AD-A094 684 CENTER FOR NAVAL ANALYSES ALEXANDRIA VA MARINE CORP--ETC F/G 5/9
A REEXAMINATION OF THE NORMALIZATION OF THE ARMED SERVICES VOCA--ETC(U)
APR 80 W H SIMS, A R TRUSS N00014-76-C-0001
UNCLASSIFIED CNS-1152 NL

1 OF 3
AD
A094684

AD A094684

(14) CNS-1152 (11) April 1980

(12)

(12) 212

LEVEL II

(9) Final Rept.,

(6)

A REEXAMINATION OF THE NORMALIZATION OF THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY (ASVAB) FORMS 6, 7, 6E, and 7E.

(10) William H. Sims, Study Director
Ann R. Truss

(15) N00014-76-C-0001

NOTICE

THIS DOCUMENT HAS BEEN APPROVED FOR PUBLIC RELEASE

CENTER FOR NAVAL ANALYSES

401327

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CNS 1152	2. GOVT ACCESSION NO. AD-A094	3. RECIPIENT'S CATALOG NUMBER 684
4. TITLE (and Subtitle) A Reexamination of the Normalization of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 6, 7, 6E, and 7E		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William H. Sims, Study Director Ann R. Truss		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Naval Analyses 2000 N. Beauregard Street Alexandria, Virginia 22311		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Deputy Chief of Staff (RD&S) Headquarters, Marine Corps Washington, D.C. 20380		12. REPORT DATE April 1980 13. NUMBER OF PAGES 193
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

The work reported here was conducted under the direction of the Center for Naval Analyses and represents the opinion of the Center for Naval Analyses at the time of issue. It does not necessarily represent the opinion of the Commandant, Marine Corps.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

AFQT (Armed Forces Qualification Test), aptitude tests, ASVAB (Armed Services Vocational Aptitude Battery), Marine Corps personnel, military operations, normalizing (statistics), personnel selection, ratings, recruiting, stratification, test methods

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study checks the normalization of the Armed Services Vocational Aptitude Battery (ASVAB) forms 6 and 7 and normalized the ASVAB forms 6E and 7E. The ASVAB measures the mental aptitude of prospective recruits. Since ASVAB 6 and 7 were first used (January 1976) questions about the correctness of the normalization have been raised. We checked the normalization of ASVAB 6 and 7 and developed the normalization of 6E and 7E using a reference test--AFQT 7. In this effort we used a large sample of Marine Corps recruits. We found that the current normalization of ASVAB 6 and 7 is too easy; it overstates the mental ability of low aptitude recruits by 15 to 17 percentiles.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DEPARTMENT OF THE NAVY
HEADQUARTERS UNITED STATES MARINE CORPS
WASHINGTON, D.C. 20380

IN REPLY REFER TO
RDS-41-eh
27 OCT 198

From: Commandant of the Marine Corps
To: Distribution List

Subj: Center for Naval Analyses Study CNS 1152, "A Reexamination of the Normalization of the Armed Services Vocational Aptitude Battery Forms 6, 7, 6E and 7E," Unclassified

Encl: (1) Subject study

1. The enclosure is the final report of a study of the normalization of the Armed Services Vocational Aptitude Battery (ASVAB).

2. The objective of the study was to examine the normalization of the ASVAB.

3. The objective of the study has been met and it is approved for distribution.

4. A copy of this letter will be affixed inside the front cover of each copy of the subject study prior to its distribution.

J W Moore
DEPUTY CH[illegible]

Distribution List Attached

Subj: Center for Naval Analyses Study 1152

DISTRIBUTION LIST

<u>Department of the Navy:</u>

21A	CINCLANTFLT, CINCPACFLT, CINCUSNAVEUR
24J	CG FMFLANT, CG FMFPAC
A1	ASST SECNAV (MRA&L)
A2A	NAVCOMPT, OPA, CNR
A5	CHNAVPERS
A6	C/S, USMC
A6	DC/S, Manpower (10 copies)
A6	DC/S, Plans and Policies
A6	DC/S, Research, Development and Studies (2 copies)
A6	DC/S, Installations and Logistics
A6	DC/S, Requirements and Programs
A6	DC/S, Operations and Training
A6	Director, C4 Systems, HqMC
B1	Ass't SecDef (MR&L)
B2A	Sec'y, Joint Chiefs of Staff
B3	National Defense University
B3	Armed Forces Staff College
FA34	Human Resource Management Center, Atlantic
FB44	Human Resource Management Center, Pacific
FF30	Naval Manpower & Material Analysis Center, Atlantic
FF30	Naval Manpower & Material Analysis Center, Pacific
FF38	U.S. Naval Academy (Nimitz Library)
FF44	Naval War College
FF48	Human Resource Management Center (OCNO)
FJ18	Navy Military Personnel Command
FJ76	Navy Recruiting Command
FKA6A16	Naval Personnel Research & Development Center
FT1	Chief of Naval Education and Training
FT73	Naval Postgraduate School
FT87	Human Resource Management School
V8	Marine Corps Recruit Depot, Parris Island
V8	Marine Corps Recruit Depot, San Diego
V12	Marine Corps Development and Education Command

OpNav: Op-09BH, Op-01

<u>Other</u>

Ass't Sec;y of Defense Program Analysis & Evaluation (2 copies)
Defense Technical Information Center
Department of the Army Library
Department of the Air Force (SAMI)
Institute for Defense Analyses
The Rand Corporation

CNS 1152 / April 1980

A REEXAMINATION OF THE NORMALIZATION OF THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY (ASVAB) FORMS 6, 7, 6E, and 7E

William H. Sims, Study Director
Ann R. Truss

Enclosure (1) to CMC ltr RDS-41-eh dated 27 October 1980

Marine Corps Operations Analysis Group

CENTER FOR NAVAL ANALYSES

2000 North Beauregard Street, Alexandria, Virginia 22311

EXECUTIVE SUMMARY

In January 1976, all branches of the armed services began using the Armed Services Vocational Aptitude Battery (ASVAB) to measure the mental aptitude of prospective recruits.

Since its first use there have been questions about whether the ASVAB had been correctly normalized--that is, whether the proper relationship had been established between the number of questions answered correctly (the raw score) and the percentile score. Some evidence suggested that the normalization[1] of the the ASVAB was too "easy"--that raw scores were being assigned percentile scores that were too high. In January 1979, the Assistant Secretary of Defense for Manpower, Reserve Affairs, and Logistics, requested that the Center for Naval Analyses (CNA) conduct an independent study of the normalization of the ASVAB.

Preliminary results from our analysis were made available[2] in May 1979. These results indicated that the current operational normalization of ASVAB was incorrect and that for this reason official reports on the mental aptitude of recruits were seriously in error. Because of the magnitude of the possible error and its attendent policy implications, the Department of Defense (DoD) set up two independent studies, one by DoD and the other by the Educational Testing Service (ETS), to determine if our results were correct. While these two additional studies were being conducted we carried out an extensive investigation of normalization methodology and further refined our results.

The data for our analyses was obtained by administering both the ASVAB and a reference test--Armed Forces Qualification Test (AFQT) 7A--to a large sample of Marine Corps recruits at two recruits depots. Testing was done under carefully controlled conditions designed to minimize any effects of test compromise and to provide equal motivation and opportunity to do well on both tests.

---

[1] Various authors use the terms "normalization", "calibration", or "equating" to describe the same procedure.

[2] Center for Naval Analyses, Memorandum (CNA)79-3059, "A Reexamination of the Normalization of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 6A, 7A, 6E, and 7E," by William H. Sims and Ann R. Truss, Unclassified, 30 May [illegible]. (This document was originally issued as a working paper.)

The findings of our study are summarized as follows:

- The current normalization of ASVAB is too easy; it overestimates the mental ability of low aptitude recruits by 15 to 17 percentiles. (For example, in figure I we show that a raw score of 31 converts to the 31st percentile by the current DoD norms but to only the 16th percentile by our norms.)

- Because the normalization has been incorrect, DoD reports have overstated the mental aptitude of recruits since January 1976. For the past 3 years, approximately 25 to 30 percent of all DoD accessions have been in mental category IV (the lowest acceptable category) rather than the 5 to 6 percent reported by DoD (see figure I).

- Although the mental quality of recruits enlisted since 1976 is lower than indicated by DoD reports, it is similar to that during the peak of the Vietnam War and better than that during the Korean War (see figure II).

- The analytical technique of sample stratification often used in the normalization of military aptitude tests will not, in general, produce correct results.

A correctly normalized test is important to managers as well as unit commanders and military trainers. The principal virtue of maintaining a correctly normalized test is that a certain score on a current version of the accession test reflects the same ability to absorb training as that same score did on previous versions of the test. Because of this continuity, managers can make informed judgments about changes over time in the aptitude of recruits. By the continued use of correctly normalized tests, a rational basis, founded on years of service experience in peace and in war, can be formed for both enlistment and job classification standards.

If the normalization of the ASVAB were changed to the one developed in this study, the supply of qualified applicants would probably decrease sharply unless compensating steps--such as a change in accession criteria, increased recruiting assets, increased enlistment incentives, or other actions--are taken.

RECOMMENDATION

We recommend that the conversion tables for the AFQT score shown in table I and for the classification composites shown in appendices M and O be used for the normalizaton of ASVAB 6, 7, 6E, and 7E.[1]

---

[1]We refer to these four tests as ASVAB 6/7/6E/7E.

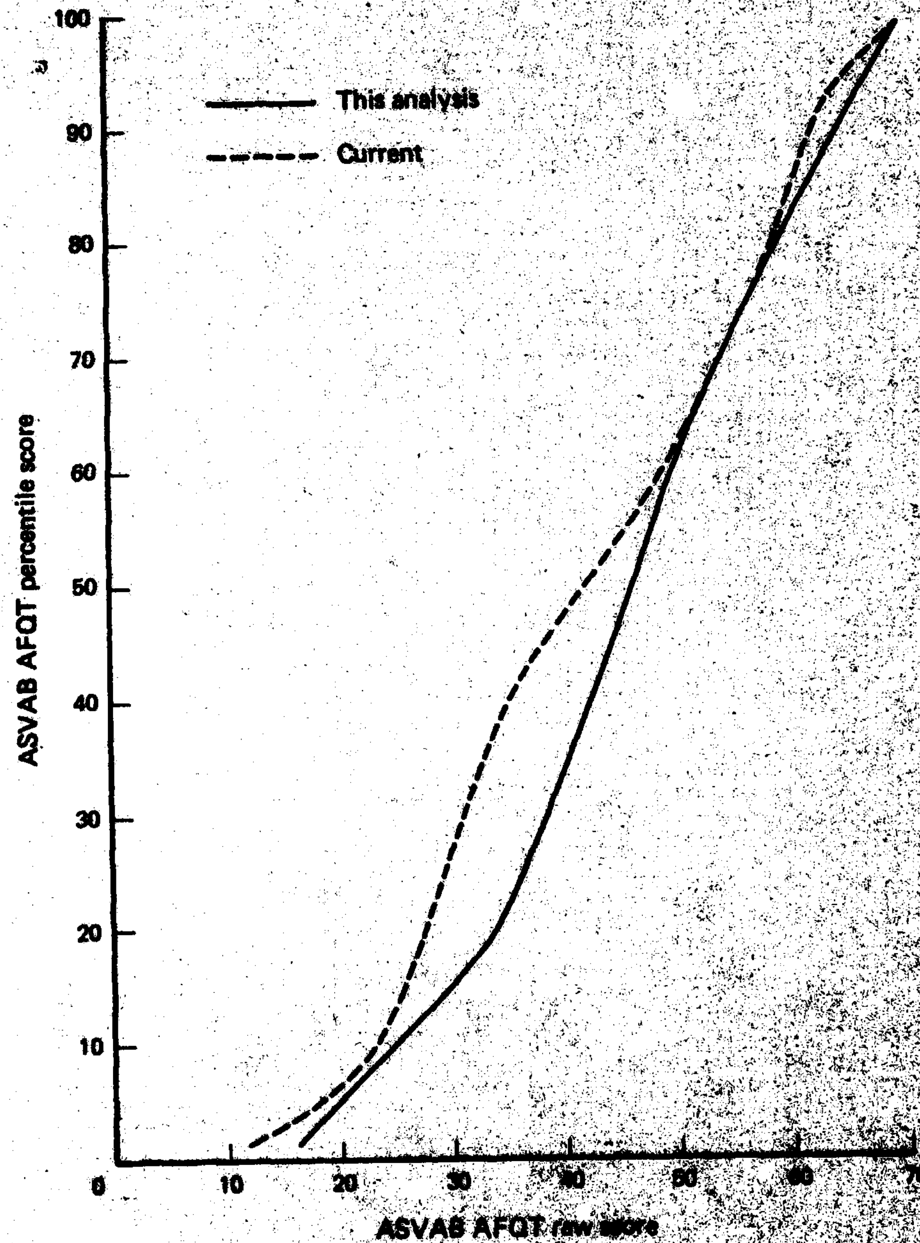


FIG. I: COMPARISON OF CURRENT OPERATIONAL ASVAB AFQT CONVERSION TABLE WITH THE ONE DEVELOPED IN THIS ANALYSIS

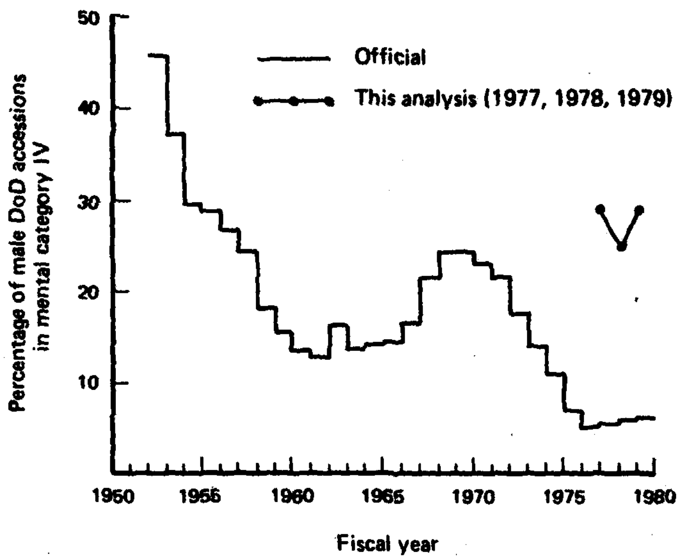


FIG. II: COMPARISON OF PERCENTAGE OF MALE DOD ACCESSIONS IN MENTAL GROUP IV AS OFFICIALLY REPORTED AND BY THIS ANALYSIS

TABLE I

RECOMMENDED CONVERSION TABLE[a] FOR
ASVAB 6/7/6E/7E AFQT SCORE

Raw score	Percentile score	Raw score	Percentile score
70	99	35	22
69	97	34	21
68	95	33	18
67	93	32	17
66	91	31	16
65	90	30	15
64	88	29	14
63	87	28	13
62	85	27	12
61	83	26	11
60	81	25	10
59	79	24	9
58	77	23	8
57	75	22	7
56	73	21	6
55	71	20	5
54	69	19	4
53	67	18	3
52	65	17	2
51	63	16	1
50	61	15	1
49	58	0-14	0
48	55		
47	52		
46	50		
45	48		
44	45		
43	42		
42	39		
41	36		
40	33		
39	31		
38	28		
37	26		
36	24		

[a]For form 7E only, add two raw score points to the AFQT raw score before using this table to convert raw score to percentile score.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont'd)

LIST OF ILLUSTRATIONS (Cont'd)

LIST OF ILLUSTRATIONS (Cont'd)

LIST OF TABLES

LIST OF TABLES (Cont'd)

LIST OF TABLES (Cont'd)

LIST OF TABLES (Cont'd)

LIST OF TABLES (Cont'd)

CHAPTER I

INTRODUCTION

BACKGROUND

The Armed Services Vocational Aptitude Battery (ASVAB) is the screening test the armed services currently use to measure the mental aptitude of prospective recruits. On 1 January 1976, two forms (6 and 7) developed by the Air Force Human Resources Laboratory (reference 1) were implemented at the Armed Forces Examining and Entrance Stations (AFEES). In this report we refer to these as ASVAB 6/7.

During the first months ASVAB 6/7 was used, an unexpectedly large number of recruits, particularly those in the Navy, were scoring high on the tests. This suggested that the normalization[1] of ASVAB 6/7 was too "easy."[2] Each of the armed services then initiated an independent analysis to examine the normalization of the test. Based on these analyses, the ASVAB Working Group[3] revised the normalization on 29 July 1976.

After the normalization of ASVAB 6/7 was revised, questions about the correctness of the revision continued among members of the ASVAB Working Group. In 1978, the Center for Naval Analyses (CNA) published a study (reference 2) that criticized the revised norms as unlikely to be correct.

In response to concern about test compromise, two additional forms, ASVAB 6E and ASVAB 7E, were scheduled for implementation in June 1979, making a total of four different forms of the test. To prepare for this implementation, CNA began a study to check the normalization of these two additional forms. The study was done at the request of the ASVAB Steering Committee,[4] through Headquarters, Marine Corps (reference 3). At about the same time that the ASVAB Steering Committee made its request, the Office of the Assistant Secretary of Defense for Manpower, Reserve Affairs, and Logistics (MRA&L) requested (reference 4) the study be

---

[1]Normalization as used here is a procedure that converts raw scores into percentile scores of a standard reference population.
[2]Easy means that a raw score has incorrectly been assigned a percentile score higher than would have been made by the proper percentage of the standard reference population.
[3]A joint service group that deals with ASVAB issues and is composed of policy and technical representatives from each service.
[4]The joint service flag officer oversight committee for the ASVAB Working Group.

expanded to include a reexamination of the norming of ASVAB 6/7. Accordingly, the study was designed to examine the normalization of the entire series--ASVAB 6, 7, 6E, and 7E. This report documents that analysis.

Because normalization information on ASVAB 6E and ASVAB 7E was needed before the scheduled June 1979 implementation, preliminary results (reference 5) of our analysis were made available in May 1979. These results showed that ASVAB 6E and ASVAB 7E could, with minor adjustments, use the same norming tables as ASVAB 6/7. The preliminary results also indicated that the normalization of the entire ASVAB series (6, 7, 6E, and 7E) was much too easy and that consequently there was a high probability that Department of Defense (DoD) reports of recruits' mental aptitude were seriously in error.

As a result of concerns raised by our preliminary report, two studies were conducted to try to verify our preliminary findings. One study was conducted by DoD[1] and used data on applicants tested at AFEES. The other study was conducted by the Educational Testing Service (ETS) and used data collected in high schools.

STRUCTURE OF THE ASVAB

The ASVAB is a group of 16 tests (sometimes referred to as subtests) that focus on different mental aptitudes. Scores from these tests are combined to form composite scores. The tests and composites are described in detail in appendix A. The Armed Forces Qualification Test (AFQT) composite is the common composite score all services use to measure general ability. Other composites are used primarily for job classification. For quality monitoring purposes, DoD reports scores of recruits in terms of broad categories known as mental groups. These mental groups are based on AFQT scores and range from I (most qualified) to V (unqualified). The normalization discussed in this report focuses on the normalization of the AFQT score, although normalizations of the other composites are also developed.

ORGANIZATION OF THE REPORT

Chapter II discusses the experimental design. In chapter III we develop our norming results from a stratified and unstratified

---

[1]This study, which initially was known as the Army Research Institute (ARI) study, is officially a Department of Defense study. It was conducted by an ARI research psychologist temporarily attached to the Office of the Assistant Secretary of Defense (MRA&L) with computational support from ARI. In this report we refer to it as the "DoD" study.

sample, respectively. In chapter IV we discuss possible problems with the results of the analysis. In chapter V we examine the effects of sample truncation, and in chapter VI we discuss whether samples should be stratified. Our recommended normalization, which is shown in chapter VII, is contrasted with alternative normalizations.

CHAPTER II

EXPERIMENTAL DESIGN

INTRODUCTION

We administered the various forms of the ASVAB along with a reference test to a large sample of Marine Corps recruits at recruit depots. We chose this approach over an administration at AFEES because it was much easier to obtain permission to give additional tests to Marine Corps recruits than to AFEES applicants. We understood that the use of a recruit sample instead of a more traditional applicant sample might cause added analytical difficulties. But, we believed these difficulties could be handled.

The testing was carried out under carefully controlled conditions designed to minimize any effects of test compromise and to provide equal motivation and opportunity for the recruits to do well on both the ASVAB[1] and the reference test. The ASVAB forms were normalized by equating ASVAB scores to scores on the reference test.

The reference test chosen for this analysis was AFQT 7A. It was used at AFEES from 1962 through 1973 and was normalized (see reference 6) to the traditional reference population according to a test known as "R-9." R-9 is an editorial revision of the Army General Classification Test used to define the World War II mobilization population.

The experimental design is discussed in detail in appendix B.

DATA SAMPLES

Our data sample consisted of test scores for 3,295 Marine Corps recruits. The tests were administered between 16 February and 3 April 1979 at the Marine Corps Recruit Depots (MCRD) located at Parris Island, South Carolina, and San Diego, California. Before enlistment, recruits took either ASVAB 6 or ASVAB 7 at the AFEES or ASVAB 5 (if they entered via the high school testing program). Once at the recruit depots, recruits took three tests: ASVAB 6 or ASVAB 7, ASVAB 6E or ASVAB 7E (AFQT parts only), and AFQT 7A (the reference test). Recruits were tested in platoon-size groups of about 60 persons.

---

[1]Definitions of ASVAB tests and composites are given in appendix A.

The order in which the three tests were given was counter-balanced; i.e., as many platoons were administered any one test first as were administered it second or third in the sequence. Total testing time was about 5 hours, and either a lunch break or overnight break separated the tests. All tests were given to recruits within a few days after their arrival at recruit depots and before they started recruit training.

In our analysis we used only results from tests administered at recruit depots. This reduced the effect of any coaching that may have occurred during testing at AFEES.

For our analysis, we separated the sample into five subsamples, as shown in table 1.

TABLE 1

DEFINITIONS OF SUBSAMPLES

Sample	Sample size	Used to normalize:	Tests used
1 (total)	3,295	Not used as a unit	AFQT 7A ASVAB 6 or 7 ASVAB 6E or 7E
2	1,634	6E	AFQT 7A ASVAB 6E
3	1,660	7E	AFQT 7A ASVAB 7E
4[a]	227	6/7[b]	AFQT 7A ASVAB 6/7
5[c]	2,208	6/7	AFQT 7A ASVAB 6/7

[a]These recruits were enlisted on the basis of scores on ASVAB 5, which they took in high school--they had not seen ASVAB 6/7 before being tested at recruit depots.

[b]Only as supporting evidence for sample 5 results.

[c]These recruits had been previously tested on ASVAB 6 or 7 at AFEES. When retested at the recruit depot they were given the opposite form to reduce the effect of practice; i.e., if they were tested at AFEES on form 6 they were given form 7 at the recruit depot and vice versa.

Samples 2 and 3 are appropriate for the normalization of ASVAB 6E and ASVAB 7E, respectively, because none of the recruits had seen AFQT 7A (the reference test) or ASVAB 6E or ASVAB 7E before being tested at the recruit depot.

Sample 5 is a good sample for the normalization of ASVAB 6/7, but has some imperfections. These imperfections result because the recruits were previously tested at AFEES on ASVAB 6 or 7. Effects from practice (taking the same test before) were eliminated because only recruits who were retested at recruit depots on the opposite form[1] of ASVAB 6/7 were included in sample 5. However, it is possible that some recruits were coached on both forms 6 and 7 before taking the test at AFEES. If this occurred and if they remembered this coaching when retested[2] at recruit depots, their scores on ASVAB 6/7 would artificially be raised. The resulting normalization would be too hard.

To control for the possibility just discussed, we used sample 4. This sample is small but very "clean."[3] All recruits in this sample were enlisted on the basis of scores on ASVAB 5, which they took in high school. That is, recruits in sample 4 had not seen ASVAB 6 or 7 before being tested at recruit depots; hence, results were not biased by a practice effect or by coaching. Our confidence in the results of our normalization for ASVAB 6/7 will be enhanced to the extent that the results from the larger sample 5 are confirmed by those of the small, but clean, sample 4.

Summary statistics for the entire data sample are given in appendix C.

---

[1]Those tested at AFEES on form 6 and at recruit depots on form 7 and vice versa.
[2]Retesting at recruit depots generally took place within 3 months of AFEES testing.
[3]Clean refers to tests on which recruits were not coached.

CHAPTER III

NORMALIZATION

INTRODUCTION

Two methods are commonly used in the normalization of military aptitude tests. We refer to one method as "unstratified graphical equating"--also known as "equipercentile equating." The other method may be called the "stratification procedure." This chapter describes normalization results obtained using both procedures.

STRATIFICATION PROCEDURE

The method of equating reported in this section was stratifying each subsample (table 1) on the AFQT 7A percentile score thereby simulating the reference mobilization population within each subsample. Cumulative frequency distributions of ASVAB 6, 7, 6E, and 7E AFQT raw scores were then made from this simulated mobilization population. The raw score-to-percentile score conversions for each form of ASVAB can be read directly from these cumulative frequency distributions.

For example, figure 1 shows the distribution of percentile scores from the reference test--AFQT 7A--in sample 5. The solid line is the distribution observed in the sample. The dashed line is that expected in the mobilization population.[1] The mobilization population is simulated in the sample by weighting individuals in the observed population in proportion to their expected occurrence in the mobilization population. The procedure is illustrated in table 2. For example, in the percentile interval 1 through 5, we observe 29 recruits. The mobilization population is expected to contain 110.4 in this interval. We calculated a weight factor, 3.807, which is the expected number divided by the observed number. We attached one of these weight factors to each recruit in the sample based on their score on the AFQT 7A reference test.

By using these weight factors we simulated the mobilization population within the sample. All distributions derived from these weighted recruits will look as they would if the mobilization population had taken the test. For example, if figure 1 were made using weighted recruits the distribution would be flat.

---

[1]The definition of percentile score is such that 5 percent of the reference population have a percentile score of five or less, 10 percent have a percentile score of 10 or less, and so on; hence, the expected distribution is flat.

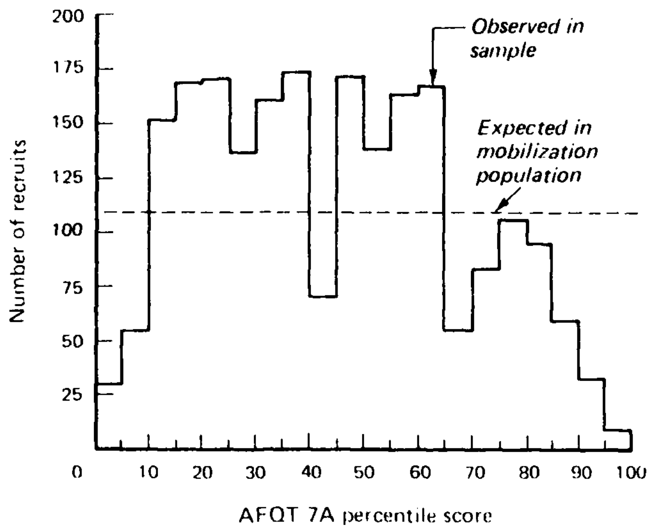


FIG. 1: COMPARISON OF PERCENTILE DISTRIBUTION OF SAMPLE 5 AND THE MOBILIZATION POPULATION

TABLE 2

CALCULATION OF WEIGHT FACTORS FOR SAMPLE 5

AFQT 7A percentile interval	Number observed in sample	Number expected in mobilization population	Weight factor[a]
(1)	(2)	(3)	(4)
1-5	29	110.4	3.807
6-10	56	110.4	1.971
11-15	152	110.4	0.726
16-20	169	110.4	0.653
21-25	171	110.4	0.646
26-30	137	110.4	0.806
31-35	162	110.4	0.681
36-40	175	110.4	0.631
41-45	71	110.4	1.555
46-50	173	110.4	0.638
51-55	138	110.4	0.800
56-60	164	110.4	0.673
61-65	167	110.4	0.661
66-70	56	110.4	1.971
71-75	84	110.4	1.314
76-80	107	110.4	1.032
81-85	96	110.4	1.150
86-90	59	110.4	1.871
91-95	33	110.4	3.345
96-100	9	110.4	12.267
Total	2,208	2,208	

[a] Column (3) divided by column (2).

To get the required conversion tables we simulated[1] the mobilization population in samples 2, 3, 4, and 5 and formed (using weighted recruits) the cumulative frequency distribution of raw AFQT scores for each form of the ASVAB.[2] The resulting tables for converting raw scores into percentile scores are tabulated in appendix D, table D-6.

EQUIPERCENTILE EQUATING

Normalization without stratification may be done by graphically equating the new test to a reference test. The procedure, known as equipercentile equating, is described in reference 7 and illustrated in figure 2. Two scores are considered equivalent if they are obtained by the same cumulative percentage of a sample (point "A" in figure 2). Hence, the raw score for the ASVAB test at point "B" would be defined as equal to the percentile score on the reference test at point "C".

ASVAB 6/7, 6E, and 7E were normalized using this procedure. Details and results are in appendix E.

DISCUSSION OF RESULTS

A comparison of the ASVAB 6/7 AFQT normalization results (conversion tables) from both the stratification and equipercentile procedures with the current operational conversion table is shown in figure 3. It shows, for example, that by the current operational norms, an ASVAB 6/7 raw score of 30 converts into a percentile score of 28. From the results of this analysis, the same raw score of 30 will convert into a percentile score of 15 using the equipercentile method, or 11, using the stratification procedure.

Figure 3 clearly shows that the results from the stratification procedure and the equipercentile method are systematically different in the highest and lowest percentiles. Because both methods have been used in the past to normalize military tests, it is important to understand the reasons for this difference and to determine which method is preferred. This question will be examined in detail in chapter VI.

It is also evident from figure 3 that regardless of which method of analysis is used, the results of this analysis strongly disagree with the current operational norms. Various services have established minimum acceptable AFQT percentile scores in the

---

[1]The calculation of the weight factors for each sample is shown in appendix D.

[2]Forms 6 and 7 are known to be similar (see references 1 and 2); hence, they are treated together.

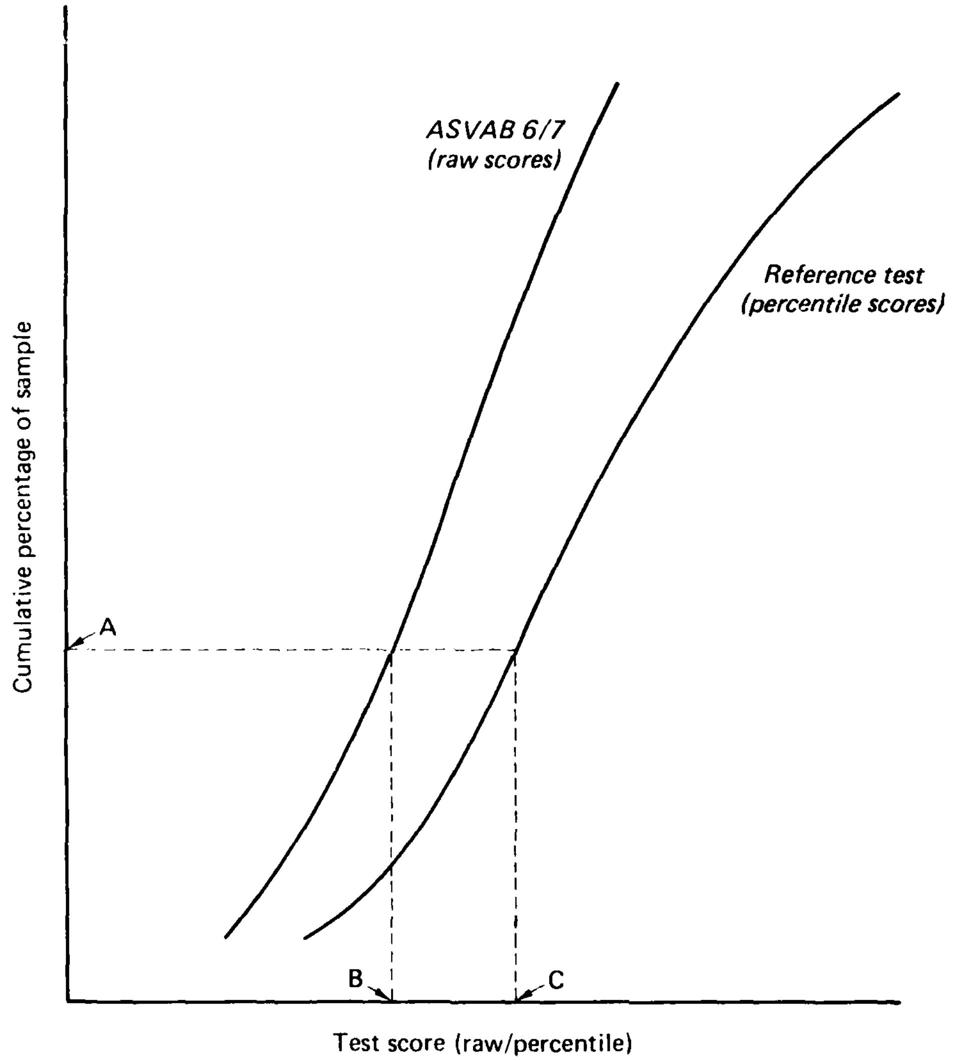


FIG. 2: ILLUSTRATION OF EQUIPERCENTILE EQUATING

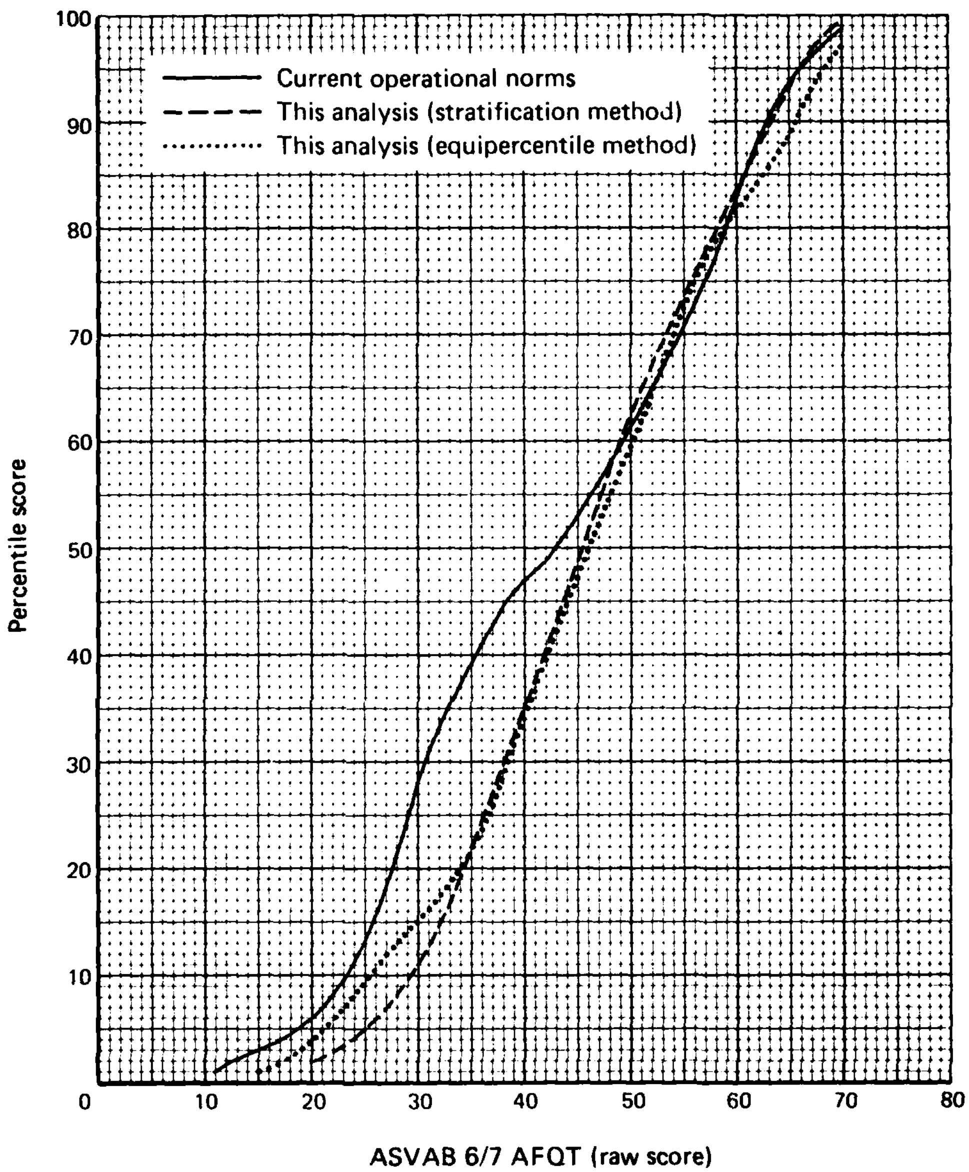


FIG. 3: COMPARISON OF OUR NORMALIZATION RESULTS FROM TWO METHODOLOGIES WITH CURRENT OPERATIONAL NORMALIZATION

area from the 16th through the 31st percentile. Our results differ from the current operational norms between these percentiles by 15 to 17 percentile points. If our norms are correct, a large percentage of current recruits would no longer be qualified for enlistment.

The seriousness of the potential error in current norms dictates that we must explore all avenues to determine if there are any flaws in our analysis. Recall that our tests were administered to Marine recruits because access was not possible to the more standard sample of applicants from all services. In chapter V we explore the question of whether this restriction could have produced a biased result.

Because ASVAB 6/7 is an operational test it is reasonable to assume that some recruits are coached on the answers. In chapter IV we examine our results for bias from this source and examine possible biases due to mistakes in administering the test and testing fatigue.

CHAPTER IV

COACHING AND ADMINISTRATIVE PROBLEMS

INTRODUCTION

We examined a number of areas in which problems could have biased our results. The areas we examined in detail are coaching, maladministration, and test fatigue. We discuss each of these areas in turn.

In our discussion of sources of possible bias, we base our conclusions on norming results using the stratification technique. Our tests for these biases are based on the observations (or lack thereof) of relative differences betwen norming results under various conditions. Hence, we believe that the conclusions reached in this chapter are insensitive to the particular normalization method used.

Coaching

As we have noted, one virtue of administering tests at recruit depots is that the effects of coaching will be smaller than at AFEES. Because the recruits are already enlisted, there is little reason for anyone to coach them on ASVAB. Moreover, most of those recruits who were enlisted based on ASVAB 6 scores are retested on ASVAB 7 and vice versa. Nonetheless, if recruits were coached on both ASVAB 6 and 7 before enlistment, they might recall enough material to bias ASVAB scores upward.

There are several ways to look at the coaching issue. For one, we compared the normalization results for ASVAB 6/7 found in sample 5 (which may be biased by coaching) with those from sample 4 (which cannot be biased by coaching). This comparison, shown in figure 4 is based on data in appendix D, table D-6. We see that results from the small, but clean, sample 4 agree very well with those from the full sample 5. This result suggests that coaching did not seriously bias the norming results for ASVAB 6/7.

For another approach to the coaching issue we removed from sample 5 those recruits who were most likely to have been coached. This procedure, which is discussed in detail in appendix F, relies on the Pseudo AFQT[1] developed by reference 2 specifically for detecting coaching.

---

[1]Pseudo AFQT = GI+GS+MC+MK,
where:
  GI = general information
  GS = general science
  MC = mechanical comprehension
  MK = mathematics knowledge.

AFQT = WK+AR+SP,
where:
  WK = word knowledge
  AR = arithmetic reasoning
  SP = space perception.

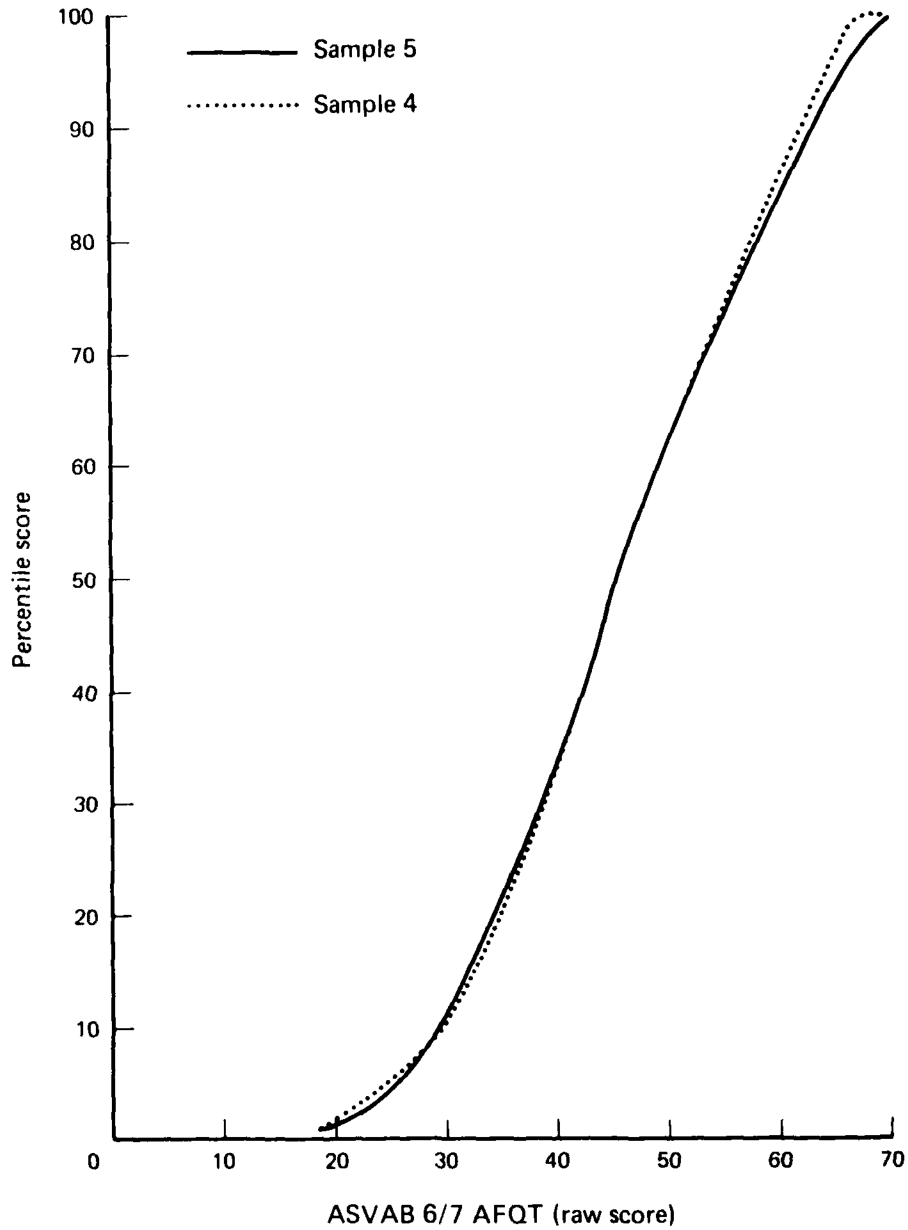


FIG. 4: COMPARISON OF ALTERNATIVE NORMALIZATIONS OF ASVAB 6/7

Referring to figure 5, we see the expected distribution in AFEES AFQT score minus the predicted AFEES AFQT score (Δ). For recruits who were not coached, this distribution is expected to be symmetric about zero. Recruits who were coached will tend to have AFEES AFQT scores higher than their predicted AFEES AFQT scores (i.e., positive values of Δ); these are shown in the shaded area of figure 5.

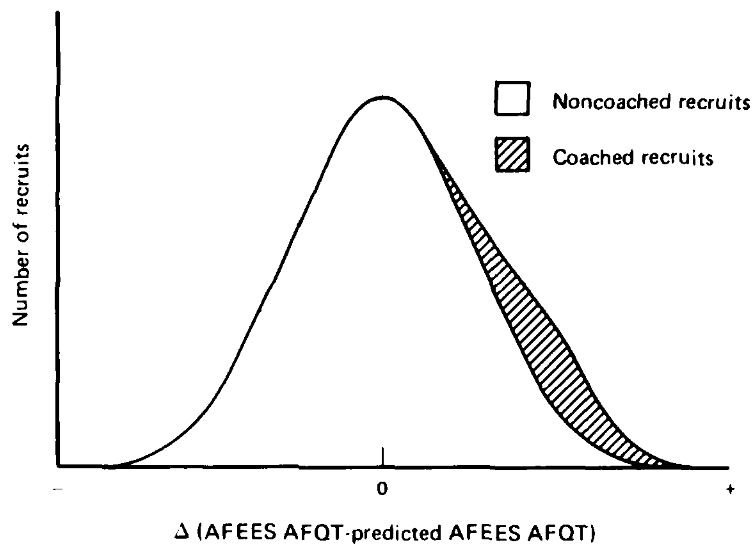


FIG. 5: ILLUSTRATION OF Δ DISTRIBUTION USED TO ESTIMATE AMOUNT OF COACHING

In appendix F we show results for removing suspect recruits from the sample. We removed suspects in two stages: the first is all recruits with $\Delta > 10$, and the second is all recruits with $\Delta > 0$. The two resulting subsamples were normalized and the results compared with those from the full sample 5. In both cases the resulting normalizations were statistically consistent[1] with the hypothesis that coaching does not distort the norming curve for ASVAB 6/7.

---

[1]Chi-squared tests for the homogeneity of parallel samples were applied to the data here and elsewhere in the report. This test is not, strictly speaking, appropriate because the samples are not completely independent, but it is useful as an approximate quantification of the homogeneity of the samples.

For a third approach to the coaching issue we compared scores on ASVAB 6/7 with those on ASVAB 6E/7E in an operational environment at AFEES. We found in this analysis that all forms of ASVAB can use the same conversion table. It has been argued that our results for forms 6/7 are biased by test compromise and that if this effect were removed then ASVAB 6/7 would have a markedly different conversion table than ASVAB 6E or 7E.[1]

Table 3 addresses this contention. It shows mean ASVAB AFQT scores for Marine Corps recruits for the months following the first use of ASVAB 6E/7E. Because ASVAB 6/7 has been used since January 1976 we assumed it was significantly compromised. ASVAB 6E/7E were first used in June 1979. We assumed they were not compromised much during the first few months of their use but have been compromised thereafter.[2]

Based on the preliminary results from our analysis (reference 5), when ASVAB 6E/7E was first used (with only minor adjustments) it would have had the same conversion table currently used for ASVAB 6/7. Note that the first line of table 3 indicates that the mean ASVAB AFQT percentile scores from both ASVAB 6/7 and ASVAB 6E/7E are identical (53.8). If one assumes that our norms for ASVAB 6/7 are grossly distorted by test compromise in our sample while the norms for ASVAB 6E/7E are not, then one should expect that over time as ASVAB 6E/7E becomes more compromised the mean score on ASVAB 6E/7E would become greater than that observed for ASVAB 6/7. But, as seen in table 3, this effect is not observed. We believe this result argues strongly that there is no significant bias in our ASVAB 6/7 norming results due to test compromise.

After examining all the material in this chapter and the details shown in appendix F, we concluded that there are indeed some cases in sample 5 that are probably distorted by coaching but that these do not seem to have had a significant effect on the normalization results for ASVAB 6/7.

1ASVAB 6/7 had been used for about 2 years when our data were collected and were certainly compromised. ASVAB 6E/7E were not in use then and were not compromised.

2In testimony before the House Armed Services Military Personnel Subcommittee, a recruiter stated that there wasn't a test devised "that I couldn't compromise in three months." (Navy Times, 7 June 1976). Other recruiters have given estimates of time required that are even shorter.

TABLE 3

COMPARISON OF MEAN ASVAB SCORES AFTER INTRODUCTION OF NEW FORMS

Dates tested at AFEES	Mean ASVAB AFQT percentile score from AFEES testing: ASVAB 6/7[a]	ASVAB 6E/7E[b]	Estimated extent of test compromise: ASVAB 6/7	ASVAB 6E/7E
June-September 1979	53.8 ± 0.2[c]	53.8 ± 0.4[d]	Significant	Negligible
October-December 1979	55.3 ± 0.3[e]	54.8 ± 0.5[f]	Significant	Significant

[a] Implemented in January 1976: assumed to be compromised as of June 1979.

[b] Implemented in June 1979: assumed to be relatively free of compromise during the first months of use (June-September 1979) and significantly compromised in later months (October-December 1979).

[c] Sample = 6,887 Marine Corps recruits.

[d] Sample = 1,755 Marine Corps recruits.

[e] Sample = 2,391 Marine Corps recruits.

[f] Sample = 1,096 Marine Corps recruits.

Maladministration

As part of the quality control procedure, one of us visited the two testing sites when testing began. Each test site seemed to conduct the testing in the same way. In appendix G we examine the data to determine if there is any difference in normalization results between the two test sites. Such a difference might indicate that at some time during the testing one of the sites may have deviated from the proper procedure.

We separated the data from sample 5 into two subsets, one from each test site, and did a separate normalization on each. Details are given in appendix G. The two resulting normalizations were very similar. (A chi-squared test indicated that the observed differences could well be due to chance.) We concluded that there was no reason to doubt that the two test sites followed the same testing procedures.

Test Fatigue

Because the sample design specified a counterbalanced series of three tests, it may be argued that norming results for the last test in the series may be biased due to recruits' fatigue. The counterbalanced design tends to reduce this problem. However, in appendix H we examine the test fatigue issue in some detail.

We selected a subsample of sample 5 consisting of those recruits who took the reference test and ASVAB 6/7 either first or second in the three-test sequence. We assumed these recruits would be less fatigued than the average recruit in sample 5. We stratified the low-fatigue subsample separately and developed a normalization curve. From our comparison of these results with those for all of sample 5 we concluded in appendix H that biases due to test fatigue, if any, are negligible.

In the next chapter we explore another source of bias due to truncation of the sample from preselection at the AFEES.

CHAPTER V

EFFECT OF TRUNCATION OF RECRUIT SAMPLE FROM PRESELECTION AT AFEES

INTRODUCTION

By necessity, we based our analysis on tests administered to recruits rather than the traditional sample of applicants for military service. Only those applicants who meet established minimums on the ASVAB are accepted for enlistment and become recruits. For this reason individuals who scored below these minimums were not present in our sample, and it may be argued that this biased our normalization results. However, our results are based on a reference test and an ASVAB both administered to recruits at recruit depots. Therefore, it may also be argued that any bias due to preselection at AFEES affects both tests equally and, in effect, the biases cancel each other out. We examine these questions in this chapter.

Our initial exploration of the truncation question involved further truncation of our data set followed by norming the residual sample to see if bias had been introduced. This analysis is described in detail in appendix I and suggests that some truncation bias may be present if the stratification method of norming is used. The limitations inherent in this already truncated data set precluded reaching a more definitive conclusion.

To fully address the effect of sample truncation on norms we used a full-range untruncated data set. First we developed norms from the full-range data set. Then we truncated this data set in the same way that our recruit data set was truncated and developed norms from the truncated data. A comparison of the norms developed from the full-range and truncated data enabled us to quantify the effect, if any, of truncation on norms.

An alternative approach would have involved using computer simulated data. We used real data to simulate the truncation rather than computer simulated data because there may be factors operating that we cannot know a priori, and hence cannot otherwise accurately simulate.

SIMULATION

We obtained a suitable full-range data set through the courtesy of the Office of the Assistant Secretary of Defense, who made available the data set[1] they were using in their study of

---

[1] This data set will be referred to as the "DoD" data set.

ASVAB norms. The data was collected at AFEES on a full-range of male applicants in June and July 1979. ASVAB forms and the reference test (AFQT 7A) were given in the counterbalanced fashion previously described. The sample sizes for the ASVAB 6/7, 6E, and 7E subsamples were 5,069, 2,870, and 2,650 cases, respectively. The data are described more fully in reference 8.

The concept behind the simulation of the truncation effect is illustrated in figure 6. The top panel of figure 6 illustrates the three-test CNA data set: the ASVAB 6/7 AFQT administered at AFEES, the ASVAB 6E AFQT administered at the Marine Corps Recruit Depots (MCRD), and the reference test also administered at the recruit depots. Note that scores on the ASVAB 6/7 AFQT administered at AFEES are directly selected (truncated) by virtue of the minimum enlistment standards at AFEES. Because scores on all three tests are highly correlated, this direct selection on the AFEES test results in an indirect selection that removed some low scoring individuals from the distribution of the two tests taken at the recruit depots.

The DoD data set consists of scores from only two (not three) separate test batteries. However, we simulated a three-test system by using the Pseudo AFQT developed in reference 2 (see chapter 4). Reference 2 found that in addition to the AFQT test embodied in the ASVAB, there is also a Pseudo AFQT. Because it is highly correlated (0.87) with the AFQT, the Pseudo AFQT is a good proxy for it.

The Pseudo AFQT may be constructed from parts of the ASVAB that do not make up the AFQT and, hence, may be viewed as a separate test. The Pseudo AFQT may be used to accurately predict an independent AFQT score for each applicant.

After simulating a three-test system we truncated the full-range sample on the predicted AFQT to simulate the truncation of the CNA data set. The real ASVAB AFQT and the reference test then show incidental selection similar to the one that occurred in the CNA data set (see the lower panel of figure 6). We normalized ASVAB 6E using those variables that are subject to incidental selection. Results were compared with those obtained from the nontruncated full-range data set.

The effects of the simulated truncation on the DoD ASVAB 6E data set and the comparison of them with the truncated CNA ASVAB 6E data set are shown in figure 7. The truncated distribution (open areas of figure 7) from the two data sets are very similar, indicating that we successfully simulated in the DoD data set a truncation like that observed in the CNA sample. The shaded areas of figure 7 represent individuals who were removed from the DoD full-range data set to simulate truncation like that observed

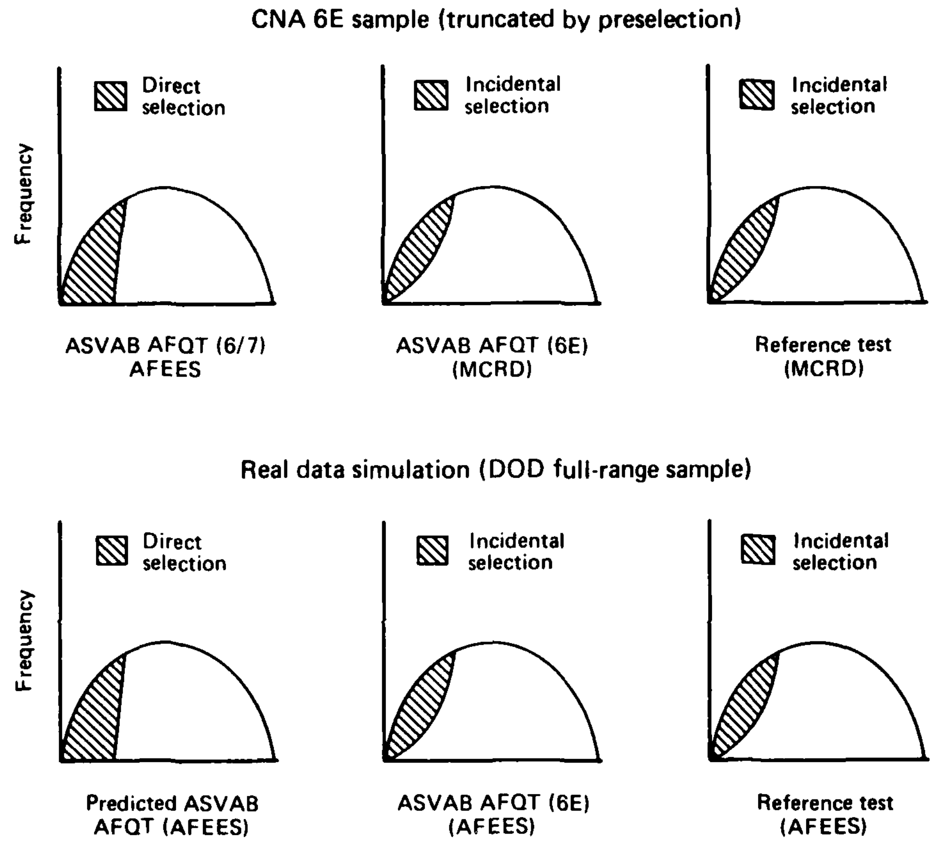
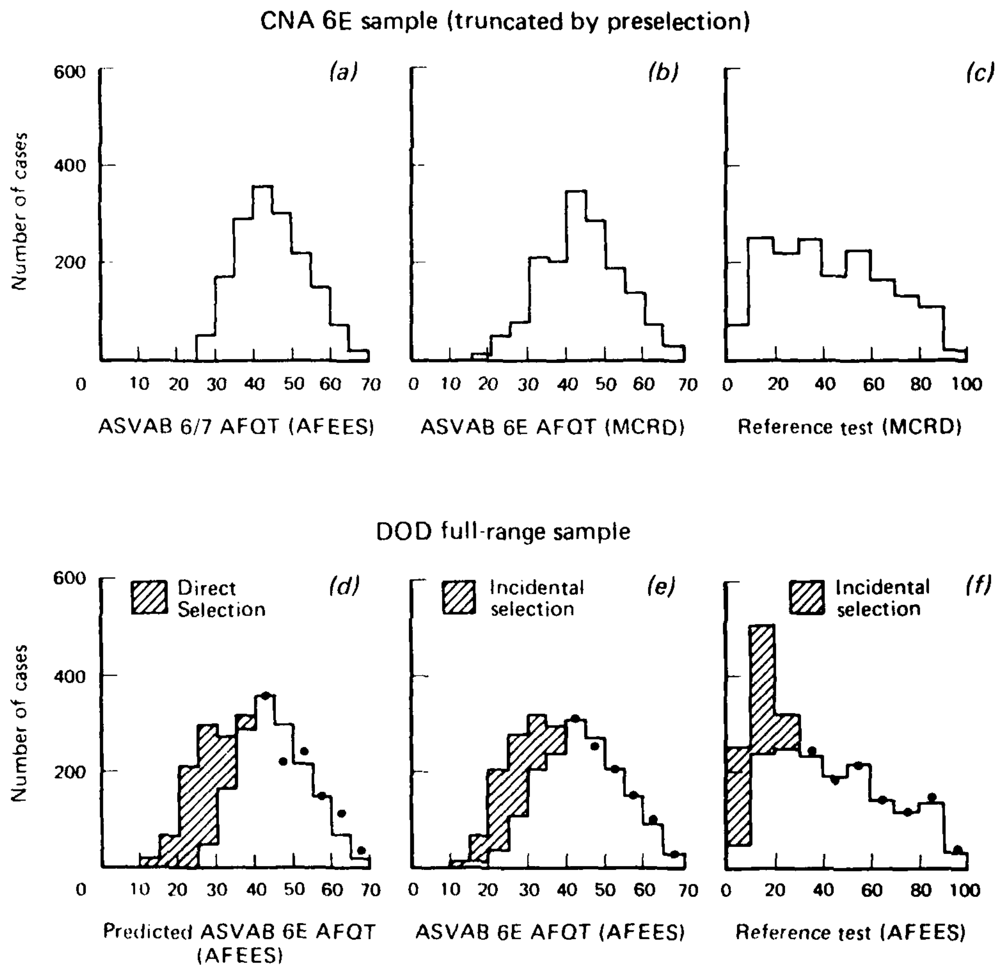


FIG. 6: ILLUSTRATION OF REAL DATA SIMULATION OF SAMPLE TRUNCATION



Note: Full-range distribution (dots plus cross-hatched area) was scaled (for illustration only) to the truncated distribution in the upper percentiles.

FIG. 7: COMPARISON OF TRUNCATED CNA DATA AND TRUNCATED DOD DATA FOR THE ASVAB 6E SAMPLE

in the CNA sample. The shaded areas of figure 7 represent individuals who were removed from the DoD full-range data set to simulate truncation like that observed in the CNA data set. Note that the direct removal of cases is carried out in figure 7(d) only--all other shaded areas represent cases that were removed by incidental selection. Further details on the simulation are given in appendix J.

Normalizations from the DoD full-range and DoD truncated data sets were made using both the stratification procedure and unstratified graphical equating. Differences between norming results from the DoD full-range and DoD truncated sample were taken as estimates of the distortions in CNA results due to the truncation effect.

STRATIFIED NORMING

Both the DoD full-range and simulated truncated data sets were normed using the stratification procedure. The samples were stratified on the reference test, and percentile equivalents of raw ASVAB AFQT scores were read directly from stratified cumulative frequencies. Details of the norming are given in appendix J.

The results from this norming for ASVAB 6E AFQT are shown in figure 8. That figure shows the comparison of norming results from the full-range and truncated DoD data set. The results from the truncated data set produce norms that are several points harder in the lower percentiles and somewhat easier in the upper percentiles. We obtained similar results (shown in appendix J) for the DoD ASVAB 6/7 and ASVAB 7E data sets.

UNSTRATIFIED NORMING

Both the DoD full-range and truncated samples were also normalized by the unstratified graphical equating (equipercentile) method. The details are given in appendix K. A comparison of the results from the full-range and truncated DoD ASVAB 6E AFQT data are shown in figure 9. The difference between the results for the truncated and full-range samples is very small and confined to the region below the 16th percentile. Similar results (shown in appendix K) were obtained for the DoD ASVAB 6/7 and 7E data sets.

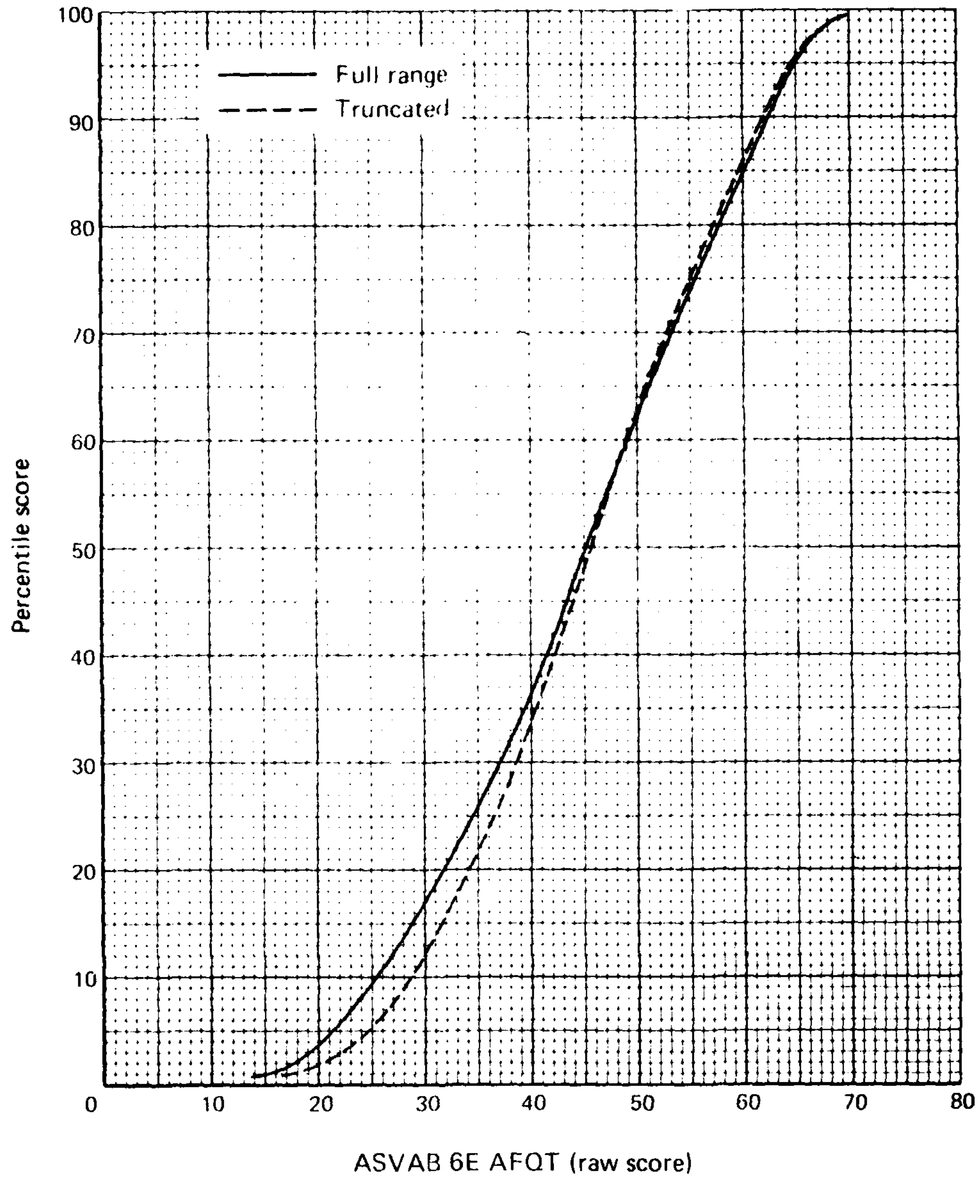


FIG. 8: EFFECT OF SIMULATED TRUNCATION ON STRATIFIED NORMING RESULTS FROM DOD GE SAMPLE

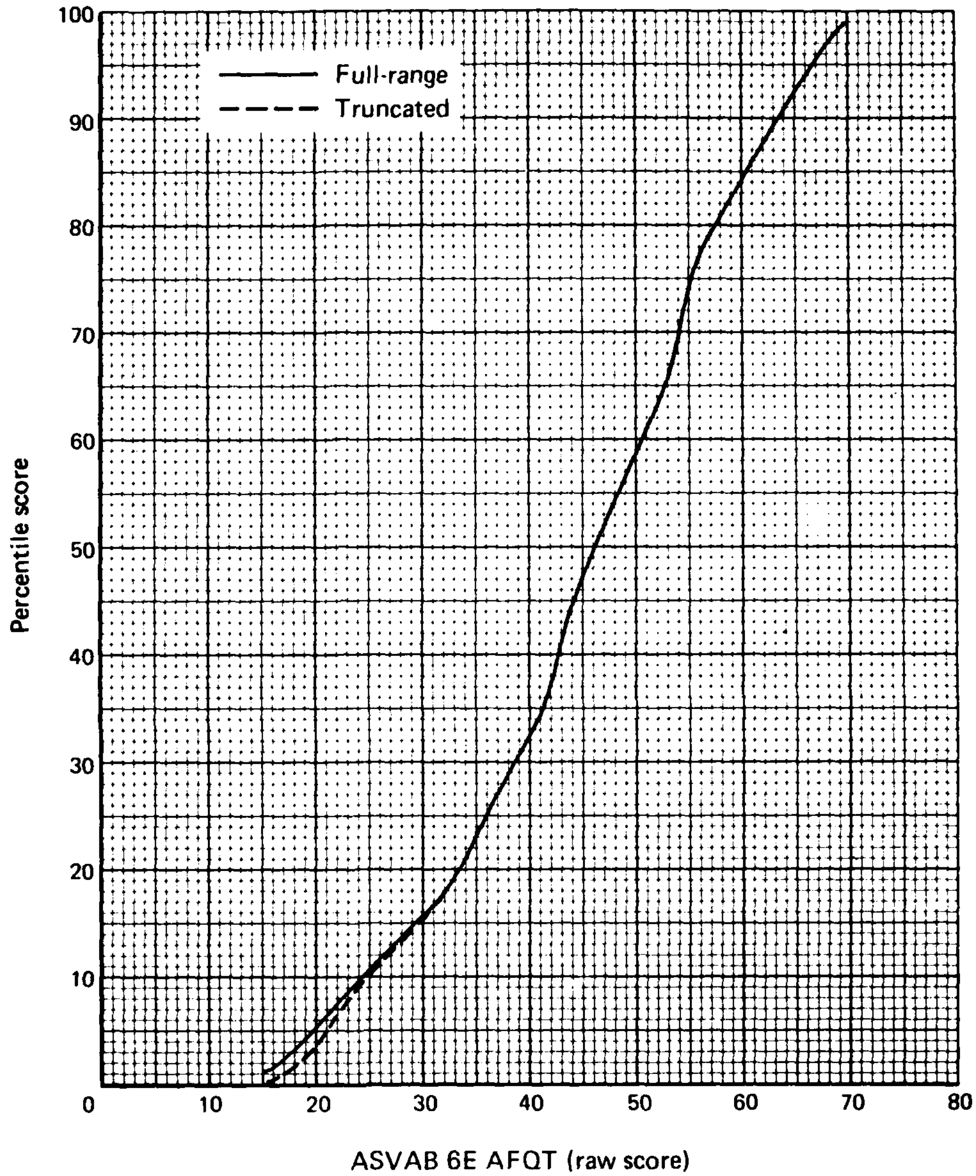


FIG. 9: COMPARISON OF ASVAB 6E NORMS FROM UNSTRATIFIED GRAPHICAL EQUATING USING FULL-RANGE DOD AND TRUNCATED DOD DATA

CHAPTER VI

TO STRATIFY OR NOT TO STRATIFY

INTRODUCTION

The stratification technique and the unstratified (equipercentile) technique each have some superficial advantages. The equipercentile technique has conceptual simplicity and is therefore intuitively appealing. However, this procedure, as carried out in this report does entail drawing and smoothing[1] graphs of cumulative frequencies and introduces some degree of subjectivity into the results. The stratified procedure superficially introduces a degree of stability in the normalization procedure by adjusting the sample so that the distribution of scores on the reference test is always flat. It also is a mechanistic procedure that introduces very little subjectivity into the normalization. The relevant criterion of whether to stratify is, however, which method produces the most accurate equating or normalization of tests. We explore this question in this chapter.

TRUNCATED DATA SETS

We have seen in figure 8 that the stratification procedure can produce significantly different results if applied to both full-range and indirectly truncated data sets. Ideally, the normalization results should be independent of the data set. For this reason, the stratification method should not be used with indirectly truncated data sets. An independent investigation[2] of the question reached the same conclusion.

Results using the unstratified graphical equating procedure were similar when applied to either full-range or indirectly truncated data sets (see figure 9). Invariance of the results with respect to truncation indicates that this method is satisfactory for indirectly truncated data sets.

---

[1]The procedure could be computerized.

[2]As this report was in final preparation, we received portions of a draft technical report based on computer simulated data stating that the "...stratified normalization technique introduces systematic biases in the estimation of population norms." Naval Personnel Research and Development Center, Draft Technical Note, "Test Norming and Equating Using Stratified Sampling: A Simulation Study," by John H. Wolfe, April 1980.

FULL-RANGE DATA SETS

We have shown that the stratified norming procedure can be unsatisfactory for indirectly truncated data sets. We now examine whether it is appropriate for full-range data sets.

In figure 10 we show a scattergram and associated projections of a typical unstratified bivariate distribution for two hypothetical parallel tests ("A" and "B") of equal difficulty.[1] The percentile distribution is peaked in the middle and depopulated on both ends. The distribution is similar to distributions of scores expected from applicant (or retested recruit) populations. The cell population, decile population, and cumulative percentage by decile are shown. Let us arbitrarily take test "A" as the reference test. Because the cumulative percentages by decile are the same for both tests, the unstratified graphical equating method would equate the 10th percentile on test "B" to the 10th percentile of the reference test "A", as we would expect.

Suppose we stratified the data in figure 10 so that the percentile distribution of the reference test, test "A", was flat. This procedure is illustrated in figure 10, and the weights necessary to force the test "A" distribution to be flat are shown.

The data set is stratified by applying the weights shown in figure 10. Results are shown in figure 11. The distribution of reference test scores is flat, as expected. However, the stratification procedure has only partially flattened the corresponding distribution of test "B" scores. The test "B" distribution is still too high in the area where it was originally high and too low in areas where it was originally low. A comparison of the resulting cumulative percentages by decile shows that scores on test "B" that should have been equated to the 10th percentile will actually be assigned to the 6th percentile. Similar distortions are observed in other score regions.

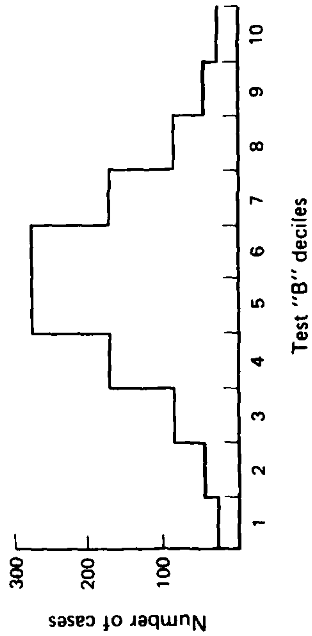
Figure 12 shows the results observed from applying the stratification procedure to our hypothetical bivariate data. The figure shows the comparison of results to the true normalization of test "B". As seen, the stratification procedure produces norms that are too hard in the lower percentiles and too easy in the upper percentiles.

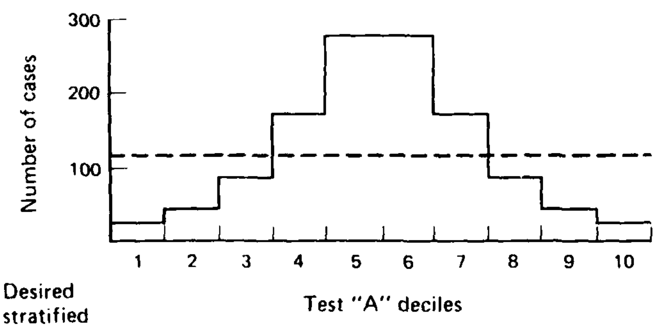
---

[1] The following example explains how to interpret figure 10. The scattergram projections of percentile scores are grouped into decile units. There are 30 cases in the first decile on test "A". These 30 cases are distributed on test "B" in the lower four deciles--10 in the first and second, 5 in the third and fourth. The cumulative percentage of the sample in test "A" is 2 percent in the first decile, 6 percent in the second, and so on.

						5	5	10	10	30	100
					5	10	10	10	10	45	98
				10	40	10	10	10	5	85	94
			20	40	40	40	10	10	5	165	87
		10	40	70	70	40	40	5		275	73
	5	40	40	70	70	40	10			275	50
5	10	10	40	40	40	20				165	27
5	10	10	10	40	10					85	13
10	10	10	10	5						45	6
10	10	5	5							30	2

Decile population	30	45	85	165	275	275	165	85	45	30
Cumulative percentage	2	6	13	27	50	73	87	94	98	100



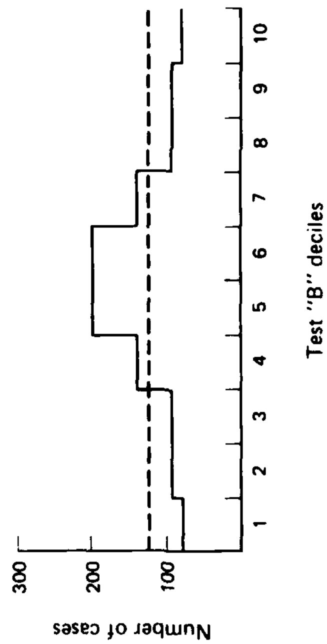


Desired stratified population	120	120	120	120	120	120	120	120	120	120
Weight	4.00	2.67	1.41	0.73	0.44	0.44	0.73	1.41	2.67	4.00

FIG. 10: ILLUSTRATION OF UNSTRATIFIED BIVARIATE DISTRIBUTION

						4	7	27	40	78	100
					2	7	14	27	40	90	94
				4	18	7	14	27	20	90	86
			15	18	18	29	14	27	20	141	79
		14	29	31	31	29	56	13		203	67
	13	56	29	31	31	29	14			203	50
20	27	14	29	18	18	15				141	33
20	27	14	7	18	4					90	22
40	27	14	7	2						90	14
40	27	7	4							78	6

Decile population	120	121	119	120	122	122	120	119	121	120
Cumulative percentage	10	20	30	40	50	60	70	80	90	100



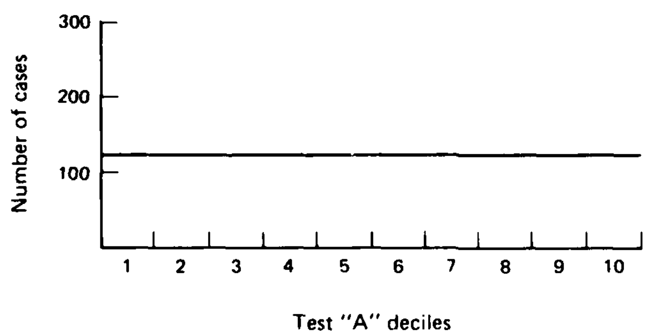


FIG. 11: ILLUSTRATION OF STRATIFIED BIVARIATE DISTRIBUTION

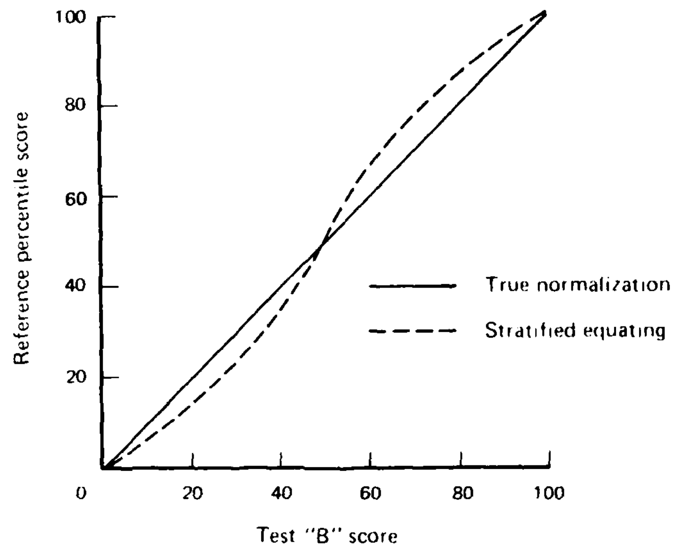


FIG. 12: COMPARISON OF STRATIFIED NORMS WITH TRUE NORMS FOR HYPOTHETICAL BIVARIATE DISTRIBUTION

Figure 13 shows a comparison of the results from stratified and unstratified equating of ASVAB 6E from the full-range DoD data sample (appendices J and K). This figure shows, as expected, that the stratified procedure produces harder norms in the lower percentiles and easier norms in the upper percentiles. The crossover point is very low in this case because the DoD data set has a reference percentile distribution peaked in the second decile (see figure 7) rather than in the fifth decile, as we assumed in our hypothetical example.

From the preceding discussion, we concluded that, in general, stratification is not appropriate for either truncated or full-range data samples. Such normalizations generally produce norms that are too hard in the lower percentiles and too easy in the upper percentiles.[1]

---

[1]This generalization will hold for all distributions of the form shown in figure 10; i.e., those that have a single maximum somewhere between the endpoints and where the endpoints tend to be depopulated.

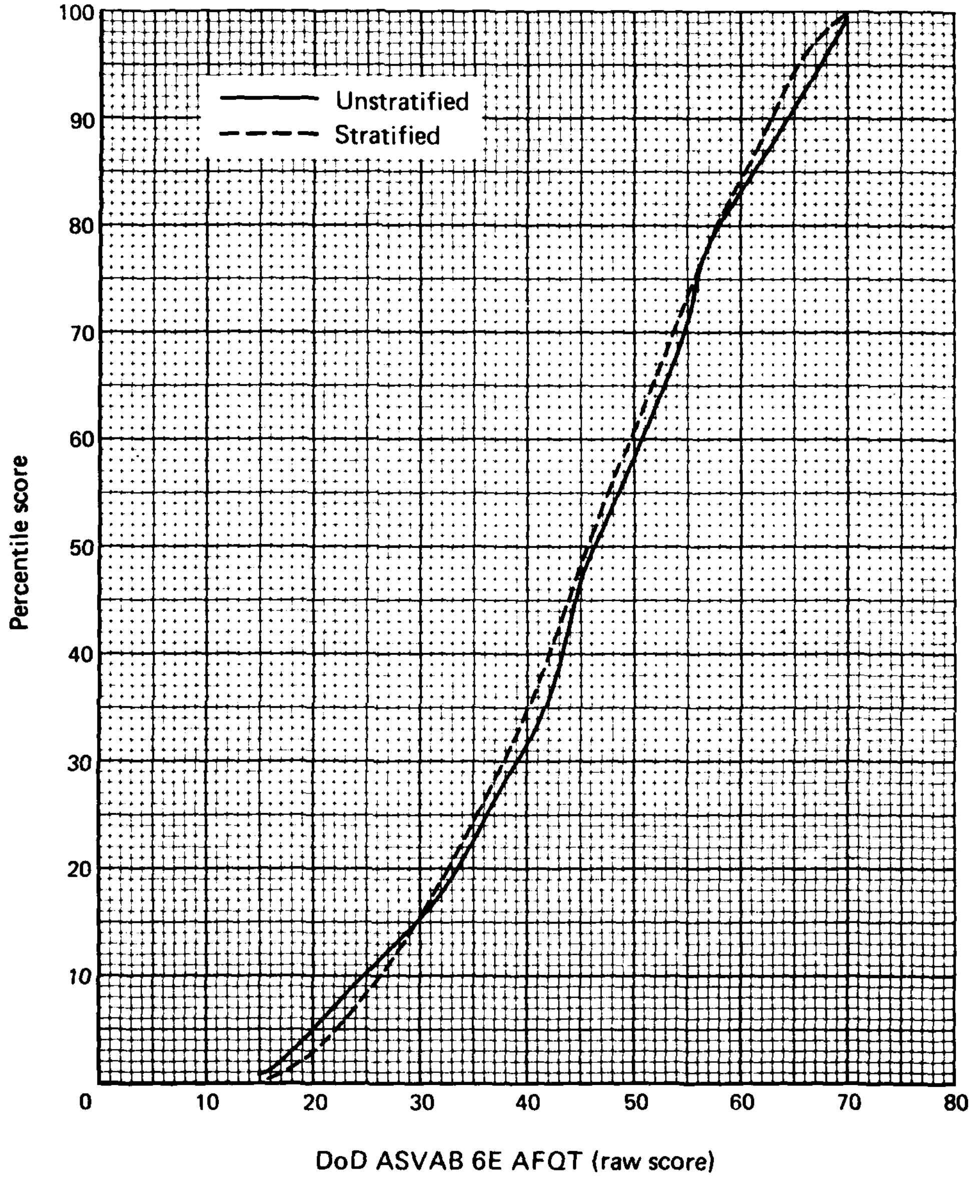


FIG. 13: COMPARISON OF NORMING RESULTS FOR FULL-RANGE DOD ASVAB 6E DATA SET USING STRATIFIED AND UNSTRATIFIED PROCEDURE

The distortion produced by stratification will be a function of the shape of the unstratified test score distributions. In general, different population groups will have different unstratified test score distributions. For this reason, stratified norming on different educational, sex, or racial groups may produce results that "show" norming bias, even if there is none.

We believe there is only one circumstance in which stratification for norming may be acceptable: where there is no reference test for direct equating. For example, in the ASVAB, the AFQT parts can be equated directly to an AFQT reference test. For the ASVAB composites and subtests this is not the case because many of them have no direct counterparts in the reference. In this case, stratification of the sample on the AFQT score and norming by cumulative frequencies may be acceptable. It should be recognized, however, that the procedure may be biased, as indicated in figure 12.

CHAPTER VII

RECOMMENDED NORMALIZATION OF ASVAB 6/7/6E/7E

ASVAB AFQT CONVERSION TABLE

In chapter VI we showed that the unstratified equating procedure is the preferred methodology for our data sample. We also showed that there will be a very small bias of 0.5 to 2.3 percentile points in the resulting norm curve below the 16th[1] percentile (see figure 9). A correction is applied for this bias and the resulting normalization curve is smoothed, as detailed in appendix L. The final smoothed set of conversion tables for ASVAB 6/7/6E/7E AFQT scores is shown in table 4.

ASVAB COMPOSITES CONVERSION TABLES

The ASVAB AFQT score is used as an overall measure of general trainability. However, to select individuals for specific military job assignments the services frequently use specific aptitude composites derived from the ASVAB. These composites are defined in appendix A and cover electrical, clerical, mechanical, and other specialties.

In order to maintain continuity of classification prequisites these composites are normalized so that their score scale is compatible with the AFQT score scale. We accomplish this by equating each composite score (in raw score form) to the ASVAB AFQT score using unstratified equipercentile equating. This approach is possible because the composites are strongly correlated to the ASVAB AFQT score.[2] Additional details and the composite conversion tables are given in appendix M.

ASVAB SUBTESTS CONVERSION TABLES

The Navy alone uses information from ASVAB subtests expressed in standard score form. The subtests are expressed in standard score form by first stratifying the sample on the ASVAB AFQT

---

[1] Because no service allows enlistments below the 16th percentile this bias has little practical significance.

[2] Alternately we could have stratified the sample on either the reference test or the ASVAB AFQT and formed cumulative frequencies of composite scores from which composite conversion tables could be constructed. We did not use this procedure due to concern about bias from stratification.

TABLE 4

RECOMMENDED CONVERSION TABLE[a] FOR
ASVAB 6/7/6E/7E AFQT SCORE

Raw score	Percentile score	Raw score	Percentile score
70	99	35	22
69	97	34	21
68	95	33	18
67	93	32	17
66	91	31	16
65	90	30	15
64	88	29	14
63	87	28	13
62	85	27	12
61	83	26	11
60	81	25	10
59	79	24	9
58	77	23	8
57	75	22	7
56	73	21	6
55	71	20	5
54	69	19	4
53	67	18	3
52	65	17	2
51	63	16	1
50	61	15	1
49	58	0-14	0
48	55		
47	52		
46	50		
45	48		
44	45		
43	42		
42	39		
41	36		
40	33		
39	31		
38	28		
37	26		
36	24		

[a]For form 7E only, two raw score points are to be added to the AFQT raw score before using this table to convert raw score to percentile score.

score and computing the mean values and standard deviations of raw scores on each subtest in the ASVAB.[1] Details of the stratification are given in appendix N. The subtest standard score conversions are then computed from the following equation:

$$\text{Standard score} = 50 + \frac{10(X_i - \overline{X})}{\sigma_x} ,$$

where:

X_i = the ith raw score of subtest X

$\overline{X}$ = the mean raw score of subtest X

σ_x = the standard deviation of raw scores on subtest X.

Resulting conversion tables are given in appendix O. Correlations and sample statistics from the stratified sample are given in appendix P.

VALIDITY OF RESULTS

We compared the norming results tabulated in table 4 with results from the DoD analysis (reference 8). The comparison is shown in figure 14. The agreement is excellent except in the upper percentiles. There the DoD results are from the stratified normalization method and differ in the expected direction.[2] In chapter VI we showed that this method generally leads to norms that are too easy in the upper percentiles. In spite of this shortcoming in the DoD study, the generally excellent agreement of these two independent studies argues strongly for the correctness of the results.

---

[1] In this case there is no viable alternative to stratification because no highly correlated reference subtests exist for equipercentile equating.

[2] The DoD results were obtained using a stratified technique in the upper percentiles and unstratified equating in the lower percentiles.

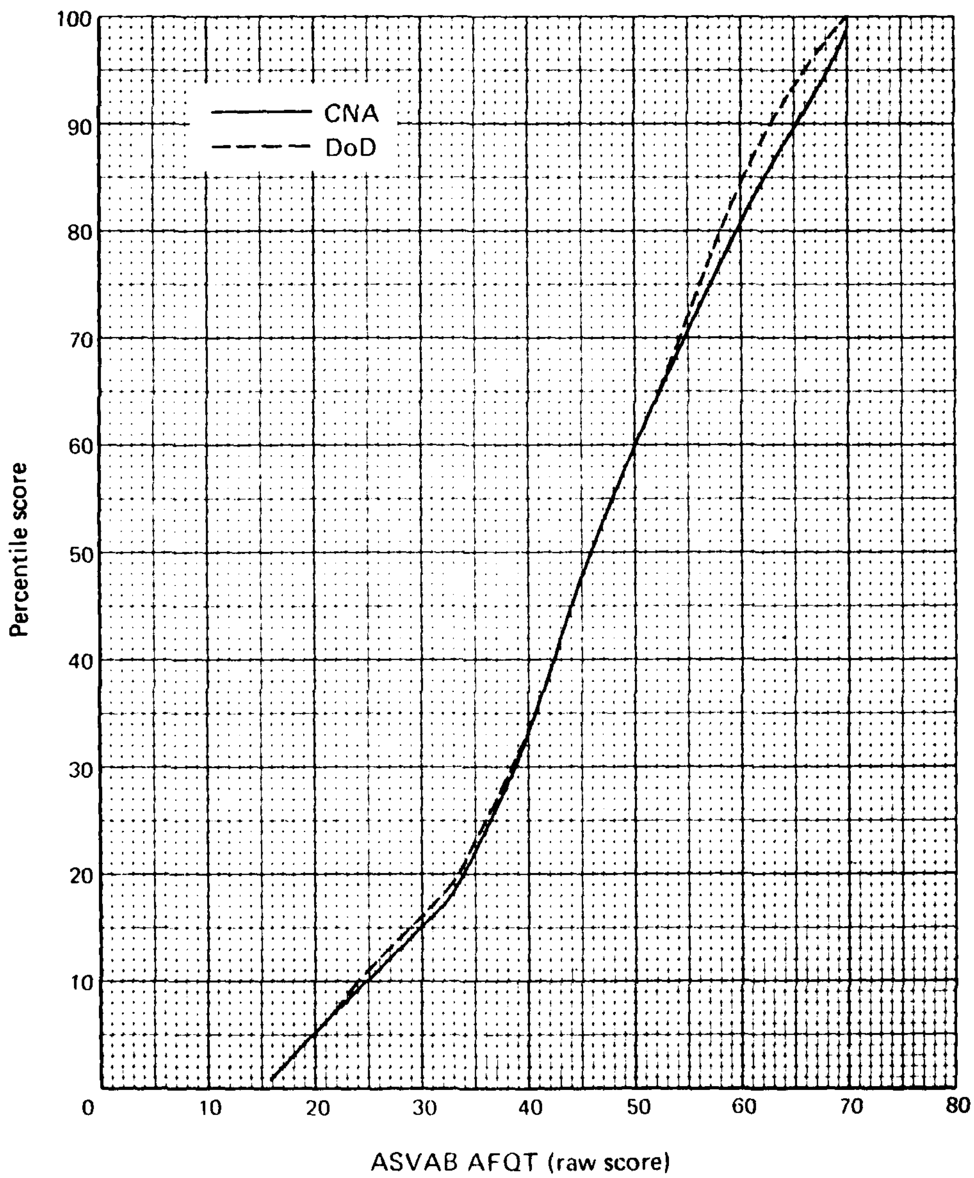


FIG. 14: COMPARISON OF CNA AND DOD NORMING RESULTS FOR ASVAB 6/7/6E/7E

Recall that preliminary results from our analysis (reference 5) were made available in May 1979. These results were based on the stratification methodology. We compared our preliminary and final results (figure 15). The differences are not great in the critical percentiles where enlistment decisions are made. At the 31st percentile, which is the breakpoint between mental categories IIIB and IV, the two results are identical. This agreement means that our preliminary estimates of the percentage of recruits in mental category IV were valid.

We compared the final results of this analysis with those from a 1978 CNA analysis (reference 2) and the current operational norms (figure 16). We see that in the critical percentiles (16th through 31st) the 1978 CNA study is closer to the correct norms (as represented by this analysis) than are the current operational norms. Nonetheless, the overall agreement of the 1978 CNA results with our current analysis is not good.

The 1978 CNA analysis was based on the best data available at the time. However, the data was a "sample of convenience" collected for other purposes by non-CNA personnel in 1970, 1974, and 1976. In addition, these data were analyzed using the stratification procedure, which we have shown is inappropriate. In contrast, the data for the current CNA analysis was collected under our supervision specifically for normalization purposes and utilized a sampling plan we designed. It was then analyzed using appropriate methodology. However, the most definitive test of correctness is reproducibility. No analysis has ever reproduced the results of the 1978 CNA analysis--in contrast, the results of the current CNA analysis have been closely reproduced by the independent DoD analysis. For these reasons we believe the results of the current CNA analysis are preferable to the 1978 CNA analysis.

DISCUSSION OF RESULTS

A comparison of our final results for the ASVAB 6/7 AFQT with the current operational norms shows that the current operational norms are 15 to 17 percentile points too easy in the critical region between the 16th and 31st percentiles (figure 17). For example, according to the current operational norms, an ASVAB 6/7 AFQT raw score of 31 should convert to the 31st percentile. Our result indicates that a raw score of 31 really corresponds to the 16th percentile--a difference of 15 percentile points. This is the area where the services have established enlistment minimums. The current norms also appear to be about 4 percentile points too easy near the 90th percentile, but this difference is not critical because no enlistment decisions are made near the 90th percentile.

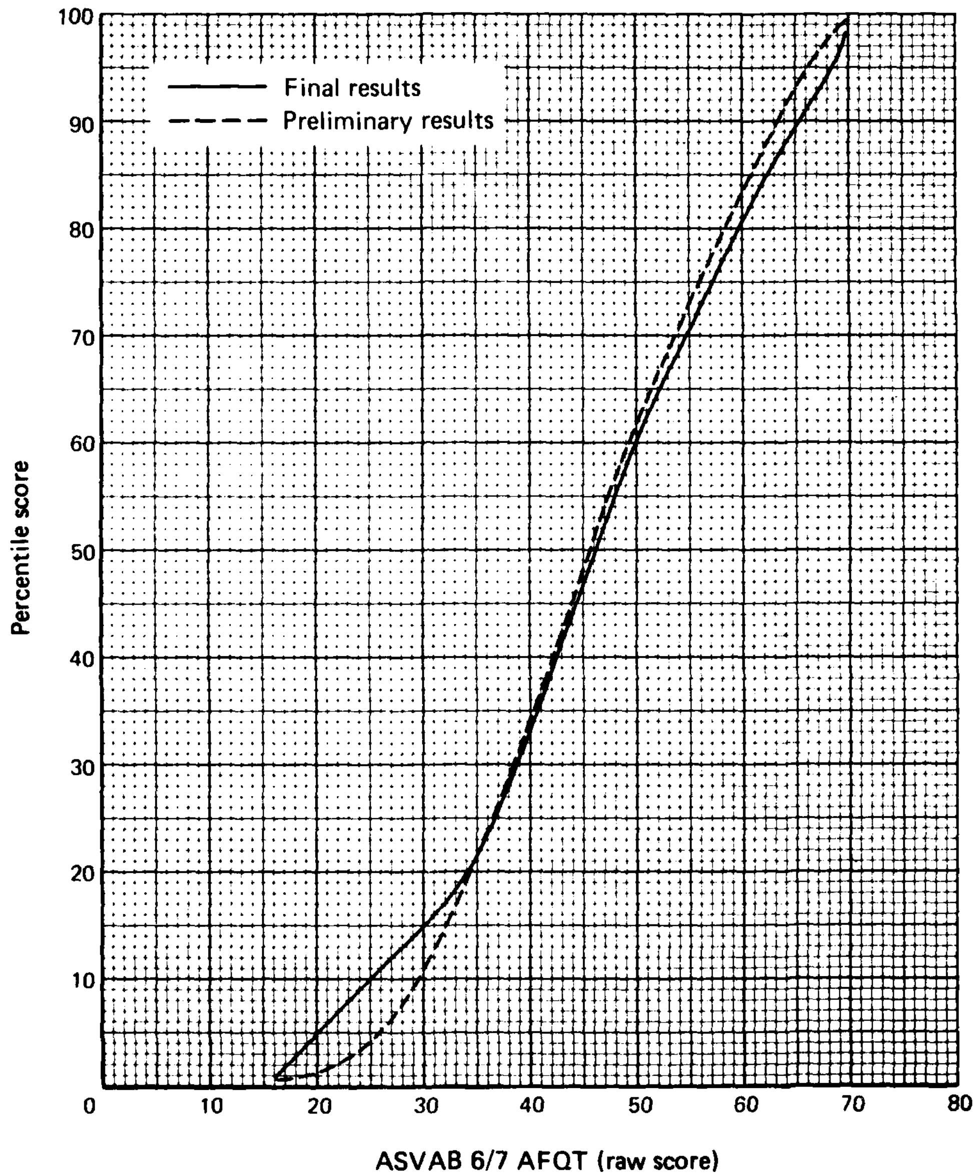


FIG. 15: COMPARISON OF PRELIMINARY AND FINAL RESULTS FROM THIS ANALYSIS

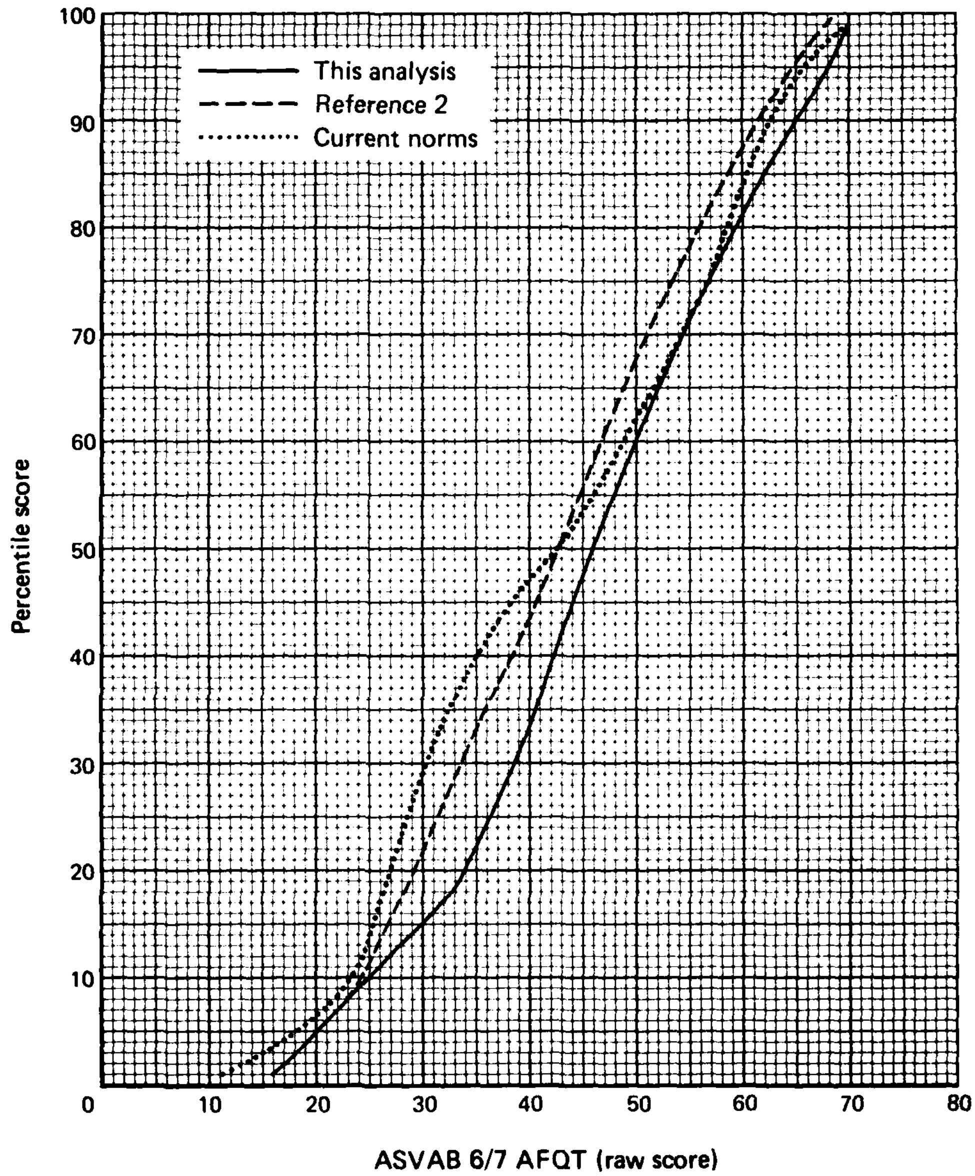


FIG. 16: COMPARISON OF RESULTS FROM THIS ANALYSIS AND AN EARLIER CNA STUDY

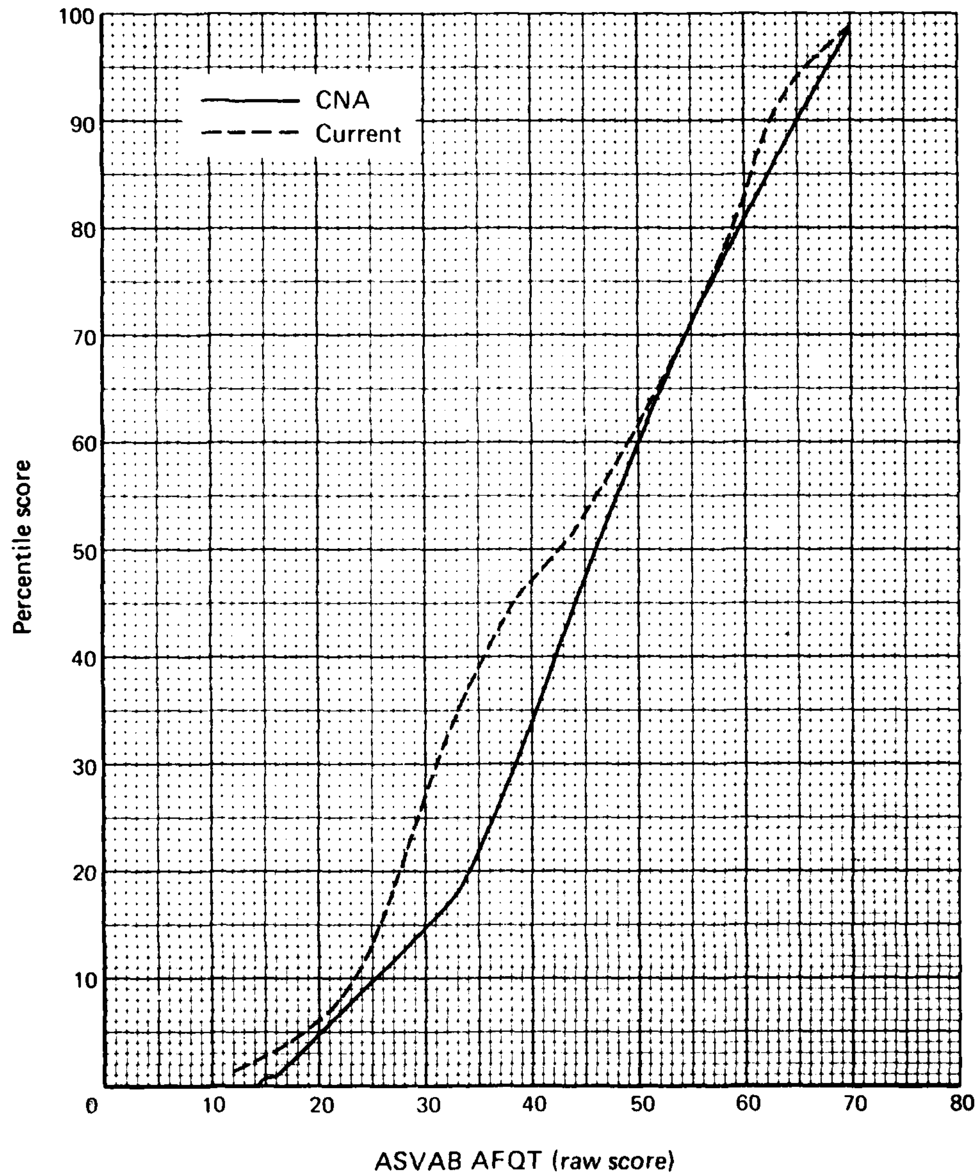


FIG. 17: COMPARISON OF CNA NORMS AND CURRENT OPERATIONAL NORMS FOR ASVAB 6/7/6E/7E AFQT

Because ASVAB aptitude composites[1] are normed to be compatible with the AFQT score scale it is not surprising that current norms for these composites are also inaccurate. Figure 18 compares the current operational norms for one of these composites (Army GT aptitude composite) with our results (see appendix M). The current operational norms lead to an aptitude composite score as much as 10 standard score points higher than would be warranted under our results.

Decades of research by the armed services have shown the AFQT score to be a good measure of general trainability. The AFQT score is frequently grouped into broad categories, called mental groups, ranging from I (highest) to V (lowest). Mental group IV is the lowest currently acceptable category. Table 5 shows the minimum percentile score that defines each mental group and the corresponding AFQT raw score from the current operational norms (reference 8). For example, the breakpoint between mental group IIIB and IVA is the 31st percentile. This percentile corresponds to an AFQT raw score of 31 by the current norms, but a 39 on either the CNA or DoD norms.

TABLE 5

MENTAL GROUP DEFINITIONS BY CURRENT AND PROPOSED NORMS

Mental group	Minimum AFQT percentile score in mental group	Minimum ASVAB AFQT raw score in mental group		
		Current operational	CNA	DoD
I	93	64	67	65
II	65	52	52	52
III A	50	42	46	46
III B	31	31	39	39
IV A	21	28	34	34
IV B	16	26	31	30
IV C	10	23	25	24

[1] These composites (defined in appendix A) are used mainly to assign recruits to suitable military jobs.

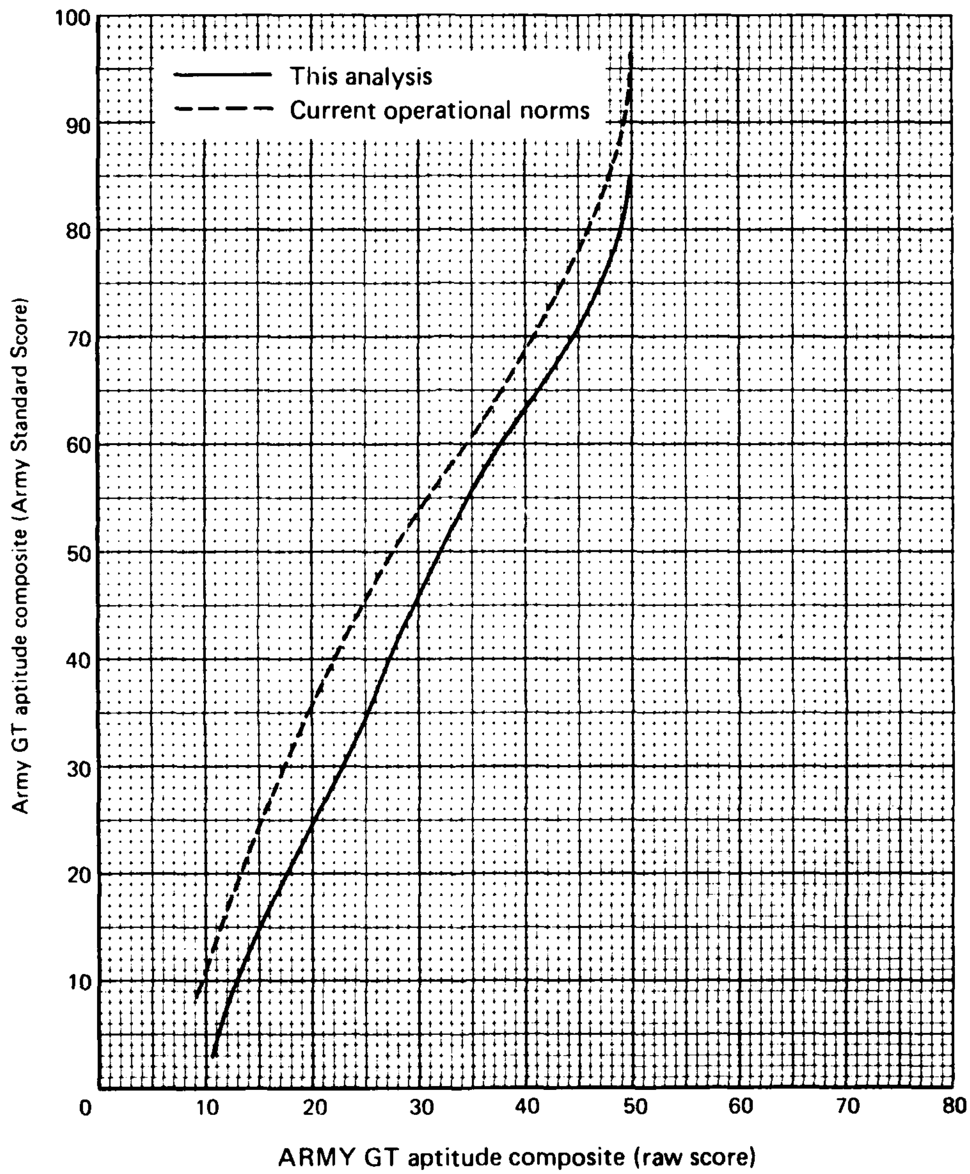


FIG. 18: COMPARISON OF OUR RESULTS AND CURRENT OPERATIONAL NORMS FOR THE ARMY GT APTITUDE COMPOSITE

If the CNA norms are correct, then a large number of recruits currently classified in category IIIB are really in category IV. Applying the CNA (or DoD) norms to distributions of AFQT raw scores from FY 1977, FY 1978, and FY 1979[1] we find that between 25 and 30 percent of accessions in these years are really in category IV rather than the 5 or 6 percent officially reported.

A historical perspective on the percentage of male DoD accessions in mental category IV is shown in figure 19. The solid line represents officially reported percentages. The dots indicate what these percentages would have been during FY 1977, 1978, and 1979 if the norms from this analysis (or the DoD analysis) had been used. Assuming our norms are correct, the 25 to 30 percent figures for mental category IV are higher than those observed during the early 1960s, similar to those during the Vietnam War, and lower than those during the Korean War.

A correctly normalized test is important to managers as well as unit commanders and military trainers. The principal virtue of maintaining a correctly normalized test is that a certain score on a current version of the accession test reflects the same ability to absorb training as that same score did on previous versions of the test. As a result, managers can make informed judgments about changes over time in the aptitude of recruits. By the continued use of correctly normalized tests, a rational basis, founded on years of service experience in peace and in war, can be formed for both enlistment and job classification standards.

---

[1]Supplied by the Department of Defense.

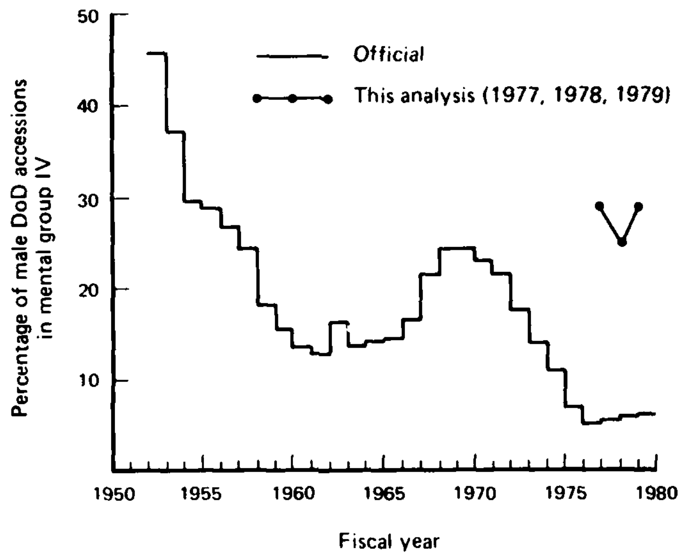


FIG. 19: COMPARISON OF PERCENTAGE OF MALE DOD ACCESSIONS IN MENTAL GROUP IV AS OFFICIALLY REPORTED AND BY THIS ANALYSIS

REFERENCES

1. Air Force Human Resources Laboratory, AFHRL-TR-76-87, "Development of the Armed Services Vocational Aptitude Battery, Forms 5, 6, and 7," by H. E. Jensen, I. H. Massey, and L. D. Valentine, Jr., Unclassified, Dec 1976

2. Center for Naval Analyses, Study 1115, "An Analysis of the Normalization and Verification of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 6 and 7," by William H. Sims, Unclassified, Apr 1978

3. Office of the Deputy Chief of Staff for Research, Development, and Studies (USMC), letter of 26 Jan 1979

4. Office of the Assistant Secretary of Defense, letter of 24 Jan 1979

5. Center for Naval Analyses, Memorandum (CNA)79-3059, "A Reexamination of the Normalization of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 6A, 7B, 6E and 7E," by William H. Sims and Ann Truss, Unclassified, 30 May 1979. (This document was originally issued as a working paper.)

6. U.S. Army Personnel Research Office, TR-1132, "Development of the Armed Forces Qualification Test 7 and 8," by A. G. Bayroff and Alan A. Anderson, Unclassified, May 1963

7. Robert L. Thorndike, "Educational Measurement," American Council on Education, Washington, D.C., LC 71-118852, 1971

8. Office of the Assistant Secretary of Defense (MRA&L), Draft Report, "Renorming ASVAB 6 and 7 at Armed Forces Examining and Entrance Stations," by Milton Maier, Unclassified, Apr 1980

APPENDIX A

DEFINITIONS OF ASVAB TESTS AND COMPOSITES

APPENDIX A

DEFINITIONS OF ASVAB TESTS AND COMPOSITES

This appendix defines the ASVAB tests and composites used in the analysis of normalization. The information is presented in tabular form.

TABLE A-1

INDIVIDUAL ASVAB 6/7 TESTS

Test symbol	Definition
GI	General Information
NO	Numerical Operations
AD	Attention to Detail
WK	Word Knowledge
AR	Arithmetic Reasoning
SP	Space Perception
MK	Mathematics Knowledge
EI	Electronics Information
MC	Mechanical Comprehension
GS[a]	General Science
SI	Shop Information
AI	Automotive Information
CC	Combat Scale
CA	Attentiveness Scale
CE	Electronics Scale
CM	Maintenance Scale

[a]Note that the full-length GS test, rather than the short General Science Biological (GSB) test, is used throughout this report.

TABLE A-2

MARINE CORPS AND ARMY ASVAB 6/7 COMPOSITES

Test composite symbol	Definition
CO	Combat
FA	Field Artillery
OF	Operators and Food Handlers
MM	Mechanical Maintenance
GM	General Maintenance
CL	Clerical
GT	General Technical
EL	Electronics
SC	Surveillance and Communications
ST	Skilled Technical
GCT[a]	General Classification Test

[a]This composite, if defined in percentile form, is referred to as the AFQT (Armed Forces Qualification Test).

TABLE A-3

FORMULAS FOR COMPUTING MARINE CORPS AND ARMY ASVAB 6/7 COMPOSITES

CO	= AR + SI + SP + AD + CC
FA	= AR + GI + MK + EI + CA
MM	= MK + SI + EI + AI + CM
GM	= AR + GS + MC + AI
CL	= AR + WK + AD + CA
GT	= AR + WK
EL[a]	= AR + GS + MK + EI
EL[b]	= AR + EI + MC + SI + CE
SC	= AR + WK + MC + SP
ST	= AR + MK + GS
OF	= GI + AI + CA
GCT[c]	= AR + WK + SP

[a] Marine Corps only.

[b] Army only.

[c] Also called the AFQT.

TABLE A-4

FORMULAS FOR COMPUTING NAVY AND AIR FORCE
ASVAB 6/7 COMPOSITES

Navy

G = WK + AR

M[a] = WK + MC + SI

E = AR + MK + EI + GS

C = NO + AD + WK

Air Force

M = MC + SI + AI

A = NO + AD + WK

G = WK + AR

E = AR + SP + EI

[a]Note that this formula is incorrectly stated in the following commonly used reference: Department of Defense, DoD 1301.12M, "Directions for Scoring the Armed Services Vocational Aptitude Battery Forms 6 and 7," Unclassified, January 1976.

APPENDIX B

EXPERIMENTAL DESIGN

APPENDIX B

EXPERIMENTAL DESIGN

The data sample on which we based our analysis consisted of 3,295 Marine Corps recruits. These recruits were given a series of aptitude tests within a few days after they arrived at the two Marine Corps Recruit Depots. Each recruit was given three tests: ASVAB 6 or 7, ASVAB 6E or 7E (AFQT part only), and AFQT 7A (a reference test). The recruits were tested in platoon-size groups of about 60 men. All recruits in a platoon were tested on the same tests in the same order. But not all platoons took the same tests in the same order. That is, the order in which the tests were given was counterbalanced so that each test was given first to a platoon as often as it was given second or third. The order of testing is shown in table B-1.[1]

Tests were administered by Marine Corps testing personnel. One of the authors of this report monitored the initial testing session at each recruit depot.

Because ASVAB 6 and 7 are routinely administered to all recruits who enter the recruit depots, the answer sheets for these tests were graded by Marine Corps personnel. Scores were then made available to use for analysis. Answer sheets for the ASVAB 6E/7E and AFQT 7A were optically scanned by the Marine Corps Institute, which produced on punched cards output with each response (A, B, C, or D) to each test item recorded. The responses for each item were compared with the correct answer by a computer program at the Center for Naval Analyses and the number of correct responses was recorded. In the case of the ASVAB 6/7/6E/7E tests, the number of correct responses (raw score) was the variable of interest for further analysis. For the AFQT 7A reference test, the number of correct answers was converted into a percentile score using the official AFQT 7/8 conversion shown in table B-2.

---

[1] Note that during the data collection phase of the experiment, ASVAB 6E and 7E were called "R" and "S", respectively.

TABLE B-1

ORDER OF TESTING

Ordered group	Platoon	Series	First test: Test	First test: Hour[a]/Date	Second test: Test	Second test: Hour[a]/Date	Third test: Test	Third test: Hour[a]/Date
______	______	______	ASVAB 6/7	______	AFQT 7A	______	R	______
______	______	______	ASVAB 6/7	______	S	______	AFQT 7A	______
______	______	______	AFQT 7A	______	ASVAB 6/7	______	S	______
______	______	______	R	______	ASVAB 6/7	______	AFQT 7A	______
______	______	______	AFQT 7A	______	R	______	ASVAB 6/7	______
______	______	______	S	______	AFQT 7A	______	ASVAB 6/7	______
______	______	______	ASVAB 6/7	______	AFQT 7A	______	R	______
______	______	______	ASVAB 6/7	______	S	______	AFQT 7A	______
______	______	______	AFQT 7A	______	ASVAB 6/7	______	S	______
______	______	______	R	______	ASVAB 6/7	______	AFQT 7A	______
______	______	______	AFQT 7A	______	R	______	ASVAB 6/7	______
______	______	______	S	______	AFQT 7A	______	ASVAB 6/7	______

[a]Hour test started (for example, 0800/30 Jan)

TABLE B-2

TABLE FOR CONVERTING RAW SCORES TO PERCENTILE SCORES ON AFQT 7 AND AFQT 8

Raw score[a]	Percentile score	Raw score	Percentile score	Raw score	Percentile score	Raw score	Percentile score
100	100	74	67	49	29	24	11
99	99	73	65	48	28	23	11
98	99	72	63	47	27	22	10
97	98	71	62	46	26	21	10
96	98	70	61	45	25	20	9
95	97	69	60	44	24	19	9
94	96	68	58	43	23	18	8
93	95	67	56	42	22	17	8
92	94	66	54	41	21	16	7
91	93	65	52	40	20	15	7
90	92	64	51	39	19	14	6
89	91	63	50	38	18	13	6
88	90	62	49	37	18	12	5
87	89	61	48	36	17	11	5
86	88	60	46	35	17	10	4
85	87	59	44	34	16	9	4
84	85	58	42	33	16	8	3
83	83	57	40	32	15	7	3
82	82	56	38	31	15	6	3
81	81	55	36	30	14	5	2
80	80	54	34	29	14	4	2
79	78	53	33	28	13	3	2
78	76	52	32	27	13	2	1
77	74	51	31	26	12	1	1
76	72	50	30	25	12	0	1
75	70						

[a] Raw score is the number right minus one-third the number wrong. Ommitted items are not counted as wrong.

APPENDIX C

SAMPLE STATISTICS

APPENDIX C

SAMPLE STATISTICS

This appendix presents various statistics for the sample and subsamples used. The information is presented in tabular form.

TABLE C-1

STATISTICS FOR TOTAL DATA SAMPLE
(3,295 recruits, unweighted)

	Item	Mean values by test site: AFEES	Mean values by test site: Depot
ASVAB 6A/7B	GI	9.4	9.5
	NO	31.0	31.2
	AD	14.2	15.2
	WK	20.2	19.5
	AR	12.7	11.8
	SP	12.6	13.3
	MK	10.9	10.4
	EI	18.7	17.8
	MC	10.3	10.3
	GS	10.6	10.5
	SI	13.6	13.3
	AI	11.1	11.3
	CM	12.3	10.8
	CA	10.3	9.5
	CE	8.6	8.7
	CC	18.6	15.1
	AFQT (raw)	45.5	44.6
	Pseudo AFQT (raw)	43.8	43.4
ASVAB 6E	WK		18.1
	AR		11.6
	SP		14.2
	AFQT		43.9
ASVAB 7E	WK		17.9
	AR		10.2
	SP		13.5
	AFQT		41.6
Other	Percentage minority		27.7
	Percentage male		100.0
	Percentage high school graduates		53.1
	Percentage Parris Island/San Diego		50.0/50.0
	Reference test (AFQT 7A) percentile score		44.4

TABLE C-2

SUBSAMPLE STATISTICS (unweighted)

<table>
<tr><th rowspan="2">Item</th><th colspan="5">Mean values</th></tr>
<tr><th>Sample 1</th><th>Sample 2</th><th>Sample 3</th><th>Sample 4</th><th>Sample 5</th></tr>
<tr><td>AFEES</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>ASVAB 6/7 AFQT[a]</td><td>45.5</td><td>45.5</td><td>45.5</td><td>--[b]</td><td>45.4</td></tr>
<tr><td>ASVAB 6/7 Pseudo AFQT[a]</td><td>43.8</td><td>43.9</td><td>43.8</td><td>--[b]</td><td>43.9</td></tr>
<tr><td>Depot</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>ASVAB 6/7 AFQT[a]</td><td>44.6</td><td>44.6</td><td>44.6</td><td>44.8</td><td>44.1</td></tr>
<tr><td>ASVAB 6/7 Pseudo AFQT[a]</td><td>43.4</td><td>43.4</td><td>43.3</td><td>45.0</td><td>43.0</td></tr>
<tr><td>ASVAB 6E AFQT[a]</td><td rowspan="2">42.7</td><td>43.9</td><td>--</td><td rowspan="2">43.7</td><td>--</td></tr>
<tr><td>ASVAB 7E AFQT[a]</td><td>--</td><td>41.6</td><td>--</td></tr>
<tr><td>AFQT 7A[a]</td><td>44.4</td><td>43.8</td><td>44.9</td><td>46.6</td><td>45.0</td></tr>
<tr><td>Sample size</td><td>3,295</td><td>1,634</td><td>1,660</td><td>227</td><td>2,208</td></tr>
</table>

[a]All test scores are expressed in raw score form except the AFQT 7A, which is in percentiles.

[b]This subsample took ASVAB 5 as an enlistment test when in high school.

APPENDIX D

STRATIFIED NORMALIZATION ANALYSIS

APPENDIX D

STRATIFIED NORMALIZATION ANALYSIS

Each of the data subsamples was stratified on the reference test, AFQT 7A, to simulate the traditional reference population. We stratified by weighting the individual recruits so that their AFQT 7A percentile score distribution was flat. The calculations of the weight factors for samples 1 through 5 are shown in tables D-1 through D-5.

By applying the weight factors to recruits depending on their AFQT 7A scores, we calculated weighted cumulative frequency distributions of the AFQT raw score of the test to be normalized (table D-6). These weighted cumulative distributions, when smoothed, become the unadjusted conversion tables between ASVAB AFQT raw scores and percentile scores shown in the main text; i.e., the normalization of the new test.

TABLE D-1

CALCULATIONS OF WEIGHT FACTORS FOR SAMPLE 1

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	44	164.75	3.744
6-10	85	164.75	1.938
11-15	228	164.75	0.723
16-20	282	164.75	0.584
21-25	252	164.75	0.654
26-30	206	164.75	0.800
31-35	248	164.75	0.664
36-40	255	164.75	0.646
41-45	111	164.75	1.484
46-50	256	164.75	0.644
51-55	202	164.75	0.816
56-60	250	164.75	0.659
61-65	248	164.75	0.664
66-70	76	164.75	2.168
71-75	115	164.75	1.433
76-80	151	164.75	1.091
81-85	146	164.75	1.128
86-90	86	164.75	1.916
91-95	44	164.75	3.744
96-100	10	164.75	16.473
Total	3,295	3,295	

[a]Column (3) divided by column (2).

TABLE D-2

CALCULATION OF WEIGHT FACTORS FOR SAMPLE 2

AFQT 'A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	23	81.7	3.552
6-10	51	81.7	1.602
11-15	106	81.7	0.771
16-20	149	81.7	0.548
21-25	129	81.7	0.633
26-30	92	81.7	0.888
31-35	130	81.7	0.628
36-40	120	81.7	0.681
41-45	52	81.7	1.571
46-50	125	81.7	0.654
51-55	116	81.7	0.704
56-60	114	81.7	0.717
61-65	128	81.7	0.638
66-70	37	81.7	2.208
71-75	66	81.7	1.238
76-80	62	81.7	1.318
81-85	77	81.7	1.061
86-90	33	81.7	2.476
91-95	19	81.7	4.300
96-100	5	81.7	16.340
Total	1,634	1,634	

[a]Column (3) divided by column (2).

TABLE D-3

CALCULATION OF WEIGHT FACTORS FOR SAMPLE 3

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	21	83.0	3.952
6-10	34	83.0	2.441
11-15	122	83.0	0.680
16-20	133	83.0	0.624
21-25	123	83.0	0.675
26-30	114	83.0	0.728
31-35	118	83.0	0.703
36-40	135	83.0	0.615
41-45	59	83.0	1.407
46-50	131	83.0	0.634
51-55	85	83.0	0.976
56-60	136	83.0	0.610
61-65	120	83.0	0.692
66-70	39	83.0	2.128
71-75	49	83.0	1.694
76-80	89	83.0	0.933
81-85	69	83.0	1.203
86-90	53	83.0	1.566
91-95	25	83.0	3.320
96-100	5	83.0	16.600
Total	1,660	1,660	

[a]Column (3) divided by column (2).

TABLE D-4

CALCULATION OF WEIGHT FACTORS FOR SAMPLE 4

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-10	8	22.7	2.838
11-20	25	22.7	0.908
21-30	29	22.7	0.783
31-40	37	22.7	0.614
41-50	38	22.7	0.597
51-60	23	22.7	0.987
61-70	25	22.7	0.908
71-80	18	22.7	1.261
81-90	21	22.7	1.081
91-100	3	22.7	7.567
Total	227	227	

[a]Column (3) divided by column (2).

TABLE D-5

CALCULATION OF WEIGHT FACTORS FOR SAMPLE 5

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	29	110.4	3.807
6-10	56	110.4	1.971
11-15	152	110.4	0.726
16-20	169	110.4	0.653
21-25	171	110.4	0.646
26-30	137	110.4	0.806
31-35	162	110.4	0.681
36-40	175	110.4	0.631
41-45	71	110.4	1.555
46-50	173	110.4	0.638
51-55	138	110.4	0.800
56-60	164	110.4	0.673
61-65	167	110.4	0.661
66-70	56	110.4	1.971
71-75	84	110.4	1.314
76-80	107	110.4	1.032
81-85	96	110.4	1.150
86-90	59	110.4	1.871
91-95	33	110.4	3.345
96-100	9	110.4	12.267
Total	2,208	2,208	

[a]Column (3) divided by column (2).

TABLE D-6

STRATIFIED CUMULATIVE FREQUENCY DISTRIBUTION
OF RAW ASVAB AFQT SCORES

Raw AFQT score	Cumulative frequency: Sample 1, form 6/7	Sample 2, form 6E	Sample 3, form 7E	Sample 4, form 6/7	Sample 5, form 6/7
0 - 15	0.4	0.4	1.5	0.0	0.5
16 - 17	0.7	0.6	2.9	0.0	0.8
18 - 19	1.2	1.2	3.6	0.4	1.2
20	1.7	1.5	4.6	1.7	1.8
21	2.2	2.1	4.9	1.7	2.1
22	2.6	3.0	5.6	4.6	2.6
23	3.2	4.1	7.1	4.6	3.4
24	3.8	4.6	8.2	5.4	4.3
25	4.7	5.9	9.6	6.2	5.2
26	5.4	6.9	10.4	6.5	5.8
27	6.3	8.2	11.8	7.3	6.9
28	7.3	9.2	13.3	8.0	7.9
29	8.4	10.8	15.2	9.1	9.2
30	10.2	12.6	17.0	10.7	11.3
31	12.0	13.7	18.7	12.2	13.1
32	13.8	15.4	20.7	15.3	15.2
33	15.6	16.5	22.5	17.2	17.3
34	17.5	19.1	25.1	18.5	19.6
35	19.6	21.1	27.5	20.5	21.9
36	21.8	23.3	30.1	22.3	24.5
37	24.0	25.4	32.6	24.7	26.6
38	26.5	27.5	35.0	28.6	28.9
39	29.4	30.0	37.9	31.7	31.9
40	32.1	32.9	41.4	36.1	34.8
41	35.0	35.8	43.8	37.3	37.5
42	37.6	38.5	46.9	41.5	40.2
43	40.1	40.8	49.4	43.1	42.9
44	43.1	43.7	52.8	45.6	45.7
45	46.4	46.3	55.3	48.4	49.2
46	49.3	49.7	58.2	51.6	52.0
47	51.8	52.4	60.7	54.0	54.3
48	55.0	55.3	62.6	57.1	57.3
49	58.3	58.6	65.5	58.9	60.0
50	61.0	61.4	68.5	62.4	62.5
51	63.1	63.8	70.7	65.3	64.4
52	65.4	66.2	73.9	66.4	66.8
53	67.8	68.8	76.9	66.9	69.2
54	70.3	71.6	78.2	69.1	71.4
55	73.2	73.4	80.1	73.4	74.6
56	75.1	75.7	81.5	76.8	76.4
57	76.7	78.0	83.0	78.6	78.0
58	78.8	80.7	84.4	79.2	79.8
59	80.5	83.1	85.7	81.9	81.5
60	83.5	85.1	87.4	86.8	84.3
61	85.4	87.1	90.9	88.6	86.1
62	87.7	88.9	92.7	89.6	88.1
63	89.2	91.0	94.3	93.3	89.6
64	91.6	91.7	96.2	93.7	92.3
65	93.2	92.7	96.9	94.3	93.9
66	95.3	94.3	98.9	99.4	96.4
67	96.8	96.3	99.7	100.0	97.2
68	98.7	98.3	99.8	100.0	98.5
69	99.9	99.7	100.0	100.0	99.9
70	100.0	100.0	100.0	100.0	100.0

APPENDIX E

UNSTRATIFIED GRAPHICAL EQUATING

APPENDIX E

UNSTRATIFIED GRAPHICAL EQUATING

In this appendix we describe the direct equating of scores on the AFQT part of ASVAB 6/7, ASVAB 6E, and ASVAB 7E to the percentile score of the reference test (AFQT 7A). The data of samples 2, 3, and 5 are used as is; i.e., the samples were not stratified on a reference test.

The procedure is illustrated in figure E-1. Cumulative percentages of the ASVAB 6E AFQT raw score and the reference test percentile score were graphed as shown. Scores on the two tests are considered to be equivalent if they are obtained by the same cumulative percentage of the sample. For example, a raw score of 30 on the ASVAB 6E AFQT was made by a cumulative 11 percent of the sample. A cumulative 11 percent of the sample also achieved the 15th percentile score on the reference test. By the definition of equivalent scores we equate a raw score of 30 on the ASVAB 6E AFQT to the 15th percentile. We used this procedure throughout the score range. (See table E-1 for results.)

In much the same way we equated ASVAB 7E AFQT and ASVAB 6/7 AFQT, as shown in figures E-2 and E-3. The results of these equations are also tabulated in table E-1.

A comparison of the results of the unstratified graphical equating with those from the stratified equating (appendix D) are shown for each form of ASVAB AFQT in figures E-4, E-5, and E-6. In general, the stratified procedure results in harder norms in the low percentiles and easier norms in the upper percentiles.

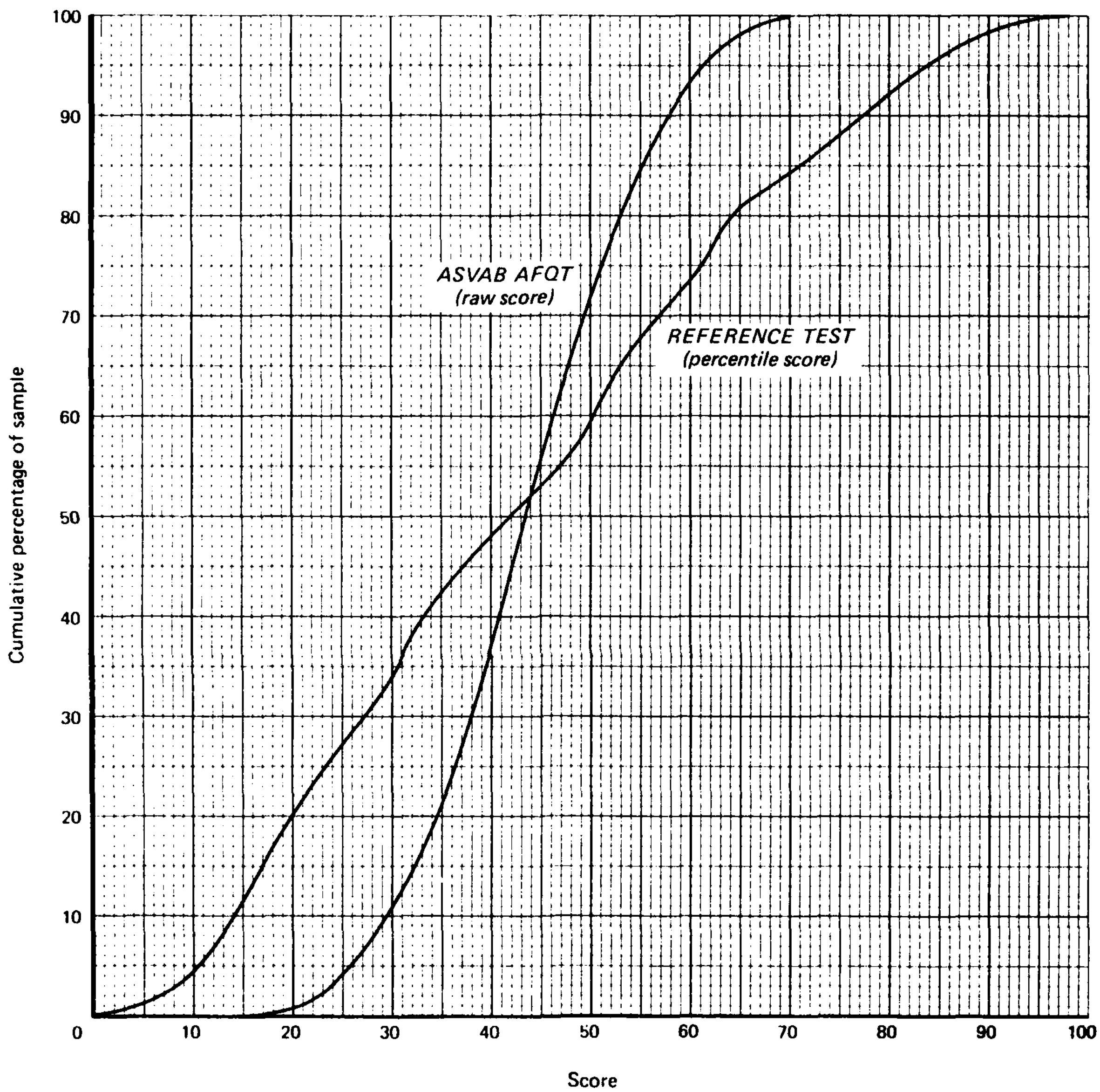


FIG. E-1: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 6E AFQT

TABLE E-1

SUMMARY OF UNSTRATIFIED GRAPHICAL EQUATING RESULTS

ASVAB AFQT interval	Percentiles					
	ASVAB 6/7	ASVAB 6E	ASVAB 7E	ASVAB $7E_2$[a]	ASVAB $7E_3$[b]	Average of forms 6/7, 6E, and $7E_2$
14-15	0.0	0.0	3.0	2.0	1.0	0.7
16-17	0.0	1.0	4.5	3.0	2.5	1.3
18-19	1.0	2.0	6.3	4.5	4.0	2.5
20	3.0	3.0	7.0	5.0	4.5	3.6
21	4.0	4.0	8.0	6.3	5.0	4.7
22	5.5	6.0	9.0	7.0	6.3	6.2
23	6.6	7.6	10.8	8.0	7.0	7.4
24	8.0	8.4	11.5	9.0	8.0	8.5
25	9.6	9.5	12.5	10.8	9.0	10.0
26	10.5	10.6	13.5	11.5	10.8	10.9
27	11.7	11.6	14.5	12.5	11.5	11.9
28	12.5	12.6	15.4	13.5	12.5	12.9
29	13.5	13.8	16.3	14.5	13.5	13.9
30	14.7	15.0	17.5	15.4	14.5	15.0
31	15.5	16.0	19.1	16.3	15.4	15.9
32	16.7	16.7	21.4	17.5	16.3	17.0
33	18.2	17.6	23.2	19.1	17.5	18.3
34	20.2	19.5	25.3	21.4	19.1	20.4
35	22.5	21.1	27.3	23.2	21.4	22.3
36	24.5	23.3	29.3	25.3	23.2	24.4
37	26.5	25.5	31.2	27.3	25.3	26.4
38	28.6	27.7	33.3	29.3	27.3	28.5
39	31.2	30.3	36.4	31.2	29.3	30.9
40	33.4	32.0	39.6	33.3	31.2	32.9
41	36.0	34.5	42.7	36.4	33.3	35.6
42	38.4	37.4	45.8	39.6	36.4	38.5
43	43.0	40.4	48.6	42.7	39.6	42.0
44	46.2	43.8	51.2	45.8	42.7	45.3
45	48.7	47.5	53.4	48.6	45.8	48.3
46	50.6	49.8	56.4	51.2	48.6	50.5
47	53.0	51.1	58.3	53.4	51.2	52.5
48	55.5	53.2	60.3	56.4	53.4	55.0
49	57.6	56.5	62.0	58.3	56.4	57.5
50	60.0	59.5	63.8	60.3	58.3	60.0
51	61.8	61.2	67.2	62.0	60.3	61.7
52	63.3	62.6	71.7	63.8	62.0	63.2
53	65.0	65.0	75.3	67.2	63.8	65.7
54	69.0	68.5	77.3	71.7	67.2	69.7
55	72.0	71.0	78.5	75.3	71.7	72.8
56	74.4	73.3	80.0	77.3	75.3	75.0
57	76.5	76.6	81.4	78.5	77.3	77.2
58	78.5	78.4	82.6	80.0	78.5	79.0
59	79.7	80.0	84.0	81.4	80.0	80.4
60	81.5	81.6	85.5	82.6	81.4	81.9
61	82.6	83.3	87.0	84.0	82.6	83.3
62	84.5	85.0	88.2	85.5	84.0	85.0
63	86.5	86.5	89.5	87.0	85.5	86.7
64	88.2	88.0	91.0	88.2	87.0	88.1
65	90.0	89.2	92.0	89.5	88.2	89.6
66	92.0	90.5	94.0	91.0	89.5	91.2
67	93.0	94.0	95.0	92.0	91.0	93.0
68	96.0	96.0	96.0	94.0	92.0	95.3
69	98.0	98.0	97.0	95.0	94.0	96.3
70	99.0	99.0	99.0	99.0	99.0	99.0

[a] Before converting to percentiles, two points are added to the raw AFQT score.

[b] Before converting to percentiles, three points are added to the raw AFQT score.

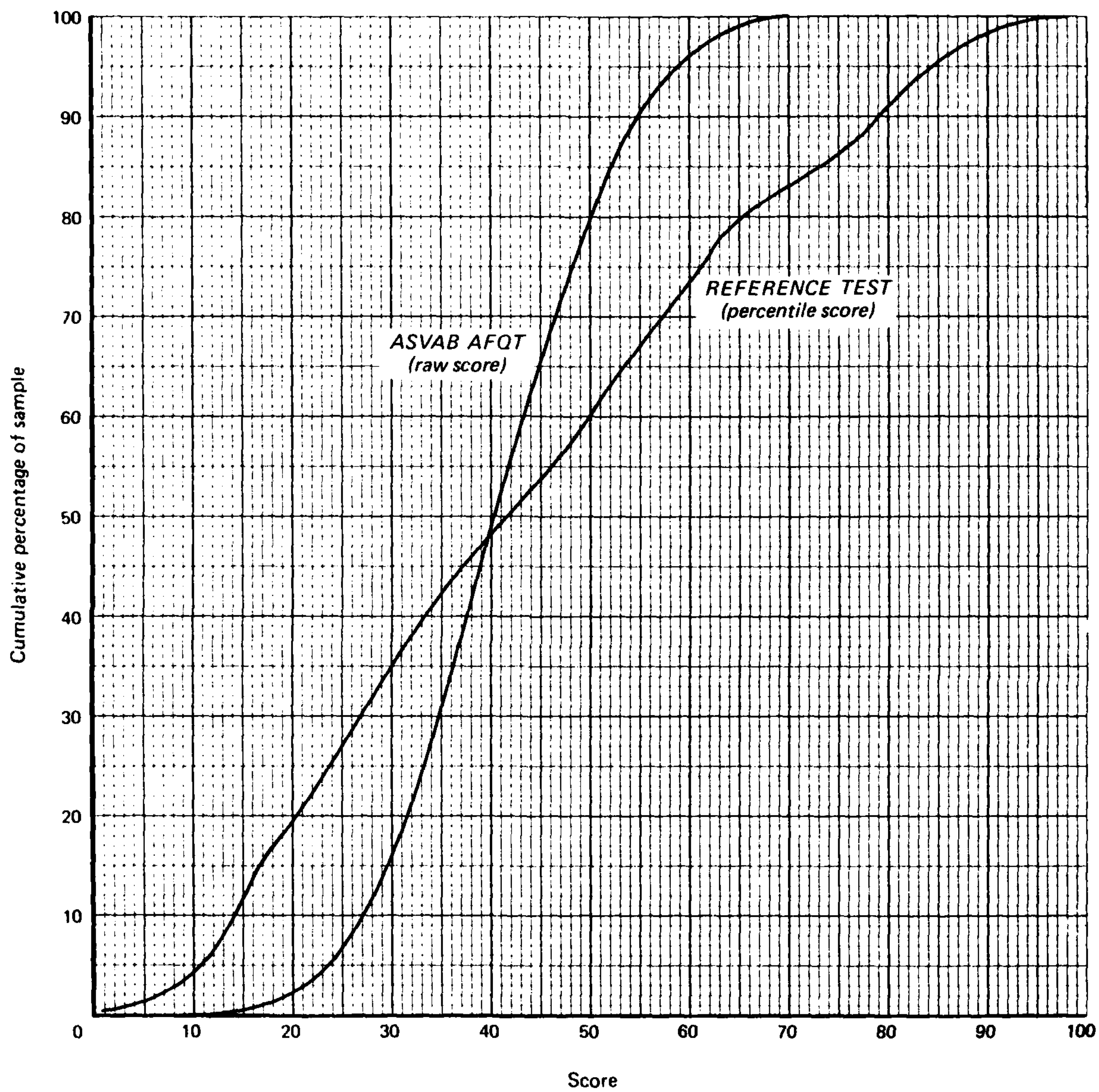


FIG. E-2: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 7E AFQT

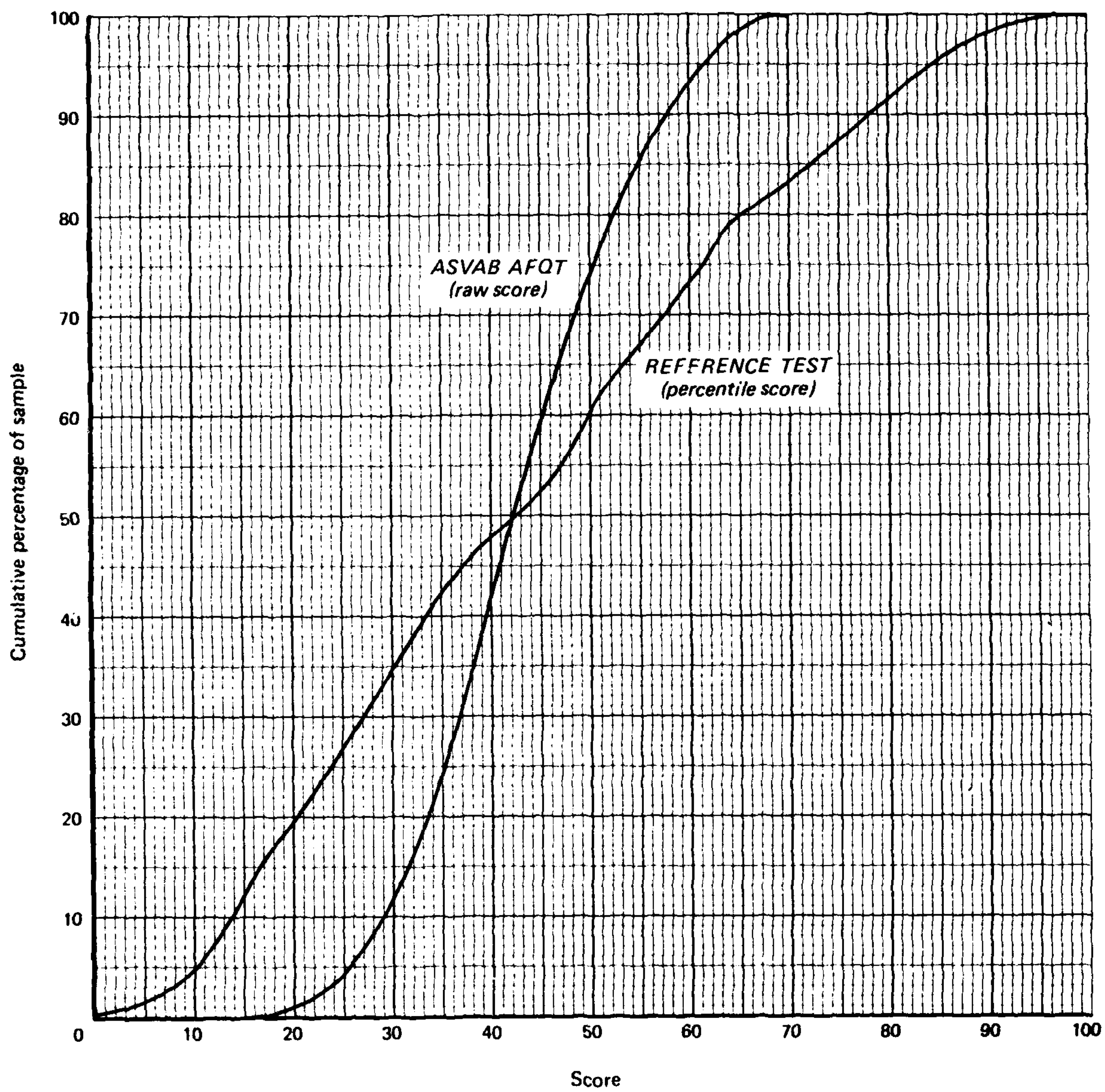


FIG. E-3: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 6/7 AFQT

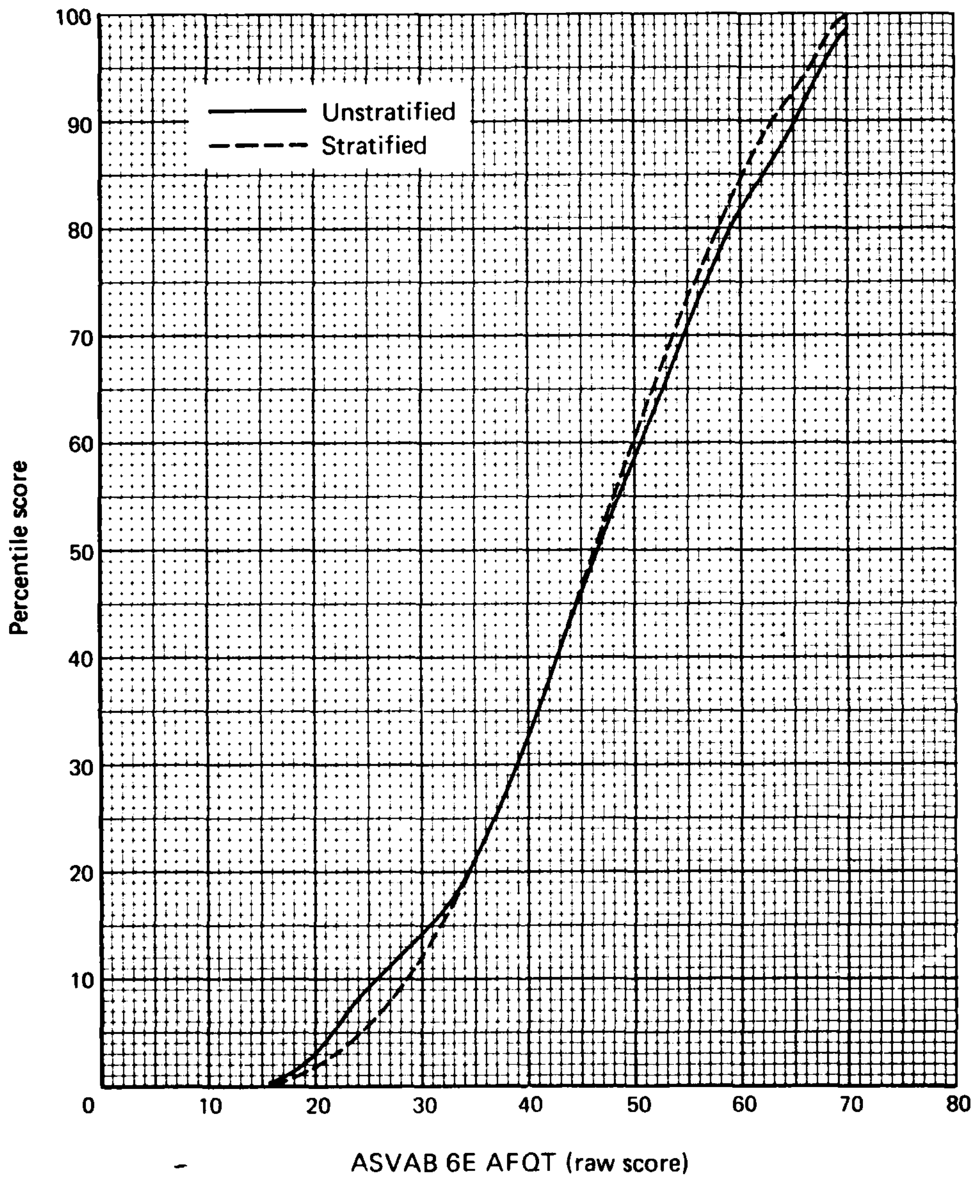


FIG. E-4: COMPARISON OF STRATIFIED AND UNSTRATIFIED NORMS FOR ASVAB 6E AFQT

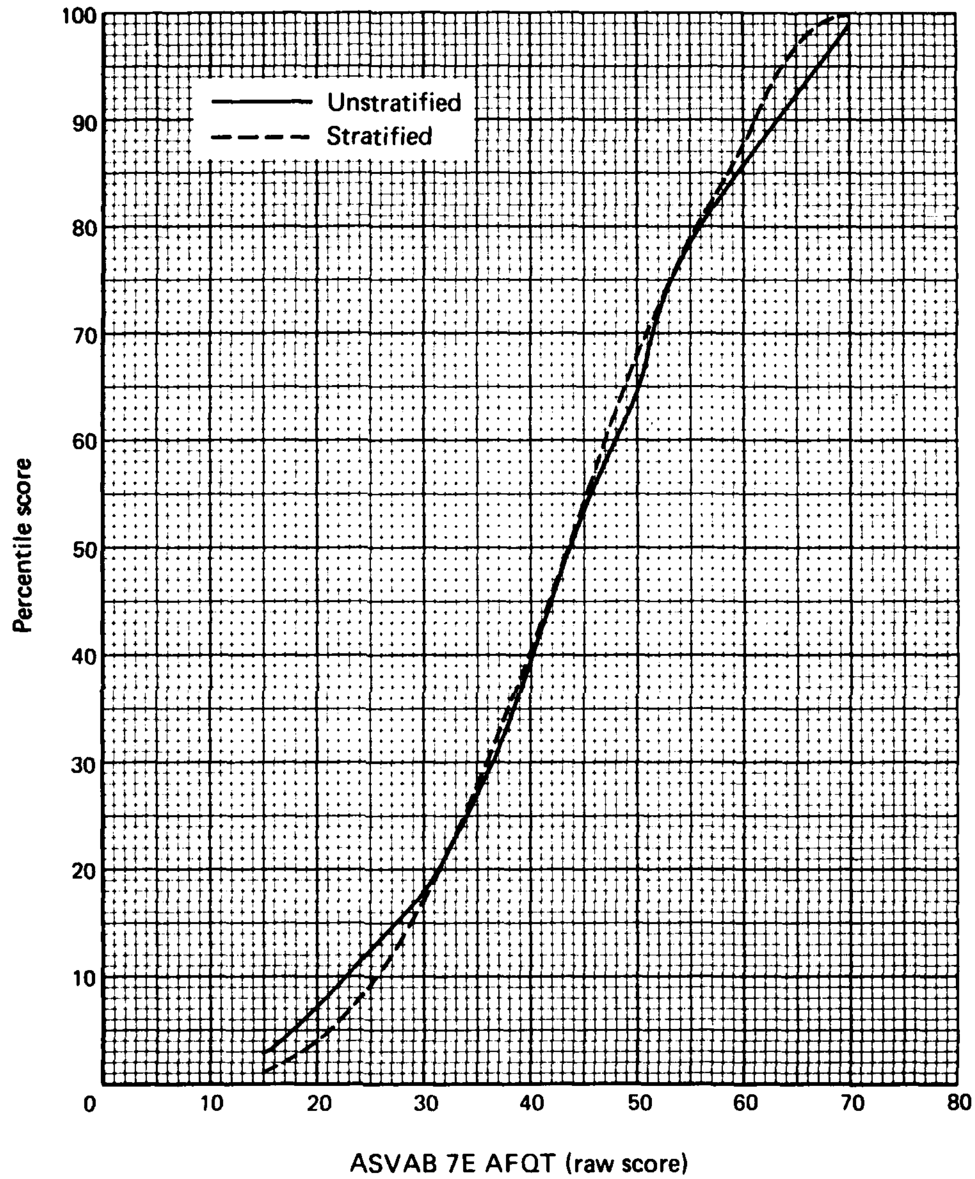


FIG. E-5: COMPARISON OF STRATIFIED AND UNSTRATIFIED NORMS FOR ASVAB 7E AFQT

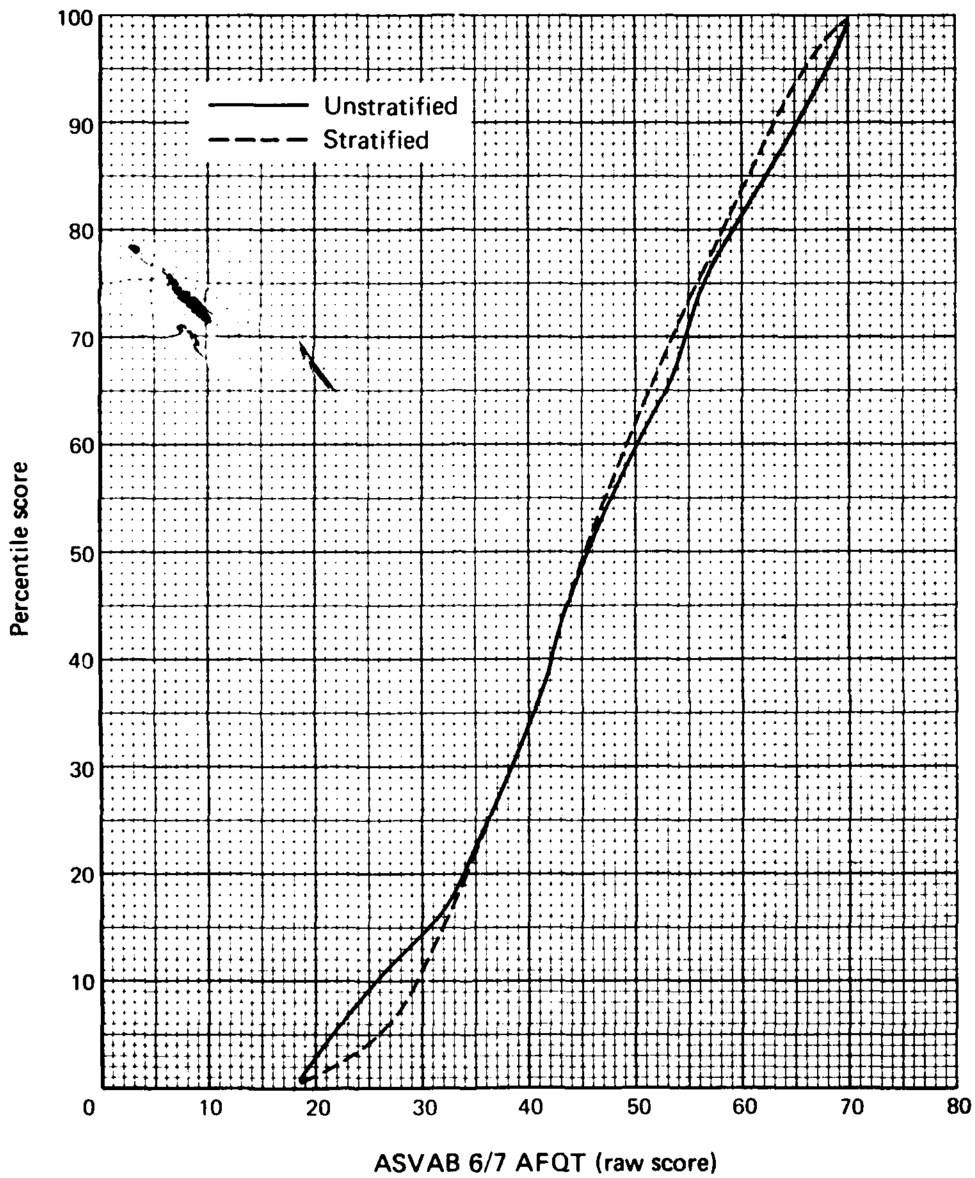


FIG. E-6: COMPARISON OF STRATIFIED AND UNSTRATIFIED NORMS FOR ASVAB 6/7 AFQT

AD-A094 684 CENTER FOR NAVAL ANALYSES ALEXANDRIA VA MARINE CORP--ETC F/G 5/9
A REEXAMINATION OF THE NORMALIZATION OF THE ARMED SERVICES VOCA--ETC(U)
APR 80 W H SIMS, A R TRUSS N00014-76-C-0001

UNCLASSIFIED CNS-1152 NL

AD-A094684

APPENDIX F

EFFECTS OF COACHING ON NORMALIZATION

APPENDIX F

EFFECTS OF COACHING ON NORMALIZATION

If recruits in the samples were coached on any of the tests used in the normalization analysis, the normalization will be biased. The effect of coaching is illustrated in figure F-1. If recruits are coached on the new test, then their raw score on the new test would be artificially high, and the curve in figure F-1 would shift to the right (in the direction of harder norms). If recruits were coached on the reference test, then they would score unexpectedly high on the reference test for a given raw score on the next test. In this case the normalization curve would shift to the left in the direction of easier norms. This appendix examines the effect of coaching on the normalization results of our analysis.

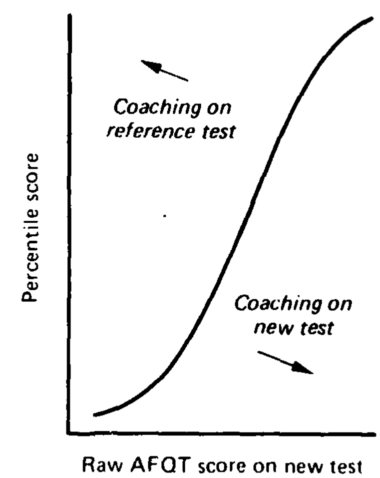


FIG. F-1: ILLUSTRATION OF EFFECT OF COACHING ON NORMALIZATION

Recruits were more likely to be coached on ASVAB 6/7 than on other tests used in this analysis. The reference test as well as ASVAB 6E/7E were not being used when we collected data; hence there would be no motivation to coach the recruits on these test forms.

Comparing scattergrams of AFQT scores from ASVAB with scores on the reference test gives some perspective on effects of coaching. Figures F-2 and F-3 show these scattergrams for ASVAB 6E/7E. These figures show the relationship expected when recruits are not coached. Figure F-4 shows a similar plot for ASVAB 6/7 sample from

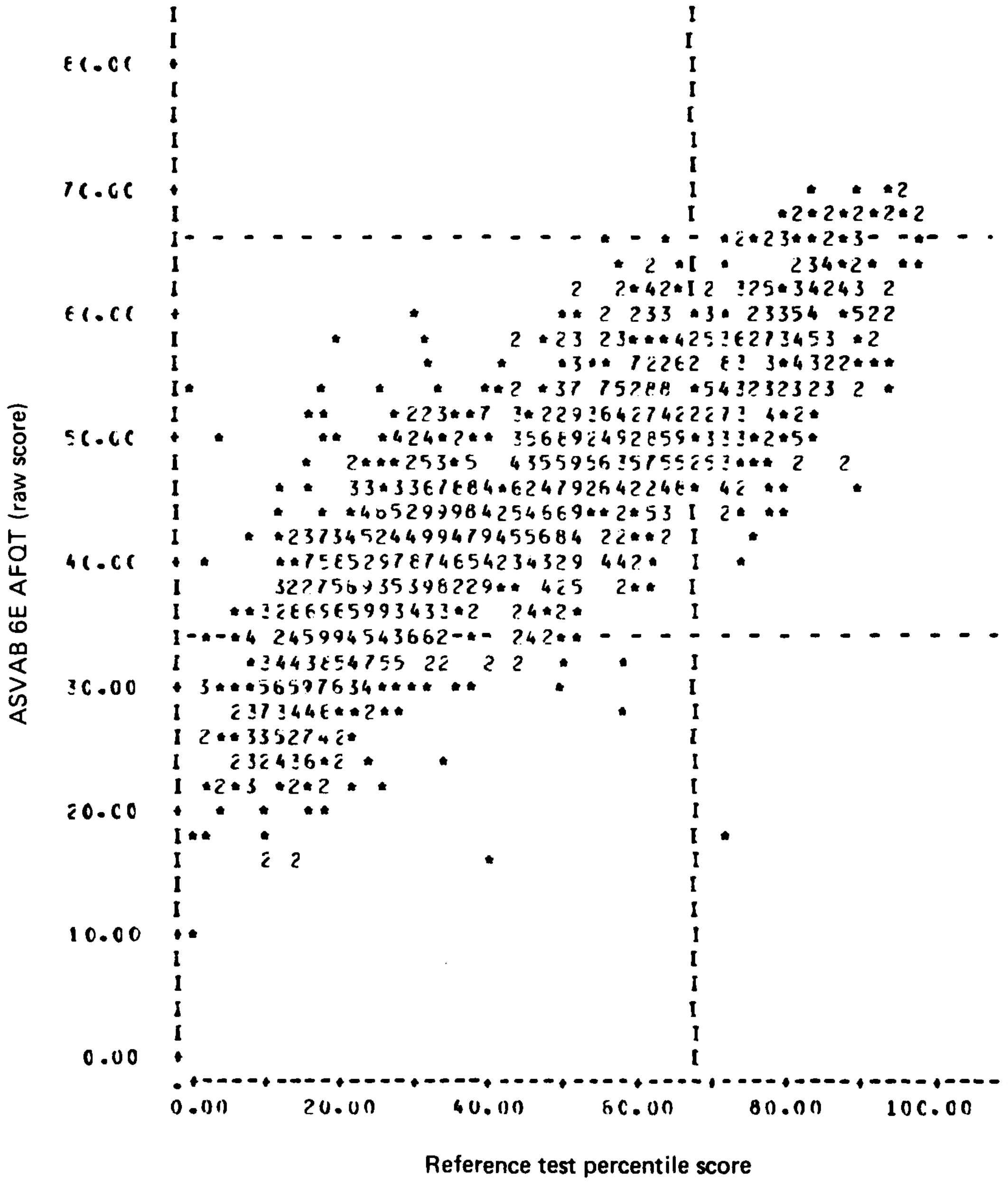


FIG. F-2: SCATTERGRAM OF ASVAB 6E AFQT SCORES VERSUS REFERENCE TEST SCORES

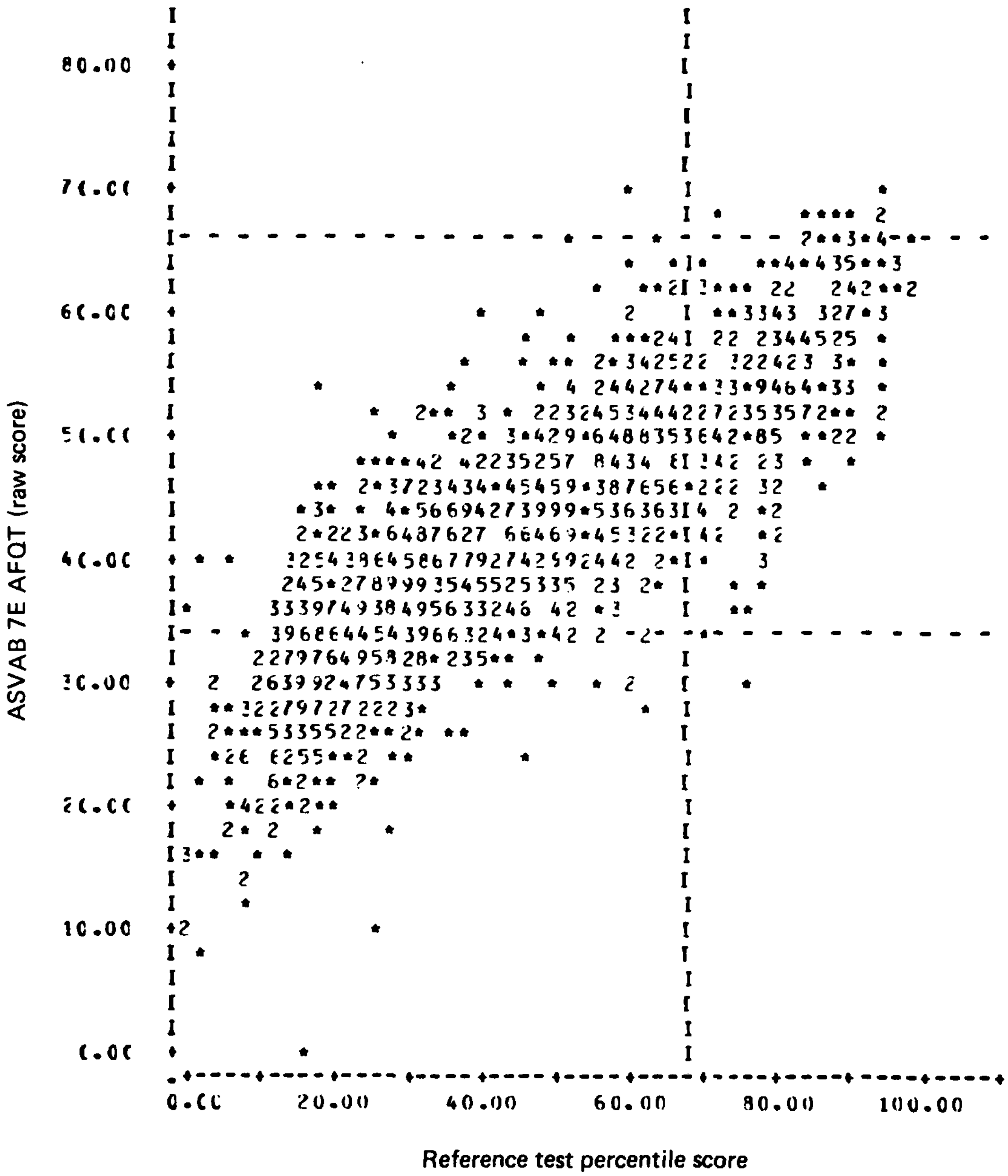


FIG. F-3: SCATTERGRAM OF ASVAB 7E AFQT SCORES VERSUS REFERENCE TEST SCORES

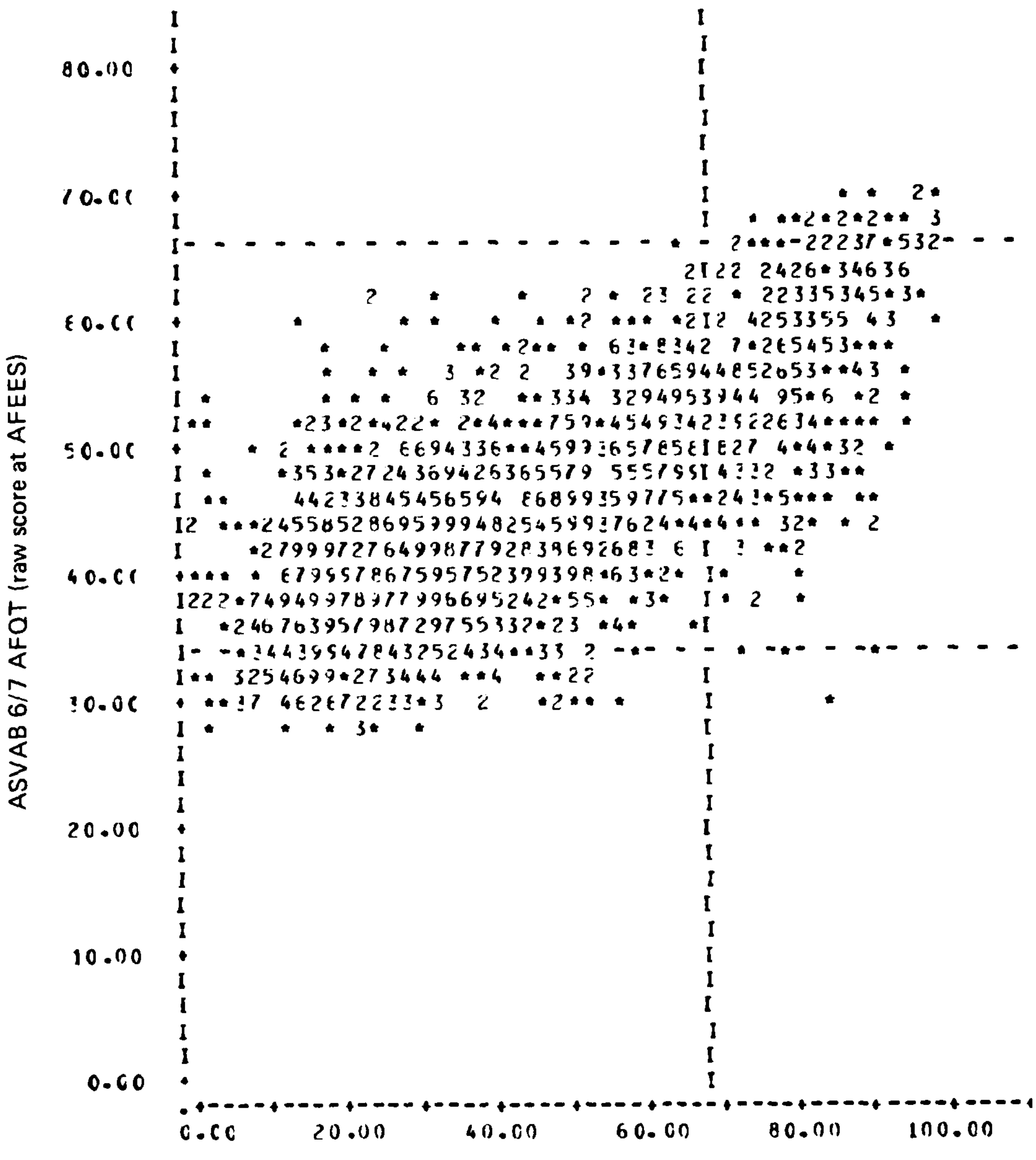


FIG. F-4: SCATTERGRAM OF ASVAB 6/7 AFQT SCORES FROM AFEES TESTING VERSUS REFERENCE TEST SCORES

AFEES testing when coaching on ASVAB might be expected. This figure shows some indication of an excess number of cases in the upper left corner, which may indicate some coaching on ASVAB 6/7.

A similar plot for ASVAB 6/7 scores from recruit depot testing is shown in figure F-5. (This is the data from which the normalization of ASVAB 6/7 is deduced.) In figure F-5, there also seems to be an excess of cases in the upper left corner, again indicating that some recruits in this sample may have been coached. Although the effect of coaching does not appear to be large, this appendix examines it in some detail.

In order to examine the coaching effect, we use an internal consistency check developed (reference F-1) for detecting coaching on ASVAB. Enlistment in the Marine Corps is determined only by performance on subtests that make up the AFQT part of ASVAB; hence, we expect coaching will focus on this part. Reference F-1 shows that scores on the AFQT part of ASVAB can be predicted with reasonable accuracy from the non-AFQT parts of ASVAB. Comparing scores on the AFQT part (on which coaching may have occurred) with predicted AFQT scores (from a part of the ASVAB on which coaching is unlikely) provides some measure of the amount of coaching that occurs. The predicted AFQT is calculated from an equation taken from reference F-1:

$$\text{Predicted AFQT} = 10.15 + 0.816\,(\text{GI+GS+MC+MK}).$$

The difference in AFQT and predicted AFQT is calculated as

$$\Delta = \text{AFQT} - \text{predicted AFQT}.$$

An illustration of the expected distribution of Δ is shown in figure F-6. In a sample containing recruits who have not been coached, the distribution is expected to be symmetric about zero. If recruits are coached on the AFQT part of ASVAB but not on the parts from which the predicted AFQT is calculated, then the values of Δ tend to be positive. The positive excess can be estimated by folding the Δ distribution about zero and subtacting the negative side from the positive side. We use the resulting excess positive group as an estimate of the cases of coaching.

The estimation technique just discussed was applied to ASVAB 6/7 scores from AFEES testing; results are given in table F-1. We estimated that 16.2 percent of the recruits were coached. We applied the same methodology to ASVAB 6/7 scores from testing at recruit depots. The results are given in table E-2 and indicate that 14.6 percent of the sample still retains effects of coaching

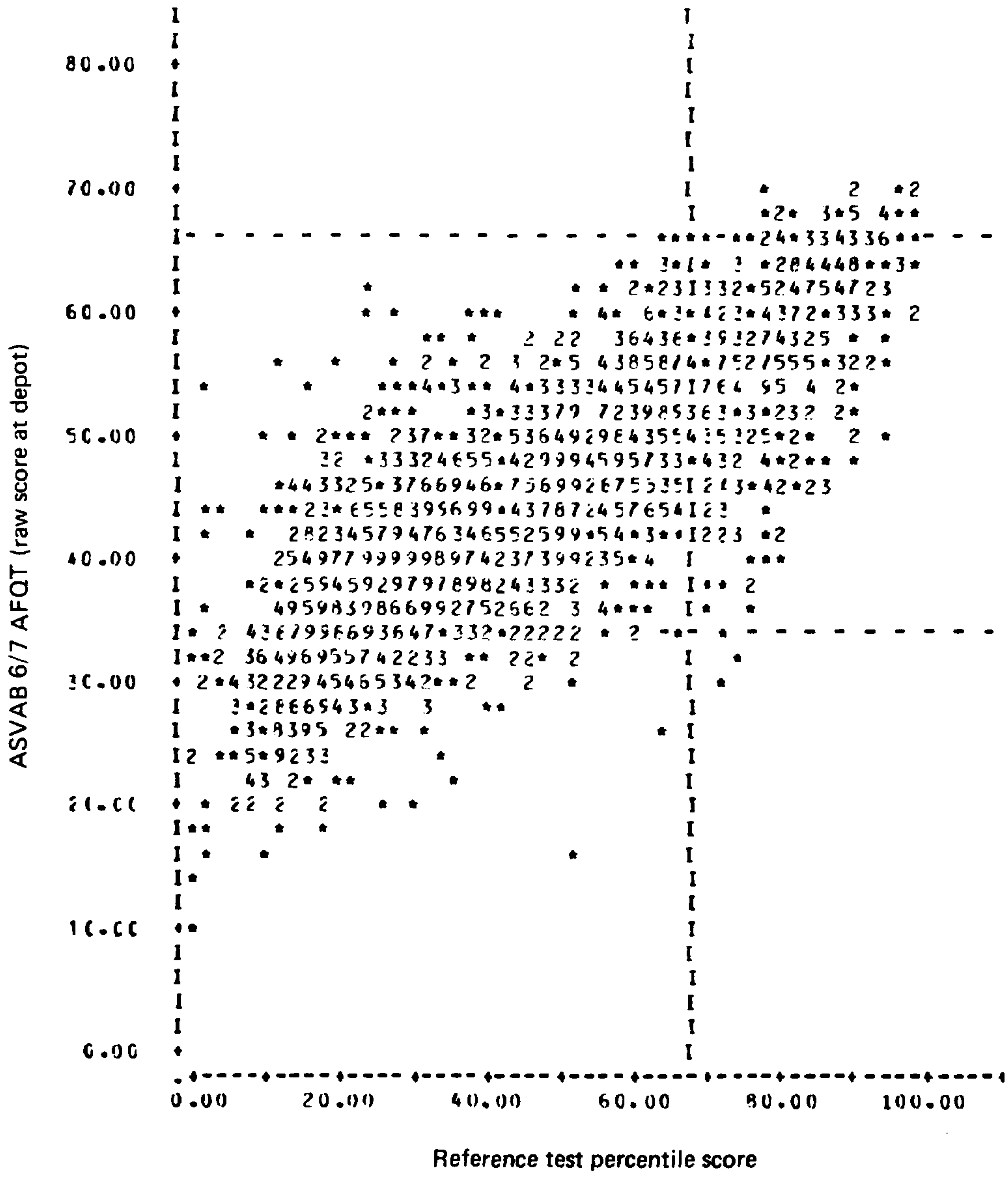


FIG. F-5: SCATTERGRAM OF ASVAB 6/7 AFQT SCORES FROM RECRUIT DEPOT TESTING VERSUS REFERENCE TEST SCORES when retested at recruit depots. This latter result is somewhat surprising because we anticipated that the recruits would not have retained the effects of coaching for the 1-to-6-month period that elapsed between testing at AFEES and at recruit depots.[1]

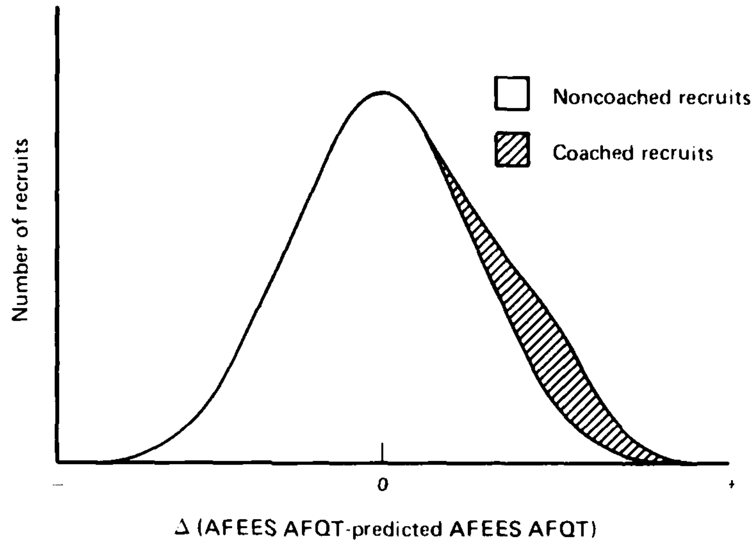


FIG. F-6: ILLUSTRATION OF USE OF Δ DISTRIBUTION TO ESTIMATE AMOUNT OF COACHING

Because it appears that some of the scores on the ASVAB 6/7 tests at recruit depots are inflated by coaching, we next address the issue of what effect this has on the normalization results. Because the coaching could only affect ASVAB 6/7 scores, we focused on the data in sample 5. In doing so we removed from sample 5 those recruits who were thought to have been coached. Then we recalculated the normalization of ASVAB 6/7. Comparing the normalization before and after removal of the suspect cases gives some indication of the effect of coaching on the norming results.

---

[1] Identical methodologies applied to larger samples of Marine Corps recruits enlisting during 1977 and 1978 have typically yielded estimates of a 3 percent coaching effect on tests taken at recruit depots. The sample used in this report was collected in February, March, and April 1979, and may be atypical in the sense that fewer of the recruits may have been in delayed entry programs; hence, only a short period of time may have elapsed between testing at AFEES and testing at recruit depots. If so, this could explain the unexpectedly high retention of coached material.

TABLE F-1

ESTIMATION OF AMOUNT OF COACHING IN AFEES TEST SCORES FROM SAMPLE 5

Δ interval	Positive half of Δ distribution	Negative half of Δ distribution	Excess positive
0	199	199	0
1	149	110	39
2	113	124	-11
3	115	90	25
4	123	98	25
5	85	80	5
6	93	49	44
7	101	61	40
8	62	31	31
9	55	38	17
10	59	23	36
11	40	21	19
12	26	13	13
13	22	12	10
14	21	9	12
15	24	7	17
16	14	5	9
17	7	2	5
18	12	1	11
19	2	1	1
≥ 20	11	1	10
Total	1,333	975	358[a]

[a]Of 2,208 recruits, 358 were coached, which is 16.2 percent.

Referring to figure F-6 and to table F-1, we see that if we excluded from the sample those cases with a large Δ from AFEES testing, we can expect to have removed a significant percentage of the contamination due to coaching. Normalization would then be carried out using scores from recruit depot testing as always. Accordingly, we formed two subsamples from sample 5--one consisting of all cases with Δ from AFEES testing < 10, and the second with Δ from AFEES testing < 0.0. In the first case we estimated that we removed 100 percent of the coached cases along with 50 percent of the noncoached cases.

TABLE F-2

ESTIMATION OF AMOUNT OF COACHING IN DEPOT TEST SCORES FROM SAMPLE 5

Δ interval	Positive half of Δ distribution	Negative half of Δ distribution	Excess positive
0	199	198	1
1	132	118	14
2	137	119	18
3	134	100	34
4	130	89	41
5	119	73	46
6	102	76	26
7	79	53	26
8	47	45	2
9	56	22	34
10	44	25	19
11	48	20	28
12	20	17	3
13	21	9	12
14	12	8	4
15	8	12	-4
16	12	4	8
17	5	1	4
18	1	0	1
19	4	1	3
≥20	5	3	2
Total	1,315	993	322[a]

[a] Of 2,208 recruits, 322 were coached, which is 14.6 percent.

Weight factors were calculated to stratify the subsamples on the reference test and thereby simulate the mobilization population. These calculations are shown in tables F-3 and F-4.

The weight factors from tables F-3 and F-4 were applied to the subsamples. The resulting stratified cumulative frequencies of ASVAB 6/7 scores are shown in table F-5. For the case of $\Delta < 10$

(40 percent of coached cases removed),[1] the subsample agrees well with the full sample. In the case of $\Delta < 0$ (100 percent of coached cases removed), the agreement is not perfect but subjectively quite close. The distributions are plotted in figure F-7.

TABLE F-3

CALCULATION OF WEIGHT FACTORS FOR $\Delta < 10$

Reference test percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	22	98.5	4.48
6-10	48	98.5	2.05
11-15	124	98.5	.79
16-20	139	98.5	.71
21-25	145	98.5	.68
26-30	120	98.5	.82
31-35	143	98.5	.69
36-40	162	98.5	.61
41-45	64	98.5	1.54
46-50	159	98.5	.62
51-55	129	98.5	.76
56-60	153	98.5	.64
61-65	149	98.5	.66
66-70	55	98.5	1.79
71-75	73	98.5	1.35
76-80	99	98.5	.99
81-85	90	98.5	1.09
86-90	58	98.5	1.70
91-95	29	98.5	3.40
96-100	9	98.5	10.94
	1,970	1,970	

[a]Column (3) divided by column (2).

[1]From table F-1 we see that restricting the sample to $\Delta < 10$ excludes 143 of the estimated 358 coached cases.

TABLE F-4

CALCULATION OF WEIGHT FACTORS FOR Δ < 0

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	7	46.15	6.59
6-10	13	46.15	3.55
11-15	41	46.15	1.13
16-20	69	46.15	.67
21-25	76	46.15	.61
26-30	65	46.15	.71
31-35	72	46.15	.64
36-40	83	46.15	.56
41-45	36	46.15	1.28
46-50	84	46.15	.55
51-55	66	46.15	.70
56-60	76	46.15	.61
61-65	57	46.15	.81
66-70	30	46.15	1.54
71-75	33	46.15	1.40
76-80	40	46.15	1.15
81-85	39	46.15	1.18
86-90	26	46.15	1.78
91-95	8	46.15	5.77
96-100	2	46.15	23.08
	923	923	

[a]Column (3) divided by column (2).

A chi-squared test (reference F-2) for the homogeneity of the frequency distributions that make up table F-5 is shown in table F-6. The result of the test shows that the probability of observing differences this large by change in parallel samples is quite large.

This analysis found that all forms of ASVAB can use the same conversion table. It has been argued that our results for forms 6/7 are biased by test compromise and that if this effect were removed then ASVAB 6/7 would have a markedly different norming

TABLE F-5

NORMALIZATION FOR DIFFERENT RESTRICTIONS ON Δ PARAMETER

ASVAB 6/7 AFQT raw score	Cumulative percentage: All (2,208 cases)	Δ < 10 (1,968 cases)	Δ < 0 (924 cases)
0-15	0.5	0.7	0.7
16-17	0.8	1.1	0.7
18-19	1.2	1.4	0.8
20	1.8	2.1	0.9
21	2.1	2.5	1.4
22	2.6	3.1	2.0
23	[illegible].4	3.6	2.6
24	[illegible]5	4.6	3.5
25	[illegible].2	5.4	4.5
26	5.8	6.0	4.8
27	6.9	7.1	5.6
28	7.9	8.0	6.0
29	9.2	9.3	8.0
30	11.3	11.6	11.4
31	13.1	13.3	13.7
32	15.2	15.4	15.8
33	17.3	17.2	17.6
34	19.6	19.7	19.7
35	21.9	22.0	22.7
36	24.5	24.8	24.8
37	26.6	27.0	27.3
38	28.9	29.6	29.5
39	31.9	32.8	33.0
40	34.8	35.7	36.0
41	37.5	38.3	39.3
42	40.2	41.1	41.6
43	42.9	43.7	43.9
44	45.7	46.5	47.0
45	49.2	49.7	50.3
46	52.0	52.3	53.7
47	54.3	54.5	55.9
48	57.3	57.5	59.2
49	60.0	60.1	61.3
50	62.5	62.4	63.9
51	64.4	64.4	66.3
52	66.8	66.8	68.7
53	69.2	69.1	70.9
54	71.4	71.3	73.3
55	74.6	74.0	75.8
56	76.4	75.9	77.5
57	78.0	77.5	79.3
58	79.8	79.3	80.5
59	81.5	81.1	82.7
60	84.3	84.0	84.0
61	86.1	85.9	85.5
62	88.1	87.8	87.5
63	89.6	89.3	89.0
64	92.3	92.1	92.5
65	93.9	93.7	94.8
66	96.4	96.2	95.4
67	97.2	97.2	97.4
68	98.5	98.5	97.5
69	99.9	99.9	100.0
70	100.0	100.0	100.0

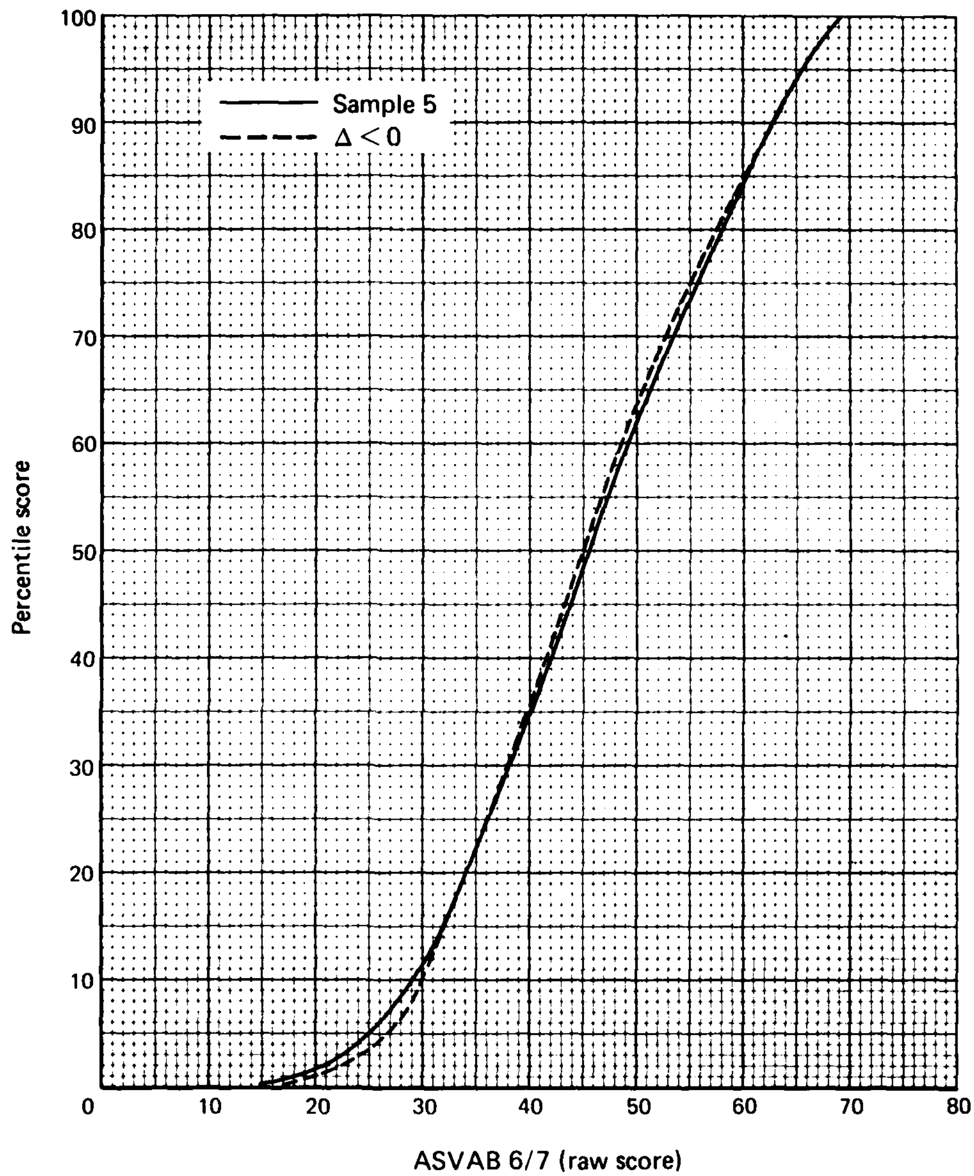


FIG. F-7: COMPARISON OF NORMALIZATION OF ASVAB 6/7 AFQT FROM FULL SAMPLE 5 AND FROM A SUBSAMPLE FROM WHICH RECRUITS SUSPECTED OF BEING COACHED ARE REMOVED

TABLE F-6

GROUPED DISTRIBUTIONS FOR HOMOGENEITY TEST[a]

ASVAB 6/7 AFQT interval (raw score) (1)	Number of cases All (2)	$\Delta < 10$ (3)[a]	$\Delta < 0$ (4)[b]
0-19	27	27	8
20-22	30	33	11
23-25	57	45	24
26-28	60	52	14
29-31	114	104	71
32-34	144	126	55
35-37	154	145	70
38-40	182	171	81
41-43	177	158	73
44-46	201	168	90
47-49	177	154	70
50-52	151	132	69
53-55	171	142	66
56-58	115	104	42
59-61	139	129	46
62-64	137	123	65
65-67	109	100	45
68-70	61	55	24
Total	2,206	1,968	924

[a]Chi-squared for a comparison of columns (2) and (3) is 2.8 for 17 degrees of freedom. The probability of differences this large by chance is about 1.00.

[b]Chi-squared for a comparison of columns (2) and (4) is 17.1 for 17 degrees of freedom. The probability of differences this large by chance is about 0.45.

[c]Does not sum to sample 5 total of 2,208 due to rounding of weighted frequency.

curve than ASVAB 6E or 7E.[1] Since June 1979, ASVAB 6/7 and ASVAB 6E/7E have all been used at AFEES, and it is reasonable to assume that by now all are equally compromised.[2] If our ASVAB 6/7 norms are seriously in error relative to our results for ASVAB 6E/7E then one would expect that mean ASVAB test scores at AFEES would be similar for ASVAB 6/7 and ASVAB 6E/7E, when the latter were first introduced in June 1979 and to diverge later as all forms become equally compromised. An examination of recent data for Marine Corps recruits on ASVAB tests administered at AFEES does not show this divergence (table F-7). We believe this observation strengthens our contention that there is no significant bias in our ASVAB 6/7 norming results due to test compromise.

TABLE F-7

COMPARISON OF MEAN SCORES FROM COMPROMISED AND UNCOMPROMISED ASVAB FORMS

	Mean ASVAB AFQT percentile score from AFEES testing	
Period tested at AFEES	ASVAB 6/7	ASVAB 6E/7E
June-September 1979	53.8 ± 0.2[a]	53.8 ± 0.4[b]
October-December 1979	55.3 ± 0.3[c]	54.8 ± 0.5[d]

[a]Sample contains 6,887 Marine Corps recruits.
[b]Sample contains 1,755 Marine Corps recruits.
[c]Sample contains 2,391 Marine Corps recruits.
[d]Sample contains 1,096 Marine Corps recruits.

[1]ASVAB 6/7 had been used for about 2 years when our data set was collected and was certainly compromised. ASVAB 6E/7E were not in use and were not compromised at that time.

[2]In testimony before the House Armed Services Military Personnel Subcommittee, a recruiter stated that there wasn't a test devised "that I couldn't compromise in three months." (Navy Times, 7 June 1976). Other recruiters have given even shorter estimates of the time required.

Based on data in this appendix, we conclude that there are some cases in sample 5 that are probably distorted by coaching but that these do not seem to have had a significant effect on the normalization results for ASVAB 6/7.

REFERENCES

F-1 Center for Naval Analyses, Study 1115, "An Analysis of the Normalization and Verification of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 6 and 7," by William H. Sims, Unclassified, Apr 1978

F-2 Rao, C.R., "Linear Statistical Inference and its Applications," John Wiley and Sons Inc., New York 1965

APPENDIX G

EQUIVALENCE OF RESULTS FROM DIFFERENT LOCATIONS

APPENDIX G

EQUIVALENCE OF RESULTS FROM DIFFERENT LOCATIONS

As part of our quality control procedure in testing recruits, one of the authors visited the two test sites for a few days when data were first collected. Each test site appeared to be conducting the testing correctly. In this appendix we examine the data to determine if there is any difference in the normalization results between the two test sites that might indicate that at some time during the testing one of the sites may have deviated from the proper procedure.

We used the data from sample 5 to look for a location effect. It was broken into two subsamples--those recruits tested at Parris Island and those tested at San Diego. Separate weight factors were calculated for each subsample to stratify them on the reference test. These calculations are shown in tables G-1 and G-2.

Using the weight factors in tables G-1 and G-2, we stratified the subsamples and made cumulative frequency distributions of the ASVAB 6/7 AFQT score, as shown in table G-3. We then ran a chi-squared test for the homogeneity of the two frequency distributions, as illustrated in table G-4. We found a chi-squared of 24.1 for 17 degrees of freedom, which indicates that the probability of differences that large occurring by change is about 0.12. There is, therefore, no compelling reason to doubt that the two test sites followed the same procedures throughout the data collection phase of the study.

TABLE G-1

CALCULATION OF WEIGHT FACTORS FOR PARRIS ISLAND SUBSAMPLE

Reference test percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	16	52.8	3.300
6-10	33	52.8	1.600
11-15	80	52.8	0.660
16-20	86	52.8	0.614
21-25	90	52.8	0.587
26-30	61	52.8	0.866
31-35	74	52.8	0.714
36-40	81	52.8	0.652
41-45	39	52.8	1.354
46-50	90	52.8	0.587
51-55	59	52.8	0.895
56-60	80	52.8	0.660
61-65	70	52.8	0.754
66-70	30	52.8	1.760
71-75	34	52.8	1.553
76-80	50	52.8	1.056
81-85	44	52.8	1.200
86-90	24	52.8	2.200
91-95	11	52.8	4.800
96-100	4	52.8	13.200
Total	1,056		

[a]Column (3) divided by column (2).

TABLE G-2

CALCULATION OF WEIGHT FACTORS FOR SAN DIEGO SUBSAMPLE

Reference test percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	13	57.6	4.431
6-10	23	57.6	2.504
11-15	72	57.6	0.800
16-20	83	57.6	0.694
21-25	81	57.6	0.711
26-30	76	57.6	0.758
31-35	88	57.6	0.655
36-40	94	57.6	0.613
41-45	32	57.6	1.800
46-50	83	57.6	0.694
51-55	79	57.6	0.729
56-60	84	57.6	0.686
61-65	97	57.6	0.594
66-70	26	57.6	2.215
71-75	50	57.6	1.152
76-80	57	57.6	1.011
81-85	52	57.6	1.108
86-90	35	57.6	1.646
91-95	22	57.6	2.618
96-100	5	57.6	11.52
Total	1,152		

[a]Column (3) divided by column (2).

TABLE G-3

TESTING LOCATION EFFECT
(sample 5)

ASVAB 6/7 AFQT raw score	Cumulative frequency	
	Parris Island	San Diego
0-15	0.3	0.8
16-17	0.4	1.4
18-19	0.8	1.7
20	1.7	1.9
21	1.9	2.5
22	2.4	2.8
23	3.1	3.9
24	3.9	4.8
25	4.8	5.6
26	5.4	6.3
27	6.9	6.9
28	8.1	7.6
29	9.4	8.8
30	11.5	11.1
31	13.3	12.8
32	15.2	15.2
33	17.1	17.5
34	18.8	20.5
35	21.2	22.7
36	24.1	24.9
37	26.1	26.9
38	28.8	28.9
39	31.9	31.9
40	34.8	34.6
41	37.1	37.9
42	39.7	40.7
43	42.7	43.1
44	46.0	45.5
45	49.4	49.1
46	52.1	51.7
47	54.5	54.0
48	57.0	57.5
49	59.3	60.4
50	61.9	62.8
51	64.2	64.5
52	66.3	67.2
53	68.4	69.8
54	70.7	72.0
55	73.7	75.4
56	75.2	77.4
57	76.6	79.1
58	78.6	80.9
59	80.7	82.1
60	85.0	83.9
61	86.3	86.0
62	87.8	88.3
63	90.0	89.4
64	92.6	91.9
65	95.8	92.6
66	97.4	95.8
67	97.7	97.0
68	99.6	97.7
69	99.8	100.0
70	100.0	100.0

TABLE G-4

GROUPED DISTRIBUTIONS FOR HOMOGENEITY TEST[a]
(sample 5)

ASVAB 6/7 AFQT (raw score)	Frequency	
	Parris Island	San Diego
0-19	8	18
20-22	16	13
23-25	25	33
26-28	35	23
29-31	55	60
32-34	58	90
35-37	77	74
38-40	92	89
41-43	83	98
44-46	100	99
47-49	75	100
50-52	75	78
53-55	78	94
56-58	52	65
59-61	81	59
62-64	68	69
65-67	53	58
68-70	24	35
Total	1,055[b]	1,155[c]

[a]Chi-squared is 24.1 witn 17 degrees of freedom. The probability of differences this large by chance is about 0.12.

[b]Does not sum to 1,056 due to rounding of weighted frequencies.

[c]Does not sum to 1,152 due to rounding of weighted frequencies.

APPENDIX H

EFFECT OF TEST FATIGUE

APPENDIX H

EFFECT OF TEST FATIGUE

This appendix looks at the effect of test fatigue on scores from the three-test series. The series consisted of AFQT-7A (1 hour), AFQT from ASVAB 6E/7E (1 hour), and the entire ASVAB 6/7 (3 hours). Breaks were given between testing sessions, and in some cases, the testing was spread over 2 days. Nonetheless, it is reasonable to think that fatigue may have contributed to lower test scores on the last test in the series and that this may have biased the normalization results. The counterbalanced design described in appendix B tends to reduce this possible source of bias. However, in this appendix we briefly examine the data to see if test fatigue biases norming results.

We examined a subsample of sample 5 that contained recruits who took the reference test and ASVAB 6/7 either first or second in the three-test series. We assumed these recruits would not suffer as much test fatigue as the average recruit in sample 5. We stratified the low-fatigue subsample on the reference test as shown in table H-1. Table H-2 shows the resulting cumulative frequency distribution of ASVAB 6/7 AFQT scores compared to those of the full sample 5. A test for the homogeneity of the low-fatigue subsample and sample 5 (see table H-3) indicated that the probability of observing differences that large by chance was about 0.04. Hence, we cannot reject with high confidence the hypothesis that the two samples are parallel. The practical consequences of any test fatigue effect, if any, is not large (table H-2).

TABLE H-1

CALCULATION OF WEIGHT FACTORS FOR LOW-FATIGUE SUBSAMPLE

Reference test percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	8	35.15	4.394
6-10	12	35.15	2.929
11-15	50	35.15	0.703
16-20	55	35.15	0.639
21-25	54	35.15	0.651
26-30	46	35.15	0.764
31-35	50	35.15	0.703
36-40	49	35.15	0.717
41-45	26	35.15	1.352
46-50	52	35.15	0.676
51-55	50	35.15	0.703
56-60	49	35.15	0.717
61-65	52	35.15	0.676
66-70	18	35.15	1.953
71-75	24	35.15	1.465
76-80	44	35.15	0.799
81-85	29	35.15	1.212
86-90	21	35.15	1.674
91-95 96-100	14	70.30	5.020
Total	703		

[a]Column (3) divided by column (2).

TABLE II-2

TEST FATIGUE EFFECTS
(sample 5)

ASVAB 6/7 AFQT raw score	Cumulative frequency: Full sample	Cumulative frequency: Low-fatigue group
0-15	0.5	0.6
16-17	0.8	0.7
18-19	1.2	1.0
20	1.8	1.0
21	2.1	1.3
22	2.6	1.8
23	3.4	2.1
24	4.3	3.3
25	5.2	4.9
26	5.8	5.2
27	6.9	6.3
28	7.9	7.7
29	9.2	9.4
30	11.3	12.3
31	13.1	14.9
32	15.2	17.0
33	17.3	18.7
34	19.6	20.8
35	21.9	23.4
36	24.5	25.3
37	26.6	27.3
38	28.9	29.7
39	31.9	33.1
40	34.8	35.6
41	37.5	37.8
42	40.2	40.0
43	42.9	42.3
44	45.7	46.1
45	49.2	49.1
46	52.0	51.8
47	54.3	53.5
48	57.3	57.3
49	60.0	60.2
50	62.5	62.2
51	64.4	64.8
52	66.8	67.0
53	69.2	69.8
54	71.4	71.4
55	74.6	76.6
56	76.4	78.8
57	78.0	80.7
58	79.8	82.6
59	81.5	85.1
60	84.3	87.8
61	86.1	88.5
62	88.1	90.8
63	89.6	91.9
64	92.3	93.5
65	93.9	95.6
66	96.4	98.3
67	97.2	99.7
68	98.5	100.0
69	99.9	100.0
70	100.0	100.0

TABLE H-3

GROUPED DISTRIBUTIONS FOR HOMOGENEITY TEST[a]
(sample 5)

ASVAB 6/7 AFQT interval (raw score) (1)	Frequency: Full sample (2)	Frequency: Low-fatigue group (3)
0-19	27	7
20-22	30	6
23-25	57	22
26-28	60	20
29-31	114	50
32-34	144	42
35-37	154	46
38-40	182	59
41-43	177	47
44-46	201	67
47-49	177	59
50_52	151	48
53-55	171	67
56-58	115	42
59-61	139	42
62-64	137	34
65-67	109	44
68-70	61	2
Total	2,206[b]	704[c]

[a]Chi-squared is 28.7 for 17 degrees of freedom. The probability of differences this large by chance is about 0.04.

[b]Does not sum to 2,208 due to rounding weighted frequencies.

[c]Does not sum to 703 due to rounding weighted frequencies.

APPENDIX I

EFFECTS OF PRESELECTION ON NORMALIZATION

APPENDIX I

EFFECTS OF PRESELECTION ON NORMALIZATION

Another possible source of bias is preselection. Ideally, the sample used for normalization should contain individuals with a wide range of mental aptitudes. In fact, because we had to use Marine Corps recruits for this analysis rather than applicants, those who made low scores[1] on the ASVAB given at AFEES would have been rejected for military service and would not be present in our sample. Figure I-1 illustrates preselection. This appendix examines the extent to which preselection biases normalization results based on scores obtained by retesting the recruits at recruit depots.

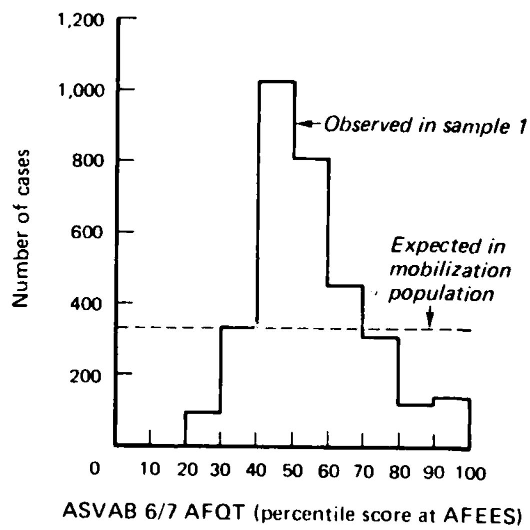


FIG. I-1: ILLUSTRATION OF PRESELECTION ON ASVAB AT AFEES

[1]The Marine Corps requires recruits to score a minimum of the 21st percentile on the AFQT part of ASVAB. Additional restrictions on the GT (general technical) composite correspond approximately to the 25th percentile for high school graduates and the 40th percentile for non-high school graduates.

To examine the preselection effect, we used sample 1 data and made successively more restrictive cuts on the AFQT score recruits made at AFEES. If further restrictions significantly change the normalization results, then we may infer that the original restriction at AFEES may have biased our normalization. Cuts on sample 1 were made to exclude cases scoring at or below the 30th, 40th, or 50th percentile on the ASVAB 6/7 AFQT at AFEES. The full sample 1 corresponds to a restriction at the 20th percentile. We then stratified the resulting subsamples on the reference tests using the weight factors shown in tables I-1, I-2, and I-3.

The cumulative frequencies of ASVAB 6/7 AFQT scores based on recruit depot testing are shown in table I-4. The results for restrictions at the 20th, 40th, and 50th percentiles are graphed and shown in figure I-2. The higher the restriction, the more the lower end of the normalization curve moves toward harder norms. The bias does not seem to be large until the restriction removes all cases below the 50th percentile. Nonetheless, the bias does seem to exist for less restrictive cuts and extends to at least the 50th percentile on the resulting normalization.

These data do not allow us to quantify the bias, but it seems reasonable to conclude that a bias exists and that although it appears to be small, it is not negligible.

TABLE I-1

CALCULATION OF WEIGHT FACTORS FOR AFEES AFQT > 30th PERCENTILE

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	39	160.05	4.104
6-10	68	160.05	2.354
11-15	206	160.05	0.777
16-20	262	160.05	0.611
21-25	240	160.05	0.667
26-30	199	160.05	0.804
31-35	245	160.05	0.653
36-40	252	160.05	0.635
41-45	111	160.05	1.442
46-50	253	160.05	0.633
51-55	201	160.05	0.796
56-60	249	160.05	0.643
61-65	248	160.05	0.645
66-70	76	160.05	2.106
71-75	115	160.05	1.392
76-80	151	160.05	1.060
81-85	146	160.05	1.096
86-90	86	160.05	1.861
91-95	44	160.05	3.638
96-100	10	160.05	16.005
Total	3,201	3,201	

[a]Column (3) divided by column (2).

TABLE I-2

CALCULATION OF WEIGHT FACTORS FOR
AFEES AFQT > 40th PERCENTILE

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	33	143.3	4.342
6-10	38	143.3	3.771
11-15	149	143.3	0.962
16-20	188	143.3	0.762
21-25	183	143.3	0.783
26-30	171	143.3	0.838
31-35	219	143.3	0.654
36-40	228	143.3	0.629
41-45	105	143.3	1.365
46-50	238	143.3	0.602
51-55	196	143.3	0.731
56-60	245	143.3	0.585
61-65	248	143.3	0.578
66-70	76	143.3	1.886
71-75	114	143.3	1.257
76-80	150	143.3	0.955
81-85	145	143.3	0.988
86-90	86	143.3	1.666
91-95	44	143.3	3.257
96-100	10	143.3	14.330
Total	2,866	2,866	

[a]Column (3) divided by column (2).

TABLE I-3

CALCULATION OF WEIGHT FACTORS FOR AFEES AFQT > 50th PERCENTILE

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
1-5	12	91.8	7.650
6-10	6	91.8	15.300
11-15	44	91.8	2.086
16-20	53	91.8	1.732
21-25	67	91.8	1.370
26-30	70	91.8	1.311
31-35	112	91.8	0.820
36-40	115	91.8	0.798
41-45	53	91.8	1.732
46-50	141	91.8	0.651
51-55	139	91.8	0.660
56-60	190	91.8	0.483
61-65	232	91.8	0.396
66-70	71	91.8	1.293
71-75	108	91.8	0.850
76-80	140	91.8	0.656
81-85	144	91.8	0.638
86-90	85	91.8	1.080
91-95	44	91.8	2.086
96-100	10	91.8	9.180
Total	1,836	1,836	

[a]Column (3) divided by column (2).

TABLE I-4

CUMULATIVE FREQUENCY OF ASVAB 6/7 AFQT FOR VARIOUS RESTRICTIONS ON AFQT SCORE AT AFEES

ASVAB 6/7 AFQT raw score	Cumulative frequency of ASVAB 6/7 AFQT			
	AFQT > 20	AFQT > 30	AFQT > 40	AFQT > 50
14-15	0.4	0.5	0.6	1.2
16-17	0.7	0.8	0.9	1.3
18-19	1.2	1.3	1.3	1.5
20	1.7	1.4	1.4	1.5
21	2.2	1.8	1.7	1.5
22	2.6	2.1	2.0	1.6
23	3.2	2.7	2.5	2.0
24	3.8	3.4	3.1	2.1
25	4.7	4.2	3.9	2.8
26	5.4	5.0	4.7	4.0
27	6.3	6.0	5.2	4.1
28	7.3	6.9	5.9	4.3
29	8.4	7.9	6.8	4.8
30	10.2	9.5	7.9	5.1
31	12.0	11.1	9.4	6.1
32	13.8	12.9	11.3	7.0
33	15.6	14.7	13.1	9.0
34	17.5	16.6	14.9	9.7
35	19.6	18.8	16.9	11.2
36	21.8	21.0	19.1	12.8
37	24.0	23.2	21.1	13.9
38	26.5	25.8	23.8	15.6
39	29.4	28.7	26.9	17.4
40	32.1	31.5	29.5	19.1
41	35.0	34.5	32.5	21.4
42	37.6	37.2	35.3	24.4
43	40.1	39.7	38.0	27.3
44	43.1	42.8	41.4	31.8
45	46.4	46.1	45.0	36.3
46	49.3	49.0	48.0	39.6
47	51.8	51.6	50.6	42.5
48	55.0	54.8	54.1	47.3
49	58.3	58.1	57.4	51.0
50	61.0	60.9	60.4	55.4
51	63.1	63.1	62.6	58.2
52	65.4	65.3	64.9	61.0
53	67.8	67.8	67.4	64.2
54	70.3	70.2	70.4	67.2
55	73.2	73.2	73.0	70.8
56	75.1	75.0	74.9	73.2
57	76.7	76.7	76.6	75.2
58	78.8	78.8	78.7	77.6
59	80.5	80.5	80.4	79.4
60	83.5	83.5	83.5	82.8
61	85.4	85.4	85.3	84.7
62	87.7	87.7	87.6	87.1
63	89.2	89.2	89.2	89.1
64	91.6	91.6	91.6	91.5
65	93.2	93.2	93.2	93.2
66	95.8	95.8	95.8	95.7
67	96.8	96.8	96.8	96.8
68	98.7	98.7	98.7	98.7
69	99.9	99.9	99.9	99.9
70	100.0	100.0	100.0	100.0

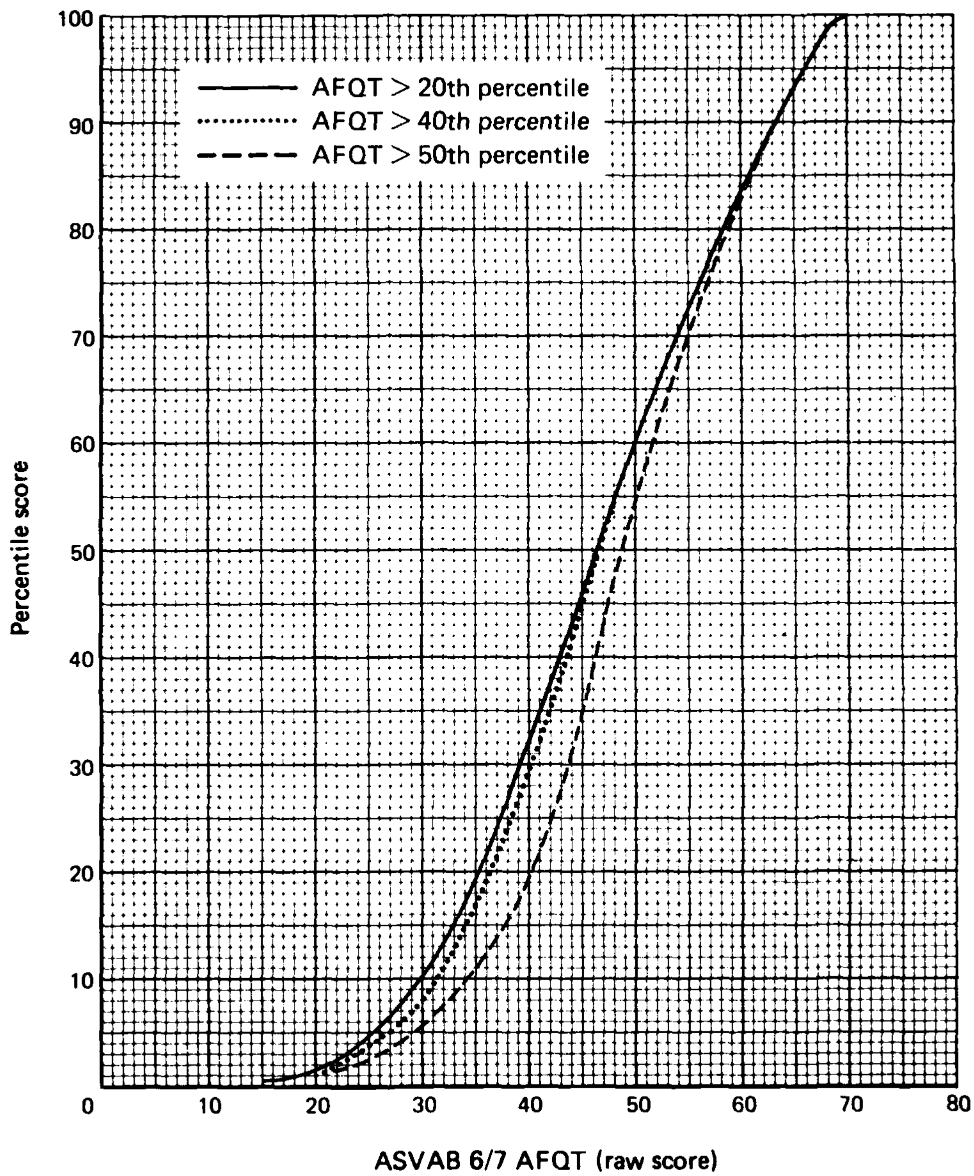


FIG. I-2: EFFECT OF PRESELECTION AT AFEES ON NORMALIZATION RESULTS

APPENDIX J

ADJUSTMENTS FOR EFFECTS OF SAMPLE TRUNCATION ON STRATIFIED NORMING RESULTS

APPENDIX J

ADJUSTMENTS FOR EFFECTS OF SAMPLE TRUNCATION ON STRATIFIED NORMING RESULTS

In this appendix we use a full-range data set to examine the effect of sample truncation on norming results from the stratification method. First, we applied the stratification method to the full-range data and obtained a normalization curve. Then, we truncated the sample to closely simulate the truncation in the CNA sample and obtained a second normalization curve. The difference in the two curves is the effect of truncation and could be used to correct for the effects of truncation on the CNA normalization curve.

The full-range data set obtained from DoD consisted of results from administering two tests to each of a sample of applicants for enlistment at AFEES. The CNA data set consisted of a test given at AFEES on which the sample was truncated, followed by two tests given to the truncated sample once they arrived at the Marine Corps Recruit Depots (MCRD). Hence, we need to simulate the effects of a three-test system using data from only two tests.

This simulation can be done by using the Pseudo AFQT developed by reference J-1. Reference J-1 finds that in addition to the AFQT test embodied in the ASVAB there is also a Pseudo AFQT. The Pseudo AFQT can be constructed from parts of the ASVAB that do not make up the AFQT and, hence, can be viewed as a separate test. However, it has a very high correlation[1] with the AFQT and may be considered a good proxy. The Pseudo AFQT is defined as

$$\text{Pseudo AFQT} = GI+GS+MC+MK.$$

It may be used to accurately predict an alternative AFQT score for each applicant. We then truncated the full-range sample on the predicted AFQT score to simulate the truncation of the CNA data set. The AFQT score and the reference test experience incidental selection similar to that occurring in the CNA data set. The ASVAB is then normed using these incidentally selected variables and compared with those from the nontruncated full-range data set. The procedure is illustrated in figure J-1.

The Pseudo AFQT and AFQT were equated by the equipercentile method using the full-range data sample. The details are given in annex J-1. In figure J-2 we show distribution of AFQT and the AFQT

[1]The correlation between the Pseudo AFQT and AFQT in the full-range data set is 0.87.

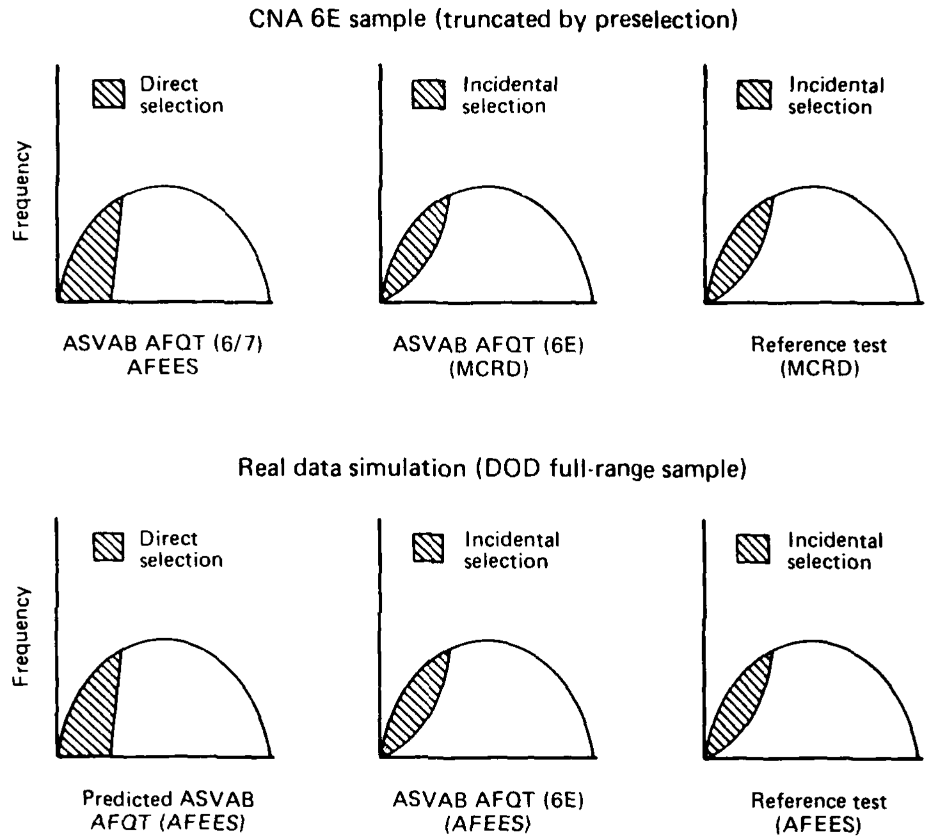


FIG. J-1: ILLUSTRATION OF REAL DATA SIMULATION OF SAMPLE TRUNCATION

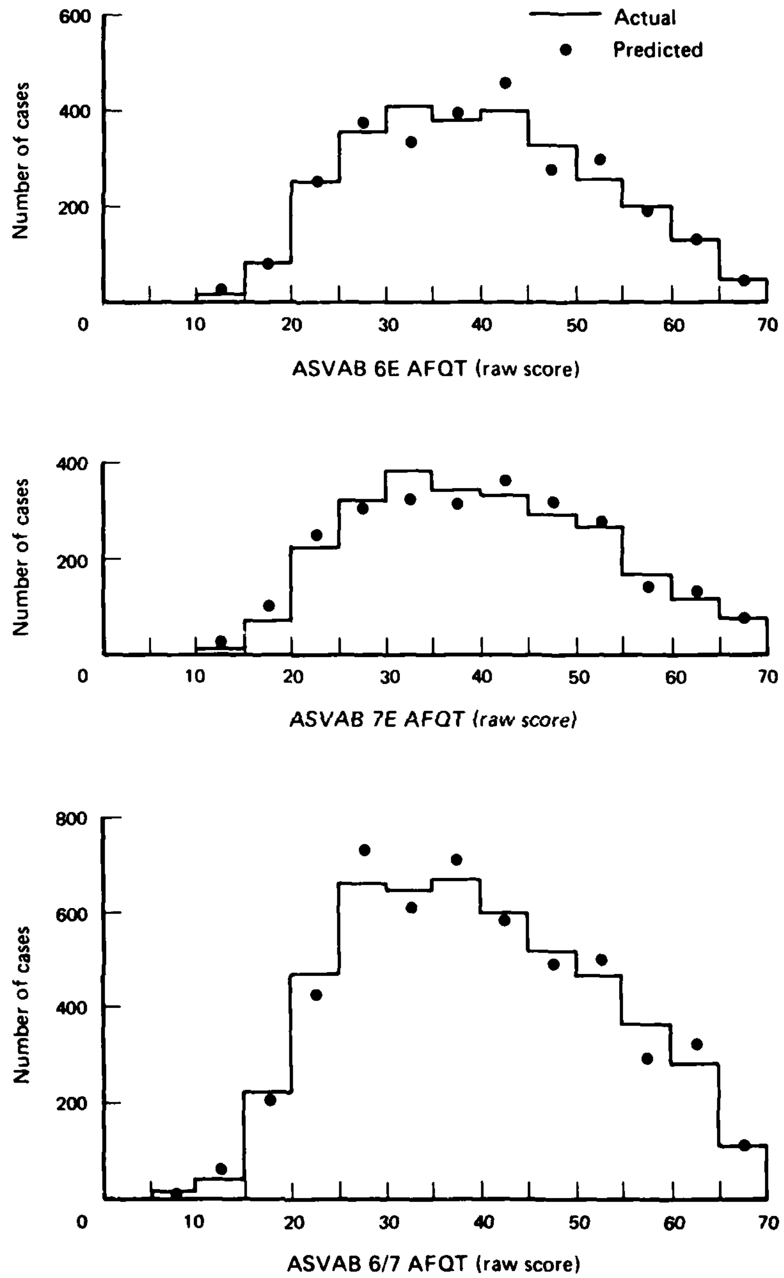


FIG. J-2: COMPARISON OF DISTRIBUTION IN ACTUAL AND PREDICTED AFQT FROM DOD FULL-RANGE SAMPLE

predicted from the Pseudo AFQT. The two distributions are very similar, which indicates that we have successfully created a three-test system from two tests.

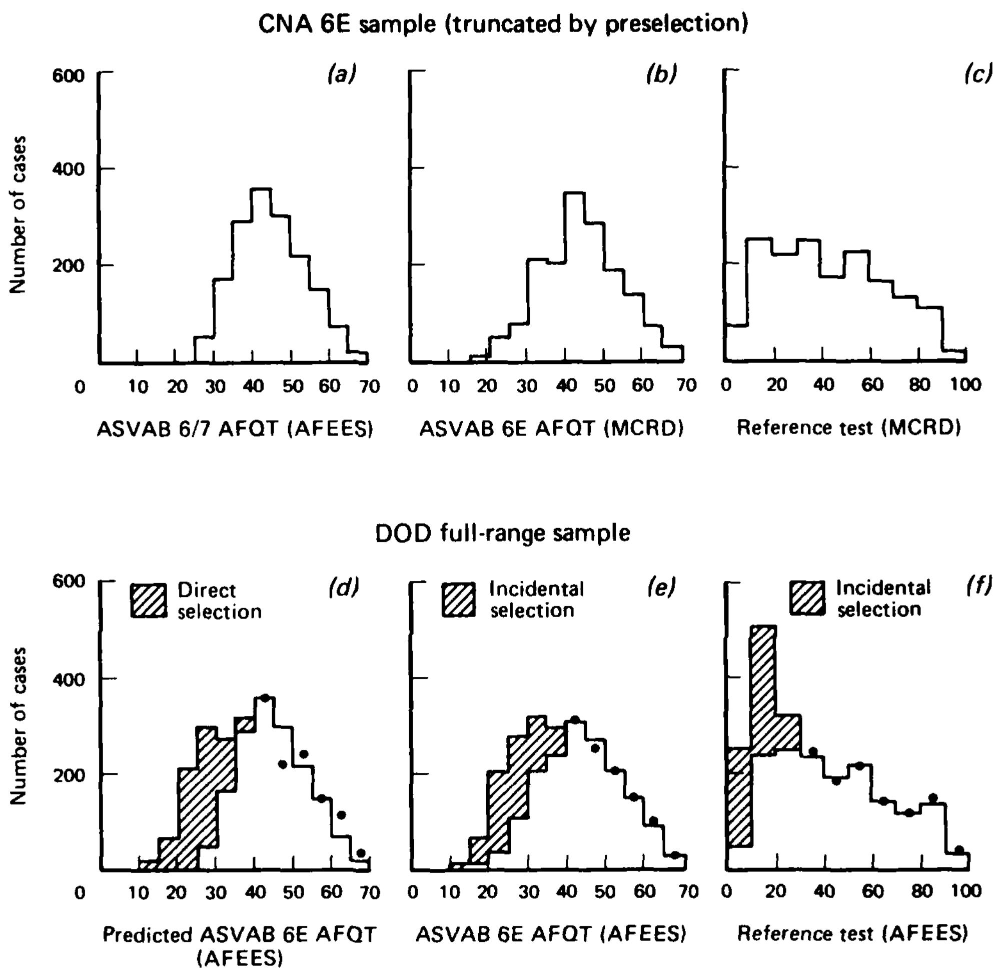
The next step was to simulate in the full-range sample the truncation of AFEES. This was accomplished by determining weights "A" such that when applied to the individuals in the DoD sample, the resulting distribution in predicted AFQT (figure J-3d) is identical to that of the AFQT taken at AFEES in the CNA sample (figure J-3a). The calculation of these weights is shown in annex J-2. When these weights are applied to the individuals in the DoD data sample the cross-hatched areas in figure J-3 are removed. In this manner, the effects of both direct and incidental selection are simulated. The truncated distributions of the relevant test scores from the DoD sample closely approximate[1] those from the CNA sample (figure J-3), which suggests that we have closely simulated the truncation of the CNA sample.

Means and correlation coefficients from the truncated CNA sample and the truncated DoD sample are compared (table J-1). The mean values are very comparable, indicating that our simulation is satisfactory. The correlation coefficients for the DoD data are somewhat higher than those for the CNA data. We believe the essential element is that the three coefficients from each CNA data set have the same relative size as the three from each DoD data set. Because the relative size of coefficients from both data sets were similar, we concluded that our simulation adequately replicated the truncation effect.

We next stratified the truncated DoD sample on the reference test and formed a cumulative frequency distribution of the scores of the test to be normed. We did this by a set of "B" weights calculated in annex J-3. These weights, applied in conjunction with the "A" weights from annex J-2, produced a stratified sample within the truncated DoD data set. The resulting cumulative frequency distribution of ASVAB 6E AFQT scores is shown in table J-2. Also shown in table J-2 is the distribution of the same variable from the full-range DoD sample stratified using the weights calculated in annex J-1. The difference in the two distributions is the result of the truncation effect. Similar results were obtained for the ASVAB 7E and ASVAB 6/7 samples and are shown in tables J-3 and J-4. The normed curves for both the full-range and truncated DoD samples are shown in figure J-4, J-5, and J-6. In each case there is a

---

[1]The full-range DoD distribution was scaled to equal the truncated DoD distribution above the 30th percentile. The "dots" in figure J-3 represent the scaled full-range distribution.



Note: Full-range distribution (dots plus cross-hatched area) was scaled (for illustration only) to the truncated distribution in the upper percentiles.

FIG. J-3: COMPARISON OF TRUNCATED CNA DATA AND TRUNCATED DOD DATA FOR THE ASVAB 6E SAMPLE

TABLE J-1

COMPARATIVE STATISTICS FOR CNA SAMPLE AND TRUNCATED DoD SAMPLE

Sample	ASVAB form	Correlation coefficient[a] between: Directly selected AFQT and indirectly selected AFQT	Directly selected AFQT and indirectly selected reference test	Indirectly selected AFQT and indirectly selected reference test	Mean value of distribution: Directly selected AFQT	Indirectly selected AFQT	Indirectly selected reference test
CNA--truncated at AFEES on AFQT	6E	0.75	0.71	.82	45.5	43.9	43.8
	7E	0.75	0.71	.82	45.5	43.6	44.9
	6/7	0.76	0.71	.79	45.4	44.1	45.0
DoD--truncated on predicted AFQT	6E	0.80[b]	0.76[b]	0.81	45.6	44.5	45.6
	7E	0.81[b]	0.76[b]	0.81	45.6	44.5	45.2
	6/7	0.79[b]	0.73	0.79	45.3	44.4	45.3

[a]All coefficients are from unstratified data.

[b]The directly selected AFQT for the DoD sample is the predicted AFQT.

TABLE J-2

CALCULATION OF TRUNCATION ADJUSTMENT FOR ASVAB 6E AFQT

ASVAB 6E raw score (1)	Cumulative percentage: DoD[a] full-range sample (2)	Cumulative percentage: DoD[b] truncated sample (3)	Adjustment[c] for truncation (4)
15	.6	.3	.3
16	.8	.3	.5
17	1.4	.4	1.0
18	1.7	.8	.9
19	2.3	1.1	1.2
20	3.0	1.2	1.8
21	3.8	1.5	2.3
22	4.9	2.0	2.9
23	6.2	3.1	3.1
24	7.3	3.8	3.5
25	8.6	4.6	4.0
26	10.0	5.8	4.2
27	11.1	6.8	4.3
28	12.6	8.2	4.4
29	14.2	9.4	4.8
30	15.8	11.2	4.6
31	17.2	12.7	4.5
32	18.9	14.5	4.4
33	20.9	17.0	3.9
34	22.8	19.0	3.8
35	24.8	21.0	3.8
36	26.6	22.8	3.8
37	28.6	25.1	3.5
38	30.7	27.2	3.5
39	32.9	29.8	3.1
40	34.8	32.1	2.7
41	37.7	35.0	2.7
42	40.0	37.6	2.4
43	42.6	40.4	2.2
44	45.7	43.9	1.8
45	48.3	47.0	1.3
46	51.0	50.0	1.0
47	54.0	53.3	.7
48	56.0	55.8	.2
49	58.5	58.8	-.3
50	61.7	62.1	-.4
51	63.5	64.0	-.5
52	66.4	67.1	-.7
53	68.8	69.5	-.7
54	71.0	71.8	-.8
55	73.6	74.5	-.9
56	76.3	77.3	-1.0
57	78.4	79.3	-.9
58	79.9	81.0	-1.1
59	82.7	83.6	-.9
60	84.2	85.1	-.9
61	86.1	87.0	-.9
62	88.1	88.9	-.8
63	90.4	91.2	-.8
64	93.2	93.7	-.5
65	94.9	95.3	-.4
66	96.2	96.5	-.3
67	97.6	97.7	-.1
68	98.2	98.3	-.1
69	98.9	98.9	0
70	100.0	100.0	0
Total	2,870	1,634	

[a] DoD sample weighted by weights in table J-1-2.

[b] DoD sample weighted by "double weights" (viz., weight "A" and weight "B") from table J-2-3 and J-3-3.

[c] Column (3) minus column (2).

TABLE J-3

CALCULATION OF TRUNCATION ADJUSTMENT FOR ASVAB 7E AFQT

	Cumulative percentage		
ASVAB 7E raw score (1)	DoD[a] full-range sample (2)	DoD[b] truncated sample (3)	Adjustment[c] for truncation (4)
11	.2	.1	.1
12	.2	.1	.1
13	.5	.1	.4
14	.9	.3	.6
15	1.3	.5	.8
16	1.7	.5	1.2
17	2.3	.6	1.7
18	3.2	1.0	2.2
19	4.1	1.5	2.6
20	5.2	2.0	3.2
21	6.3	2.7	3.6
22	7.5	3.3	4.2
23	8.9	4.6	4.3
24	10.4	5.9	4.5
25	11.7	7.2	4.5
26	13.4	8.6	4.8
27	14.5	9.5	5.0
28	16.1	10.7	5.4
29	17.6	12.3	5.3
30	19.5	14.2	5.3
31	21.4	16.4	5.0
32	23.3	18.5	4.8
33	25.2	20.7	4.5
34	26.7	22.8	3.9
35	28.7	25.1	3.6
36	30.4	27.0	3.4
37	33.0	30.0	3.0
38	35.3	32.3	3.0
39	38.0	35.2	2.8
40	40.2	38.0	2.2
41	43.0	41.1	1.9
42	45.5	43.7	1.8
43	48.2	46.9	1.3
44	50.4	49.3	1.1
45	53.3	52.4	.9
46	55.5	54.8	.7
47	58.4	58.0	.4
48	60.7	60.5	.2
49	63.3	63.3	0
50	65.7	66.1	-.4
51	68.1	68.6	-.5
52	71.2	71.9	-.7
53	74.5	75.3	-.8
54	76.3	77.3	-1.0
55	78.4	79.3	-.9
56	80.0	80.9	-.9
57	82.0	83.0	-1.0
58	83.5	84.5	-1.0
59	85.0	85.8	-.8
60	86.3	87.0	-.7
61	88.2	89.0	-.8
62	90.1	90.9	-.8
63	91.2	92.0	-.8
64	93.7	94.2	-.5
65	95.0	95.3	-.3
66	97.2	97.5	-.3
67	99.6	99.7	-.1
68	100.0	100.0	0
69	100.0	100.0	0
70	100.0	100.0	0
Total	2,650	1,660	

[a] DoD sample weighted by weights in table J-1-2.

[b] DoD sample weighted by "double weights" (viz., weight "A" and weight "B") from table J-2-3 and J-3-3.

[c] Column (3) minus column (2).

TABLE J-4

CALCULATION OF TRUNCATION ADJUSTMENT FOR ASVAB 6/7 AFQT

ASVAB 6/7 raw score	Cumulative percentage: DoD[a] full-range sample	Cumulative percentage: DoD[b] truncated sample	Adjustment[c] for truncation
15	.9	.5	.4
16	1.2	.6	.6
17	1.7	1.0	.7
18	2.2	1.2	1.0
19	3.0	1.3	1.7
20	4.1	1.6	2.5
21	4.8	2.0	2.8
22	5.8	2.5	3.3
23	6.9	3.1	3.8
24	8.4	4.1	4.3
25	9.9	5.3	4.6
26	11.1	6.3	4.8
27	12.6	7.5	5.1
28	14.1	8.9	5.2
29	15.8	10.5	5.3
30	17.5	12.1	5.4
31	19.2	13.9	5.3
32	20.6	15.6	5.0
33	22.2	17.6	4.6
34	24.0	19.6	4.4
35	25.9	21.6	4.3
36	27.4	23.5	3.8
37	29.3	25.5	3.8
38	31.6	28.0	3.6
39	34.2	31.2	3.0
40	36.7	33.8	2.8
41	38.9	36.2	2.7
42	41.2	39.1	2.1
43	43.7	42.0	1.7
44	45.9	44.7	1.2
45	48.2	47.1	1.1
46	50.4	49.7	.7
47	52.8	52.4	.4
48	55.2	55.1	.1
49	57.3	57.4	-.1
50	59.5	59.9	-.4
51	62.0	62.6	-.6
52	64.4	65.5	-1.1
53	66.8	68.0	-1.2
54	69.3	70.8	-1.5
55	71.3	73.0	-1.7
56	73.1	74.7	-1.6
57	75.4	77.3	-1.9
58	77.7	79.6	-1.9
59	79.4	81.3	-1.9
60	81.8	83.3	-1.5
61	84.1	85.4	-1.3
62	86.0	87.1	-1.1
63	88.3	89.3	-1.0
64	90.5	91.1	-.6
65	92.8	93.3	-.5
66	93.6	94.0	-.4
67	95.7	96.0	-.3
68	97.6	97.6	.0
69	98.8	98.7	.1
70	100.0	100.0	.0
Total	5,070	2,208	

[a] DoD sample weighted by weights in table J-1-2.

[b] DoD sample weighted by "double weights" (viz., weight "A" and weight "B") from table J-2-3 and J-3-3.

[c] Column (3) minus column (2).

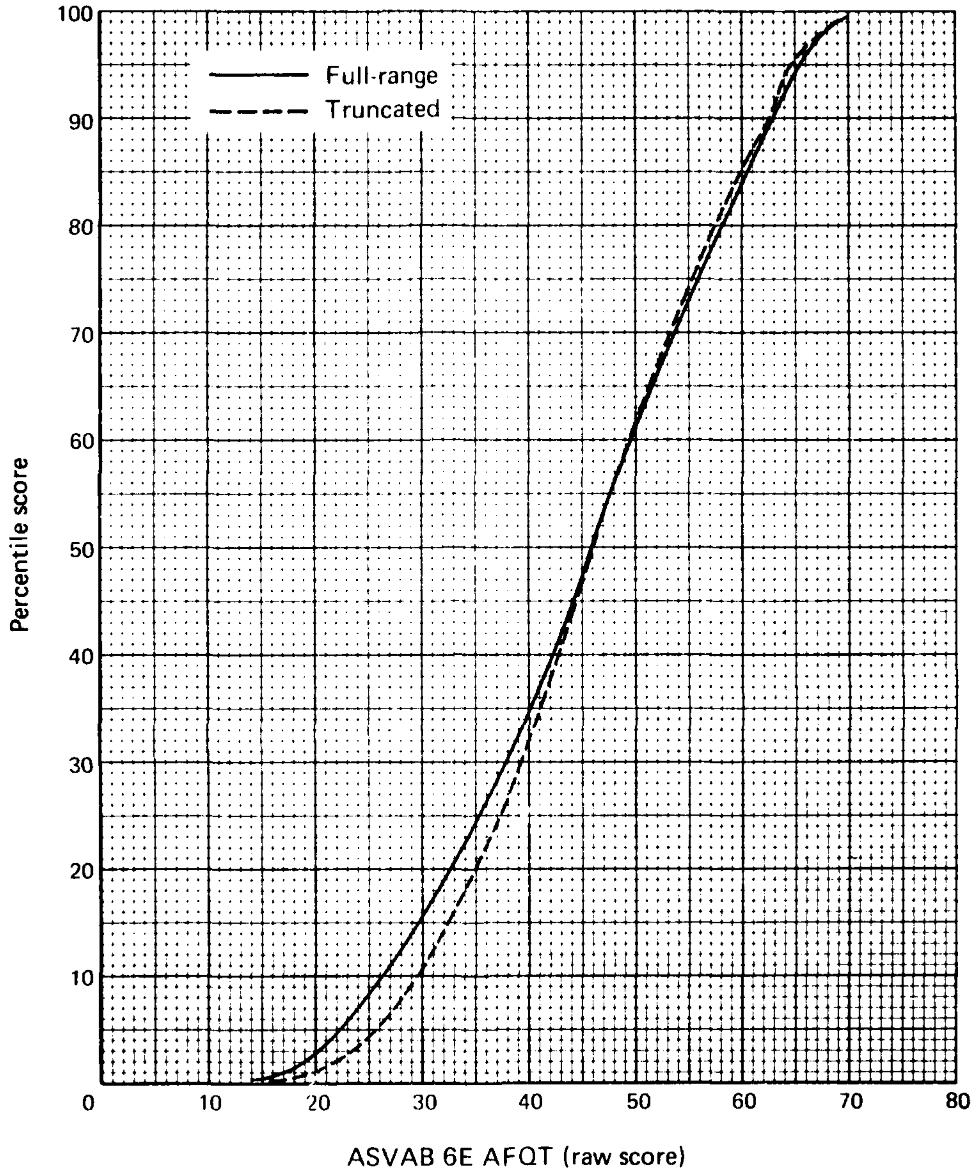


FIG. J-4: EFFECT OF SIMULATED TRUNCATION ON NORMING RESULTS FOR DOD ASVAB 6E SAMPLE

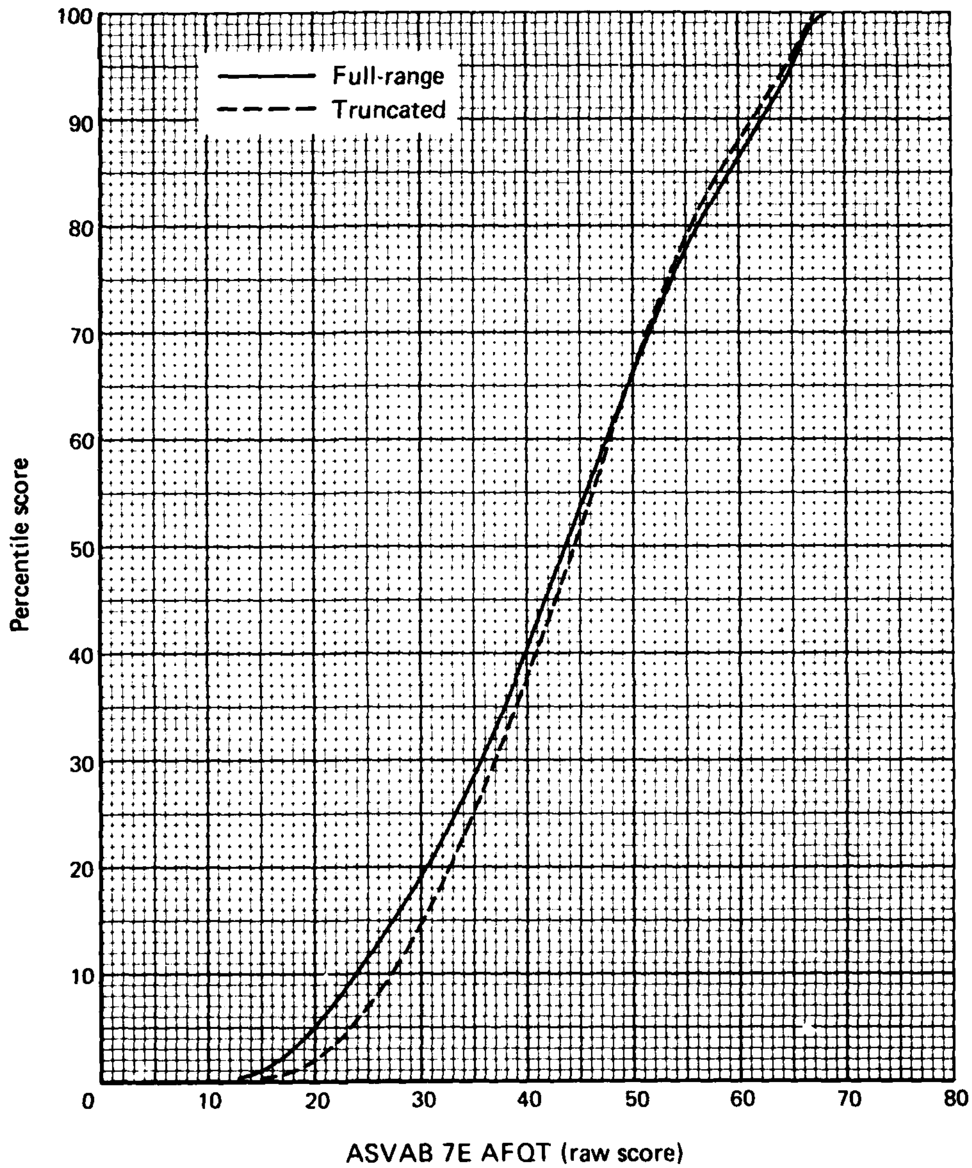


FIG. J-5: EFFECT OF SIMULATED TRUNCATION ON NORMING RESULTS FOR DOD ASVAB 7E SAMPLE

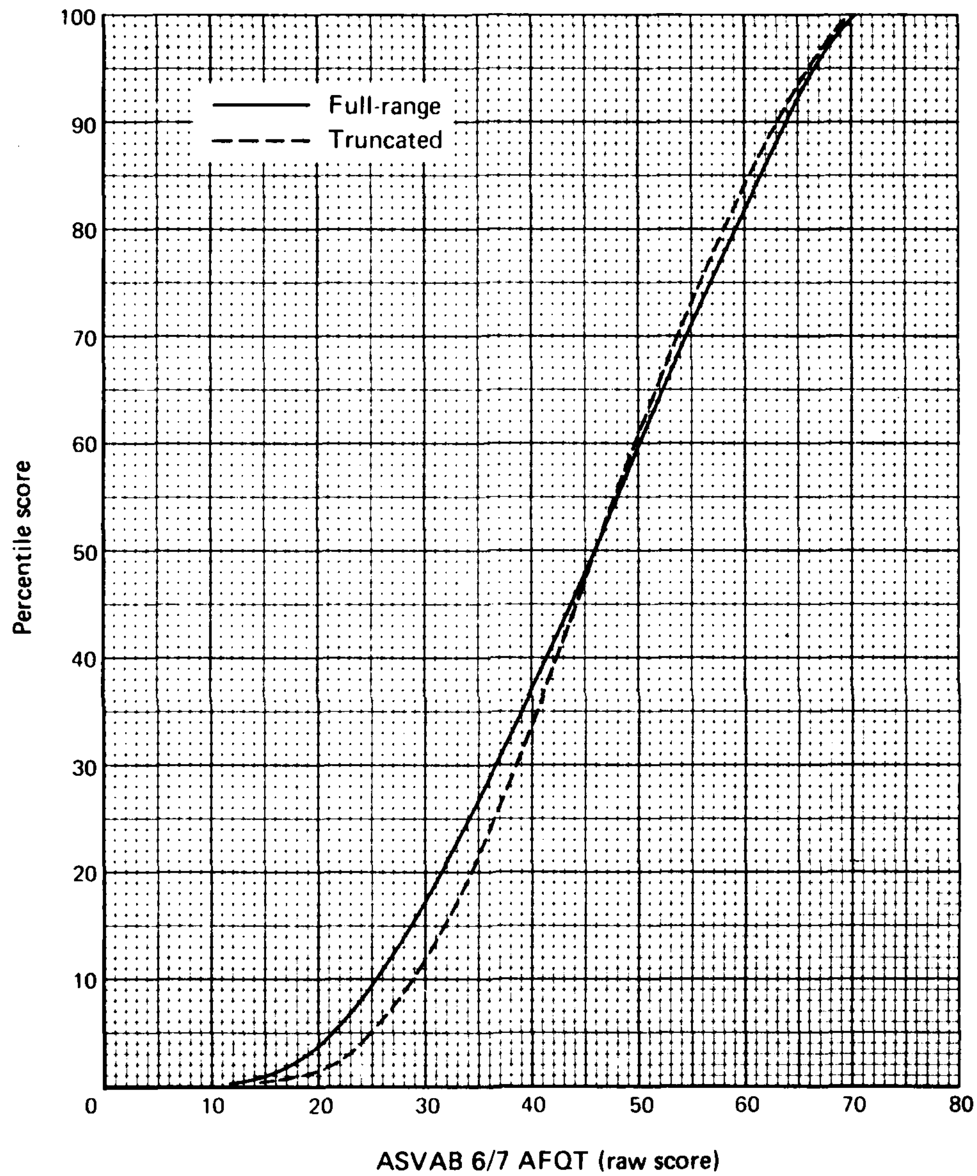


FIG. J-6: EFFECT OF SIMULATED TRUNCATION ON NORMING RESULTS FOR DOD ASVAB 6/7 SAMPLE bias toward harder norms in the low percentiles and a bias toward easier norms in the higher percentiles. The maximum extent of the bias appears to be about 5 percentile points near the 20th percentile.

REFERENCE

J-1 Center for Naval Analyses, Study 1115, "An Analysis of the Normalization and Verification of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 6 and 7," by William H. Sims, Unclassified, Apr 1978

ANNEX J-1

EQUIPERCENTILE TABLES FOR PREDICTED AFQT

Equipercentile tables were constructed from the DoD sample to equate ASVAB AFQT scores with Pseudo[1] AFQT scores. Reference J-1 indicated that the Pseudo AFQT is an excellent predictor of AFQT scores.

We calculated weight factors (tables J-1-1, J-1-2, and J-1-3) to stratify the three DoD samples on the reference test (AFQT 7A). We applied these weight factors to everyone in the DoD samples depending on their AFQT 7A score to simulate the standard mobilization population. The cumulative percentages of each sample using weighted individuals is shown in figures J-1-1, J-1-2, and J-1-3.

Raw scores on the ASVAB AFQT and Pseudo AFQT were equated by the standard graphical equipercentile method. Raw scores on the two tests were considered to be equivalent if they were obtained by the same cumulative percentage of the sample. Equivalent ASVAB AFQT and Pseudo AFQT raw scores were read directly from figures J-1-1, J-1-2, and J-1-3 and are recorded in table J-1-4.

---

[1] Pseudo AFQT = GI+GS+MC+MK, where:

GI = general information
GS = general science
MC = mechanical comprehension
MK = mathematical knowledge.

TABLE J-1-1

CALCULATION OF WEIGHT FACTORS FOR DoD ASVAB 6E SAMPLE

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
0-5	111	143.5	1.293
6-10	222	143.5	.646
11-15	356	143.5	.403
16-20	312	143.5	.460
21-25	236	143.5	.608
26-30	189	143.5	.759
31-35	163	143.5	.880
36-40	162	143.5	.886
41-45	70	143.5	2.050
46-50	166	143.5	.864
51-55	114	143.5	1.259
56-60	161	143.5	.891
61-65	139	143.5	1.032
66-70	48	143.5	2.990
71-75	74	143.5	1.939
76-80	83	143.5	1.729
81-85	113	143.5	1.270
86-90	88	143.5	1.631
91-95	50	143.5	2.870
96-100	13	143.5	11.038
Total	2,870	2,870	

[a]Column (3) divided by column (2).

TABLE J-1-2

CALCULATION OF WEIGHT FACTORS FOR DoD ASVAB 7E SAMPLE

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
0-5	104	132.5	1.274
6-10	179	132.5	.740
11-15	341	132.5	.389
16-20	297	132.5	.446
21-25	199	132.5	.666
26-30	177	132.5	.749
31-35	152	132.5	.872
36-40	153	132.5	.866
41-45	57	132.5	2.325
46-50	170	132.5	.779
51-55	98	132.5	1.352
56-60	123	132.5	1.077
61-65	136	132.5	.974
66-70	53	132.5	2.500
71-75	76	132.5	1.743
76-80	99	132.5	1.338
81-85	97	132.5	1.366
86-90	78	132.5	1.699
91-95	49	132.5	2.704
96-100	12	132.5	11.042
Total	2,650	2,650	

[a]Column (3) divided by column (2).

TABLE J-1-3

CALCULATION OF WEIGHT FACTORS FOR DoD ASVAB 6/7 SAMPLE

AFQT 7A percentile interval (1)	Number observed in sample (2)	Number expected in mobilization population (3)	Weight factor[a] (4)
0-5	215	253.45	1.179
6-10	368	253.45	.689
11-15	662	253.45	.383
16-20	587	253.45	.432
21-25	327	253.45	.775
26-30	344	253.45	.737
31-35	290	253.45	.874
36-40	284	253.45	.892
41-45	127	253.45	1.996
46-50	295	253.45	.859
51-55	203	253.45	1.249
56-60	239	253.45	1.060
61-65	275	253.45	.922
66-70	99	253.45	2.560
71-75	153	253.45	1.657
76-80	148	253.45	1.713
81-85	198	253.45	1.280
86-90	148	253.45	1.713
91-95	89	253.45	2.848
96-100	18	253.45	14.081
Total	5,069	5,069	

[a] Column (3) divided by column (2).

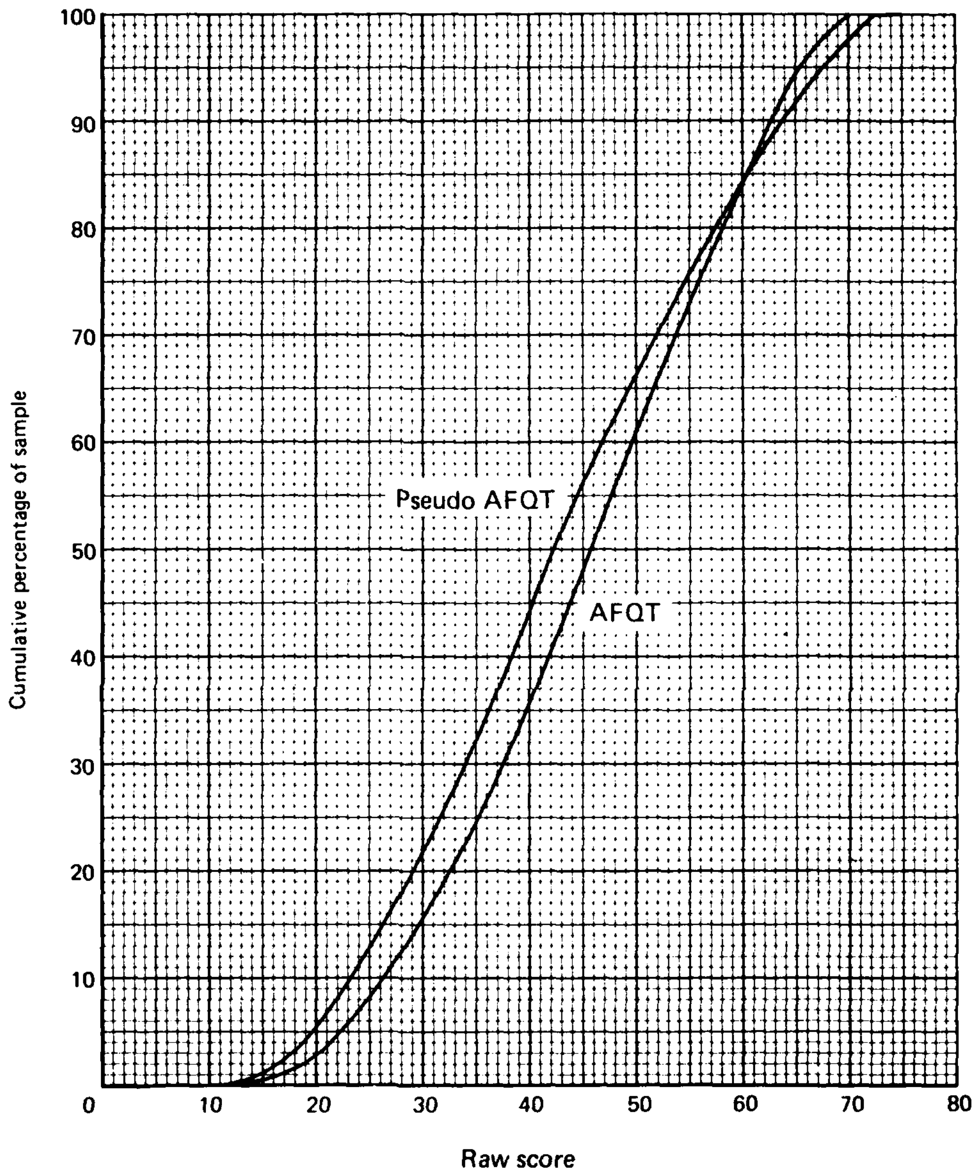


FIG. J-1-1: CUMULATIVE PERCENTAGE OF STRAFIFIED DOD SAMPLE ON ASVAB 6E AFQT AND PSEUDO AFQT

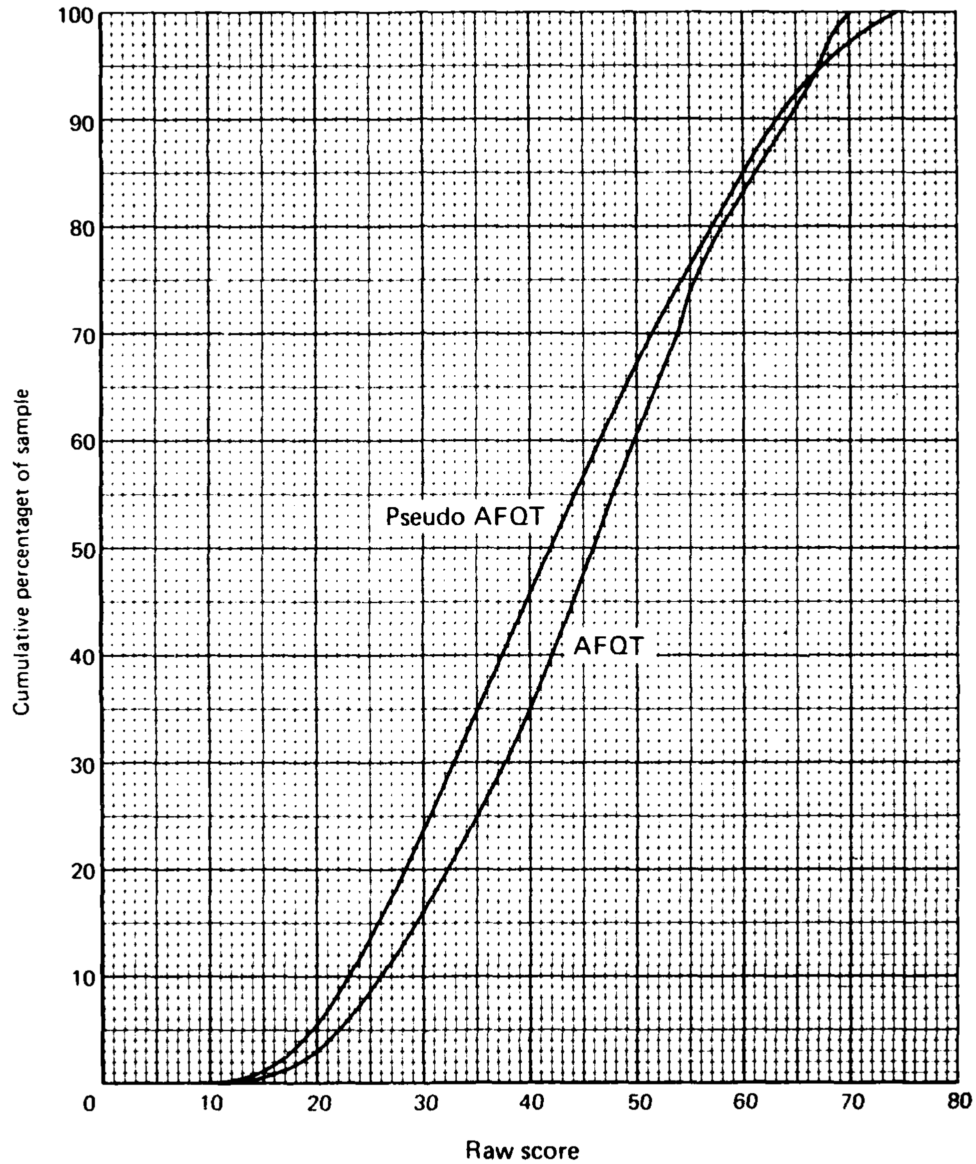


FIG. J-1-2: CUMULATIVE PERCENTAGE OF STRATIFIED DOD SAMPLE ON ASVAB 7E AFQT AND PSEUDO AFQT

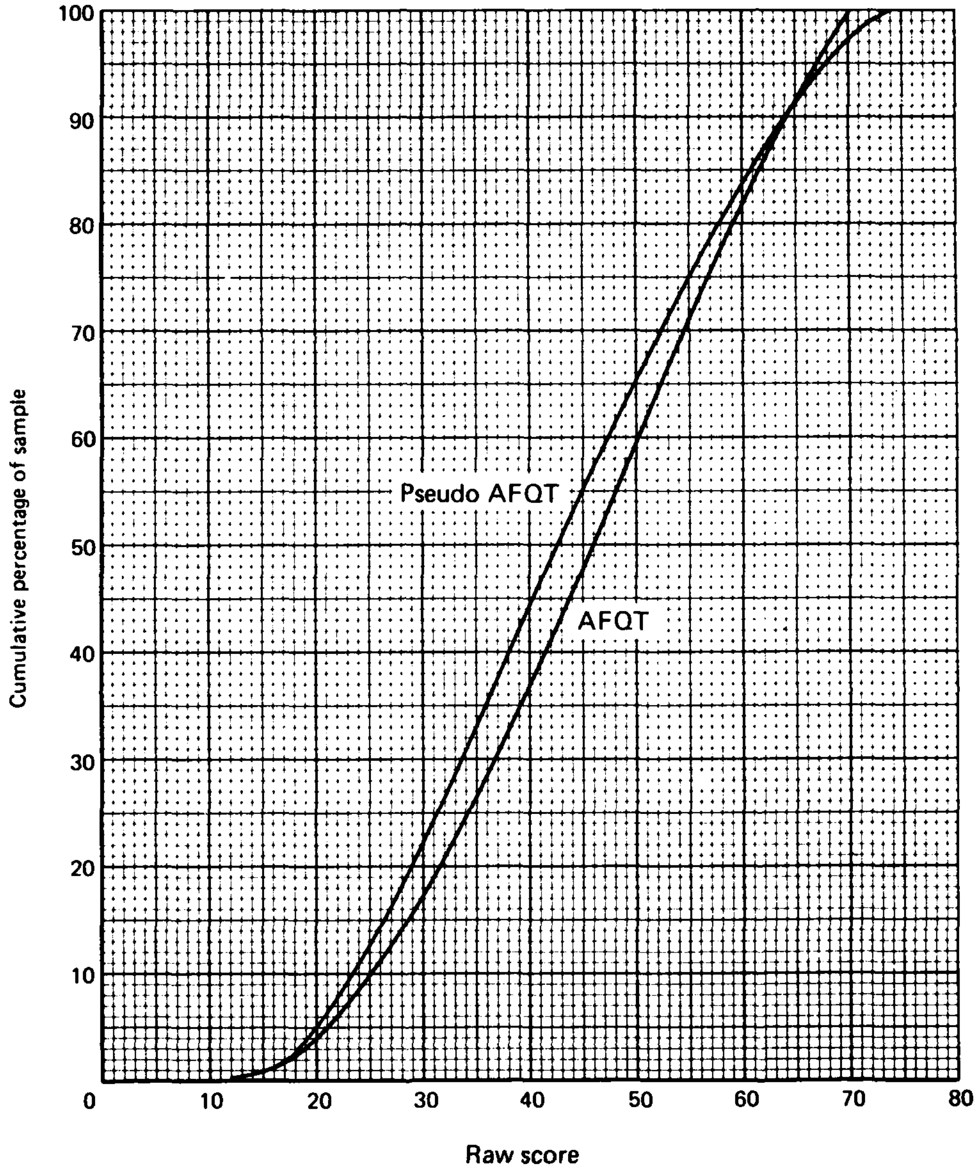


FIG. J-1-3: CUMULATIVE PERCENTAGE OF STRATIFIED DOD SAMPLE ON ASVAB 6/7 AND PSEUDO AFQT

TABLE J-1-4

EQUIPERCENTILE CONVERSION TABLE FOR PSEUDO AFQT

Pseudo AFQT	ASVAB 6E AFQT	ASVAB 7E AFQT	ASVAB 6/7 AFQT
12	12	15	12
13	13	15	13
14	15	16	14
15	16	17	15
16	18	18	16
17	20	19	17
18	21	20	18
19	22	21	20
20	23	22	21
21	24	23	22
22	25	24	23
23	26	25	25
24	27	26	26
25	28	28	27
26	29	29	28
27	30	30	29
28	32	32	30
29	33	33	31
30	34	34	33
31	35	35	34
32	36	37	35
33	37	38	36
34	38	39	37
35	39	40	38
36	40	41	39
37	41	42	40
38	42	42	41
39	43	43	42
40	44	44	43
41	45	45	44
42	45	46	45
43	46	47	46
44	47	47	47
45	48	48	48
46	49	49	49
47	50	50	50
48	51	51	51
49	52	52	52
50	52	53	52
51	53	53	53
52	54	54	54
53	55	55	55
54	55	55	56
55	56	56	57
56	57	57	58
57	57	58	59
58	58	59	60
59	59	60	61
60	60	61	61
61	61	62	62
62	61	63	62
63	62	63	63
64	63	64	64
65	63	65	65
66	64	66	65
67	65	67	66
68	66	67	67
69	66	68	67
70	67	68	68
71	69	68	69
72	70	68	69
73	70	69	69
74	70	70	70
75	70	70	70

ANNEX J-2

CALCULATION OF "A" WEIGHTS FOR DoD DATA

The purpose of this annex is to show how to simulate the same truncation or preselection in the DoD data sample as occurred in the CNA sample. This is accomplished by calculating weights "A", which will force the distribution of predicted DoD ASVAB AFQT scores to look like those of the CNA AFEES ASVAB AFQT scores. The calculation of the weight factors is shown in tables J-2-1, J-2-2, and J-2-3. When these weights are attached to individuals in the DoD sample (as a function of their predicted ASVAB AFQT score), the resulting distribution will be identical to that of the truncated CNA ASVAB AFQT scores based on AFEES testing.

TABLE J-2-1

CALCULATION OF "A" WEIGHTS FOR DoD 6E SAMPLE

	Number of cases (unweighted)		
AFQT interval (1)	Predicted ASVAB AFQT DoD 6E sample (2)	AFEES ASVAB AFQT CNA 6E sample (3)	Weight factor[a] (4)
0-15	23	0	0.000
16-20	85	0	0.000
21-25	258	0	0.000
26-30	373	54	0.145
31-35	337	166	0.493
36-40	399	288	0.722
41-45	447	358	0.801
46-50	274	304	1.109
51-55	302	217	0.719
56-60	188	148	0.787
61-65	140	74	0.529
66-70	44	25	0.568
	2,870	1,634	

[a]Column (3) divided by column (2).

TABLE J-2-2

CALCULATION OF "A" WEIGHTS FOR DoD 7E SAMPLE

AFQT interval (1)	Number of cases (unweighted): Predicted ASVAB AFQT DoD 7E sample (2)	Number of cases (unweighted): AFEES ASVAB AFQT CNA 7E sample (3)	Weight factor[a] (4)
0-15	23	0	0.000
16-20	104	0	0.000
21-25	253	0	0.000
26-30	309	45	0.146
31-35	331	163	0.492
36-40	320	305	0.953
41-45	367	383	1.044
46-50	320	304	0.950
51-55	276	215	0.779
56-60	142	127	0.894
61-65	130	89	0.685
66-70	75	29	0.387
	2,650	1,660	

[a]Column (3) divided by column (2).

TABLE J-2-3

CALCULATION OF "A" WEIGHTS FOR DoD 6/7 SAMPLE

	Number of cases (unweighted)		
AFQT interval (1)	Predicted ASVAB AFQT DoD 6/7 sample (2)	AFEES ASVAB AFQT CNA 6/7 sample (3)	Weight factor[a] (4)
0-15	65	0	0.000
16-20	210	0	0.000
21-25	415	0	0.000
26-30	739	68	0.092
31-35	616	232	0.377
36-40	712	403	0.566
41-45	584	494	0.846
46-50	491	396	0.807
51-55	502	290	0.578
56-60	297	179	0.603
61-65	330	111	0.336
66-70	108	35	0.324
	5,069	2,208	

[a]Column (3) divided by column (2).

ANNEX J-3

CALCULATION OF "B" WEIGHTS FOR DoD DATA

Annex J-2 showed that "A" weights enable us to simulate the effect of preselection in the DoD sample. This annex shows how to calculate "B" weights to stratify that truncated sample on the reference test (AFQT 7A).

The distribution of the "B" weighted sample on the reference test is given in tables J-3-1, J-3-2, and J-3-3 for the three DoD samples as is the calculation of the weights necessary to stratify the sample.

TABLE J-3-1

CALCULATION OF "B" WEIGHTS FOR DoD 6E SAMPLE

AFQT 7A percentile interval (1)	Number[a] observed in "A" weighted DoD 6E sample (2)	Number expected in mobilization population (3)	Weight factor[b] (4)
0-5	15.00	81.718	5.4479
6-10	32.20	81.718	2.5378
11-15	100.28	81.718	.8149
16-20	135.37	81.718	.6037
21-25	133.82	81.718	.6107
26-30	118.66	81.718	.6887
31-35	113.24	81.718	.7216
36-40	123.13	81.718	.6637
41-45	53.23	81.718	1.5352
46-50	132.41	81.718	.6172
51-55	92.91	81.718	.8795
56-60	130.85	81.718	.6245
61-65	114.72	81.718	.7123
66-70	39.01	81.718	2.0948
71-75	59.03	81.718	1.3843
76-80	61.47	81.718	1.3294
81-85	79.61	81.718	1.0265
86-90	60.81	81.718	1.3438
91-95	31.50	81.718	2.5942
96-100	7.15	81.718	11.4291
Total	1,634.40	1,634.36	

[a]This column has fractional frequency distributions because it is the result of weighting the DoD sample by weight "A".

[b]Column (3) divided by column (2).

TABLE J-3-2

CALCULATION OF "B" WEIGHTS FOR DoD 7E SAMPLE

AFQT 7A percentile interval (1)	Number[a] observed in "A" weighted DoD 7E sample (2)	Number expected in mobilization population (3)	Weight factor[b] (4)
0-5	13.98	83.005	5.9374
6-10	30.29	83.005	2.7403
11-15	113.79	83.005	.7295
16-20	149.43	83.005	.5555
21-25	120.04	83.005	.6915
26-30	124.05	83.005	.6691
31-35	117.80	83.005	.7046
36-40	129.01	83.005	.6434
41-45	49.37	83.005	1.6813
46-50	151.11	83.005	.5493
51-55	87.07	83.005	.9533
56-60	111.65	83.005	.7434
61-65	113.64	83.005	.7304
66-70	44.07	83.005	1.8835
71-75	64.42	83.005	1.2885
76-80	76.93	83.005	1.0790
81-85	73.80	83.005	1.1247
86-90	52.91	83.005	1.5688
91-95	30.54	83.005	2.7179
96-100	6.25	83.005	13.2808
Total	1,660.15	1,660.10	

[a]This column has fractional frequency distributions because it is the result of weighting the DoD sample by weight "A".

[b]Column (3) divided by column (2).

TABLE J-3-3

CALCULATION OF "B" WEIGHTS FOR DoD 6/7 SAMPLE

AFQT 7A percentile interval (1)	Number[a] observed in "A" weighted DoD 6/7 sample (2)	Number expected in mobilization population (3)	Weight factor[b] (4)
0-5	16.19	110.43	6.8210
6-10	46.51	110.43	2.3743
11-15	135.23	110.43	.8166
16-20	185.35	110.43	.5958
21-25	149.38	110.43	.7393
26-30	172.58	110.43	.6399
31-35	158.87	110.43	.6951
36-40	169.91	110.43	.6499
41-45	79.69	110.43	1.3858
46-50	191.18	110.43	.5776
51-55	134.65	110.43	.8201
56-60	157.24	110.43	.7023
61-65	180.43	110.43	.6120
66-70	61.37	110.43	1.7994
71-75	88.68	110.43	1.2453
76-80	81.84	110.43	1.3494
81-85	97.34	110.43	1.1345
86-90	62.11	110.43	1.7780
91-95	33.95	110.43	3.2528
96-100	6.17	110.43	17.8981
Total	2,208.67	2,208.63	

[a]This column has fractional frequency distributions because it is the result of weighting the DoD sample by weight "A".

[b]Column (3) divided by column (2).

APPENDIX K

ADJUSTMENTS FOR EFFECTS OF SAMPLE TRUNCATION ON UNSTRATIFIED NORMING RESULTS

APPENDIX K

ADJUSTMENTS FOR EFFECTS OF SAMPLE TRUNCATION ON UNSTRATIFIED NORMING RESULTS

In this appendix we carry out unstratified graphical equating using the full-range DoD sample and the truncated DoD sample described in appendix J. Neither sample was stratified on the reference test.

Cumulative frequency distributions of the reference test scores and ASVAB 6E AFQT scores were graphed for the full-range sample (figure K-1) and for the truncated sample (figure K-2). Scores made by the same cumulative frequency of each sample were equated. The percentile scores equated to each ASVAB 6E AFQT raw score are shown in table K-1. Similar calculations were made for ASVAB 7E and ASVAB 6/7; these are shown in figures K-3, K-4, K-5, and K-6. The results are recorded in tables K-2 and K-3.

The difference between the norming curves for the truncated and full-range samples is very small (figures K-7, K-8, and K-9) and confined mainly to the region below the 10th percentile.

The comparison of these norming curves constructed from unstratified data with those in appendix J using stratified data indicates that using unstratified graphical equating produces much less bias in a truncated sample.

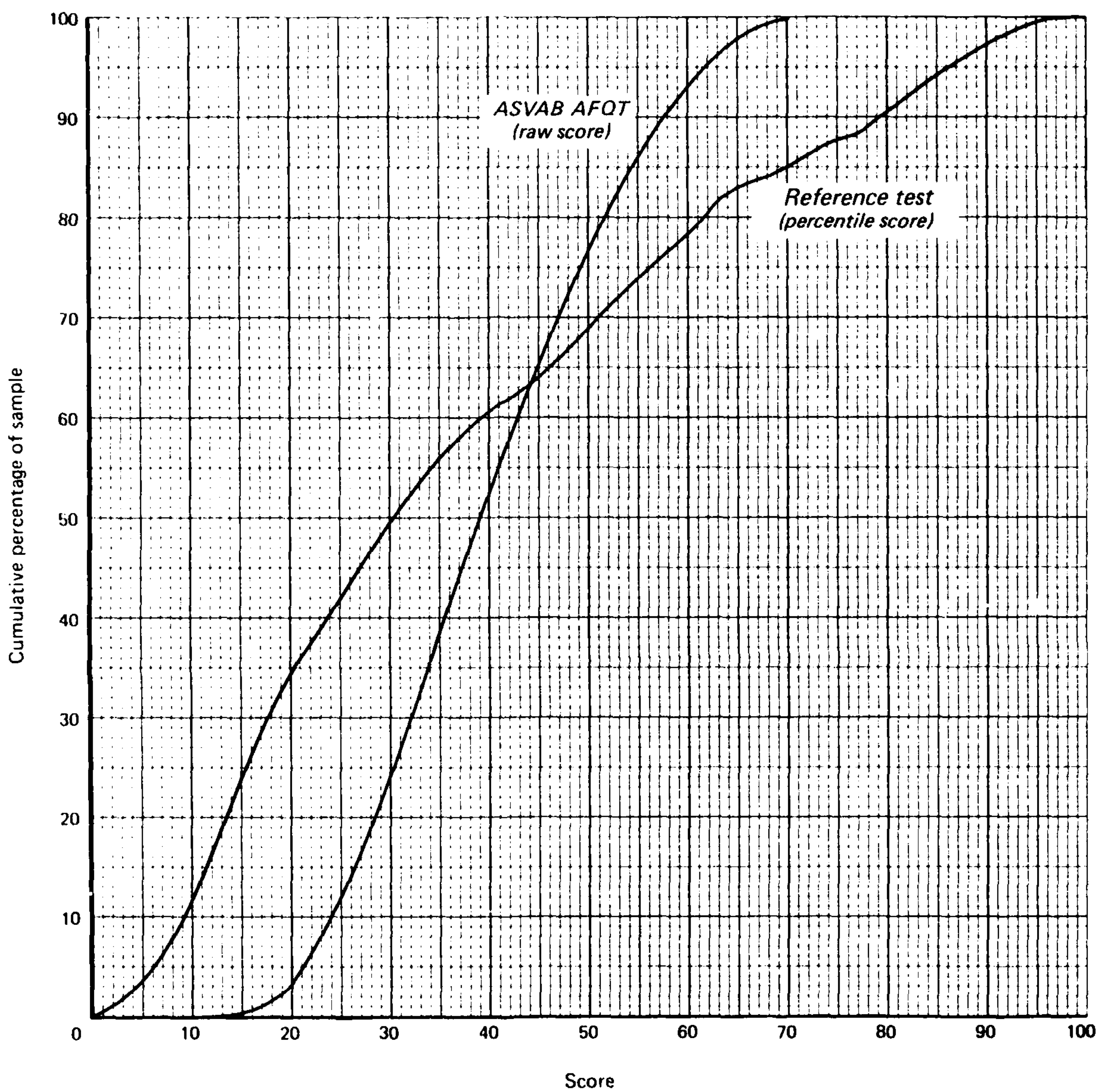


FIG. K-1: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 6E AFQT (DOD FULL-RANGE SAMPLE)

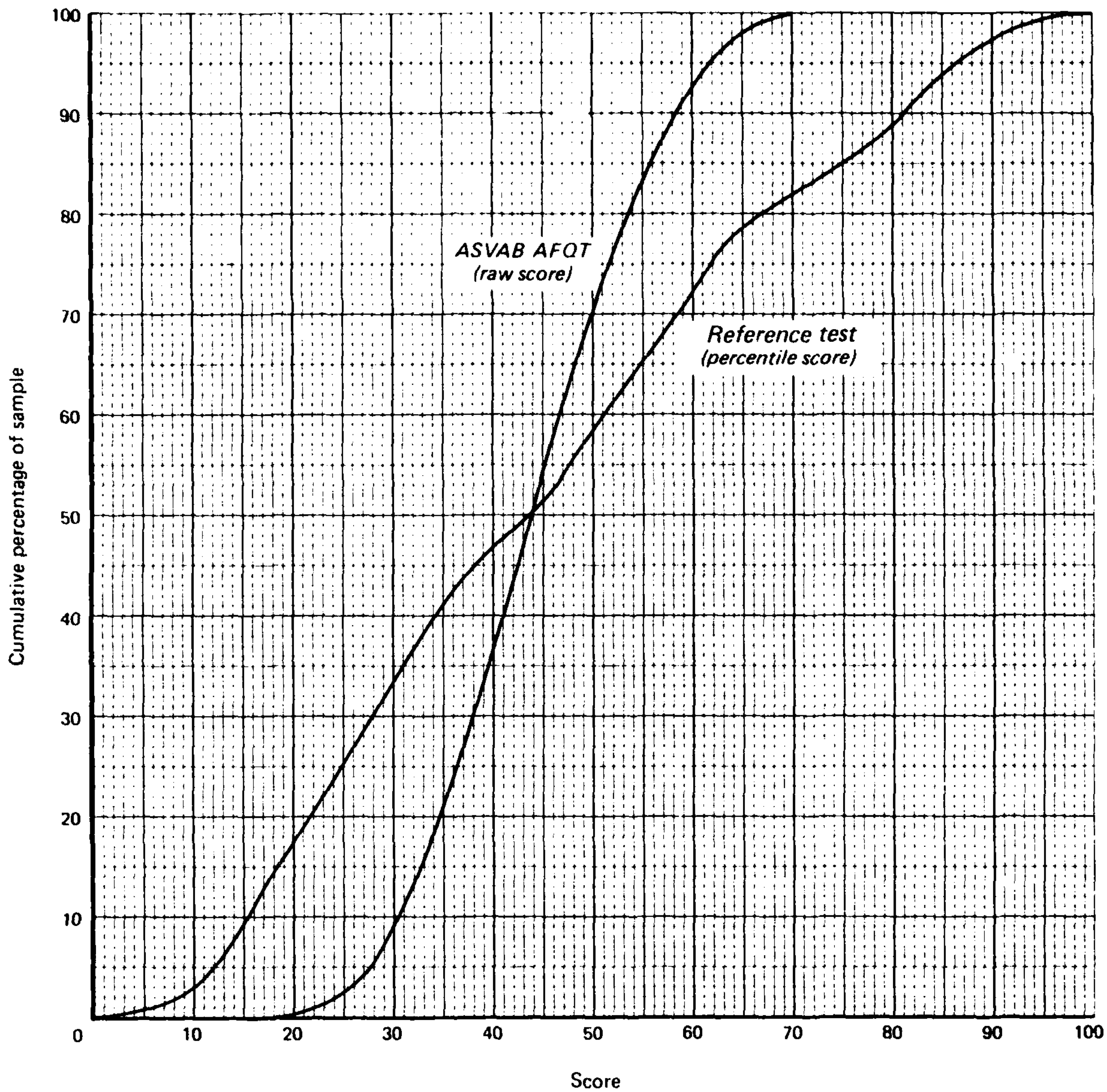


FIG. K-2: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 6E AFQT (DOD TRUNCATED SAMPLE)

TABLE K-1

COMPARISON OF EQUATING TECHNIQUES ON DOD 6E SAMPLE

ASVAB 6E AFQT raw score	Percentiles: Truncated sample	Percentiles: Full-range sample	Full-range case minus truncated case
14-15	0.0	1.0	1.0
16-17	1.0	2.0	1.0
18-19	2.0	4.0	2.0
20	3.0	5.0	2.0
21	4.0	6.0	2.0
22	6.0	7.0	1.0
23	7.8	8.5	0.7
24	8.5	9.3	0.8
25	9.5	10.4	0.9
26	10.5	11.3	0.8
27	11.5	12.0	0.5
28	12.5	13.3	0.8
29	13.6	14.2	0.6
30	15.1	15.5	0.4
31	16.2	16.5	0.3
32	17.2	17.5	0.3
33	19.0	19.0	0.0
34	21.0	21.0	0.0
35	23.0	23.0	0.0
36	24.5	24.5	0.0
37	26.2	26.5	0.3
38	28.0	28.4	0.4
39	29.5	30.0	0.5
40	31.5	31.6	0.1
41	33.5	33.0	-0.5
42	36.3	36.3	0.0
43	39.5	38.8	-0.7
44	44.0	44.0	0.0
45	47.3	47.0	-0.3
46	50.0	49.4	-0.6
47	52.3	51.3	-1.0
48	54.5	54.0	-0.5
49	56.5	56.0	-0.5
50	59.3	58.5	-0.8
51	61.0	60.5	-0.5
52	62.5	62.0	-0.5
53	65.0	64.3	-0.7
54	68.5	69.2	0.7
55	72.4	71.6	-1.0
56	76.0	76.5	0.5
57	78.5	78.8	0.3
58	80.5	80.0	-0.5
59	82.0	82.0	0.0
60	83.5	83.7	0.2
61	84.5	85.0	0.5
62	87.0	86.6	-0.4
63	88.5	88.4	-0.1
64	90.0	90.0	0.0
65	91.5	91.5	0.0
66	93.0	93.0	0.0
67	94.0	93.5	-0.5
68	97.0	97.0	0.0
69	98.0	98.0	0.0
70	100.0	100.0	0.0

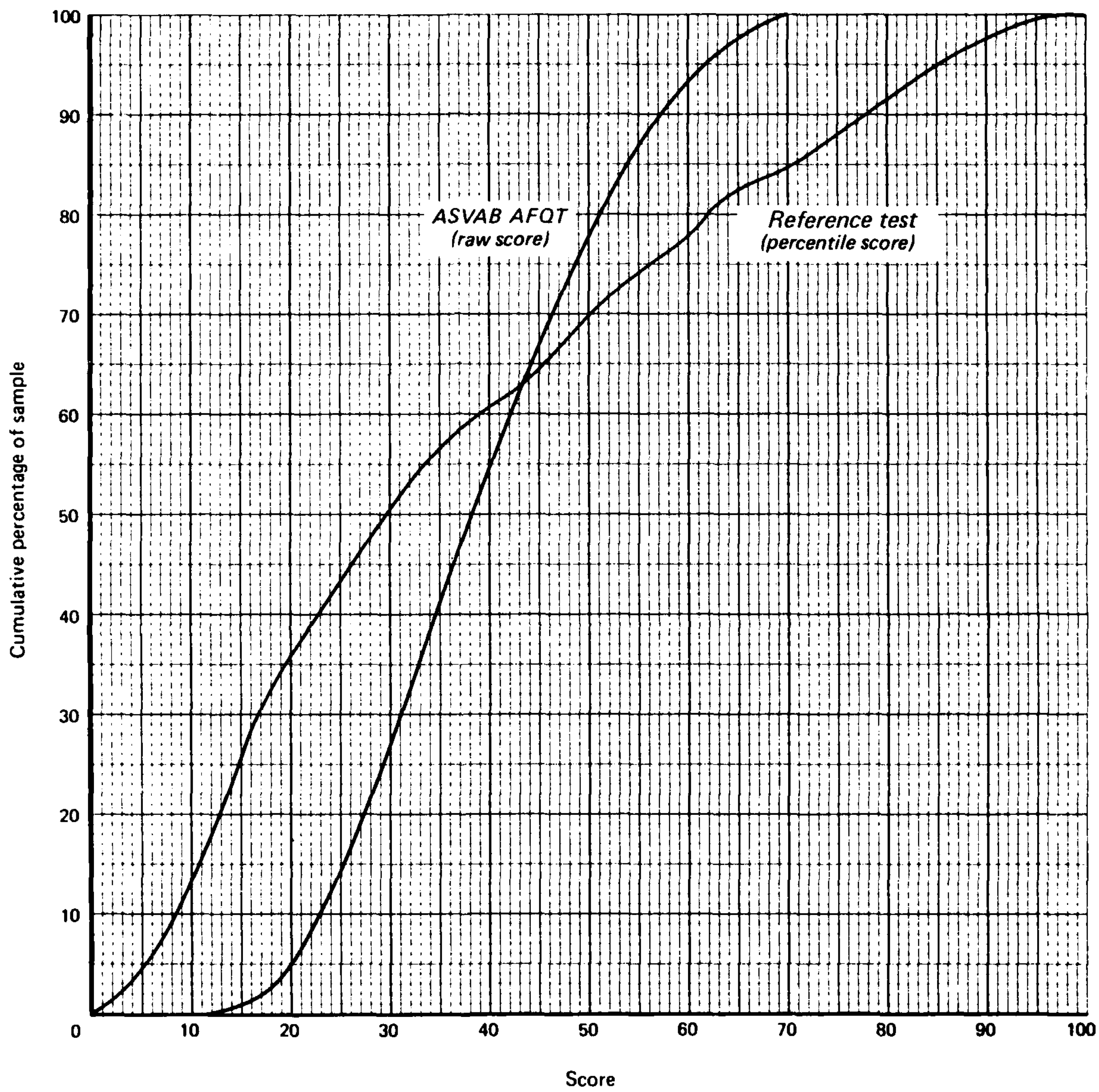


FIG. K-3: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 7E AFQT (DOD FULL-RANGE SAMPLE)

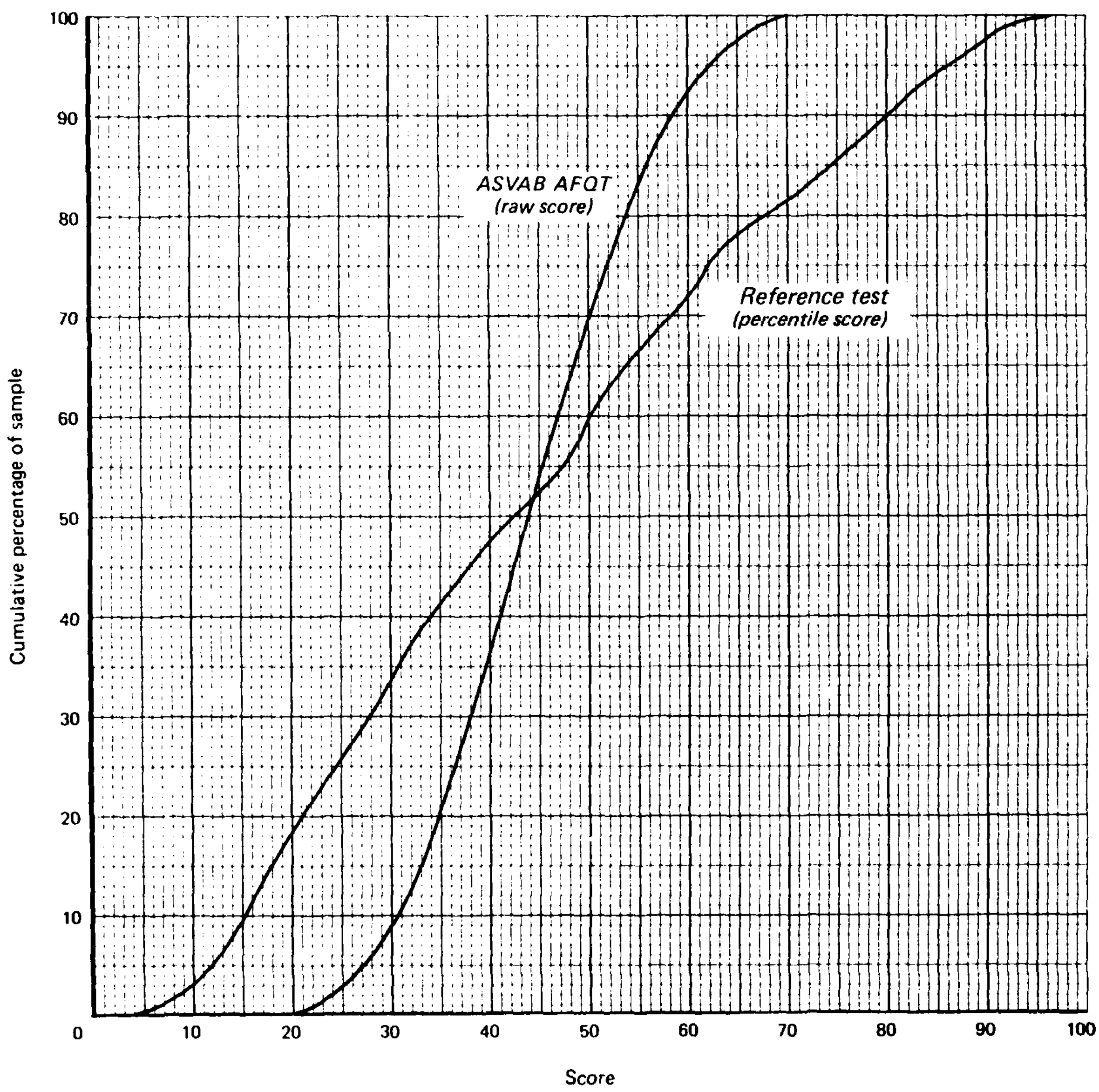


FIG. K-4: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 7E AFQT (DOD TRUNCATED SAMPLE)

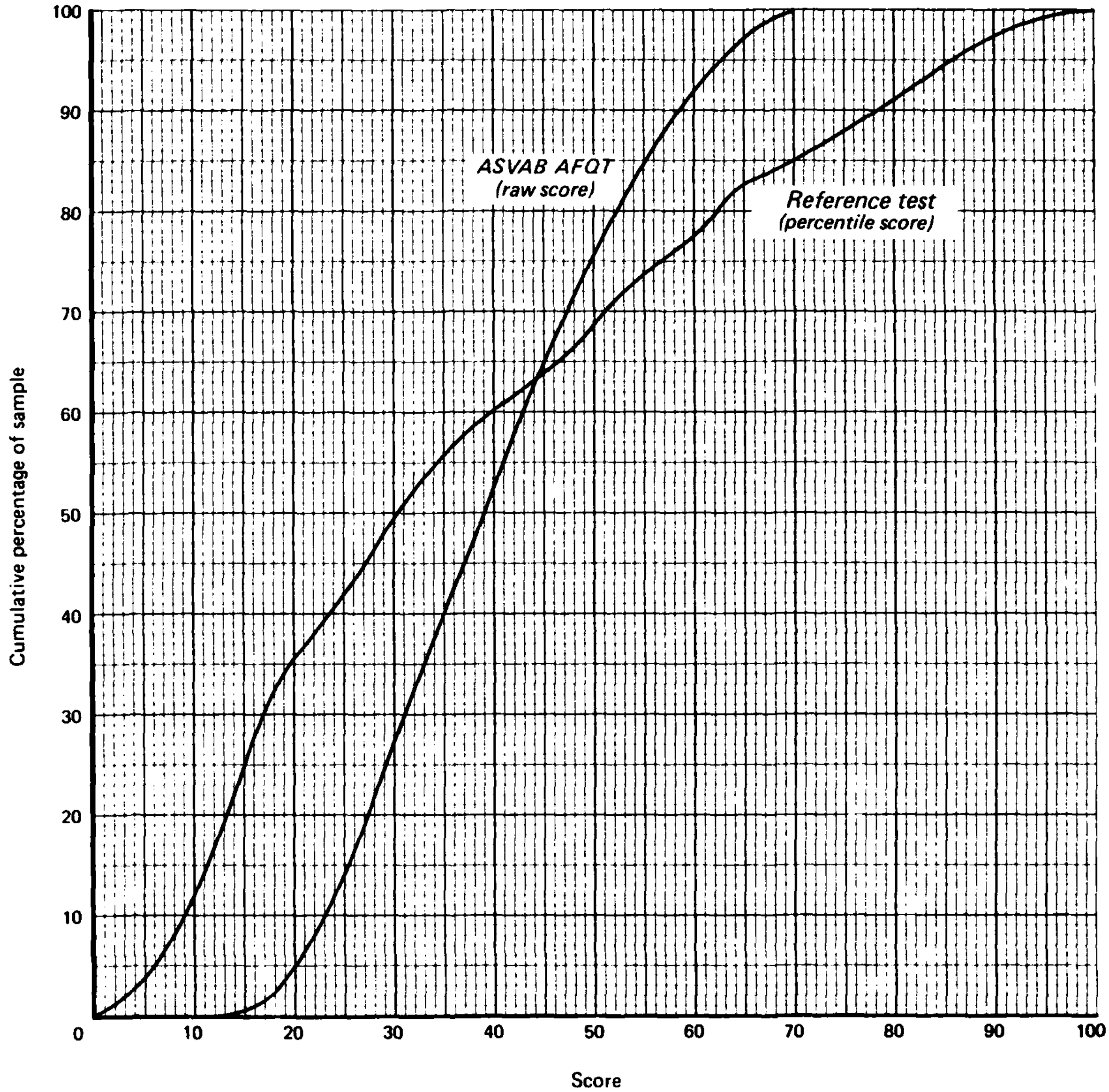


FIG. K-5: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 6/7 AFQT (DOD FULL-RANGE SAMPLE)

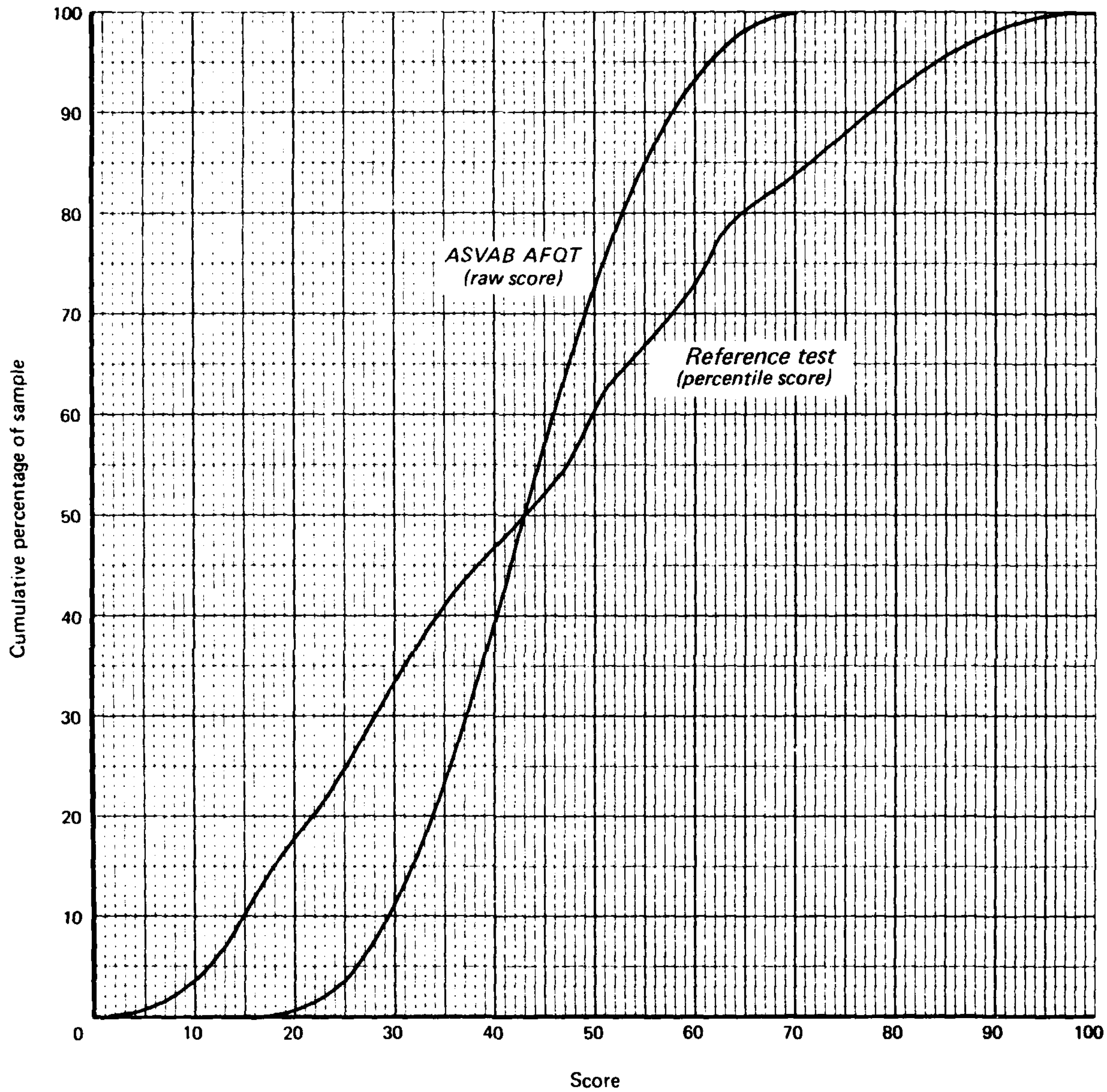


FIG. K-6: UNSTRATIFIED GRAPHICAL EQUATING FOR ASVAB 6/7 AFQT (DOD TRUNCATED SAMPLE)

TABLE K-2

COMPARISON OF EQUATING TECHNIQUES ON DOD 7E SAMPLE

ASVAB 7E AFQT raw score	Percentiles: Truncated sample	Percentiles: Full-range sample	Full-range case minus truncated case
14-15	1.0	1.0	0.0
16-17	1.0	2.3	1.3
18-19	2.0	4.4	2.2
20	3.5	5.2	1.7
21	4.0	6.3	2.3
22	5.6	7.6	2.0
23	7.5	8.6	1.1
24	9.0	9.6	0.6
25	10.0	10.5	0.5
26	10.2	11.7	1.5
27	11.0	12.5	1.5
28	13.1	13.6	0.5
29	14.0	14.5	0.5
30	14.8	15.5	0.7
31	16.0	16.5	0.5
32	17.0	17.5	0.5
33	18.5	19.5	1.0
34	20.0	21.3	1.3
35	22.3	23.4	1.1
36	24.2	25.0	0.8
37	26.2	27.1	0.9
38	27.8	28.3	0.5
39	29.8	30.3	0.5
40	31.8	32.3	0.5
41	34.4	34.5	0.1
42	36.7	36.7	0.0
43	39.8	39.3	-0.5
44	43.0	44.0	1.0
45	46.6	46.5	-0.1
46	48.7	48.6	-0.1
47	50.3	50.5	0.2
48	52.7	53.0	0.3
49	55.6	55.5	-0.1
50	58.3	58.5	0.2
51	60.9	61.0	0.1
52	62.5	62.5	0.0
53	65.6	64.5	-1.1
54	69.5	69.0	-0.5
55	73.0	73.0	0.0
56	75.5	75.5	0.0
57	77.6	77.0	-0.6
58	79.6	79.0	-0.6
59	81.2	80.6	-0.6
60	82.3	82.0	-0.3
61	83.5	83.0	-0.5
62	85.0	84.7	-0.3
63	87.0	86.5	-0.5
64	88.5	88.5	0.0
65	89.7	89.5	-0.2
66	91.3	91.5	0.2
67	92.5	93.0	0.5
68	94.5	95.0	0.5
69	97.0	97.0	0.0
70	99.0	99.0	0.0

TABLE K-3

COMPARISON OF EQUATING TECHNIQUES ON DoD 6/7 SAMPLE

ASVAB 6/7 AFQT raw score	Percentiles: Truncated sample	Percentiles: Full-range sample	Full-range case minus truncated case
14-15	0.0	1.0	1.0
16-17	1.0	2.5	1.5
18-19	2.0	4.6	2.6
20	4.5	6.0	1.5
21	5.5	6.9	1.4
22	6.7	8.0	1.3
23	8.0	9.3	1.3
24	9.5	10.2	0.7
25	10.3	11.2	0.9
26	11.3	12.2	0.9
27	12.5	13.2	0.7
28	13.8	14.1	0.3
29	14.5	15.0	0.5
30	16.0	16.1	0.1
31	17.0	17.2	0.2
32	18.2	18.3	0.1
33	19.5	19.5	0.0
34	21.5	22.0	0.5
35	23.5	23.6	0.1
36	25.0	25.5	0.5
37	26.5	27.1	0.6
38	28.4	28.7	0.3
39	30.5	30.9	0.4
40	32.7	32.6	-0.1
41	35.0	35.0	0.0
42	37.6	37.5	-0.1
43	41.0	40.7	-0.3
44	43.8	44.0	0.2
45	46.8	46.6	-0.2
46	49.0	48.6	-0.4
47	50.5	50.6	0.1
48	53.0	52.5	-0.5
49	55.5	55.0	0.0
50	58.0	57.5	0.5
51	60.5	60.2	-0.3
52	62.1	62.0	-0.1
53	63.8	62.8	-1.0
54	67.5	66.0	-1.5
55	70.5	69.4	-1.1
56	73.5	72.0	-1.5
57	75.8	74.5	-1.3
58	78.2	77.0	-1.2
59	79.6	79.0	-0.6
60	81.6	80.5	-1.1
61	82.8	82.4	-0.4
62	84.4	84.0	-0.4
63	86.0	86.3	0.3
64	88.0	88.3	0.3
65	90.0	90.0	0.0
66	91.0	91.0	0.0
67	93.0	93.0	0.0
68	94.0	94.0	0.0
69	95.0	97.0	2.0
70	98.0	98.0	0.0

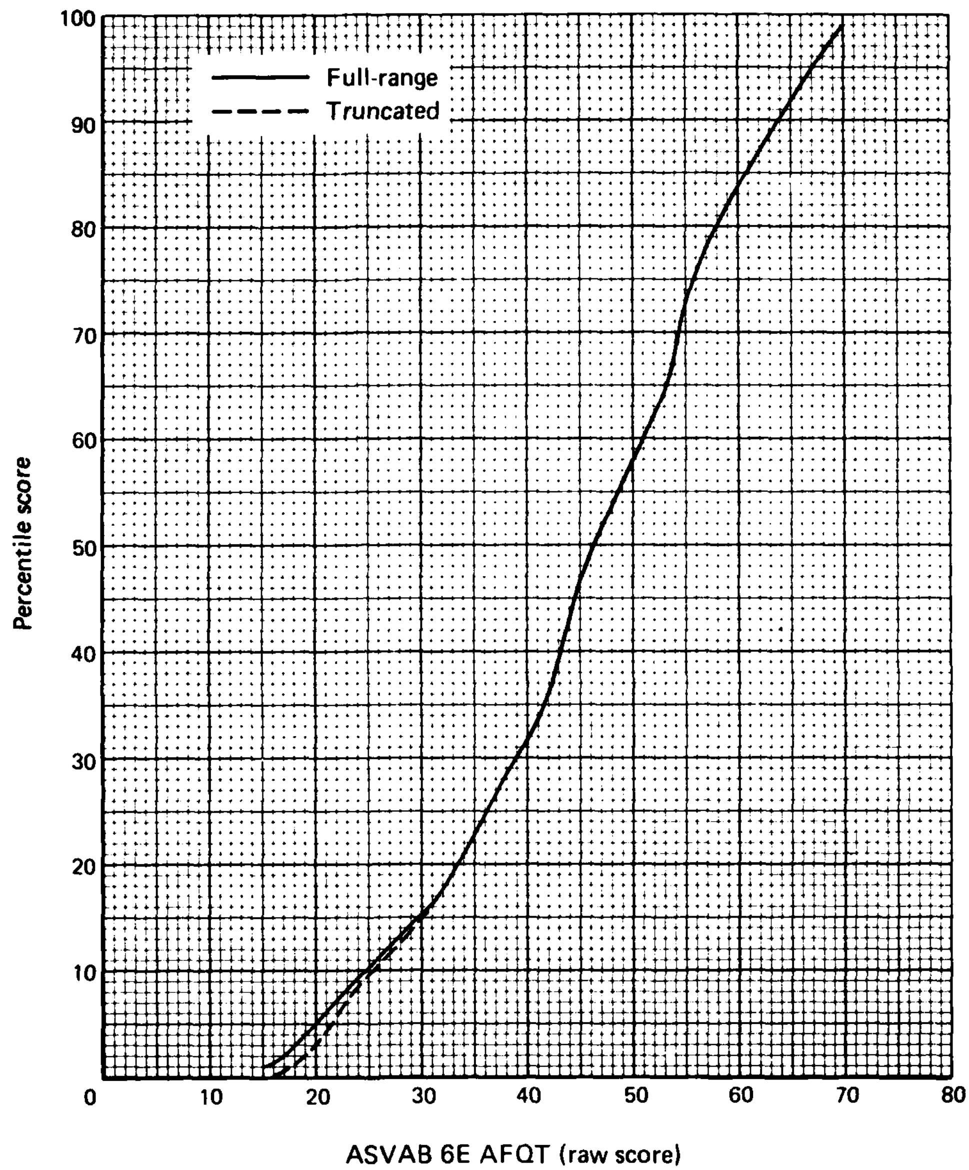


FIG. K-7: COMPARISON OF ASVAB 6E NORMS FROM UNSTRATIFIED GRAPHICAL EQUATING USING FULL-RANGE DOD AND TRUNCATED DOD DATA

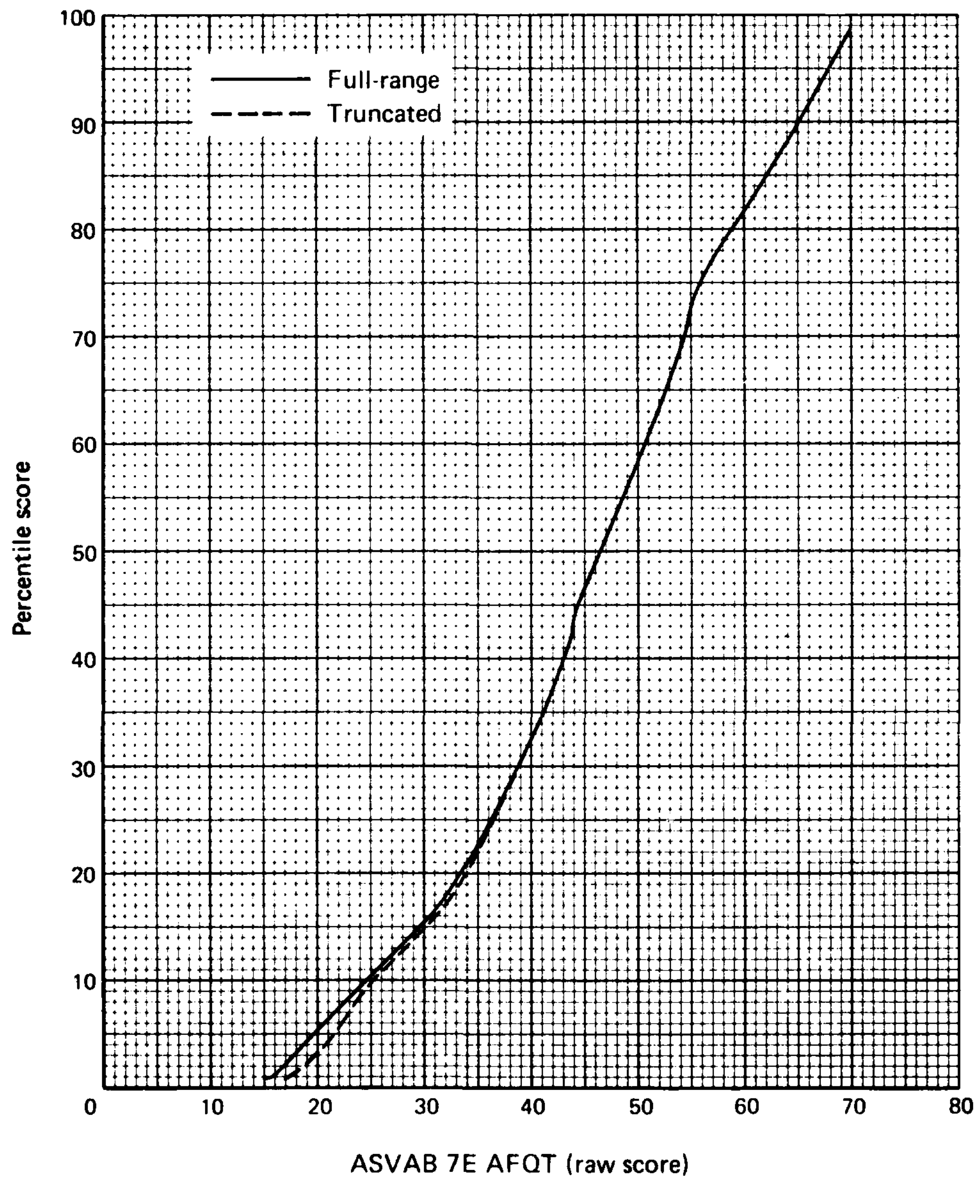


FIG. K-8: COMPARISON OF ASVAB 7E NORMS FROM UNSTRATIFIED GRAPHICAL EQUATING USING FULL-RANGE DOD AND TRUNCATED DOD DATA

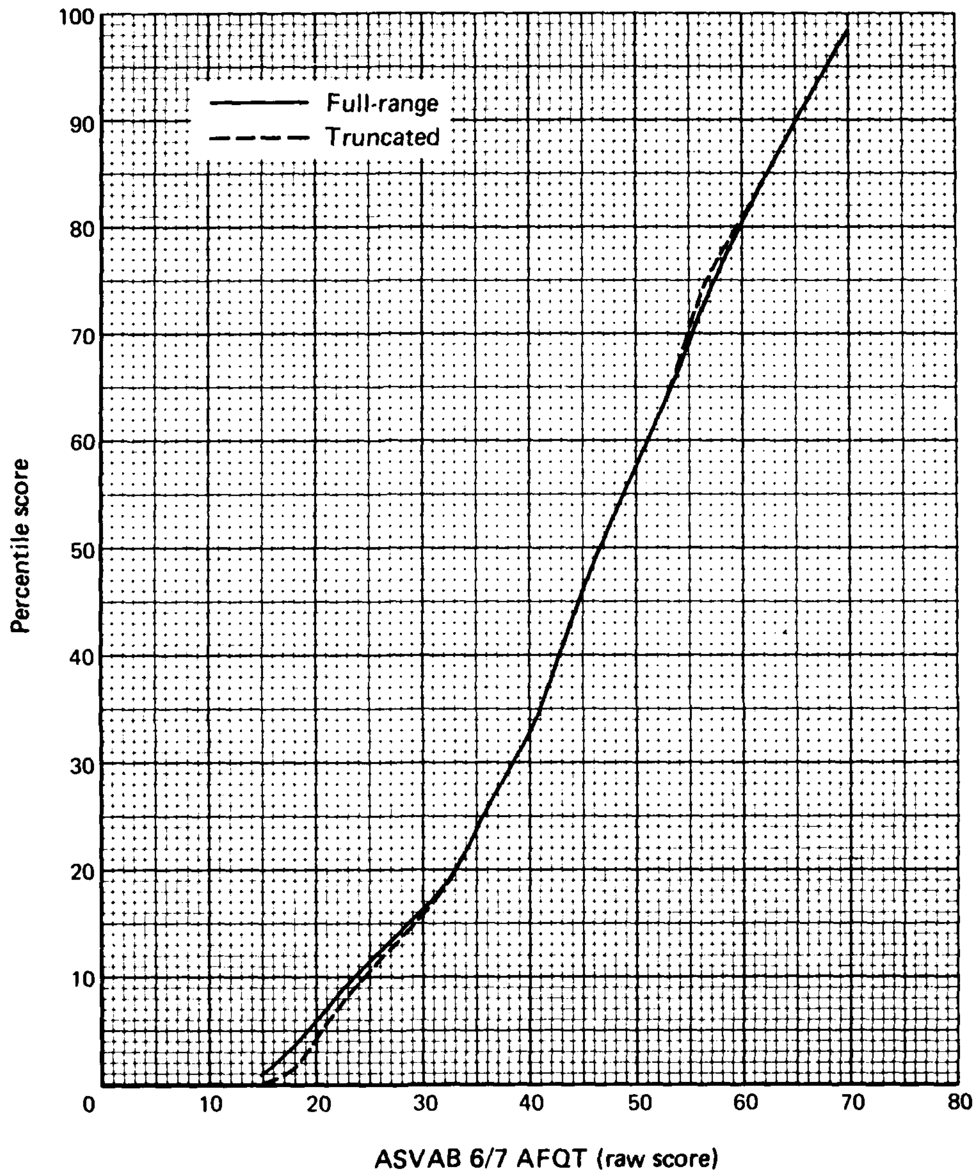


FIG. K-9: COMPARISON OF ASVAB 6/7 NORMS FROM UNSTRATIFIED GRAPHICAL EQUATING USING FULL-RANGE DOD AND TRUNCATED DOD DATA

APPENDIX L

SMOOTHING OF FINAL CONVERSION TABLES

APPENDIX L

SMOOTHING OF FINAL CONVERSION TABLES

A summary of the results from unstratified graphical equating applied to the CNA data sample in appendix E is reproduced in table L-1. The percentiles equated to each raw score are shown for ASVAB forms 6/7, 6E, and 7E. These distributions are shown graphically (figure L-1). This figure shows that the conversion tables for ASVAB 6/7 and ASVAB 6E are very similar, which suggests that a common conversion table can be used for both forms. The curve for ASVAB 7E is displaced to the left by about two raw score points in the central region and about three raw score points in the higher and lower percentiles. It appears that a constant could be added to each applicant's ASVAB 7E score that would slide the curve to the right and enable using a common conversion table for all current forms of ASVAB.

Table L-1 shows the result of adding two and three raw score points to the ASVAB 7E AFQT raw score before converting to percentiles. As seen, either system produces a norm table for ASVAB 7E that is more closely compatible with those for ASVAB 6/7 and ASVAB 6E. Table L-1 and figure L-1 confirm that adding two points appears to be the best approach over most of the percentile range of interest. To statistically test the compatibility of the separate norms for the three forms of ASVAB, we applied the conversion tables (table L-1) to an assumed mobilization population of the same size as our CNA subsamples and calculated the expected frequency distribution of applicants. The resulting distributions are shown in table L-2. We made a test for the homogeneity of parallel samples (table L-3). We see that the probability of observing differences as large as between ASVAB 6/7 and ASVAB 6E by chance if the two samples were parallel is about 0.07. We believe this is a good reason to use a common conversion table for ASVAB 6/7 and ASVAB 6E.

Similar comparisons for ASVAB 6/7 with ASVAB 7E and with variants of form 7E made by adding two and three points are also shown in table L-3. The chance probabilities are less than 0.00, and statistically the case for using the same conversion tables for forms 6/7 and 7E is not compelling. We do see that the chi-squared value is most favorable (i.e., lowest) for the case when two points were added to the ASVAB 7E score. Referring to figure L-1, we concluded that the practical difference between forms is small if two points are added to ASVAB 7E scores before converting to percentile scores. On this basis, we believe that a common conversion table is practical and construct (table L-1) the percentile associated with each raw score for mean of the three common forms--6/7, 6E, and 7E--.

TABLE L-1

SUMMARY OF UNSTRATIFIED GRAPHICAL EQUATING RESULTS

ASVAB AFQT Interval	Percentiles					
	ASVAB 6/7	ASVAB 6E	ASVAB 7E	ASVAB $7E_2$ [a]	ASVAB $7E_3$ [b]	Average of forms 6/7, 6E and $7E_2$
14-15	0.0	0.0	3.0	2.0	1.0	0.7
16-17	0.0	1.0	4.5	3.0	2.5	1.3
18-19	1.0	2.0	6.3	4.5	4.0	2.5
20	3.0	3.0	7.0	5.0	4.5	3.6
21	4.0	4.0	8.0	6.3	5.0	4.7
22	5.5	6.0	9.0	7.0	6.3	6.2
23	6.6	7.6	10.8	8.0	7.0	7.4
24	8.0	8.4	11.5	9.0	8.0	8.5
25	9.6	9.5	12.5	10.8	9.0	10.0
26	10.5	10.6	13.5	11.5	10.8	10.9
27	11.7	11.6	14.5	12.5	11.5	11.9
28	12.5	12.6	15.4	13.5	12.5	12.9
29	13.5	13.8	16.3	14.5	13.5	13.9
30	14.7	15.0	17.5	15.4	14.5	15.0
31	15.5	16.0	19.1	16.3	15.4	15.9
32	16.7	16.7	21.4	17.5	16.3	17.0
33	18.2	17.6	23.2	19.1	17.5	18.3
34	20.2	19.5	25.3	21.4	19.1	20.4
35	22.5	21.1	27.3	23.2	21.4	22.3
36	24.5	23.3	29.3	25.3	23.2	24.4
37	26.5	25.5	31.2	27.3	25.3	26.4
38	28.6	27.7	33.3	29.3	27.3	28.5
39	31.2	30.3	36.4	31.2	29.3	30.9
40	33.4	32.0	39.6	33.3	31.2	32.9
41	36.0	34.5	42.7	36.4	33.3	35.6
42	38.4	37.4	45.8	39.6	36.4	38.5
43	43.0	40.4	48.6	42.7	39.6	42.0
44	46.2	43.8	51.2	45.8	42.7	45.3
45	48.7	47.5	53.4	48.6	45.8	48.3
46	50.6	49.8	56.4	51.2	48.6	50.5
47	53.0	51.1	58.3	53.4	51.2	52.5
48	55.5	53.2	60.3	56.4	53.4	55.0
49	57.6	56.5	62.0	58.3	56.4	57.5
50	60.0	59.5	63.8	60.3	58.3	60.0
51	61.8	61.2	67.2	62.0	60.3	61.7
52	63.3	62.6	71.7	63.8	62.0	63.2
53	65.0	65.0	75.3	67.2	63.8	65.7
54	69.0	68.5	77.3	71.7	67.2	69.7
55	72.0	71.0	78.5	75.3	71.7	72.8
56	74.4	73.3	80.0	77.3	75.3	75.0
57	76.5	76.6	81.4	78.5	77.3	77.2
58	78.5	78.4	82.6	80.0	78.5	79.0
59	79.7	80.0	84.0	81.4	80.0	80.4
60	81.5	81.6	85.5	82.6	81.4	81.9
61	82.6	83.3	87.0	84.0	82.6	83.3
62	84.5	85.0	88.2	85.5	84.0	85.0
63	86.5	86.5	89.5	87.0	85.5	86.7
64	88.2	88.0	91.0	88.2	87.0	88.1
65	90.0	89.2	92.0	89.5	88.2	89.6
66	92.0	90.5	94.0	91.0	89.5	91.2
67	93.0	94.0	95.0	92.0	91.0	93.0
68	96.0	96.0	96.0	94.0	92.0	95.3
69	98.0	98.0	97.0	95.0	94.0	96.3
70	99.0	99.0	99.0	99.0	99.0	99.0

[a] Before converting to percentiles, two points are added to the raw AFQT score.

[b] Before converting to percentiles, three points are added to the raw AFQT score.

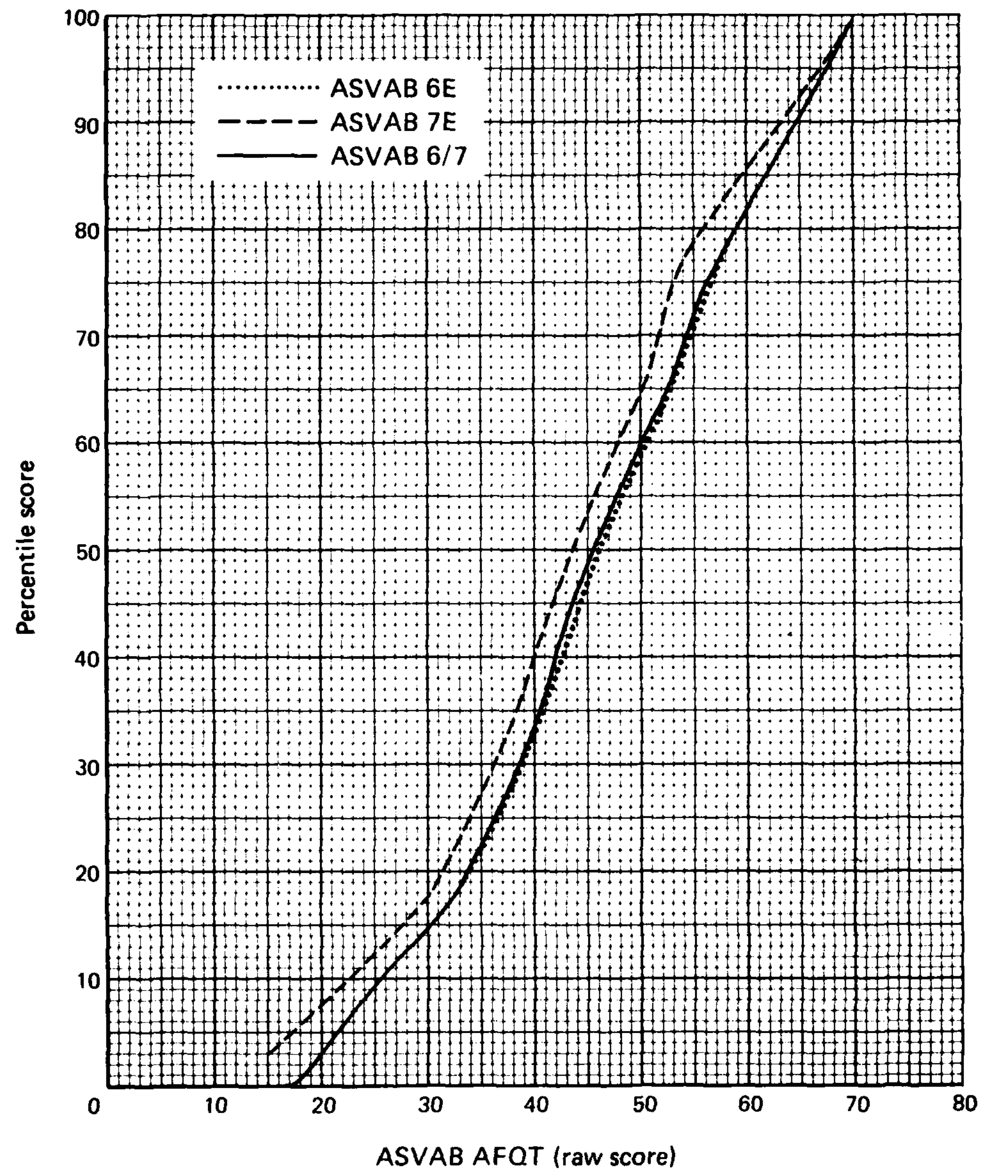


FIG. L-1: COMPARISON OF NORMS FOR VARIOUS FORMS OF ASVAB

TABLE L-2

INFERRED FREQUENCY DISTRIBUTIONS FOR SEPARATE NORMS OF EACH FORM OF ASVAB

ASVAB AFQT interval (raw score) (1)	Percentage of sample inferred to be in indicated interval				
	Form 6/7 (2)	Form 6E (3)	Form 7E (4)	Form $7E_2$[a] (5)	Form $7E_3$[b] (6)
0-19	1.0	2.0	6.3	4.5	4.0
20-22	4.5	4.0	2.7	2.5	2.3
23-25	4.1	3.5	3.5	3.8	2.7
26-28	2.9	3.1	2.9	2.7	3.5
29-31	3.0	3.4	3.7	2.8	2.9
32-34	4.7	3.5	6.2	5.1	3.7
35-37	6.3	6.0	5.9	5.9	6.2
38-40	6.9	6.5	8.4	6.0	5.9
41-43	9.6	8.4	9.0	9.4	8.4
44-46	7.6	9.4	7.8	8.5	9.0
47-49	7.0	6.7	5.6	7.1	7.8
50-52	5.7	6.1	9.7	5.5	5.6
53-55	8.7	8.4	6.8	11.5	9.7
56-58	6.5	7.4	4.1	4.7	6.8
59-61	4.1	4.9	4.4	4.0	4.1
62-64	5.6	4.7	4.0	4.2	4.4
65-67	4.8	6.0	4.0	3.8	3.0
68-70	6.0	5.0	4.0	7.0	8.0
Sample size	2,208	1,634	1,660	1,660	1,660

[a]Two points added to each form 7E raw AFQT score.

[b]Three points added to each form 7E raw AFQT score.

TABLE L-3

TEST FOR EQUIVALENCE OF SEPARATE NORMS FOR EACH FORM OF ASVAB

ASVAB forms compared	Chi-squared[a]	Degrees of freedom	Probability of chance difference this large
6/7 with 6E	26.2	17	0.07
6/7 with 7E	152.3	17	0.00
6/7 with $7E_2$[b]	81.2	17	0.00
6/7 with $7E_3$[c]	82.7	17	0.00

[a]For a test of the homogeneity of parallel samples.

[b]Two points added to each form 7E raw AFQT score.

[c]Three points added to each form 7E raw AFQT score.

In table L-4 (from appendix K) we show the estimated adjustments (based on DoD data) that would be necessary to completely remove any effects due to truncation of the CNA sample. The mean of the adjustments is also shown. The adjustments are very small (less than 1 percentile) except in the region below the 8th percentile. In the region above the 15th percentile, the adjustments are generally less than 0.5 percentile and are not consistent from form to form. This may suggest inaccuracies in estimation more than a real bias that needs an adjustment. Accordingly, we smoothed the adjustments in the region below the 15th percentile and added the adjustment to get the corrected mean percentile shown in table L-5.

The corrected mean percentiles were smoothed in two stages, as shown in table L-5. In the first stage, percentiles were rounded off to whole numbers paying attention to the need for raw scores to correspond to percentiles at critical points that separate official mental groups (16, 21, 31, 50, 65, and 93rd percentiles). The result of this partial smoothing is shown in column 5 of table L-5. Further smoothing was done to eliminate an atypical progression of scores as shown in figure L-2. In our opinion this unnatural score distribution is the result of anomalies in the official conversion table for the reference test AFQT 7A shown in figure L-3. An examination of figure L-3 discloses a number of unusual undulations in the curve. Most of these undulations were removed in the hard smoothing carried out during the graphical equating in appendix H. However, it appears from figure L-2 that one anomaly remained. We

TABLE L-4

SUMMARY OF ADJUSTMENTS FOR TRUNCATION EFFECT ON NORMS PRODUCED BY UNSTRATIFIED GRAPHICAL EQUATING

ASVAB AFQT raw score	Adjustment: ASVAB 6E sample	Adjustment: ASVAB 7E sample	Adjustment: ASVAB 6/7 sample	Adjustment: Mean value	Adjustment: Smoothed	Approximate percentile[a]
14-15	1.0	0.0	1.0	0.7	0.7	0.7
16-17	1.0	1.3	1.5	1.3	1.3	1.3
18-19	2.0	2.2	2.6	2.3	2.3	2.5
20	2.0	1.7	1.5	1.7	1.8	3.6
21	2.0	2.3	1.4	1.9	1.8	4.7
22	1.0	2.0	1.3	1.4	1.4	6.2
23	0.7	1.1	1.3	1.0	1.0	7.4
24	0.8	0.6	0.7	0.7	0.9	8.5
25	0.9	0.5	0.9	0.8	0.9	10.0
26	0.8	1.5	0.9	1.1	0.9	10.9
27	0.5	1.5	0.7	0.9	0.9	11.9
28	0.8	0.5	0.3	0.5	0.5	12.9
29	0.6	0.5	0.5	0.5	0.5	13.9
30	0.4	0.7	0.1	0.4	0.0	15.0
31	0.3	0.5	0.2	0.3	0.0	15.9
32	0.3	0.5	0.1	0.3	0.0	17.0
33	0.0	1.0	0.0	0.3	0.0	18.3
34	0.0	1.3	0.5	0.6	0.0	20.4
35	0.0	1.1	0.1	0.4	0.0	22.3
36	0.0	0.8	0.5	0.4	0.0	24.4
37	0.3	0.9	0.6	0.6	0.0	26.4
38	0.4	0.5	0.3	0.4	0.0	28.5
39	0.5	0.5	0.4	0.5	0.0	30.9
40	0.1	0.5	-0.1	0.2	0.0	32.9
41	-0.5	0.1	0.0	-0.1	0.0	35.6
42	0.0	0.0	-0.1	0.0	0.0	38.5
43	-0.7	-0.5	-0.3	-0.5	0.0	42.0
44	0.0	1.0	0.2	0.4	0.0	45.3
45	-0.3	-0.1	-0.2	-0.2	0.0	48.3
46	-0.6	-0.1	-0.4	-0.4	0.0	50.5
47	-1.0	0.2	0.1	-0.1	0.0	52.5
48	-0.5	0.3	-0.5	-0.2	0.0	55.0
49	-0.5	-0.1	0.0	-0.2	0.0	57.5
50	-0.8	0.2	0.5	0.0	0.0	60.0
51	-0.5	0.1	-0.3	-0.2	0.0	61.7
52	-0.5	0.0	-0.1	-0.2	0.0	63.2
53	-0.7	-1.1	-1.0	-0.9	0.0	65.7
54	0.7	-0.5	-1.5	-0.4	0.0	69.7
55	-1.0	0.0	-1.1	-0.7	0.0	72.8
56	0.5	0.0	-1.5	-0.3	0.0	75.0
57	0.3	-0.6	-1.3	-0.5	0.0	77.2
58	-0.5	-0.6	-1.2	-0.8	0.0	79.0
59	0.0	-0.6	-0.6	-0.4	0.0	80.4
60	0.2	-0.3	-1.1	-0.4	0.0	81.9
61	0.5	-0.5	-0.4	-0.3	0.0	83.3
62	-0.4	-0.3	-0.4	-0.4	0.0	85.0
63	-0.1	-0.5	0.3	-0.1	0.0	86.7
64	0.0	0.0	0.3	0.1	0.0	88.1
65	0.0	-0.2	0.0	-0.1	0.0	89.6
66	0.0	0.2	0.0	0.1	0.0	91.2
67	-0.5	0.5	0.0	0.0	0.0	93.0
68	0.0	0.5	0.0	0.2	0.0	95.3
69	0.0	0.0	2.0	0.7	0.0	96.3
70	0.0	0.0	0.0	0.0	0.0	99.0

[a]Mean value from table L-1.

TABLE L-5

SMOOTHED CONVERSION TABLE FOR ASVAB 6/7/6E/7E

ASVAB AFQT interval	Percentile: Mean of CNA results for forms 6E, 7E$_2$, and 6/7 (1)	Percentile: Correction for preselection (2)	Percentile: Corrected mean (3)	Partially smoothed percentiles (4)	Fully smoothed percentiles (5)
14-15	0.7	0.7	1.4	0-1	0-1
16-17	1.3	1.3	2.6	1-2	1-2
18-19	2.5	2.3	4.8	3-4	3-4
20	3.6	1.8	5.4	5	5
21	4.7	1.8	6.5	6	6
22	6.2	1.4	7.6	7	7
23	7.4	1.0	8.4	8	8
24	8.5	0.9	9.4	9	9
25	10.0	0.9	10.9	10	10
26	10.9	0.9	11.8	11	11
27	11.9	0.9	12.8	12	12
28	12.9	0.5	13.4	13	13
29	13.9	0.5	14.4	14	14
30	15.0		15.0	15	15
31	15.9		15.9	16	16
32	17.0		17.0	17	17
33	18.3		18.3	18	18
34	20.4		20.4	21	21
35	22.3		22.3	22	22
36	24.4		24.4	24	24
37	26.4		26.4	26	26
38	28.5		28.5	28	28
39	30.9		30.9	31	31
40	32.9		32.9	33	33
41	35.6		35.6	36	36
42	38.5		38.5	39	39
43	42.0		42.0	42	42
44	45.3		45.3	45	45
45	48.3		48.3	48	48
46	50.5		50.5	50	50
47	52.5		52.5	52	52
48	55.0		55.0	55	55
49	57.5		57.5	57	58
50	60.0		60.0	60	61
51	61.7		61.7	62	63
52	63.2		63.2	63	65
53	65.7		65.7	65	67
54	69.7		69.7	69	69
55	72.8		72.8	72	71
56	75.0		75.0	75	73
57	77.2		77.2	77	75
58	79.0		79.0	79	77
59	80.4		80.4	80	79
60	81.9		81.9	82	81
61	83.3		83.3	83	83
62	85.0		85.0	85	85
63	86.7		86.7	87	87
64	88.1		88.1	88	88
65	89.6		89.6	90	90
66	91.2		91.2	91	91
67	93.0		93.0	93	93
68	95.3		95.3	95	95
69	96.3		96.3	97	97
70	99.0		99.0	99	99

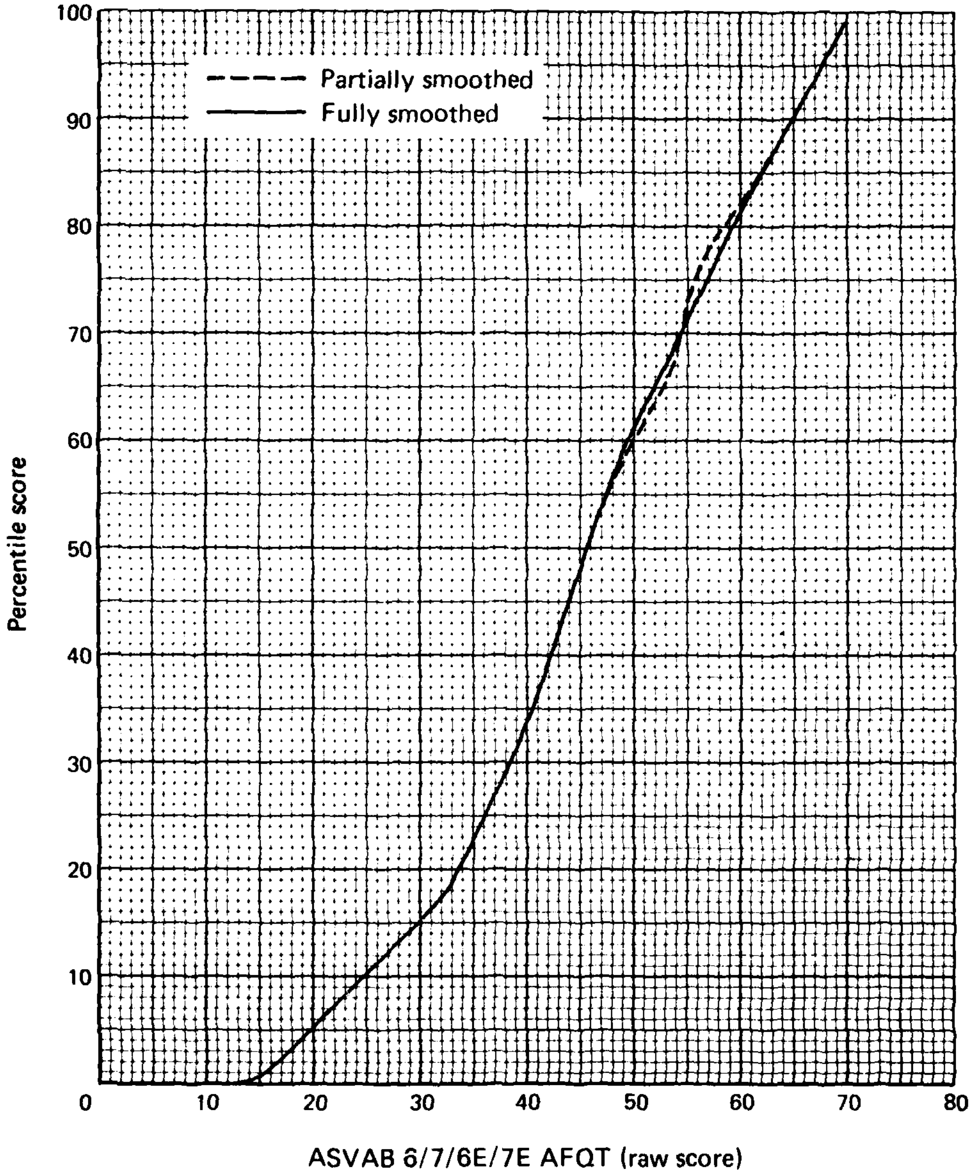


FIG. L-2: COMPARISON OF PARTIALLY SMOOTHED AND FULLY SMOOTHED NORMS

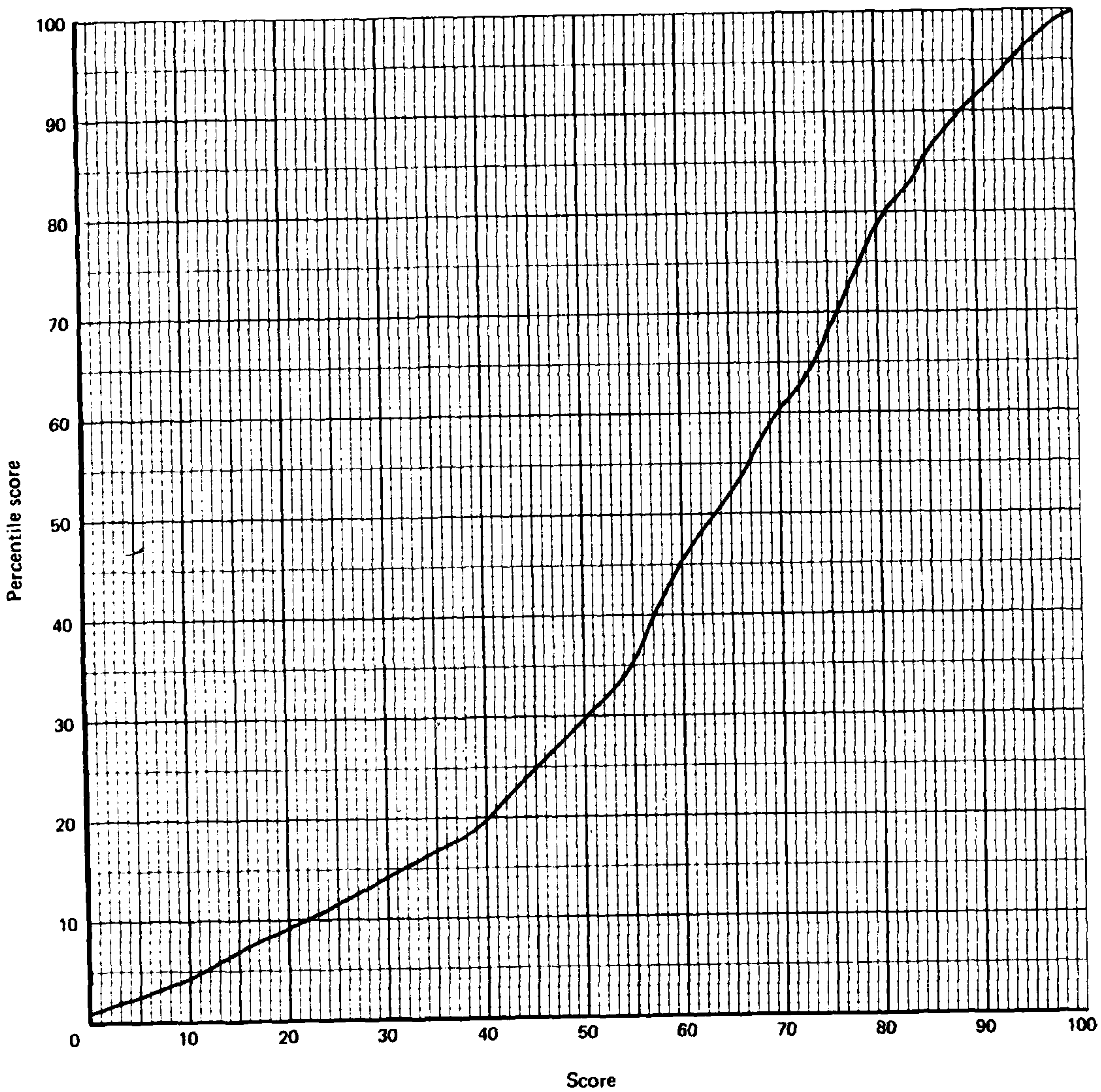


FIG. L-3: OFFICIAL CONVERSION TABLE FOR REFERENCE TEST AFQT 7A removed the anomaly by the smoothing shown in column 5 of table L-5 and illustrated in figure L-2. The ancestry of the resultant curve with its sharp break at the 20th percentile is apparent by examining the conversion table for the reference test shown in figure L-3.

APPENDIX M

CONVERSION TABLES FOR COMPOSITES

APPENDIX M

CONVERSION TABLES FOR COMPOSITES

To avoid bias from sample stratification, we used the unstratified equipercentile equating technique to build composite conversion tables. Each composite in raw score form was equated to the ASVAB 6/7 AFQT score. This procedure is possible because of the high correlations between the AFQT score and the composite scores. The definitions of the composites are given in appendix A. We used sample 5 as the data set.

For purposes of this equating the ASVAB 6/7 AFQT score was expressed in percentiles (for Air Force composites) and Army Standard Scores (for Army and Marine Corps composites). Traditional conversion tables (annex M-1) were used to convert AFQT from percentile form to Army Standard Score form.

The resulting conversion tables are given in tables M-1 through M-6.

TABLE M-1

ARMY AND MARINE CORPS ASVAB 6E/7E/6/7 CONVERSION
TABLES FOR COMPOSITES

Raw score	Composite score: GT	GM	EL[a]	CL	MM	Raw score
105-110	-	-	-	-	135	105-110
104					132	104
103					132	103
102					131	102
101					130	101
100	-	-	-	135	129	100
99				135	128	99
98				135	127	98
97				135	126	97
96				135	126	96
95	-	-	-	135	125	95
94				135	124	94
93				135	123	93
92				135	122	92
91				135	121	91
90	-	-	135	135	121	90
89			135	133	120	89
88			135	132	119	88
87			135	132	118	87
86			131	132	117	86
85	-	-	130	131	116	85
84			129	130	115	84
83			128	130	114	83
82			128	129	113	82
81			127	128	113	81
80	-		126	127	112	80
79			125	126	111	79
78		135	124	125	110	78
77		132	124	125	109	77
76		131	123	124	108	76
75	-	131	122	123	107	75
74		130	121	122	106	74
73		128	120	120	105	73
72		128	119	119	104	72
71		127	118	117	103	71
70	-	126	117	117	102	70
69		126	117	115	101	69
68		125	116	114	100	68
67		124	115	113	99	67
66		123	114	111	98	66
65	-	122	113	110	97	65
64		120	113	109	96	64
63		119	112	108	95	63
62		118	111	107	94	62
61		117	110	105	93	61

[a]Marine Corps only.

AD-A094 684 CENTER FOR NAVAL ANALYSES ALEXANDRIA VA MARINE CORP--ETC F/G 5/9
A REEXAMINATION OF THE NORMALIZATION OF THE ARMED SERVICES VOCA--ETC(U)
APR 80 W H SIMS, A R TRUSS N00014-76-C-0001
UNCLASSIFIED CNS-1152 NL

3 OF 3
AD
A094684

END
DATE
FILMED
3 81
DTIC

TABLE M-1 (Cont'd)

Raw score	GT	GM	EL[a]	CL	MM	Raw score
	Composite score					
60	-	116	109	103	92	60
59		115	108	102	91	59
58		114	107	101	90	58
57		113	106	99	89	57
56		112	104	98	88	56
55	-	111	103	96	86	55
54		110	102	94	85	54
53		109	101	92	84	53
52		108	100	90	82	52
51		107	99	88	81	51
50	135	106	98	86	81	50
49	131	104	97	84	80	49
48	126	103	95	82	79	48
47	124	102	94	81	78	47
46	123	101	92	79	77	46
45	121	99	91	78	76	45
44	119	98	90	77	75	44
43	117	97	88	75	73	43
42	116	96	86	73	72	42
41	115	94	85	71	71	41
40	113	93	83	70	70	40
39	112	92	82	68	68	39
38	111	90	81	67	67	38
37	109	89	80	65	66	37
36	108	87	79	64	65	36
35	106	85	78	63	64	35
34	104	83	76	62	63	34
33	102	82	75	61	62	33
32	100	81	74	61	61	32
31	98	79	72	59	60	31
30	96	78	71	57	59	30
29	94	76	69	55	57	29
28	92	75	68	54	55	28
27	89	73	67	53	53	27
26	87	72	65	53	53	26
25	85	70	64	53	53	25
24	82	68	63	53	53	24
23	80	66	62	53	53	23
22	78	65	61	53	53	22
21	77	64	58	53	53	21
20	75	62	55	53	53	20
19	73	61	53	53	53	19
18	71	58	53	53	53	18
17	69	53	53	53	53	17
16	67	53	53	53	53	16
15	65	53	53	53	53	15
14	64	53	53	53	53	14
13	61	53	53	53	53	13
12	60	53	53	53	53	12
11	55	53	53	53	53	11
0-10	53	53	53	53	53	0-10

TABLE M-2

ARMY AND MARINE CORPS ASVAB 6E/7E/6/7 CONVERSION TABLE FOR COMPOSITES

Raw score	SC	CO	FA	OF	ST	Raw score
	Composite score					
110-117	-	135	-	-	-	110-117
109		135				109
108		132				108
107		130				107
106		130				106
105	-	130	135	-	-	105
104		129	135			104
103		129	135			103
102		129	135			102
101		128	135			101
100	-	128	135	-	-	100
99		128	135			99
98		127	135			98
97		127	135			97
96		126	132			96
95	-	125	131	-	-	95
94		124	130			94
93		123	130			93
92		122	129			92
91		121	128			91
90	135	120	127	-	-	90
89	135	118	127			89
88	135	117	126			88
87	135	115	125			87
86	132	114	125			86
85	132	113	124	-	-	85
84	130	112	123			84
83	128	111	122			83
82	127	110	122			82
81	126	109	121			81
80	125	108	120	-	-	80
79	124	106	119			79
78	123	105	118			78
77	122	104	117			77
76	121	102	116			76
75	120	100	115	-	-	75
74	119	99	114			74
73	118	98	113			73
72	117	96	112			72
71	116	95	111			71
70	115	94	111	-	-	70
69	113	93	110			69
68	112	91	109			68
67	112	90	107			67
66	111	88	106			66

TABLE M-2 (Cont'd)

Raw score	SC	CO	FA	OF	ST	Raw score
	Composite score					
65	110	86	105	-	-	65
64	109	85	103			64
63	108	83	102			63
62	107	82	101			62
61	106	82	100			61
60	105	79	99	-	135	60
59	103	78	98		135	59
58	102	77	96		135	58
57	101	76	95		131	57
56	100	75	92		130	56
55	99	74	92	135	128	55
54	98	72	91	135	126	54
53	96	70	89	135	125	53
52	95	70	88	135	123	52
51	94	68	86	134	122	51
50	92	66	84	133	121	50
49	91	66	83	133	120	49
48	89	65	82	132	119	48
47	87	64	81	132	118	47
46	85	63	79	130	116	46
45	84	62	78	127	115	45
44	82	61	77	126	114	44
43	81	59	76	125	113	43
42	80	57	74	123	112	42
41	78	57	73	121	111	41
40	77	55	72	118	109	40
39	76	55	71	115	108	39
38	75	53	69	114	107	38
37	73	53	67	112	105	37
36	72	53	66	110	104	36
35	70	53	65	108	102	35
34	69	53	64	105	101	34
33	67	53	62	102	100	33
32	66	53	61	100	98	32
31	65	53	60	98	96	31
30	64	53	58	95	95	30
29	63	53	57	93	93	29
28	61	53	54	90	91	28
27	59	53	53	88	89	27
26	57	53	53	84	87	26
25	55	53	53	82	84	25
24	54	53	53	80	82	24
23	53	53	53	78	80	23
22	53	53	53	76	78	22
21	53	53	53	73	76	21
20	53	53	53	71	75	20
19	53	53	53	68	72	19
18	53	53	53	66	69	18
17	53	53	53	64	66	17
16	53	53	53	62	64	16
15	53	53	53	60	62	15
14	53	53	53	58	60	14
13	53	53	53	55	58	13
12	53	53	53	53	54	12
11	53	53	53	53	53	11
0-10	53	53	53	53	53	0-10

TABLE M-3

ARMY ONLY CONVERSION TABLES FOR ASVAB 6E/7E/6/7 EL COMPOSITE

Raw score	EL	Raw score	EL
103-110	135	64	100
102	134	63	99
101	133	62	98
100	132	61	97
99	132	60	96
98	129	59	95
97	129	58	93
96	128	57	92
95	127	56	91
94	127	55	90
93	126	54	89
92	126	53	87
91	125	52	86
90	124	51	85
89	123	50	83
88	123	49	82
87	122	48	81
86	121	47	80
85	120	46	78
84	119	45	77
83	118	44	76
82	117	43	75
81	116	42	74
80	116	41	73
79	115	40	72
78	114	39	70
77	113	38	69
76	112	37	68
75	112	36	66
74	111	35	65
73	110	34	64
72	109	33	62
71	108	32	61
70	106	31	60
69	105	30	57
68	104	29	56
67	103	28	56
66	102	27	55
65	101	26	54
		0-25	53

TABLE M-4

AIR FORCE CONVERSION TABLE
ASVAB 6E/7E/6/7

Mechanical aptitude index		Administrative aptitude index		General aptitude index		Electronic aptitude index	
Raw score	AI	Raw score	AI	Raw score	AI	Raw score	AI
57 & above	95	102 & above	95	49 & above	95	67 & above	95
54-56	90	96-101	90	47-48	90	63-66	90
52-53	85	89-95	85	46	85	60-62	85
50-51	80	86-88	80	44-45	80	58-59	80
48-49	75	82-85	75	41-43	75	56-57	75
45-47	70	79-81	70	40	70	53-55	70
43-44	65	76-78	65	38-39	65	51-52	65
42	60	74-75	60	36-37	60	49-50	60
40-41	55	72-73	55	35	55	47-48	55
38-39	50	69-71	50	33-34	50	45-46	50
35-37	45	67-68	45	31-32	45	43-44	45
34	40	64-66	40	30	40	42	40
32-33	35	62-63	35	29	35	40-41	35
30-31	30	60-61	30	27-28	30	38-39	30
27-29	25	57-59	25	25-26	25	36-37	25
24-26	20	53-56	20	23-24	20	32-35	20
20-23	15	48-52	15	20-22	15	28-31	15
16-19	10	38-47	10	15-19	10	23-27	10
11-15	05	29-37	05	11-14	05	18-22	05
10 & below	01	28 & below	01	10 & below	01	17 & below	01

TABLE M-5

MARINE CORPS ONLY ASVAB 6E/7E/6/7 CONVERSION
TABLE FOR GCT COMPOSITE
(expressed in Army Standard Score)

Raw score	GCT	Raw score	GCT
70	135	45	99
69	135	44	97
68	131	43	96
67	129	42	94
66	128	41	93
65	127	40	91
64	124	39	89
63	123	38	86
62	122	37	84
61	120	36	83
60	119	35	81
59	117	34	79
58	116	33	78
57	115	32	76
56	114	31	75
55	113	30	73
54	112	29	71
53	111	28	70
52	109	27	68
51	108	26	66
50	107	25	65
49	105	24	63
48	104	23	62
47	102	22	61
46	101	21	59
		20	55
		19	54
		0-18	53

TABLE M-6

ARMY ONLY CONVERSION TABLE FOR ASVAB 6/7/6E/7E WST[a]

Raw score	Percentile
50	97
49	95
48	91
47	89
46	86
45	84
44	82
43	79
42	77
41	74
40	72
39	69
38	67
37	64
36	61
35	58
34	54
33	51
32	48
31	45
30	42
29	37
28	34
27	31
26	29
25	26
24	23
23	21
22	19
21	17
20	16
19	15
18	13
17	12
16	11
15	10
14	9
13	7
12	6
11	5
10	3
0 -9	1

[a]WST is GT expressed in percentile form.

ANNEX M-1

TRADITIONAL CONVERSIONS FROM PERCENTILES TO ARMY STANDARD SCORE FROM AFQT 1 AND AFQT 2

TABLE M-1-1

CONVERSION TABLE: AFQT 1 OR AFQT 2
PERCENTILE SCORES TO ARMY STANDARD SCORES

Percentile	Standard score	Percentile	Standard score
100	164	28	86
100	157	27	85
100	151	26	84
100	146	24	83
99	142	23	82
98	139	22	81
97	137	21	80
96	134	20	79
95	131	19	78
93	130	18	77
92	128	17	76
90	126	16	75
89	125	15	73
87	123	14	71
85	122	13	70
84	121	12	69
82	120	12	68
80	118	11	66
78	117	10	65
76	116	9	64
74	115	9	63
73	114	8	62
71	113	7	61
69	112	7	60
67	111	6	59
65	110	5	57
63	109	5	56
61	107	4	55
59	106	4	53
57	105	3	52
55	104	3	50
53	103	2	48
51	101	2	47
49	100	2	45
47	99	2	43
45	98	2	42
43	97	2	42
41	96	1	41
39	95	1	41
37	94	1	40
36	93	1	39
34	92	1	39
32	91	1	39
31	90	1	39
30	88	1	39

APPENDIX N

STRATIFICATION ON ASVAB 6/7 PERCENTILE SCORE

APPENDIX N

STRATIFICATION ON ASVAB 6/7 PERCENTILE SCORE

We stratified sample 5 on ASVAB 6/7 AFQT percentiles by using the weights calculated in table N-1. Subtests were normed from this stratified sample.

TABLE N-1

CALCULATION OF WEIGHTS TO STRATIFY SAMPLE 5 ON ASVAB 6/7 AFQT SCORE

ASVAB 6/7 AFQT percentile (1)	Observed cases (2)	Expected cases (3)	Weight[a] (4)
0-5	20	110.4	5.520
6-10	61	110.4	1.810
11-15	138	110.4	.800
16-20	141	110.4	.783
21-25	195	110.4	.566
26-30	128	110.4	.863
31-35	180	110.4	.613
36-40	144	110.4	.767
41-45	155	110.4	.712
46-50	172	110.4	.642
51-55	143	110.4	.772
56-60	64	110.4	1.725
61-65	163	110.4	.677
66-70	99	110.4	1.115
71-75	138	110.4	.800
76-80	69	110.4	1.600
81-85	89	110.4	1.240
86-90	66	110.4	1.673
91-95	37	110.4	2.984
96-100	6	110.4	18.400
Total		2,208	

[a]Column (3) divided by column (2).

APPENDIX O

CONVERSION TABLES FOR SUBTESTS

APPENDIX O

CONVERSION TABLES FOR SUBTESTS

To build the subtest conversion tables, we chose ASVAB 6/7 (sample 5) as a representative sample of ASVAB 6E, 7E, and 6/7. Using the weights developed in appendix N we simulated the mobilization population and obtained mean and standard deviation statistics for all subsets. These statistics are computed from the following equation for all possible scores (x) on each subtest:[1]

$$\text{Navy Standard Score } (x) = 50 + \frac{10(x-\bar{x})}{\sigma_x} .$$

The resultant conversion tables for all subtests are shown in tables O-1 through O-4.

---

[1] $\bar{x}$ denotes mean value of y and σ_x denotes the standard deviation.

TABLE O-1

ASVAB 6E/7E/6/7 SUBTEST CONVERSION TABLES
(in Navy Standard Score)

Raw score	General information (GI)	Numerical operations (NO)	Attention to detail (AD)	Word knowledge (WK)	Raw score
50	-	69	-	-	50
49		68			49
48		67			48
47		65			47
46		64			46
45	-	63	-	-	45
44		62			44
43		61			43
42		60			42
41		59			41
40	-	58	-	-	40
39		57			39
38		56			38
37		55			37
36		54			36
35	-	53	-	-	35
34		52			34
33		51			33
32		50			32
31		49			31
30	-	48	81	64	30
29		47	79	63	29
28		46	77	61	28
27		45	75	60	27
26		44	73	58	26
25	-	43	71	57	25
24		42	68	55	24
23		41	66	54	23
22		40	64	53	22
21		39	62	51	21
20	-	38	62	50	20
19		37	60	48	19
18		36	58	47	18
17		35	56	45	17
16		34	53	44	16
15	66	33	51	42	15
14	63	32	49	41	14
13	60	31	47	40	13
12	57	30	45	38	12
11	54	29	43	37	11
10	51	28	41	35	10
9	48	27	38	34	9
8	45	26	36	32	8
7	41	25	34	31	7
6	38	24	32	29	6
5	35	23	30	28	5
4	32	22	28	27	4
3	29	21	26	25	3
2	26	20	23	24	2
1	23	19	19	22	1
0	20	18	17	21	0

TABLE O-2

ASVAB 6E/7E/6/7 SUBTEST CONVERSION TABLES
(in Navy Standard Score)

Raw score	Arithmetic reasoning (AR)	Space perception (SP)	Math knowledge (MK)	Electric information (EI)	Raw score
30	-	-	-	69	30
29				67	29
28				66	28
27				64	27
26				62	26
25	-	-	-	61	25
24				59	24
23				58	23
22				56	22
21				54	21
20	66	64	68	53	20
19	64	62	66	51	19
18	62	60	64	50	18
17	60	58	62	48	17
16	58	56	60	46	16
15	55	53	58	45	15
14	53	51	56	43	14
13	51	49	54	42	13
12	49	47	52	40	12
11	47	45	50	38	11
10	45	42	48	37	10
9	53	40	46	35	9
8	41	38	44	34	8
7	39	36	42	32	7
6	37	34	40	30	6
5	35	32	38	29	5
4	33	29	36	27	4
3	30	27	34	26	3
2	28	25	32	24	2
1	26	23	30	22	1
0	24	21	27	21	0

TABLE O-3

ASVAB 6E/7E/6/7 SUBTEST CONVERSION TABLES
(in Navy Standard Score)

Raw score	Mechanical comprehension (MC)	General science (GS)	Shop information (SI)	Automotive information (AI)	Raw score
20	72	70	65	67	20
19	70	68	63	65	19
18	68	66	61	63	18
17	65	63	59	61	17
16	63	61	56	59	16
15	60	59	54	57	15
14	58	57	52	55	14
13	56	54	49	53	13
12	53	52	47	51	12
11	51	50	45	49	11
10	49	47	42	47	10
9	46	45	40	44	9
8	44	43	38	42	8
7	41	40	35	40	7
6	39	38	33	38	6
5	37	36	31	36	5
4	34	34	28	34	4
3	32	31	26	32	3
2	29	29	24	30	2
1	27	27	21	28	1
0	25	24	19	26	0

TABLE O-4

ASVAB 6E/7E/6/7 SUBTEST CONVERSION TABLES

Raw score	Maintenance scale (CM)	Attentiveness scale (CA)	Electronics scale (CE)	Combat scale (CC)	Raw score
27	-	-	-	74	27
26				71	26
25	-	-	-	69	25
24				67	24
23				65	23
22				63	22
21				61	21
20	69	85	76	59	20
19	67	82	74	57	19
18	65	79	72	55	18
17	63	75	69	53	17
16	61	72	67	51	16
15	59	68	64	49	15
14	57	65	62	47	14
13	54	61	60	45	13
12	52	58	57	43	12
11	50	55	55	41	11
10	48	51	53	39	10
9	46	48	50	37	9
8	44	44	48	35	8
7	42	41	45	33	7
6	39	37	43	31	6
5	37	34	41	29	5
4	35	31	38	27	4
3	33	27	36	24	3
2	31	24	34	22	2
1	29	20	31	20	1
0	27	17	29	18	0

APPENDIX P

CORRELATIONS AND SAMPLE STATISTICS

APPENDIX P

CORRELATIONS AND SAMPLE STATISTICS

From sample 5 stratified on ASVAB 6/7 AFQT percentile scores in appendix M we calculated mean values, standard deviations, and correlation coefficients of ASVAB subtests and composites. Statistics for the subtests are shown in tables P-1 and P-2. Correlations for the composites are shown in table P-3. Refer to appendix A for definitions of the subtests and composites.

TABLE P-1

MEAN VALUES AND STANDARD DEVIATIONS OF ASVAB 6/7 SUBTESTS

Variable	Mean value	Standard deviation
GI	9.76	3.23
NO	31.70	9.87
AD	15.39	4.67
WK	20.07	6.97
AR	12.40	4.79
SP	13.55	4.64
MK	11.03	4.89
EI	18.31	6.26
MC	10.65	4.18
GS	11.16	4.35
SI	13.33	4.31
AI	11.72	4.85
CM	10.97	4.69
CA	9.68	2.91
CE	8.92	4.21
CC	15.51	4.89
AFQT 7A	49.3	27.42
ASVAB AFQT	50.4	28.84

TABLE P-2

CORRELATION COEFFICIENTS[a] OF ASVAB SUBTESTS

	GI	NO	AD	WK	AR	SP	MK	EI	MC	GS	SI	AI	CM	CA	CE	CC	ASVAB AFQT	AFQT 7A
GI	-	50	23	72	62	46	56	63	57	67	58	54	12	10	20	30	71	68
NO	50	-	50	48	60	39	58	49	45	51	40	35	09	17	26	22	55	55
AD	23	50	-	24	31	29	31	28	28	24	24	18	07	15	20	13	30	29
WK	72	48	24	-	69	51	62	67	62	77	58	54	12	15	22	33	89	75
AR	62	60	31	69	-	56	73	65	63	68	56	52	16	20	30	33	87	78
SP	46	39	29	51	56	-	56	61	57	53	50	44	12	05	24	24	76	67
MK	56	58	31	62	73	56	-	63	56	66	46	43	10	21	29	28	74	70
EI	63	49	28	67	65	61	63	-	68	71	67	67	24	16	31	35	74	71
MC	57	45	28	62	63	57	56	68	-	68	67	65	30	12	26	41	71	74
GS	67	51	24	77	68	53	66	71	68	-	63	61	16	16	26	33	79	76
SI	58	40	24	58	56	50	46	67	67	63	-	73	30	07	22	35	63	70
AI	54	35	18	54	52	44	43	67	65	61	73	-	37	12	24	31	59	66
CM	12	09	07	12	16	12	10	24	30	16	30	37	-	30	41	50	15	22
CA	10	17	15	15	20	05	21	16	12	16	07	12	30	-	48	30	17	15
CE	20	26	20	22	30	24	29	31	26	26	22	24	41	48	-	24	29	29
CC	30	22	13	33	33	24	28	35	41	33	35	31	50	30	24	-	35	35
ASVAB AFQT	71	55	30	89	87	76	74	74	71	79	63	59	15	17	29	35	-	86
AFQT 7A	68	55	29	75	78	67	70	71	74	76	70	66	22	15	29	35	86	-

[a]Decimal points omitted.

TABLE P-3

CORRELATION COEFFICIENTS[a] OF ASVAB COMPOSITES

	CO	MM	GM	CL	GT	USMC EL	USA EL	SC	ST	OF	GCT	USAF M	USAF A	USAF G	USAF E	USN G	USN M	USN E	USN C
CO	-	81	84	85	80	84	86	88	82	71	87	79	80	80	88	80	83	84	80
MM	81	-	89	69	73	87	92	80	79	83	77	92	63	73	86	73	85	87	63
GM	84	89	-	80	86	92	92	92	90	84	87	92	72	86	88	86	91	92	72
CL	85	69	80	-	92	84	78	88	85	73	90	64	88	92	80	92	82	84	88
GT	80	73	86	92	-	89	81	95	90	72	96	70	79	99	84	99	90	89	79
USMC EL	84	87	92	84	89	-	91	91	97	75	90	77	76	89	94	89	85	99	76
USA EL	86	92	92	78	81	91	-	89	85	79	84	88	70	81	91	81	89	91	70
SC	88	80	92	88	95	91	89	-	90	74	99	80	78	95	93	95	93	91	78
ST	82	79	90	85	90	97	85	90	-	71	90	70	77	90	88	90	82	97	77
OF	71	83	84	73	72	75	79	74	71	-	72	83	62	72	72	72	78	75	62
GCT	87	77	87	90	96	90	84	99	90	72	-	73	78	96	92	96	90	90	78
USAF M	79	92	92	64	70	77	88	80	70	83	73	-	59	70	77	70	90	77	59
USAF A	80	63	72	88	79	76	70	78	77	62	78	59	-	79	72	79	74	76	99
USAF G	80	73	86	92	99	89	81	95	90	72	96	70	79	-	84	99	90	89	79
USAF E	88	86	88	80	84	94	91	93	88	72	92	77	72	84	-	84	83	94	72
USN G	80	73	86	92	99	89	81	95	90	72	96	70	79	99	84	-	90	89	79
USN M	83	85	91	82	90	85	89	93	82	78	90	90	74	90	83	90	-	85	74
USN E	84	87	92	84	89	99	91	91	97	75	90	77	76	89	94	89	85	-	76
USN C	80	63	72	88	79	76	70	78	77	62	78	59	99	79	72	79	74	76	-

[a]Decimal points omitted.

MED

8