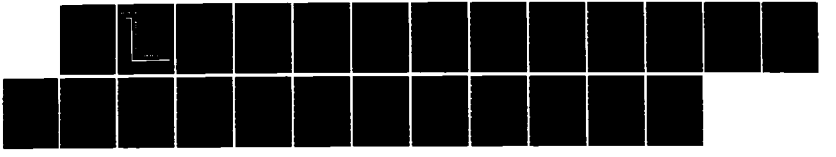
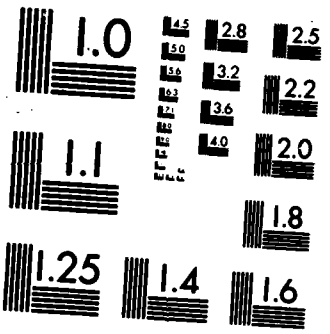
AD-A173 027 ALTERNATIVE ARMED FORCES QUALIFICATION TEST COMPOSITES 1/1
(U) AIR FORCE HUMAN RESOURCES LAB BROOKS AFB TX
T G WEGNER ET AL. SEP 86 AFHRL-TP-86-27

UNCLASSIFIED F/G 5/9 NL





AFHRL-TP-86-27

(12)

AIR FORCE HUMAN RESOURCES LABORATORY

ALTERNATIVE ARMED FORCES QUALIFICATION TEST COMPOSITES

Toni G. Wegner
Malcolm James Ree

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

September 1986
Interim Technical Paper for Period July 1983 - January 1986

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS A173 027
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFHRL-TP-86-27	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Manpower and Personnel Division 6b. OFFICE SYMBOL (If applicable) AFHRL/MOAE	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235-5601	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Human Resources Laboratory 8b. OFFICE SYMBOL (If applicable) HQ AFHRL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) Brooks Air Force Base, Texas 78235-5601	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62703F; PROJECT NO. 7719; TASK NO. 18; WORK UNIT ACCESSION NO. 46

11. TITLE (Include Security Classification)
Alternative Armed Forces Qualification Test Composites

12. PERSONAL AUTHOR(S)
Wegner, Toni G.; Ree, Malcolm James

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM Jul 83 TO Jan 86	September 1986	24

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
05	09	
05	08	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
AFQT category; Armed Forces Qualification Test (AFQT); aptitude tests; classification; Armed Services Vocational Aptitude Battery (ASVAB) (Continued)

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

Alternatives to the present Armed Forces Qualification Test (AFQT) composite, a subset of the Armed Services Vocational Aptitude Battery, were explored to examine the acceptability of a composite that did not include speeded tests. Fifteen alternative composites were analyzed for their predictive validity of training course scores, how their use would affect AFQT category classifications for different gender and ethnic subgroups relative to the present AFQT composite, and whether adequate pseudo-AFQT composites could be created to detect compromise. Results revealed some alternatives that compared favorably with the present AFQT composites on the different criteria.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	
22a. NAME OF RESPONSIBLE INDIVIDUAL Nancy A. Perrigo, Chief, STINFO Office	22b. TELEPHONE (Include Area Code) (512) 536-3877 22c. OFFICE SYMBOL AFHRL/TSR

Item 18 (Concluded):

enlisted selection, ~~and classification~~
mental abilities testing,
power tests,
predictive validity, Air Force training, Air Force personnel.
selection
speeded tests
test bias
test compromise
test validity

AFHRL Technical Paper 86-27

September 1986

ALTERNATIVE ARMED FORCES QUALIFICATION TEST COMPOSITES

Toni G. Wegner
Malcolm James Ree

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

Reviewed by

John R. Welsh
Chief, Enlisted Selection and Classification Function

Submitted for publication by

Lonnie D. Valentine, Jr.
Chief, Force Acquisition Branch

This publication is primarily a working paper. It is published solely to document work performed.

SUMMARY

When the norming population for the Armed Services Vocational Aptitude Battery (ASVAB) was changed from a 1944 metric to a 1980 metric, problems were discovered with the speeded tests of the ASVAB that led to questions about their use for enlistment qualification. One speeded test is currently part of the Armed Forces Qualification Test (AFQT), the qualification composite used by all services. The AFQT is important because it is used by all services for qualification and it is used to report one measure of quality of the enlisted forces to the Congress. The purpose of this effort was to explore the acceptability of alternatives to the present AFQT that do not include speeded tests. Fifteen composites were considered with respect to three issues: (a) how successful they are in predicting training course grades relative to the present AFQT, (b) how AFQT category classifications are affected for different gender and ethnic groups, and (c) whether adequate means exist for detecting compromise on these composites. Results revealed some alternative composites that compared favorably with the present AFQT composite on the different criteria.

PREFACE

This work was accomplished under Project 7719, "Development and Validation of Selection Methodologies" and executed as part of the Air Force Human Resources Laboratory's responsibility to provide technical information to aid policy decisions.

Special appreciation is extended to Ms. Doris Black, Mr. Jim Friemann, SSgt Harry Loveland, SrA Dave LeBrun, SrA Tom Sackett, AIC Steve Hoffer and Mr. Bill Glasscock of the Technical Services Division. Their efforts were noteworthy in completing the analyses necessary for this research in a timely manner.

TABLE OF CONTENTS

LIST OF TABLES

ALTERNATIVE ARMED FORCES QUALIFICATION TEST COMPOSITES

I. INTRODUCTION

The present effort was conducted to examine alternative military qualification composites from tests of the Armed Services Vocational Aptitude Battery (ASVAB). Since the ASVAB became the single Department of Defense enlistment test in 1976, a portion of the battery called the Armed Forces Qualification Test (AFQT) has been a common qualification composite for all services. The AFQT is also used to report on the quality of the enlisted forces to the United States Congress and has been linked to the 1944 World War II reference population. For these and other reasons, the AFQT is one of the most important portions of the ASVAB. The impetus for this investigation came from an effort to establish the AFQT on a score scale based on the 1980 population of American youth, ages 18 through 23. The AFQT is currently composed of the raw scores of the ASVAB tests Arithmetic Reasoning (AR), Word Knowledge (WK), Paragraph Comprehension (PC), and Numerical Operations (NO). AFQT scores are computed by summing unit-weighted AR, WK, and PC and half-weighted NO. It is, by design, a measure of general trainability and is reported in percentile scores.

The AFQT is composed of three power tests (AR, WK, and PC), and one speeded test (NO). A number of concerns about the speeded tests of the ASVAB arose in an effort (Maier & Sims, 1982; Sims & Maier, 1983) to establish a reference scale for the AFQT based on the 1980 population of American youth (ages 18 through 23). In the effort to generate the new score scale for the AFQT and other classification composites, differences in performance on the speeded tests between military applicants and the sample of 1980 youth were noted. In general, military applicants and recruits scored about 3 to 5 raw score points higher on the speeded tests than did the sample of 1980 youth. These differences had considerable impact on the scaling of AFQT scores. For example, if the differences were attributable to nonstandard testing conditions including inaccurate timing of the speeded tests, practice effects, and/or motivational differences between the 1980 youth sample and service applicants, then the score scale for the AFQT derived from the 1980 youth study using speeded tests would be inappropriate. Under these circumstances, a new AFQT composite, as well as service-unique occupational classification composites, would have to be considered for use without the speeded tests. On the other hand, the differences might have been attributable to a factor for which it would be possible to make corrections. A study by Earles, Giuliano, Ree, and Valentine (1983) indicated that the use of non-operational answer sheets in the original study to generate the 1980 youth population could account for the observed differences in test performance. This was confirmed by Wegner and Ree (1985). Based on this, the scores would be incorrect, but "corrections" in the scaling of the AFQT in the 1980 metric could be made. The present study was conducted in order to prepare for the possibility that an alternative AFQT might have to be constructed without speeded tests. A number of possible AFQTs are examined below according to several different criteria.

The development of alternative composites was limited to existing ASVAB tests, with the requirement that no speeded tests (Coding Speed, CS, and NO) be included. The eligible tests, then, were AR, WK, PC, General Science (GS), Auto and Shop Information (AS), Mathematics Knowledge (MK), Mechanical Comprehension (MC), and Electronics Information (EI). Based on previous experience with these tests, 15 composites were constructed which were expected to measure aptitudes similar to those measured by the current AFQT and to be measures, to a varying degree, of general trainability. All composites were computed using raw test scores. These composites are presented in Table 1.

Table 1. Alternative Qualification Composites

AFQT-a	= AR + MK + PC + MC
AFQT-b	= GS + MK
AFQT-c	= AR + MK + PC + MK
AFQT-d	= GS + AR + MK + PC
AFQT-e	= 2(AR + MK + PC) + MK
AFQT-f	= GS + (2AR) + MK + PC
AFQT-g	= MK + PC + MK + MC
AFQT-h	= GS + AR + AS
AFQT-i	= AR + MK + PC
AFQT-j	= GS + AR + MK + PC + AS + MK + MC + EI
AFQT-k	= GS + AR + MK + PC + MK
AFQT-l	= GS + AR + MK + PC + MK + MC
AFQT-m	= GS + AR + MK + PC + MK + MC + EI
AFQT-n	= GS + AR + MK + (2PC) + AS + MK + EI
AFQT-o	= GS + AR + MK + PC + AS + MK + EI

Three criteria were selected to compare the alternatives and to determine the appropriateness of each. The first criterion was the predictive validity of the composites. It was reasoned that because the AFQT is intended as a measure of general trainability, any composite designed to replace it should be predictive of military training performance. The second criterion was the impact of the new AFQT on qualification rates for various demographically defined subgroups. This involved comparing each of the qualification rates of the alternative composites with the present AFQT qualification rates to examine implications for applicants as a whole and for applicants of specific subgroups. The third criterion involved exploring the possibility of generating appropriate pseudo-AFQTs for each of the alternative AFQTs (see Sims & Truss, 1982, for related research) to see if it would be possible to detect certain types of test compromise.

II. PREDICTIVE VALIDITY

The purpose of these analyses was to determine how well the alternative composites measure general trainability. The validity of the composites was examined by correlating scores on composites with final scores obtained in Air Force technical training courses.

Method

Samples

Data were obtained from 154,788 recruits in 211 Air Force technical training courses from October 1980 through August 1984. Courses were included if enlistment scores were obtained from ASVAB Form 8, 9, or 10 and subtest scores were available. ASVAB Forms 8, 9, and 10 are parallel and can be used interchangeably. In addition, final course grades on a continuous scale with a 100-point maximum had to be available for a least 100 airmen. Only course graduates were included in the analyses, which tends to restrict the range of final course scores to between 60 and 100 points.

Procedure

Means and standard deviations for each of the alternative AFQT composites were computed separately for each technical training course. Validities were computed separately for each technical training course by correlating each alternate AFQT composite (computed using test scores from the parallel ASVAB Forms 8/9/10) with final course grade. Standard errors of estimate were computed based on these uncorrected validities. Validities were then corrected for restriction in range. Unrestricted estimates for this correction were obtained from data based on the 1980 American youth reference population for military applicants (McWilliams, 1980).

Results and Discussion

The range, median, and standard deviation of the means across different technical training courses for each alternate AFQT are presented in Table 2. Table 2 also contains the maximum raw score possible for each AFQT composite.

Table 2. Maximum Possible Score and Range, Median, and Standard Deviation of Course Mean Scores for the Alternative Composites

Composite	Maximum	Range	Median	SD
AFQT	105	76.17 - 99.31	85.78	5.63
AFQT-a	105	72.27 - 97.64	82.45	6.05
AFQT-b	50	29.37 - 45.31	34.74	4.05
AFQT-c	105	70.27 - 99.51	81.23	7.33
AFQT-d	105	74.06 - 98.69	83.51	6.25
AFQT-e	185	127.66 - 175.95	146.98	12.13
AFQT-f	135	93.24 - 127.52	106.54	8.68
AFQT-g	100	66.50 - 91.87	75.64	6.14
AFQT-h	80	49.29 - 70.88	59.51	4.95
AFQT-i	80	57.27 - 76.44	65.20	4.82
AFQT-j	200	127.26 - 179.13	149.46	12.08
AFQT-k	130	87.23 - 121.75	99.69	8.77
AFQT-l	155	101.64 - 142.95	117.06	10.04
AFQT-m	175	113.12 - 159.32	131.39	11.11
AFQT-n	190	124.97 - 171.96	143.73	11.20
AFQT-o	175	113.32 - 157.93	131.80	10.61

Table 3 contains the range and median of the uncorrected validities across technical training schools and the mean standard errors of estimate for each composite averaged across the schools. Table 4 contains the range and median of the validities corrected for restriction in range. These were computed separately for each technical training course.

The results reveal variability among the technical training courses, as well as among the different composites. All of the uncorrected validities for the alternative AFQTs are better than the validity of the present AFQT for these data. This may be because the present AFQT is the only composite containing a speeded subtest, and speeded subtests are traditionally low in reliability. Or it may be because Air Force recruits are directly selected on the General composite, which contains the same three power tests as appear in the present AFQT. This type of

Table 3. Validity Coefficients and Standard Errors of Estimate (SEE) of Alternative Composites with Technical Training Course Grades

Composite	Validity coefficient		Mean SEE
	Range	Median	
AFQT	.0634 - .6183	.3754	5.48
AFQT-a	.0261 - .6162	.4123	5.40
AFQT-b	.0440 - .6093	.3931	5.44
AFQT-c	.1097 - .6458	.4232	5.36
AFQT-d	.0699 - .6259	.4069	5.42
AFQT-e	.1115 - .6416	.4162	5.38
AFQT-f	.0755 - .6341	.4066	5.41
AFQT-g	.0500 - .6147	.4241	5.38
AFQT-h	-.0895 - .6165	.3959	5.44
AFQT-i	.0947 - .6217	.3903	5.45
AFQT-j	-.0438 - .6704	.4484	5.31
AFQT-k	.0968 - .6562	.4312	5.34
AFQT-l	.0538 - .6374	.4436	5.32
AFQT-m	.0163 - .6551	.4498	5.31
AFQT-n	-.0331 - .6442	.4467	5.31
AFQT-o	-.0342 - .6468	.4449	5.30

Table 4. Validity Coefficients of Alternative Composites with Air Force Technical Training Grades Corrected for Restriction in Range

Composite	Validity coefficient	
	Range	Median
AFQT	.2628 - .8807	.7015
AFQT-a	.0898 - .8663	.6974
AFQT-b	.1083 - .8009	.6129
AFQT-c	.1970 - .9033	.7160
AFQT-d	.2598 - .9021	.7013
AFQT-e	.2270 - .9213	.7235
AFQT-f	.2699 - .9009	.7040
AFQT-g	.1548 - .8739	.6946
AFQT-h	-.2012 - .8528	.6323
AFQT-i	.2596 - .9168	.6932
AFQT-j	-.1164 - .9235	.7145
AFQT-k	.2383 - .8985	.7284
AFQT-l	.1719 - .8880	.7228
AFQT-m	.0489 - .9212	.7304
AFQT-n	-.0987 - .9159	.7359
AFQT-o	-.0971 - .9177	.7272

curtailment often leads to a reduced validity coefficient. This would also explain why the second lowest composite is AFQT-i, which contains only those subtests in the General composite. The standard errors of estimate show little variability. When the correction for restriction in range is made for the curtailment at the lower end of the distribution, much higher validities are found between the composites and technical training course grades. Several alternatives still show stronger correlations than the present AFQT, but some show weaker correlations. The alternative composites with the weakest correlations are AFQT-b and AFQT-h. These composites have fewer items than most of the other composites. The alternative composites with the strongest correlations are AFQT-m and AFQT-n. These composites are among those with the greatest number of items. Except for AFQT-b and AFQT-h, the median validities fall in the range of .69 to .73. There is, therefore, little practical reason to distinguish among the rest of the alternatives on this criterion. The negative validities were mainly among foreign language courses with composites containing the Automotive and Shop Information test.

III. COMPARISON OF PRESENT AND ALTERNATIVE COMPOSITES

An important criterion for acceptability of the AFQT is how applicants would be distributed according to AFQT-derived ability categories, or simply AFQT categories. Military manpower managers have used the AFQT category boundaries established with the 1944 reference scale and have traditionally based many personnel and manpower decisions on the percentages of applicants expected in the various categories. Table 5 shows the AFQT categories by corresponding percentile score range and distribution for the 1944 score scale. In 1983, the Manpower and Accession Policy Steering Committee (which is composed of general officers who oversee the ASVAB) decided that the 1980 youth population reference scale will retain the same category percentile score boundaries as with the 1944 scale. This portion of the study will examine the percentages of applicants who would shift AFQT categories on the various alternate AFQTs.

Table 5. AFQT Raw Score Category Boundaries for Qualification Composites

Category	V	IV	IIIB	IIIA	II	I
	Percentiles					
	1 - 9	10 - 30	31 - 49	50 - 64	65 - 92	93 - 100
Composite	Raw scores					
AFQT	0.0 - 41.0	41.5 - 65.0	65.5 - 77.5	78.0 - 84.5	85.0 - 98.5	99.0 - 105.0
AFQT-a	0 - 37	38 - 58	59 - 71	72 - 80	81 - 95	96 - 105
AFQT-b	0 - 15	16 - 22	23 - 28	29 - 33	34 - 44	45 - 50
AFQT-c	0 - 35	36 - 56	57 - 70	71 - 79	80 - 98	99 - 105
AFQT-d	0 - 38	39 - 61	62 - 74	75 - 81	82 - 97	98 - 105
AFQT-e	0 - 64	65 - 104	105 - 128	129 - 144	145 - 174	175 - 185
AFQT-f	0 - 47	48 - 74	75 - 91	92 - 103	104 - 126	127 - 135
AFQT-g	0 - 35	36 - 55	56 - 66	67 - 74	75 - 90	91 - 100
AFQT-h	0 - 25	26 - 38	39 - 48	49 - 54	55 - 70	71 - 80
AFQT-i	0 - 28	29 - 47	48 - 57	58 - 64	65 - 75	76 - 80
AFQT-j	0 - 69	70 - 104	105 - 127	128 - 143	144 - 175	176 - 200
AFQT-k	0 - 44	45 - 70	71 - 86	87 - 97	98 - 120	121 - 130
AFQT-l	0 - 53	54 - 82	83 - 100	101 - 113	114 - 139	140 - 155
AFQT-m	0 - 60	61 - 92	93 - 112	113 - 127	128 - 156	157 - 175
AFQT-n	0 - 66	67 - 103	104 - 125	126 - 139	140 - 168	169 - 190
AFQT-o	0 - 60	61 - 93	94 - 113	114 - 126	127 - 155	156 - 175

To examine the impact of a new qualification composite, scores on the present AFQT and each of the alternative AFQTs were computed for a subset of American youth who had taken the ASVAB as part of this 1980 youth aptitude profile study. AFQT categories were computed for subjects on each composite, and category switches between the present AFQT and each alternative composite were noted. Category changes were examined across all subjects, for subjects of each sex, and for

subjects of different population groups (Whites, Blacks, and Hispanics). Additional analyses compared the percentage of people in each of these subgroups who would score below the 50th percentile, and the percentage who would score below the 21st percentile. The 21st percentile is currently the minimum qualification for enlistment in any of the services.

Method

Sample

Analyses were conducted using data from the 1980 Profile of American Youth sample gathered by the National Opinion Research Center (see Bock & Mislevy, 1981). This sample contains 9,173 youth, ages 18 through 23 (4,550 males and 4,623 females). The sample was statistically weighted to be representative of 1980 American youth (approximately 25 million).

Procedure

Boundaries for six major AFQT categories were designed for each composite, based on the percentile boundaries for the current AFQT. That is, for example, Category II consists of applicants who are in the 65th through 92d percentile. Using this sample and the present AFQT, those percentiles were associated with raw scores on ASVAB Form 8a of 85.0 through 98.5, respectively. Because of its importance to manpower planners, AFQT Category III was divided in half, with Category IIIA extending from the 50th percentile to the 64th percentile and Category IIIB extending downward from the 49th percentile to the 31st percentile. For the alternative AFQTs, the raw scores associated with the same percentiles (as determined using cumulative frequency distributions) were designated as the raw score boundaries for that category. The category boundaries for each composite are listed in Table 5.

The measure chosen to look at the impact of the proposed qualification composites compared to the present AFQT was the percentage of individuals who would be classified in each AFQT category on the proposed versus present composite. The percentage of examinees who change categories when compared to the present AFQT is useful for understanding how enlistment qualification and selection would be affected by adopting an alternate AFQT. The magnitude of AFQT category changes caused by a new AFQT for different subgroups is particularly important to military manpower planners.

For these analyses, cross-tabulations were computed showing subjects' AFQT category classification according to the current versus the alternate AFQT composites. These crossover analyses were done for females, males, Whites, Blacks, Hispanics, and for the total sample. Percentages were then computed to show how the AFQT category classification of different groups of subjects was influenced by the proposed composites.

Additional analyses were done to compare the percentage of individuals in each subgroup who would be disqualified on each composite because of failing to meet the minimal qualification requirement at the 21st percentile. No service is enlisting recruits with scores below the 21st percentile at this time. This allows a comparison of the most basic and immediate impact of a change to a new qualification composite. Finally, analyses were conducted comparing the percentage of individuals who would fall below the 50th percentile on each composite. This is the cutoff currently used by the Army to determine who will get an enlistment bonus, and the number of recruits in the upper half of the AFQT distribution is seen as a measure of quality of the force.

Results and Discussion

Tables 6 through Table 11 present the summarized results of the crossover classification analyses for the total sample, females, males, Whites, Blacks, and Hispanics, respectively. The tables show, for each proposed composite, the proportion of subjects whose AFQT category remained the same as it is for the present AFQT, the proportion who fell to one category lower (e.g., from Category II to Category IIIa), the proportion who advanced to one category higher, the proportion who declined two or more categories, and the proportion who advanced two or more categories.

Tables 6 through 11 reveal some interesting patterns concerning the alternative qualification composites. For the entire sample as well as each of the subgroups, the composites that consistently classified the most people into the same mental category as under the present AFQT were AFQT-e, AFQT-i, and AFQT-c. The composites that showed the most discrepancies were AFQT-b, AFQT-h, and AFQT-j. The direction of the discrepancies varied by subgroup. The majority of females who were affected moved to a lower mental category, and the majority of changes by males were to a higher category. Whites showed mixed changes. Hispanics who switched categories were also split, with a slight tendency to move down. Blacks too showed mixed changes, with AFQT-b producing the biggest move up and AFQT-j resulting in the biggest move down.

Table 12 shows the percentage of individuals in each subgroup who fall below the 21st percentile on the various composites and would therefore be disqualified for any of the services.

This analysis generally concurred with the results of the classification analyses. AFQT-i, AFQT-e, and AFQT-c disqualified subgroups at approximately the same proportions as the present AFQT. At this level, AFQT-j, AFQT-h AFQT-m, and AFQT-o disqualified more females, Blacks and Hispanics, and qualified more Whites and males. AFQT-b disqualified more females and Hispanics, but increased the qualification percentage of males, Whites, and Blacks.

Table 6. Percentage Category Reclassification on Proposed Versus Present Qualification Composites: All Subjects[a]

	Category shift				
Composite	Two or more lower	One lower	No change	One higher	Two or more higher
AFQT-a	.3	15.3	68.2	15.6	.5
AFQT-b	2.8	19.2	52.5	<u>22.7</u>	2.8
AFQT-c	.1	12.5	75.1	12.2	.2
AFQT-d	.2	14.6	70.8	13.9	.4
AFQT-e	.0	11.9	<u>77.8</u>	10.3	.1
AFQT-f	.0	13.8	73.7	12.2	.2
AFQT-g	1.1	17.4	62.9	17.2	1.4
AFQT-h	<u>3.0</u>	<u>21.4</u>	51.0	19.9	<u>4.7</u>
AFQT-i	.0	10.5	77.5	11.8	.1
AFQT-j	1.6	20.5	58.1	18.1	1.8
AFQT-k	.2	13.6	72.2	13.8	.3
AFQT-l	.4	15.9	67.4	15.6	.7
AFQT-m	.8	18.5	63.7	16.0	.9
AFQT-n	.8	18.2	63.2	16.7	1.1
AFQT-o	1.0	19.2	61.2	17.3	1.3

[a]Greatest value in each column is underlined.

Table 7. Percentage Category Reclassification on Proposed Versus Present Qualification Composites: Females[a]

Composite	Category shift				
	Two or more lower	One lower	No change	One higher	Two or more higher
AFQT-a	.5	22.9	69.8	6.7	.1
AFQT-b	3.9	24.6	52.2	17.3	2.0
AFQT-c	.1	15.0	74.3	10.4	.1
AFQT-d	.3	19.3	70.7	9.6	.1
AFQT-e	.0	14.3	77.0	8.6	.0
AFQT-f	.1	20.0	72.9	7.0	.0
AFQT-g	1.6	23.3	63.4	11.0	.7
AFQT-h	5.4	35.6	52.6	6.2	.2
AFQT-i	.0	12.4	77.4	10.1	.0
AFQT-j	2.9	33.2	58.9	5.0	.0
AFQT-k	.3	17.8	71.5	10.2	.1
AFQT-l	.9	23.4	68.0	7.7	.0
AFQT-m	1.5	28.6	63.4	6.5	.1
AFQT-n	1.4	28.8	63.8	5.9	.1
AFQT-o	1.7	31.0	61.5	5.7	.1

[a]Greatest value in each column is underlined.

Table 8. Percentage Category Reclassification on Proposed Versus Present Qualification Composites: Males[a]

Composite	Category shift				
	Two or more lower	One lower	No change	One higher	Two or more higher
AFQT-a	.1	8.0	66.7	24.2	1.0
AFQT-b	1.8	14.0	52.8	27.8	3.6
AFQT-c	.1	10.0	75.8	13.9	.2
AFQT-d	.0	10.1	71.0	18.1	.8
AFQT-e	.0	9.5	78.5	11.9	.1
AFQT-f	.0	7.8	74.5	17.3	.4
AFQT-g	.6	11.7	62.3	23.2	2.1
AFQT-h	.6	7.5	49.5	33.2	9.1
AFQT-i	.0	8.7	77.6	13.5	.2
AFQT-j	.3	4.1	57.4	30.8	3.5
AFQT-k	.1	9.5	72.8	17.2	.5
AFQT-l	.0	8.5	66.9	23.2	1.4
AFQT-m	.2	8.7	64.1	25.2	1.8
AFQT-n	.2	7.9	62.7	27.2	2.1
AFQT-o	.2	7.7	60.9	28.6	2.5

[a]Greatest value in each column is underlined.

Table 9. Percentage Category Reclassification on Proposed Versus Present Qualification Composites: Whites[a]

Composite	Category shift				
	Two or more lower	One lower	No change	One higher	Two or more higher
AFQT-a	.3	16.0	66.1	17.0	.6
AFQT-b	3.3	20.0	51.5	22.5	2.6
AFQT-c	.1	13.6	73.9	12.3	.2
AFQT-d	.2	15.3	69.0	15.0	.5
AFQT-e	.0	12.8	76.6	10.5	.1
AFQT-f	.0	14.4	72.4	12.9	.2
AFQT-g	1.2	18.2	60.8	18.3	1.5
AFQT-h	3.2	21.8	48.8	20.7	5.5
AFQT-i	.0	11.2	76.5	12.2	.1
AFQT-j	1.7	20.8	55.4	20.1	2.0
AFQT-k	.2	14.5	70.9	14.1	.3
AFQT-l	.5	16.5	65.4	16.8	.8
AFQT-m	1.0	19.3	61.2	17.5	1.0
AFQT-n	.9	18.8	60.8	18.3	1.2
AFQT-o	1.1	19.8	58.6	19.1	1.5

[a]Greatest value in each column is underlined.

Table 10. Percentage Category Reclassification on Proposed Versus Present Qualification Composites: Blacks[a]

Composite	Category shift				
	Two or more lower	One lower	No change	One higher	Two or more higher
AFQT-a	.4	12.7	77.5	9.2	.1
AFQT-b	.7	13.8	57.5	24.4	3.6
AFQT-c	.0	7.3	80.1	12.5	.1
AFQT-d	.1	11.3	78.8	9.7	.1
AFQT-e	.0	7.5	83.0	9.5	.0
AFQT-f	.1	11.1	79.4	9.4	.0
AFQT-g	.4	13.4	72.2	13.0	1.0
AFQT-h	2.2	19.3	61.0	16.7	.8
AFQT-i	.0	7.5	82.2	10.3	.0
AFQT-j	.8	18.6	70.3	9.8	.4
AFQT-k	.0	8.5	78.5	12.8	.2
AFQT-l	.1	12.6	76.3	10.8	.2
AFQT-m	.3	14.6	74.6	10.3	.2
AFQT-n	.2	15.1	74.3	10.1	.3
AFQT-o	.3	16.5	73.0	9.9	.3

[a]Greatest value in each column is underlined.

Table 11. Percentage Category Reclassification on Proposed Versus Present Qualification Composites: Hispanics[a]

Composite	Category shift				
	Two or more lower	One lower	No change	One higher	Two or more higher
AFQT-a	.1	13.1	75.6	10.6	.6
AFQT-b	1.6	20.5	54.2	20.5	3.2
AFQT-c	.0	9.8	79.6	10.5	.1
AFQT-d	.1	12.8	77.4	8.9	.8
AFQT-e	.0	9.6	81.7	8.6	.1
AFQT-f	.1	11.9	78.6	9.1	.3
AFQT-g	.6	15.7	69.9	12.5	1.4
AFQT-h	1.3	20.5	58.7	16.2	3.4
AFQT-i	.0	9.0	81.2	9.8	.0
AFQT-j	.8	19.8	66.8	10.7	1.9
AFQT-k	.1	12.6	75.2	11.8	.3
AFQT-l	.3	14.2	74.7	10.0	.9
AFQT-m	.4	16.9	72.2	9.1	1.4
AFQT-n	.5	17.6	69.9	10.7	1.4
AFQT-o	.6	17.6	69.6	10.7	1.5

[a]Greatest value in each column is underlined.

Table 12. Percentage of Youth Population Below the 21st Percentile on Each Qualification Composite[a]

Composite	Subgroup				
	Females	Males	Whites	Blacks	Hispanics
AFQT-Present	19.2	21.4	12.5	56.1	43.8
AFQT-a	20.7	18.0	11.1	57.0	43.9
AFQT-b	21.1	17.1	11.5	51.6	46.1
AFQT-c	19.4	19.7	11.8	54.4	44.0
AFQT-d	20.5	20.1	12.1	57.2	46.0
AFQT-e	20.3	20.3	12.3	56.5	45.0
AFQT-f	20.8	19.6	12.1	56.7	45.5
AFQT-g	20.4	18.8	11.7	55.3	44.4
AFQT-h	25.7	14.6	11.1	60.9	47.1
AFQT-i	20.4	20.6	12.3	57.7	45.2
AFQT-j	23.2	16.9	11.3	59.1	47.3
AFQT-k	20.6	19.9	12.2	55.8	46.3
AFQT-l	21.6	18.4	11.7	56.9	46.1
AFQT-m	22.5	18.3	11.8	59.0	47.6
AFQT-n	21.8	18.1	11.5	57.4	47.8
AFQT-o	23.3	17.7	11.7	59.5	48.7

[a]Smallest value in each column is underlined.

The percentage of individuals in each subgroup who fall below the 50th percentile on the alternative composites is shown in Table 13. These are the people who would not qualify for an enlistment bonus in the Army. Bonuses would be given at approximately the same as the present rate to each subgroup with the use of AFQT-c or AFQT-e. The biggest differences in the distribution of bonuses by sex (males, females) would occur for AFQT-h, AFQT-j, AFQT-n, and AFQT-o. These composites affect population groups in a similar way as the present AFQT. Population group (Whites, Blacks, Hispanics) differences in the distribution of bonuses are relatively unaffected by the alternatives, except AFQT-b which qualifies more Blacks and Hispanics and fewer Whites.

Table 13. Percentage of Youth Population Below the 50th Percentile on Each Qualification Composite[a]

	Subgroup				
Composite	Females	Males	Whites	Blacks	Hispanics
AFQT-Present	50.6	48.0	40.7	87.5	77.0
AFQT-a	54.5	41.9	39.2	88.0	76.1
AFQT-b	53.8	43.8	41.1	82.8	73.7
AFQT-c	51.9	46.6	40.9	85.9	76.2
AFQT-d	53.5	45.2	40.7	87.2	76.7
AFQT-e	51.9	46.4	40.7	86.6	76.5
AFQT-f	53.3	44.5	40.3	87.0	76.1
AFQT-g	53.2	42.5	39.3	85.4	74.9
AFQT-h	63.3	35.6	40.4	89.3	76.5
AFQT-i	50.6	45.4	39.4	85.8	75.4
AFQT-j	59.6	39.2	40.6	88.3	76.3
AFQT-k	52.0	45.2	40.2	85.6	75.5
AFQT-l	54.4	42.7	39.8	87.2	75.5
AFQT-m	56.4	41.7	40.2	87.5	76.4
AFQT-n	57.2	41.0	40.3	87.5	76.6
AFQT-o	57.9	40.3	40.3	87.8	76.6

[a]Smallest value in each column is underlined.

IV. PSEUDOS FOR THE ALTERNATIVE QUALIFICATION COMPOSITES

A pseudo-composite of the AFQT is currently used as an internal consistency check to indicate potential compromise on the ASVAB. The pseudo is made up of ASVAB tests not used in the current AFQT that together correlate highly with the current AFQT. Some prevalent forms of compromise can be detected by comparing AFQT scores with a highly correlated composite composed of non-AFQT tests. Tables are developed that define the relationship between scores on the pseudo-AFQT and expected scores on the AFQT. This information is used to flag scores of applicants whose actual AFQT scores are considerably higher than would be predicted by their pseudo-AFQT scores. Applicants with extreme scores are required to retest. In addition to flagging suspect individual scores, the pseudo is used to identify recruiters who may be involved with compromise. The ASVAB Forms 8, 9, and 10 pseudo-AFQT contains unit-weighted GS and MK, and one-fifth-weighted CS; its correlation with the AFQT was reported by Sims and Truss (1982) to be .824. According to Sims and Truss, the first pseudo-AFQT, used with ASVAB Forms 6 and 7, was developed because of suspected compromise on the ASVAB. Since, for some services, qualification for the Armed Services is based only on the AFQT composite, service testing experts believe compromise is concentrated on these tests.

Finding adequate pseudo-AFQTs for alternative qualification composites is an important consideration in selecting a new composite in order to retain the ability to detect compromise. The purpose of this phase of the study was to investigate possible pseudo-AFQTs for each of the alternative composites. In making comparisons between the alternative AFQTs and their pseudos, it should be noted that the importance of a pseudo decreases as the number of subtests in the AFQT increases. For an applicant to be helped to compromise on an AFQT composed of all of the power subtests, for example, he or she must demonstrate proficiency in all areas being tested.

Method

Sample

The 1980 youth sample used in Section III was used for these analyses.

Procedure

Stepwise regressions were conducted to explore pseudos for each of the composites. For each regression, one of the new composites was the criterion variable and all power tests not included in that composite were used as predictor variables in raw score form. It is not possible to form a pseudo for AFQT-j using this criterion because all power tests are included in the composite.

Pseudos were computed using the combination of tests for each composite with the largest R^2, provided each test explained at least one percent of the variance in the composite. One pseudo was formed for each composite by unit-weighting the chosen tests. A second pseudo was formed for each composite using the raw score least squares (called "b-weight") weights (rounded to thousandths) obtained in the regressions of the chosen tests on the composites. Correlations between the composites and their pseudos were then computed.

Results and Discussion

The pseudos and their relationship with the composites from which they were developed are presented in Table 14.

Table 14 reveals that the correlations between the alternative qualification composites and their pseudos range from .66 to .93. Pseudo-AFQT with AFQT correlations were generally high ($r >$ .81) when four or fewer power tests were contained in the alternative composite and there were, therefore, at least four power tests available to be used in the pseudo. Satisfactory pseudos are generally available, then, among AFQT alternatives a through i. The weakest correlations between the pseudos and alternatives for this group were with AFQT-c and AFQT-e. These are also the only two of this group for which the regression weights resulted in a noticeably increased correlation over the unit weights. For the remaining alternatives, pseudos were limited by the restriction in available subtests. Nevertheless, correlations with the alternatives still ranged from .77 to .78, except for the pseudo of AFQT-m which showed a correlation of .66 and used only the technical knowledge subtest AS. There were no appreciable differences for these subtests between the unit-weighted and regression-weighted pseudos.

Table 14. Pseudos for Alternative Qualification Composites

Composite	Pseudo	r[a]
AFQT[b]	GS + MK + (CS x .2)	.90
AFQT[c]	GS + MK[d]	.87
	(GS x 1.934) + (MK x 1.567)[e]	.87
AFQT-a	GS + MK + EI	.91
	(GS x 1.584) + (MK x 1.253) + (EI x 1.253)	.91
AFQT-b	AR + MK + EI	.89
	(AR x .741) + (MK x .423) + (EI x .407)	.90
AFQT-c	GS + MC	.82
	(GS x 2.964) + (MC x .916)	.84
AFQT-d	MK + EI	.87
	(MK x 1.812) + (EI x 2.144)	.87
AFQT-e	GS + MC	.82
	(GS x 5.243) + (MC x 1.561)	.84
AFQT-f	MK + EI + MC	.88
	(MK x 2.416) + (EI x 1.983) + (MC x .823)	.89
AFQT-g	GS + AR	.91
	(GS x 1.926) + (AR x 1.317)	.92
AFQT-h	MK + MK + MC + EI	.92
	(MK x .470) + (MK x .622) + (MC x .902) + (EI x 1.040)	.93
AFQT-i	GS + MK + EI	.89
	(GS x 1.366) + (MK x 1.108) + (EI x .638)	.89
AFQT-k	MC + EI	.77
	(MC x 1.730) + (EI x 3.001)	.78
AFQT-l	EI	.77
	(EI x 5.567)	.77
AFQT-m	AS	.66
	(AS x 4.037)	.66
AFQT-n	MC	.77
	(MC x 5.185)	.77
AFQT-o	MC	.78
	(MC x 4.858)	.78

[a] All correlations are statistically significant at $p < .001$.

[b] Pseudo for ASVAB Forms 8, 9, and 10.

[c] Proposed pseudo given constraint of not using speeded tests.

[d] The first pseudo for each composite is unit-weighted.

[e] The second pseudo for each composite is regression-weighted.

V. GENERAL DISCUSSION

This investigation was designed to explore alternative AFQT composites that do not contain speeded tests. The purpose was to see if an alternative composite could be identified that would satisfy various criteria at least as well as the present AFQT composite. Fifteen alternatives were chosen.

Three different criteria for the acceptability of an alternative AFQT composite were explored. First, predictive validity was explored using correlations of the alternative composites with Air Force technical training course grades. These correlations were corrected for restriction in range. The validities of most of the composites were in a narrow range (.69-.72) similar to that of the present AFQT (.70). The validities for only two composites, AFQT-b and AFQT-h, are significantly different from the others and are therefore the ones that are distinguishable on this criterion. These two stand out in that they are the composites with the fewest items.

The second criterion was the extent to which classification into mental categories would vary from the classifications produced by the present AFQT, especially with regard to how various subgroups (males, females, Whites, Blacks, and Hispanics) would be affected. Of the AFQT category crossovers that did occur, all of the composites negatively affected females and positively affected males. The composites that had the least impact on sex subgroups were AFQT-i, AFQT-c, and AFQT-e, which were also the composites that showed the least category changes (i.e., the most stability) relative to the present AFQT. These composites all contain the same three power tests as are contained in the present AFQT (AR, WK, PC) and would be expected to have a similar impact. The composites that showed the greatest fluctuations among sex subgroups, in that 10% more females moved down a mental category than moved up a mental category, were AFQT-a, AFQT-d, AFQT-f, AFQT-g, AFQT-h, AFQT-j, AFQT-l, AFQT-m AFQT-n, and AFQT-o. AFQT-k also resulted in considerably more women moving down than up. The composites affected the population groups (Whites, Blacks, and Hispanics) in mixed ways. The composites with the greatest negative impact on minorities were AFQT-h, AFQT-j, AFQT-m, AFQT-n, and AFQT-o.

Exploration of the third criterion (i.e., whether an adequate pseudo could be found for the composites) revealed a direct relationship between the adequacy of the pseudo and the number of power tests used in the composite (hence restricting those available for the pseudo). The pseudos for AFQTs a through i were about the same in appropriateness, with AFQT-c and AFQT-e having slightly lower correlations with their pseudos. There was a drop in correlations for the remaining composites, especially AFQT-m, which correlated .66. AFQT-j had no pseudo since it contained all the power tests.

Recommendations

Several factors must be taken into consideration to choose the most appropriate replacement for the present AFQT. Although each of the three criteria will be considered, the importance of each of these criteria to what the AFQT should accomplish will determine which alternative is ultimately selected. For the validity criterion, for example, should differences of .02 be used to separate one composite from another? Or is any composite within the .69-.72 range acceptable? For the classification criterion, is the goal to find the composite which classifies individuals in a manner most like the present composite, or should a composite be chosen which qualifies more minorities? Also, the criterion of having an acceptable pseudo may not be important if there are other ways to detect test compromise or if so many power tests are included that compromise would be minimized anyway. Finally, the question of ease of computation during hand-scoring should be a secondary consideration.

Recommendations of the composites can be made based on the following assumptions: (a) The validity of a new AFQT should be at least similar to that of the present AFQT, but small differences are not critical; (b) mental categories should remain as stable as possible, with females and minorities not being negatively impacted; and (c) the pseudo for composites with fewer than five power tests should correlate with its composite at about the same level as the current pseudo. Based on these assumptions, AFQT-b and AFQT-h are eliminated as contenders on the basis of their validities; AFQT-a, AFQT-d, AFQT-f, AFQT-g, AFQT-h, AFQT-j, AFQT-k, AFQT-l, AFQT-m, AFQT-n, and AFQT-o are eliminated on the basis of their negative impact on females, with some of those additionally having negative impact on minorities.

The remaining alternative composities are AFQT-c, AFQT-e, and AFQT-i. All three of these composites have adequate validities, adequate pseudos, and minimal change from the present AFQT in classifying applicants among the AFQT categories. AFQT-i qualifies more individuals for bonuses in the Army than does the current AFQT, has the least negative impact on females, and has a slightly positive impact on minorities. AFQT-c and AFQT-e have a greater negative impact on females, but a greater positive impact on minorities. Alternative AFQT-e requires double-weighting of the sum of AR, WK, and PC subtests, which may cause errors during hand-scoring. These three composites are all acceptable alternatives based on the criteria explored here, and a choice among them needs to be based on more specific determinations of the purpose of an alternative AFQT.

REFERENCES

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

Earles, J.A., Giuliano, T., Ree, M.J., & Valentine, L.D., Jr. (1983). The 1980 youth population: An investigation of speeded subtests. Unpublished manuscript. Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

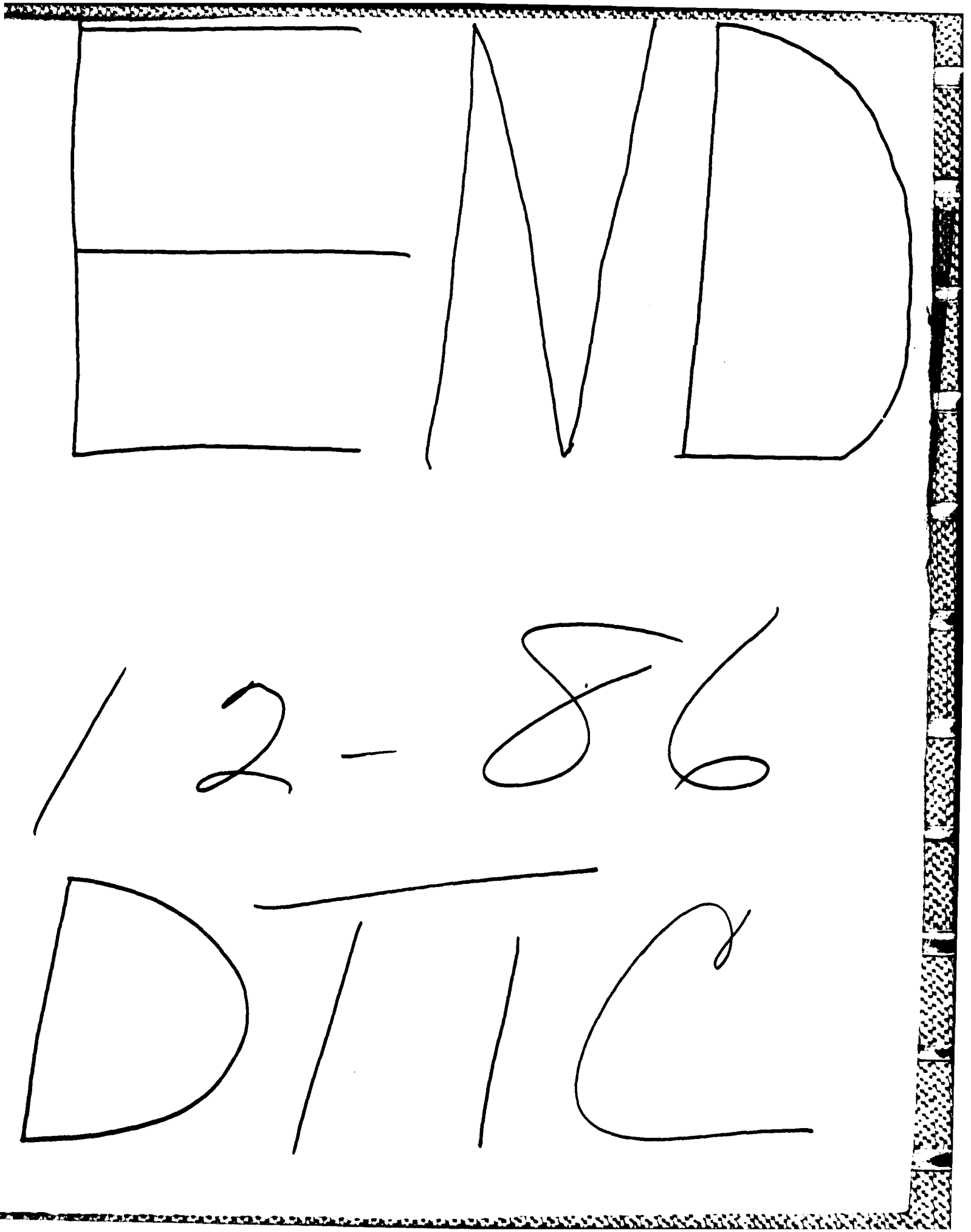
Maier, M.H., & Sims, W.H. (1982). Constructing an ASVAB score scale in the 1980 reference population (Memorandum 82-3288). Alexandria, VA: Center for Naval Analyses.

McWilliams, H.A. (1980). Profile of American youth: Field report. Chicago, IL: National Opinion Research Center.

Sims, W.H., & Maier, M.H. (1983). The appropriateness for military applications of the ASVAB subtests and score scale in the new 1980 reference population (Memorandum 83-3102). Alexandria, VA: Center for Naval Analyses.

Sims, W.H., & Truss, A.R. (1982). The development and application of a pseudo AFQT for ASVAB 8/9/10 (CRC 470). Alexandria, VA: Center for Naval Analyses.

Wegner, T.G., & Ree, M.J. (1985). Armed Services Vocational Aptitude Battery: Correcting the speeded subtests for the 1980 youth population (AFHRL-TR-85-14, AD-A158 823). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.


END
12-86
DTIC