	-.01	-[a]
Auto-Shop Information (AS)	25	18.24	5.84	-.53	-.69	.88
Mathematics Knowledge (MK)	25	12.19	5.93	.49	-.75	.87
Mechanical Comprehension (MC)	25	15.24	5.68	-.27	-.81	.88
Electronics Information (El)	20	12.20	4.45	-.38	-.75	.83

Note. From Armed Services Vocational Aptitude Battery: Item and Factor Analysis of Forms 8, 9 and 10 (AFHRL-TR-81-55) by M. J. Ree, C. J. Mullins, J. J. Mathews and R. H. Massey, 1882, Brooks AFB, TX: Air Force Human Resources Laboratory. Reprinted by permission.

[a] Internal consistency reliability not computed for speeded tests.

Table C-2

Subtest Analysis of Form 9

Subtest	Number of Items	Mean	Standard Deviation	Skew	Kurtosis	KR 20 Reliability
	Form 9a: N = 2,590 Service applicants					
General Science (GS)	25	15.52	5.73	-.29	-.88	.88
Arithmetic Reasoning (AR)	30	18.22	7.32	.08	-1.09	.91
Word Knowledge (WK)	35	24.72	7.87	-.53	-.64	.92
Paragraph Comprehension (PC)	15	9.81	3.58	-.40	-.85	.81
Numerical Operations (NO)	50	35.04	10.70	-.62	-.18	-[a]
Coding Speed (CS)	84	42.78	15.22	-.17	-.13	-[a]
Auto-Shop Information (AS)	25	16.71	5.85	-.66	-.28	.89
Mathematics Knowledge (MK)	25	12.42	5.88	.43	-.63	.07
Mechanical Comprehension (MC)	25	15.29	5.51	-.34	-.82	.85
Electronics Information (EI)	20	12.65	4.26	-.37	-.41	.82
	Form 9b: N = 2,500 Service applicants					
General Science (GS)	25	15.49	5.70	-.25	-.91	.87
Arithmetic Reasoning (AR)	30	18.43	7.21	.03	-1.12	.91
Word Knowledge (WK)	35	24.83	7.89	-.67	-.52	.92
Paragraph Comprehension (PC)	15	10.41	3.33	-.74	-.18	.80
Numerical Operations (NO)	50	35.37	10.37	-.50	-.25	-[a]
Coding Speed (CS)	84	43.04	14.86	-.14	-.07	-[a]
Auto-Shop Information (AS)	25	16.75	5.73	-.52	-.50	.81
Mathematics Knowledge (MK)	25	12.27	0.02	.51	-.85	.88
Mechanical Comprehension (MC)	25	15.26	5.29	-.23	-.72	.84
Electronics Information (EI)	20	12.72	4.07	-.35	-.35	.81

Note. From Armed Services Vocational Aptitude Battery: Item and Factor Analysis of Forms 8, 9 and 10 (AFHRL-TR-81-55) by M. J. Ree, C. J. Mullins, J. J. Mathews and R. H. Massey, 1982, Brooks AFB, TX: Air Force Human Resources Laboratory. Reprinted by permission.

[a] Internal consistency reliability not computed for speeded tests.

Table C-3

Subtest Analysis of Form 10

Subtest	Number of Items	Mean	Standard Deviation	Skew	Kurtosis	KR 20 Reliability
	Form 10a: N = 2,480 Service applicants					
General Science (GS)	25	15.49	5.33	-.34	-.63	.88
Arithmetic Reasoning (AR)	30	19.12	6.97	-.17	-1.10	.90
Word Knowledge (WK)	35	24.20	8.09	-.38	-.87	.93
Paragraph Comprehension (PC)	15	10.10	3.86	-.51	-.81	.84
Numerical Operations (NO)	50	35.80	10.12	-.57	-.24	-[a]
Coding Speed (CS)	84	43.71	15.25	-.12	.01	-[a]
Auto-Shop Information (AS)	25	16.59	5.87	-.57	-.44	.87
Mathematics Knowledge (MK)	25	12.42	5.88	.43	-.63	.87
Mechanical Comprehension (MC)	25	13.35	5.65	.38	-.86	.86
Electronics Information (EI)	20	15.43	5.48	-.29	-.72	.88
	Form 10b N = 2,420 Service applicants					
General Science (GS)	25	15.46	5.43	-.35	-.70	.88
Arithmetic Reasoning (AR)	30	18.24	7.26	.08	-1.13	.91
Word Knowledge (WK)	35	24.41	7.90	-.53	-.73	.92
Paragraph Comprehension (PC)	15	10.61	3.24	-.69	-.32	.80
Numerical Operations (NO)	50	35.20	10.53	-.58	-.20	-[a]
Coding Speed (CS)	84	43.33	14.76	-.05	-.11	-[a]
Auto-Shop Information (AS)	25	16.66	5.89	-.53	-.50	.88
Mathematics Knowledge (MK)	25	13.32	5.89	.30	-.89	.87
Mechanical Comprehension (MC)	25	15.13	5.47	-.23	-.81	.85
Electronics Information (EI)	20	12.35	4.11	-.43	-.28	.80

Note. From Armed Services Vocational Aptitude Battery: Item and Factor Analysis of Forms 8, 9 and 10 (AFHRL-TR-81-55) by M. J. Ree, C. J. Mullins, J. J. Mathews and R. H. Massey, 1982, Brooks AFB, TX: Air Force Human Resources Laboratory. Reprinted by permission.

[a] Internal consistency reliability not computed for speeded tests.

Table C-4

Item Analytic Statistics for Forms 8a, 8b, and 9a

Subtest	Number of Items	Number of Items in Range: Difficulty (p) .25-.49	.50-.74	.75-.89	Discrimination (r) biserial .10-.29	.30-.59	.80-.99
		Form 8a: N = 2,820 Service applicants					
General Science (GS)	25	0	11	8	0	10	15
Arithmetic Reasoning (AR)	30	10	15	5	0	3	27
Word Knowledge (WK)	35	2	14	10	0	4	31
Paragraph Comprehension (PC)	15	1	8	6	0	2	13
Auto-Shop Information (AS)	25	3	17	5	0	4	21
Mathematics Knowledge (MK)	25	14	8	2	0	7	18
Mechanical Comprehension (MC)	25	4	18	5	0	9	18
Electronics Information (EI)	20	6	7	7	0	7	13
		Form 8b: N = 2,510 Service applicants					
General Science (GS)	25	7	10	8	0	9	18
Arithmetic Reasoning (AR)	30	7	18	5	0	6	24
Word Knowledge (WK)	35	5	14	18	0	4	31
Paragraph Comprehension (PC)	15	2	7	8	0	2	13
Auto-Shop Information (AS)	25	3	18	8	0	9	18
Mathematical Knowledge (MK)	25	15	9	1	0	7	18
Mechanical Comprehension (MC)	25	8	18	3	0	8	17
Electronics Information (EI)	20	7	7	8	1	5	14
		Form 8a: N = 2,590 Service applicants					
General Science (GS)	25	7	10	8	0	2	22
Arithmetic Reasoning (AR)	30	8	15	7	0	3	27
Word Knowledge (WK)	35	4	17	14	0	5	30
Paragraph Comprehension (PC)	15	3	7	5	0	2	13
Auto-Shop Information (AS)	25	4	14	7	0	4	21
Mathematics Knowledge (MK)	26	14	8	3	0	7	18
Mechanical Comprehension (MC)	25	5	15	5	0	11	14
Electronics Information (EI)	20	8	8	8	0	8	14

Table C-4 (cont.)

Subtest	Number of Items	Number of Items in Range: Difficulty (p) .25-.49	.50-.74	.75-.99	Discrimination (r) biserial .10-.29	.30-.59	.60-.99
		Form 9b: N = 2,500 Service applicants					
General Science (GS)	25	6	12	7	0	4	21
Arithmetic Reasoning (AR)	30	8	16	0	0	4	26
Word Knowledge (WK)	35	3	17	15	0	5	30
Paragraph Comprehension (PC)	15	2	6	7	0	3	12
Auto-Shop Information (AS)	25	4	14	7	0	4	21
Mathematics Knowledge (MK)	25	14	9	2	0	6	19
Mechanical Comprehension (MC)	25	4	17	4	0	8	17
Electronics Information (EI)	20	5	6	9	0	7	13
		Form 10a: N = 2,480 Service applicants					
General Science (GS)	25	8	11	6	0	7	18
Arithmetic Reasoning (AR)	30	6	17	7	0	8	22
Word Knowledge (WK)	35	7	14	14	0	0	35
Paragraph Comprehension (PC)	15	0	11	4	0	1	14
Auto-Shop Information (AS)	25	2	17	6	0	3	22
Mathematical Knowledge (MK)	25	2	17	3	0	3	15
Mechanical Comprehension (MC)	25	6	12	7	0	10	15
Electronics Information (EI)	20	6	6	8	1	5	14
		Form 10b: N = 2,420 Service applicants					
General Science (GS)	25	7	12	6	0	6	19
Arithmetic Reasoning (AR)	30	10	15	5	0	5	25
Word Knowledge (WK)	35	5	17	13	0	8	27
Paragraph Comprehension (PC)	15	3	3	9	0	2	13
Auto-Shop Information (AS)	25	2	17	6	0	4	21
Mathematics Knowledge (MK)	25	13	10	2	0	7	18
Mechanical Comprehension (MC)	25	7	13	5	0	10	15
Electronics Information (EI)	20	7	6	7	0	7	13

Note. From Armed Services Vocational Aptitude Battery: Item and Factor Analysis of Forms 8, 9 and 10 (AFHRL-TR-81-55) by M. J. Ree, C. J. Mullins, J. J. Mathews and R. H. Massey, 1982: Brooks AFB, TX: Air Force Human Resources Laboratory. Reprinted by permission.

Table C-5

Intercorrelation Matrix of Forms 8, 9 and 10 Subtests

Subtest	GS	AR	WK	PC	NO	CS	AS	MK	MC	EI
Form 8										
General Science (GS)	1.00	.71	.83	.74	.46	.43	.70	.65	.71	.78
Arithmetic Reasoning (AR)	.71	1.00	.70	.70	.59	.52	.60	.79	.69	.68
Word Knowledge (WK)	.83	.73	1.00	.82	.52	.48	.68	.52	.67	.76
Paragraph Comprehension (PC)	.75	.71	.81	1.00	.55	.49	.63	.60	.64	.69
Numerical Operations (NO)	.51	.64	.56	.55	1.06	.64	.40	.58	.45	.46
Coding Speed (CS)	.42	.51	.47	.48	.65	1.00	.42	.51	.45	.46
Auto-Shop Information (AS)	.68	.61	.65	.62	.43	.42	1.00	.52	.75	.79
Mathematics Knowledge (MK)	.63	.78	.62	.62	.57	.50	.53	1.00	.64	.61
Mechanical Comprehension (MC)	.71	.69	.67	.66	.47	.45	.78	.63	1.60	.75
Electronics Information (EI)	.76	.66	.74	.89	.45	.43	.78	.61	.75	1.00
Form 9										
General Science (GS)	1.00	.73	.86	.78	.47	.45	.73	.67	.73	.74
Arithmetic Reasoning (AR)	.71	1.00	.74	.74	.61	.56	.64	.81	.72	.65
Word Knowledge (WK)	.85	.74	1.00	.82	.51	.49	.71	.66	.71	.73
Paragraph Comprehension (PC)	.76	.70	.82	1.00	.53	.50	.66	.65	.69	.67
Numerical Operations (NO)	.52	.64	.57	.56	1.66	.87	.41	.55	.44	.41
Coding Speed (CS)	.45	.54	.50	.49	.66	1.00	.45	.52	.48	.45
Auto-Shop Information (AS)	.72	.63	.69	.65	.41	.43	1.00	.57	.77	.86
Mathematics Knowledge (MK)	.67	.81	.68	.61	.59	.51	.54	1.60	.89	.62
Mechanical Comprehension (MC)	.71	.71	.68	.64	.46	.45	.76	.67	1.00	.76
Electronics Information (EI)	.75	.64	.72	.66	.46	.44	.81	.59	.75	1.00
Form 10										
General Science (GS)	1.00	.74	.63	.78	.49	.46	.66	.70	.75	.74
Arithmetic Reasoning (AR)	.73	1.00	.75	.72	.60	.54	.61	.78	.73	.67
Word Knowledge (WK)	.83	.73	1.00	.83	.51	.56	.65	.68	.71	.73
Paragraph Comprehension (PC)	.78	.73	.81	1.00	.52	.52	.61	.65	.67	.67
Numerical Operations (NO)	.52	.65	.54	.57	1.00	.68	.37	.57	.45	.43
Coding Speed (CS)	.47	.50	.50	.53	.69	1.00	.42	.52	.46	.48
Auto-Shop Information (AS)	.68	.66	.68	.62	.41	.42	1.60	.51	.74	.77
Mathematics Knowledge (MK)	.71	.68	.66	.68	.80	.55	.55	1.00	.69	.83
Mechanical Comprehension (MC)	.75	.71	.71	.68	.48	.48	.73	.72	1.00	.77
Electronics Information (EI)	.75	.72	.72	.67	.47	.47	.75	.66	.75	1.00

Note. From Armed Services Vocational Aptitude Battery: Item and Factor Analysis of Forms 8, 9 and 10 (AFHRL-TR-81-55) by M. J. Ree, C. J. Mullins, J. J. Mathews and R. H. Massey, 1982, Brooks AFB, TX: Air Force Human Resources Laboratory. Reprinted by permission.

Table C-8

Subtest Analysis of Form 11a

Subtest	Number of Items	Mean	Standard Deviation	Skew	Kurtosis	KR 21[a] Reliability
Form 11a: N range (each subtest) = 17,198-26,373 Service applicants						
General Science (GS)	25	16.18	5.08	-.25	-.89	.81
Arithmetic Reasoning (AR)	30	18.90	6.92	-.18	-1.02	.88
Word Knowledge (WK)	35	25.33	7.08	-.53	-.47	.89
Paragraph Comprehension (PC)	15	11.02	3.08	-.58	-.42	.74
Numerical Operations (NO)	50	33.42	8.71	-.08	-.21	_[b]
Coding Speed (CS)	84	44.71	13.14	-.04	.24	_[b]
Auto-Shop Information (AS)	25	15.86	5.62	-.28	-.90	.85
Mathematics Knowledge (MK)	25	12.88	5.91	.40	-.83	.86
Mechanical Comprehension (MC)	25	15.48	4.99	-.22	-.76	.80
Electronics Information (EI)	20	11.70	4.12	.05	-.82	.75

Note. From The Development of ASVAB Forms 11, 12, and 13 by J. S. Prestwood, C. D. Vale and R. H. Massey (in press). Reprinted by permission.

[a] KR20 reliabilities were not available. KR21 reliabilities, a lower bound estimate of KR20 (Kudor & Richardson, 1937), were computed from the data concerning each subtest's mean, standard deviation and length.

[b] Internal consistency reliability not computed for speeded tests.

U.S. GOVERNMENT PRINTING OFFICE: 1984-450-684