UNCLASSIFIED

AD NUMBER
ADB951321
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited. Document partially illegible.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; 24 MAY 1955. Other requests shall be referred to Army Personnel Research Office, Washington, DC. Document partially illegible.**
AUTHORITY
APRO per DTIC form 55

THIS PAGE IS UNCLASSIFIED

THIS REPORT HAS BEEN DELIMITED
AND CLEARED FOR PUBLIC RELEASE
UNDER DOD DIRECTIVE 5200.20 AND
NO RESTRICTIONS ARE IMPOSED UPON
ITS USE AND DISCLOSURE.

DISTRIBUTION STATEMENT A

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

Morton

UNANNOUNCED

(1)

(9) RESEARCH MEMORANDUM

NUMBER 55-2 May 1955

(14) APRO-RM-55-2

(6) INTERRELATIONSHIPS OF ARMED FORCES QUALIFICATION TEST (AFQT-3 AND -4) SUBTESTS.

A-1-162-33.

Released for Distribution
Date 24 May 1955

(12) 8

(11)

AD B951321

DDC
RECEIVED
JAN 24 1980
A

DDC FILE COPY

THIS DOCUMENT IS BEST QUALITY PRACTICABLE.
THE COPY FURNISHED TO DDC CONTAINED A
SIGNIFICANT NUMBER OF PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

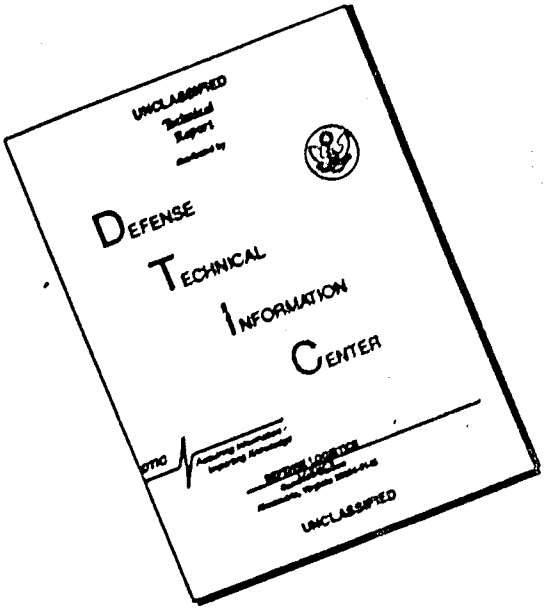
APRO

U. S. ARMY
PERSONNEL
RESEARCH
OFFICE

79 12 19 066

040 650

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

INTERRELATIONSHIPS OF ARMED FORCES QUALIFICATION TEST (AFQT-3 AND -4) SUBTESTS

I. STATEMENT OF THE PROBLEM

A. BACKGROUND

The Armed Forces Qualification Test, AFQT-3 and AFQT-4 was constructed for use at the Armed Forces Examining Stations as an instrument for the selection of enlisted personnel and the allocation of this personnel among the Services. The construction of the test was based upon the assumption that a combination of verbal analogy, arithmetic reasoning, spatial relations, and tool usage items would result in an instrument having maximal validity for performance in service.

During the item selection phase of construction, biserial correlations were obtained between each item and the reference test scores of like and unlike types, but no attempt was made to separately score the AFQT subtests nor to make a study of their interrelationships.

B. OBJECTIVES

The objective of the present project has been to determine the relationships among the separately scored subtests, and to study the differences among them for the various mental levels. In addition to seeking general pertinent information, this project undertook the testing of two specific hypotheses:

1. That significant differentiation occurs, in terms of a hierarchy of means through the mental grades, within each subtest.

2. That the subtests contribute equally to the total score through the five mental grades.

II. METHOD OF INVESTIGATION

A. POPULATION AND SAMPLING

The population was composed of all CONUS men taking the AFQT during the months March through December, 1953. The AFQT answer sheets of part of this population were available in the Branch. This subpopulation consists of 2% of the persons classified in Mental Grades I through IV (AFQT passers) and 10% of the persons classified in Grade V (AFQT failures). From this subpopulation of AFQT answer sheets, two samples were drawn:

1. One sample consisted of 1000 AFQT-3 and 1000 AFQT-4 answer sheets with 200 in each Mental Grade for each form. (Administrative acceptee answer sheets were not included--their scores would place them in Grade V, although they are supposed to be in Mental Grade IV.) This is not a rectangular distribution because of the unequal assignment of percentiles to Mental Grades (See Table 1). This group will be referred to as Sample 1.

2. Sample 2 was comprised of 1000 cases drawn from Sample 1. The purpose was to compose a sample which had in each Mental Grade a number of cases proportional to the number expected on the basis of the percentile scores which define the Mental Grade (See Table 1). From Sample 1, cases were drawn at random from each Mental Grade until the totals in the "N" column remained.

Table 1

DEFINITION OF SAMPLE 2 IN TERMS OF AFQT PERCENTILES IN EACH MENTAL GRADE

Mental Grade	Percentile Interval	Percent Included	N
I	93 - 100	8	80
II	65 - 92	28	280
III	31 - 64	34	340
IV	10 - 30	21	210
V	1 - 9	9	90
			1000

B. VARIABLES

1. The variables with which this project is concerned are the raw scores on AFQT-3 and -4 subtests (verbal, arithmetic, tool usage, and spatial) and the total AFQT raw score (the sum of the four subtest scores).

2. Each individual's answer sheet yielded five scores: one for each subtest and his total score. The scoring formula was R - 1/3W. Negative as well as positive scores were used.

III. ANALYSIS AND RESULTS

A. INITIAL PROCEDURE

Sample 1 was scored in preparation for analysis of variance (multiple classification design) and then was sorted into the following cross-classifications:

1. AFQT-3 cases and AFQT-4 cases. (1000 in each group)

2. Each of these two into the five Mental Grades. (200 in each group)

3. Each of these 10 groups, at random, into 4 equal parts, each part used as the source of scores for one subtest. (50 per group)

This established 40 groups of 50 cases each, in which no individual appeared in more than one group or cell. (The 40 cells are conceptually arranged as a 2 x 4 x 5 cube.)

B. RESULTS

1. Bartlett's test for homogeneity of variance was applied to the 40 cells. This resulted in a χ^2 of 566.9, representing a P of less than .01, indicating heterogeneity of variance. It is not a surprising outcome since this could appear solely as a result of the definitions (ranges included) of the mental grades. (See Table 1). The effort to transform the data appeared to be unwarranted; the analysis of variance was completed, however, for sake of interest. The .001 point was set as required for significance in the subsequent analysis. Table 2 shows the means of the components of cross-classifications. Table 3 shows the patterns of means involved in the only significant interaction, subtests with mental grades.

Table 2

MEANS OF SUBTEST, FORMS, MENTAL GRADES, AND TOTAL
(N = 2000)

Classification	Means
Subtests (500 in each)	
Verbal	15.2
Arithmetic	14.8
Tool Usage	15.3
Spatial	13.1
Forms (1000 in each)	
AFQT-3	58.1
AFQT-4	58.8
Mental Grades (400 in each)	
I	92.8
II	81.6
III	62.6
IV	41.7
V	13.5
Total	58.4

Accession For
NTIS GRA&I
DDC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or special
2 | 23

Table 3

MEANS AND STANDARD DEVIATIONS IN CELLS OF SUBTEST-MENTAL GRADES CROSS-CLASSIFICATION*
(In Each Cell N = 100)

	I	II	III	IV	V
Means					
Verbal	23.8	22.1	18.4	10.6	1.3
Arithmetic	23.8	21.6	16.3	9.8	2.4
Tool Usage	22.1	18.8	15.1	13.7	7.0
Spatial	23.1	19.1	12.9	7.6	2.8
Standard Deviations					
Verbal	1.3	2.4	4.2	5.8	2.9
Arithmetic	1.4	2.4	4.2	4.5	3.5
Tool Usage	2.4	4.3	4.9	5.2	5.6
Spatial	1.9	3.2	4.2	3.7	3.1

*This is the only interaction which was statistically significant (tested against Residual).

2. The results of the first phase of the analysis are shown in Table 4 and Table 5.

Table 4

ANALYSIS OF VARIANCE OF AFQT SUBTESTS, MENTAL GRADES, AND FORMS -3 AND -4

Source of Variation	Sum of Squares	df	Mean Square	F	Signif.
Subtests	1602.75	3	534.24	37.18	*
Grades	100811.73	4	25202.93	1753.86	*
Forms	14.96	1	14.96	1.04	
Interactions					
Subtests x Grades	4723.88	12	393.66	27.39	*
Subtests x Forms	7.80	3	2.60	-	
Grades x Forms	47.36	4	11.84	-	
Subtests x Grades x Forms	198.19	12	16.52	1.15	
Residual	28169.02	1960	14.37		
Total	135575.69	1999			

*Significant beyond the .001 point.

Table 5

ANALYSIS OF VARIANCE OF MENTAL GRADES SEPARATELY FOR EACH SUBTEST

Subtest	Source of Variation	Sum of Squares	df	Mean Squares	F*
Verbal	Between	34464	4	8616.00	644.42
	Within	6621	495	13.37	
	Total	41085	499		
Arithmetic	Between	30872	4	7718.00	676.42
	Within	5648	495	11.41	
	Total	36520	499		
Tool Usage	Between	13050	4	3262.50	152.16
	Within	10615	495	21.44	
	Total	23665	499		
Spatial	Between	27151	4	6787.75	608.76
	Within	5523	495	11.15	
	Total	32674	499		

*All F's significant beyond the .001 point.

3. A new series of F's based on the pooling of nonsignificant interactions, was calculated; but, because of the large number of degrees of freedom available in the error term, the differences were almost imperceptible.

4. The inferences drawn from these findings are as follows:

a. The means of subtests are statistically significantly different. However, these differences are not great, as can be seen in Table 3.

b. Means of mental grades are very different. This is to be expected from definitions of their ranges.

c. The difference of 0.7 between forms is not significant.

d. There is a very significant interaction between subtests and mental grades. This is observable in Tables 2, 3, and 4. For example, as grade level goes down, the proportion of total score due to the Tool Usage subtest increases from 24% to 52% (comparing means). The implication is that not only various ability levels, but also various hierarchies of the abilities place men in the several mental levels.

e. The other three interactions are not significant. All three of them involve the forms classification.

5. The results of the other principal method of analysis, correlational, based on Sample 2, are presented in Table 6.

Table 6

MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS
OF SUBTESTS BASED ON SAMPLE 2
(N = 1000)

Subtest	Mean	SD	r's 1	2	3	4	(1+2)
Verbal	16.7	7.6					
Arithmetic	16.1	7.1	.80				
Tool Usage	15.5	6.0	.38	.39			
Spatial	13.4	6.9	.67	.71	.47		
Total	61.7	23.0	.88	.89	.65	.86	(.93)

It can be seen in Table 6 that the range of intercorrelations extends from .38 to .80 (excluding r's with total). The lower r's involve tool usage, as suggested by the earlier analysis of variance results. The median is similar to the .55 cited for the Army Classification Battery. The implication of these findings is that the AFQT, like the ACB, is measuring several abilities; although unlike the ACB, the AFQT yields a single score. Since the AFQT is not intended to classify, it is a shorter test, and so can be expected to have less reliability in the separate subtests. However, there is sufficient differential performance in the AFQT to offer some assurance that men are not being screened on one tight cluster of abilities.

IV. SUMMARY

The interrelationships involved in parts of AFQT-3 and -4 performance were investigated, using a sample of 2000 CONUS men.

1. Analysis of variance showed (in addition to significant mental grade differences) that:

a. The subtests were not equally difficult.

b. There was an interaction between the mental grade of individuals, as determined by total score, and performance on the separate subtests.

c. On all subtests there is a hierarchy of means through the mental grades, but this is less evident for tool usage.

2. Correlational analysis showed AFQT to be operating as a screening battery. The lowest intercorrelations were those with tool usage.

PERSONNEL

Program Coordinator: Dr. J. P. Mundy

Project Director: Dr. K. F. Schenkel

Statistical Adviser: Mrs. L. K. Burke