AD-781 033

RELATIONSHIPS AMONG AN INDIVIDUAL INTELLIGENCE TEST AND TWO AIR FORCE SCREENING AND SELECTION TESTS

David F. McGrevy, et al

Air Force Human Resources Laboratory
Brooks Air Force Base, Texas

March 1974

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Personnel Research Division, Air Force Human Resources Laboratory, Lackland Air Force Base, Texas 78236, under project 7719, with Hq Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235.

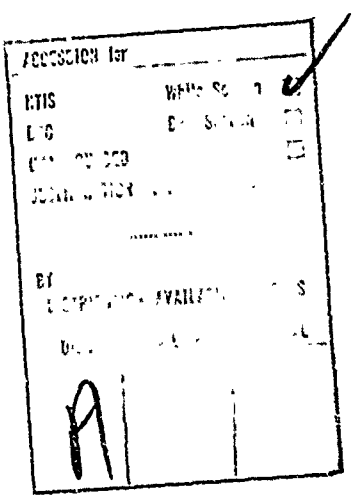
This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved.

LELAND D. BROKAW, Chief
Personnel Research Division

Approved for publication.

HAROLD E. FISCHER, Colonel, USAF
Commander



Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFHRL-TR-74-25	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD 781033
4. TITLE *(and Subtitle)* RELATIONSHIPS AMONG AN INDIVIDUAL INTELLIGENCE TEST AND TWO AIR FORCE SCREENING AND SELECTION TESTS		5. TYPE OF REPORT & PERIOD COVERED Final February 1972–July 1973 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) David F. McGrevy Stephen B. Knouse Ronnie A. Thompson		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Personnel Research Division Air Force Human Resources Laboratory Lackland Air Force Base, Texas 78236		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62703F 77191106
11. CONTROLLING OFFICE NAME AND ADDRESS Hq Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235		12. REPORT DATE March 1974 13. NUMBER OF PAGES 18
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

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

Armed Forces Qualification Test
Airman Qualifying Examination
Wechsler Adult Intelligence Scale
race
minority groups

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

With the implementation of the All Volunteer Force concept, the Air Force must insure that the objectively measurable range of ability in its manpower pool is being utilized. This is especially true for minority groups who have been categorized and channeled into military career areas based upon their performance on two selection tests: the Armed Forces Qualification Test (AFQT) and the Airman Qualifying Examination (AQE). To investigate the relationship of the AFQT and AQE to the general mental ability of different racial groups of airmen, a sample of 100 black and 100 white Air Force basic trainees was administered an established test of general mental ability, the Wechsler Adult Intelligence Scale (WAIS). The Verbal, Performance, and Full Scale IQ scores of the white and black airmen were compared to their AFQT scores and their four AQE Aptitude Indexes. Significant differences between

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

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

Item 20 (Continued)

black and white airmen were found on the AFQT, the four AQF Aptitude Indexes, and the three WAIS IQs. Regression analyses demonstrated most differences to be interaction effects between race and Air Force test scores. Implications for further research into Air Force airmen selection tests were discussed.

PREFACE

This research was accomplished under Project 7719, Air Force Personnel System Development on Selection, Assignment, Evaluation, Quality Control, Retention, Promotion, and Utilization; Task 771911, Selection and Classification Instruments for Airman Personnel Programs.

TABLE OF CONTENTS

LIST OF TABLES

RELATIONSHIPS AMONG AN INDIVIDUAL INTELLIGENCE TEST AND TWO AIR FORCE SCREENING AND SELECTION TESTS

I. INTRODUCTION

The aptitude assessment program of a military organization largely defines the effectiveness of the organization in satisfying its manpower needs. The initial screening and subsequent classification of applicants must occur in such a manner as to define accurately the abilities of the manpower resource, while maintaining fair and unbiased selection practices. In terms of the individual, selection and classification testing determines the areas of his career opportunity which later affect his motivation, job satisfaction, and career retention. In terms of the military organization, the objective of the selection program is the economical utilization of its human resources.

To optimize the usefulness of any selection device, an effort must be made to determine its validity in the particular selection process. In addition to aptitude test validation, much current emphasis has also focused upon developing or maintaining unbiased or "culture free" assessment. In the past, enlistees, both draft and non-draft motivated, produced a manpower pool with a sufficiently wide range in aptitude to provide adequately for personnel needs. Now, in the absence of draft pressure, the Air Force may be restricted in its manpower input and must insure that its selection tests tap the range of ability accurately, particularly in minority groups where possible test bias can confound test results.

To meet the needs of the enlisted manpower pool, the two main aptitude measures utilized by the Air Force in selection and assignment are the Airman Qualifying Examination (AQE) and the Armed Forces Qualification Test (AFQT). The AQE was adopted for use by the Air Force in 1958 (On 1 July 1973 the Armed Services Vocational Aptitude Battery (ASVAB) replaced the AQE in the Air Force testing inventory. Like the AQE, the ASVAB will provide the four Aptitude Indexes.) The development and standardization of the last AQE, Form J, has been described by Vitola, Massey, and Wilbourn (1971). Administered and scored at the recruiting level, the AQE yields four Aptitude Indexes (AIs)–Administrative (AQE-A), Electronic (AQE-E), General (AQE-G), and Mechanical (AQE-M). These AIs have been used in the selective recruiting and initial assignment of basic airmen. Although the AQE was originally developed as a military classification tool, in 1962 the High School Testing Program was established and the AQE was administered to large numbers of students as part of the schools' counseling program.

A second selection instrument administered to all pre-inductees and enlistees is the AFQT. (In late 1973 the ASVAB replaced the AFQT. An AFQT composite was drawn from ASVAB subtests.) This Department of Defense test simultaneously evaluates verbal, numerical, spatial, and mechanical abilities. Its single composite score is primarily used to determine mental categories and service eligibility of candidates for military service. Administration and scoring usually occur at the Armed Forces Entrance and Examining Stations (AFEES).

A majority of previous research has focused on the development and periodic validation of the AQE on samples of technical school graduates, although several studies were initiated to compare the AQE with commercial aptitude batteries, or to investigate the extent of aptitude test differences among various subgroups of the population. Considering their impact upon the career aspirations of thousands of potential Air Force Recruits and upon high school counseling, such investigations were justified. Madden, Valentine, and Tupes (1966) compared the AQE with the Differential Aptitude Tests and found that both batteries measure essentially the same factors or abilities. Madden and Tupes (1966) described the relationships among the AQE, the California Achievement Test, and the Davis Reading Test in the estimation of reading achievement. Madden and Valentine (1967) found a moderate positive relationship between ability measures obtained from the AQE and the Employee Aptitude Survey. From these studies, the AQE compares favorably with other aptitude batteries.

Comparative studies of the AFQT by the Air Force have mainly involved other Air Force ability tests, such as the Airman Classification Test (Thompson, 1958), the Airman Proficiency Tests (Brokaw, 1959), and the AQE (Valentine, 1968). It should be noted that Anastasi (1968) lists the AFQT under "Group Intelligence Tests and Developmental Scales" (p. 640).

The possibility of bias against minority groups in Air Force selection instruments has recently been investigated by Guinn, Tupes, and Alley (1970a). They have reported a study exploring the relationships of demographic variables to several Air Force aptitude tests. They found that a majority of test scores was significantly related to race, education level, and geographic area of the trainee. Racial differences were most prominent in tests requiring specific prior knowledge, such as verbal and mechanical tests, while situationally defined tests measuring non-verbal, spatial, and psychomotor abilities revealed less pronounced racial effects. A second study by Guinn, Tupes, and Alley (1970b) investigated racial bias in the AQE as a predictor of final school grade in 10 Air Force technical school courses. Bias against minority groups, defined as underprediction of final school grade, was not found in any of the 10 courses. There was, however, a tendency for the AQE to overpredict black performance in six of the courses (three of these race differences were found to be statistically significant). Possible racial bias in the AQE was also investigated by Shore and Marion (1972) who employed the AQE to predict performance on the Specialty Knowledge Test (SKT), a test of job knowledge used in determining an airman's promotability to a higher grade. Bias, defined as underprediction of the SKT by the AQE, was not found in the 16 career areas investigated. No Air Force research has directly investigated the question of bias against minority group members in the AFQT.

The present study extends the exploration of the relationships of AQE and AFQT scores to other ability measures. A test of general mental ability, the Wechsler Adult Intelligence Scale (WAIS), was employed to accomplish this objective.

II. METHOD

Subjects

The subjects were 200 non-prior service male Air Force basic trainees. One hundred black and one hundred white airmen, at Lackland AFB, Texas, were randomly selected over a 12 month period (February 1972–February 1973) during routine testing on their 6th day of basic training. The age of the subjects ranged from 17 to 25 years with a mean age of 19 years (black 19.13, white 18.95 years). Education ranged from 8 to 16 years of schooling completed. Mean education level was 12 years completed (black 11.96, white 12.15).

Procedure

The complete Wechsler Adult Intelligence Scale was individually administered to the trainees by the authors. Because the authors were members of the military and could possibly elicit anxiety in the trainees, military aspects of the situation were deemphasized. Each subject was asked to divorce himself as much as possible from the military situation (i.e., relax, smoke, drink a beverage). Each subject was further told to do as well as he could, that he was in an experimental situation, and that his test scores would not be entered into his records or used for job placement in the Air Force. Subjects were not told that they were taking an intelligence test. When questions about the nature of the test arose, subjects were told that they were taking a test of general mental ability.

Scores on the AQE and AFQT were obtained from basic training records after the WAIS was scored. The WAIS provided three scores, Verbal IQ (VIQ), Performance IQ (PIQ), and Full Scale IQ (FSIQ), against which the four AQE AIs and AFQT scores were contrasted.

III. RESULTS

To insure that the sample was representative of the airman basic population, a series of z tests was performed between the sample and the total 1972 Air Force accession data. Black and white AQE and AFQT data were analyzed separately. There were no statistical differences between any AQE or AFQT means. The z tests are summarized in Tables 1 and 2.

Means and standard deviations of the four AQE AIs, the AFQT score, the three WAIS IQs, and WAIS subtests for the black and white samples appear in Table 3. The t tests between the black and white airman means revealed significant differences on all measures except for the Digit Span and Digit Symbol subtests of the WAIS. In each case, black airmen scored lower than white airmen. Also presented in Table 3 are the non-significant tests for age and education differences.

A correlation matrix of AFQT, AQE AIs, and WAIS IQs appears in Table 4 for the black sample and in Table 5 for the white sample. Correlations were not corrected for range restriction effects. All correlations for the black sample were positive and significant ($p < .05$), except for PIQ vs. AQE-A, VIQ vs. AQE-M, and FSIQ vs. AQE-M, where correlations were positive but nonsignificant. All correlations for the white sample were positive and

Table 1. Comparison of AFQT and AQE Aptitude Indexes for 100 Black Enlistees and the Total 1972 Black Enlistees Accessions

	Black Enlistee Sample			1972 Black Accessions			
Measure	N	M	SD	N	M	SD	z[a]
AFQT	100	41.68	13.83	10,456	43.32	15.57	−1.05
AQE-A	100	44.55	16.82	10,472	45.90	18.76	− .72
AQE-E	100	45.25	18.57	10,472	47.10	18.00	−1.03
AQE-G	100	48.26	16.73	10,471	50.65	15.56	−1.54
AQE-M	100	44.20	18.56	10,467	44.06	18.29	.08

[a]For a discussion of the use of the *z* score to indicate the likelihood of a sample mean differing from the population mean, see Hays, *Statistics,* 1963, p. 203-204. In this situation,

$$z = \frac{M - m}{\frac{\sigma}{\sqrt{N}}}$$

where M = sample mean, m = population mean, σ = population standard deviation, and N = number of cases in the sample. Applying Tchebycheff's inequality to the above *z*'s reveals the probability of these sample means deviating as much as they do from the population means approaches 1.00 for all tests for both racial groups, given the size of the groups, and the population standard deviations.

Table 2. Comparison of AFQT and AQE Aptitude Indexes for 100 White Enlistees and the Total 1972 White Enlistees Accessions

	White Enlistee Sample			1972 White Accessions			
Measure	N	M	SD	N	M	SD	z[a]
AFQT	100	64.98	20.67	70,614	64.42	19.90	.28
AQE-A	100	56.01	21.47	70,919	58.05	20.73	− .98
AQE-E	100	63.65	20.06	70,920	65.19	19.77	− .78
AQE-G	100	60.90	19.41	70,916	63.33	18.31	−1.33
AQE-M	100	62.75	18.95	70,901	62.28	19.73	.23

[a]For a discussion of the use of the *z* score to indicate the likelihood of a sample mean differing from the population mean, see Hays, *Statistics,* 1963, p. 203-204. In this situation,

$$z = \frac{M - m}{\frac{\sigma}{\sqrt{N}}}$$

where M = sample mean, m = population mean, σ = population standard deviation, and N = number of cases in the sample. Applying Tchebycheff's inequality to the above *z*'s reveals the probability of these sample means deviating as much as they do from the population means approaches 1.00 for all tests for both racial groups, given the size of the groups, and the population standard deviations.

Table 3. WAIS Subtests and IQs, AFQT Scores, AQE Aptitude Indexes, and Age and Education Levels for 100 Black and 100 White Air Force Enlistees

Measures	Black Sample (N=100) M	Black Sample (N=100) SD	White Sample (N=100) M	White Sample (N=100) SD	t	p
Information	8.87	2.08	10.77	2.34	6.03	.01
Comprehension	9.21	2.67	10.83	2.74	4.21	.01
Arithmetic	8.68	2.09	10.70	2.48	6.20	.01
Similarities	9.61	2.39	10.46	2.59	2.39	.05
Digit Span	9.19	2.76	9.67	2.46	1.29	ns
Vocabulary	8.74	1.84	10.24	2.15	5.26	.01
Digit Symbol	9.76	2.16	10.33	2.17	1.85	ns
Picture Completion	9.69	1.75	11.46	2.40	5.96	.01
Block Design	9.21	2.27	12.22	2.66	8.58	.01
Picture Arrangement	9.97	2.11	11.12	2.45	3.53	.01
Object Assembly	9.01	2.31	11.73	3.08	7.03	.01
Verbal IQ	96.74	8.78	104.99	10.47	6.02	.01
Performance IQ	97.41	9.65	109.30	10.57	8.26	.01
Full Scale IQ	96.72	8.33	107.17	9.90	8.04	.01
AFQT	41.68	13.83	64.98	20.67	9.32	.01
AQE-A	44.55	16.82	56.01	21.47	4.18	.01
AQE-E	45.25	18.57	63.65	20.06	6.70	01
AQE-G	48.26	16.73	60.90	19.41	4.91	.01
AQE-M	44.20	18.56	62.75	18.95	6.96	.01
Age	19.13	1.45	18.95	1.42	.88	ns
Education	11.96	.92	12.15	1.00	1.42	ns

Table 4. Intercorrelations of WAIS IQs, AFQT Scores, and AQE Aptitude Indexes for 100 Black Air Force Enlistees

Test	1	2	3	4	5	6	7	8
1. WAIS Verbal IQ								
2. WAIS Performance IQ	.46							
3. WAIS Full Scale IQ	.88	.83						
4. AFQT	.38	.44	.48					
5. AQE-Admin	.27	.19	.27	.26				
6. AQE-Elec	.26	.26	.30	.37	.31			
7. AQE-Gen	.23	.29	.30	.30	.62	.55		
8. AQE-Mech	.10	.23	.19	.30	.29	.40	.63	

$p < .05 = .195$ $p < .01 = .254$

Table 5. **Intercorrelations of WAIS IQs, AFQT Scores, and AQE Aptitude Indexes for 100 White Air Force Enlistees**

Test	1	2	3	4	5	6	7	8
1. WAIS Verbal IQ								
2. WAIS Performance IQ	.54							
3. WAIS Full Scale IQ	.91	.84						
4. AFQT	.64	.62	.71					
5. AQE-Admin	.60	.28	.52	.51				
6. AQE-Elec	.69	.60	.74	.75	.56			
7. AQE-Gen	.77	.44	.71	.64	.71	.73		
8. AQE-Mech	.45	.50	.53	.67	.34	.69	.56	

$p < .05 = .195$ $p < .01 = .254$

significant ($p < .05$). The intercorrelations of WAIS IQs with AQE AIs and AFQT scores, and AFQT scores with AQE AIs, were consistently lower for black airmen than for white airmen. A correlation matrix for each sample which includes demographic data and WAIS subtests appears in Tables 6 and 7.

Regression analyses were run on the 200 case sample utilizing the WAIS IQs as criterion scores and the AFQT and AQE AIs as individual predictors to determine racial differences in test scores. The regression design and models appear in Table 8. Table 9 presents a summary of the regression analyses. Where VIQ served as the criterion, significant race interactions were indicated for all AQE AIs as predictors. There were no significant race effects where AFQT served to predict VIQ. Where PIQ served as the criterion score, significant level or intercept differences were indicated for the AFQT, AQE-A, and AQE-G. Significant interactions (or slope differences) were found for AQE-E and AQE-M. Where FSIQ served as the criterion score, significant intercept differences were found for AFQT and AQE-A. Significant interactions were found when AQE-E, AQE-G, and AQE-M served as predictors of FSIQ.

IV. DISCUSSION

Consistent with the results of previous studies where Air Force selection tests have been compared with other ability measures, significant positive relationships were found between AQE and AFQT performance and performance on a test of general mental ability. Also consistent with more recent minority group test performance analyses is the finding of racial differences within Air Force tests. In the present study, black airmen consistently scored significantly lower than white airmen on all Air Force selectors.

Analyses of WAIS performance revealed that black and white enlistees had statistically different Verbal, Performance, and Full Scale IQs. An examination of WAIS subtests shows that black airmen scored significantly lower on all subtests except Digit Span and Digit Symbol. Although mean IQs for both samples fell within the normal range (90-110), black airmen scored 8 to 12 IQ points lower on the average. Score variance was less than the theoretical 15 point IQ standard deviation for both groups. The small IQ standard deviations obtained may be due to the restricted population of Air Force basic trainees. Recruiting policy (e.g., high school diploma requirement) cuts off the lower tail of the ability distribution of American youth, while the limited appeal of military service as an enlisted man cuts off the higher tail, the college graduates.

The relationships between the AFQT, AQE AIs, and WAIS IQs were explored both by simple correlational analyses and by regression analyses. For white enlistees, the correlations show that more test variance is shared by Air Force tests and the intelligence test than for black airmen. Regression analyses employing WAIS IQs as criteria and selection test scores as predictors demonstrated some type of racial effect in all but one case. In that case, black and white regression lines shared a common intercept and slope when the AFQT was used to predict WAIS Verbal IQ. Regression lines differing both in intercept and slope occurred most frequently. Implications of these complex interactions of race and selector score are not definitive, and would require an investigation of the AQE subtests to determine what specific items or group of items are causing the racial differences.

Table 6. Intercorrelations of WAIS Subtests and IQs, AFQT Scores, AQE Aptitude Indexes, and Age and Education Levels for 100 Black Enlistees

Variable	Intercorrelation of Variables																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. Age	1.00																				
2. Education	.21	1.00																			
3. AFQT	.07	.07	1.00																		
4. AQE-A	−.01	.13	.26	1.00																	
5. AQE-E	−.04	.00	.37	.31	1.00																
6 AQE-G	−.13	−.02	.36	.62	.55	1.00															
7. AQE-M	−.12	−.17	.30	.29	.40	.63	1.00														
8. VIQ	−.07	.25	.43	.27	.26	.23	.10	1.00													
9. PIQ	−.04	.10	.44	.19	.26	.29	.23	.46	1.00												
10. FSIQ	−.07	.21	.48	.27	.30	.30	.19	.88	.83	1.00											
11. Information	.04	.25	.28	.09	.09	.07	.06	.67	.33	.60	1.00										
12. Comprehension	.18	.20	.29	.17	.19	.12	.06	.71	.37	.65	.46	1.00									
13. Arithmetic	-.20	.21	.29	.25	.29	.18	.06	.57	.29	.52	.25	.22	1.00								
14. Similarities	.17	.21	.26	.20	.10	.12	.01	.67	.22	.55	.39	.47	.30	1.00							
15. Digit Span	−.13	−.02	.02	.10	.07	.13	.04	.39	.16	.34	.03	.06	.11	.03	1.00						
16. Vocabulary	.13	.29	.42	.13	.24	.19	.05	.72	.39	.66	.60	.58	.31	.51	−.05	1.00					
17. Digit Symbol	.01	.18	.11	.11	.04	.03	01	.27	.54	.45	.27	.24	.13	.15	.06	.18	1.00				
18. Picture Completion	.06	.08	.43	.20	.25	.30	.25	.39	.67	.61	.37	.32	.18	.17	.08	.45	.29	1.00			
19. Block Design	−.15	.04	.43	.23	.36	.37	.36	.30	.74	.59	.08	.17	.30	.14	.12	.27	.18	.44	1.00		
20. Picture Arrangement	.08	.12	.21	.01	.06	.01	.02	.33	.63	.55	.20	.33	.21	.24	.07	.29	.19	.32	.33	1.00	
21. Object Assembly	.00	−.01	.25	.05	.14	.19	.14	.20	.67	.49	.22	.25	−.01	.11	.01	.18	.18	.35	.42	.31	1.00

Table 7. Intercorrelations of WAIS Subtests and IQs, AFQT Scores, AQE Aptitude Indexes, and Age and Education Levels for 100 White Enlistees

Variable	Intercorrelation of Variables																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. Age	1.00																				
2. Education	.42	1.00																			
3. AFQT	.07	.18	1.00																		
4. AQE-A	.14	.41	.51	1.00																	
5. AQE-E	.11	.28	.75	.56	1.00																
6. AQE-G	.09	.36	.64	.71	.73	1.00															
7. AQE-M	.13	.25	.67	.34	.69	.56	1.00														
8. VIQ	−.01	.33	.64	.60	.69	.77	.45	1.00													
9. PIQ	.09	.19	.62	.28	.60	.44	.50	.54	1.00												
10. FSIQ	.04	.30	.71	.52	.74	.71	.53	.91	.84	1.00											
11. Information	.19	.46	.45	.58	.60	.71	.38	.77	.40	.69	1.00										
12. Comprehension	.11	.26	.47	.27	.44	.50	.24	.69	.44	.66	.47	1.00									
13. Arithmetic	−.03	.11	.54	.52	.54	.55	.34	.68	.33	.60	.45	.28	1.00								
14. Similarities	.06	.34	.53	.41	.56	.52	.43	.74	.51	.72	.53	.42	.42	1.00							
15. Digit Span	.03	.16	.17	.40	.30	.43	.16	.62	.23	.52	.37	.25	.35	.29	1.00						
16. Vocabulary	.13	.35	.63	.51	.63	.72	.52	.85	.48	.78	.69	.59	.44	.57	.47	1.00					
17. Digit Symbol	.26	.41	.22	.48	.32	.40	.20	.31	.40	.40	.34	.21	.28	.14	.30	.29	1.00				
18. Picture Completion	.15	.21	.37	.14	.34	.26	.27	.30	.66	.51	.23	.28	.11	.36	.10	.29	.17	1.00			
19. Block Design	.07	.13	.57	.22	.52	.36	.46	.43	.71	.63	.34	.32	.33	.41	.12	.36	.11	.33	1.00		
20. Picture Arrangement	−.03	.05	.27	.02	.24	.17	.25	.25	.53	.42	.13	.24	.11	.26	.12	.24	.01	.28	.16	1.00	
21. Object Assembly	.07	−.04	.45	.08	.42	.21	.36	.36	.78	.61	.25	.30	.21	.36	.19	.33	.18	.33	.54	.22	1.00

Table 8. **Regression Design**

Regression Models

Model 1. WAIS IQ = B + W + (B x Sel Test) + (W x Sel Test)
Model 2: WAIS IQ = B + W + Sel Test
Model 3: WAIS IQ = Sel Test

F - Tests

$$F = \frac{(R_F^2 - R_R^2)/df_1}{(1 - R_1^2)/df_2}$$

	Full Model		Restricted Model	Testing for:
Analysis I	Model 1	vs	Model 3	Racial Difference
Analysis II	Model 1	vs	Model 2	Slope Difference
Analysis III	Model 2	vs	Model 3	Intercept Difference

Definitions

WAISIQ = VIQ, PIQ, or FSIQ score
B = 1 if black, 0 if not black
W = 1 if white, 0 if not white
Sel Test = AFQT or AQE AI percentile score

R_F^2 = Squared multiple correlation of full model

R_R^2 = Squared multiple correlation of restricted model

R_1^2 = Squared multiple correlation of model 1

df_1 = Independent vectors in full model minus independent vectors in restricted model

df_2 = Elements in vectors minus independent vectors in model 1

Table 9. Summary of Regression Analyses

F tests for Presence of Racial Difference (slope or intercept differences)[a] Model 1 vs Model 3							
Criterion	Predictor	R_F^2	R_R^2	df_1	df_2	F	prob
VIQ	AFQT	.40742	.40058	2	196	1.13	NS
VIQ	AQE-A	.35580	.27294	2	196	12.61	.001
VIQ	AQE-E	.41354	.34079	2	196	12.16	.001
VIQ	AQE-G	.46722	.36258	2	196	19.25	.001
VIQ	AQE-M	.25999	.17376	2	196	11.55	.001
PIQ	AFQT	.47611	.44930	2	196	5.01*	.01
PIQ	AQE-A	.29991	.11748	2	196	25.54*	.001
PIQ	AQE-E	.42403	.31719	2	196	18.18	.001
PIQ	AQE-G	.36390	.22263	2	196	21.76*	.001
PIQ	AQE-M	.37527	.26173	2	196	17.81	.001
FSIQ	AFQT	.54021	.52540	2	196	3.16*	.05
FSIQ	AQE-A	.38758	.24062	2	196	23.52*	.001
FSIQ	AQE-E	.51365	.41136	2	196	13.74	.001
FSIQ	AQE-G	.49822	.36716	2	196	25.60	.001
FSIQ	AQE-M	.38220	.26353	2	196	18.82	.001

[a]Starred (*) *F* tests denote race main effects (intercept differences but no slope difference). All other significant differences were shown to be interaction effects (both slope and intercept differences).

Having found test performance differences between black and white enlistees when age and level of education were taken into account, the question may be raised as to what factors contributed to differential test performance. One possible explanation of racial differences has been described by Guinn, Tupes, and Alley (1970a). They have categorized memory, spatial, and psychomotor skills as situational abilities, as opposed to prior knowledge abilities which are largely a function of an individual's formal educational background. Racial differences in test performance are least pronounced when situational abilities are required, as in the Digit Span and Digit Symbol subtests of the WAIS. Racial differences in test performance are most pronounced when prior knowledge abilities are required; thus, it may be hypothesized that there is an overbundance of prior knowledge items th- in both Air Force selection tests and the WAIS. An overbundance of prior knowledge items may be masking the identification of true abilities and may account, in part, for the differences in test performance of black and white airmen.

Future revisions of selection tests, especially the ASVAB now coming into use, should include investigations into the degree of importance assigned to situational and prior knowledge items in the construction of tests. Such research may discover a control for the differential quality of education of black and white recruits who possess the same quantitative level of education (e.g., years of education or grade point average).

Additional racial test variance may be contributed by other test specific factors. The results of the present study show a much closer relationship between the selection tests and the intelligence test for white airmen than for black airmen. The consistently high intercorrelations of the white sample indicate a good deal of common variance shared between "intelligence" and the indicators of success in the Air Force training situation. The much lower intercorrelations of the black sample suggest that the selection tests and the WAIS are measuring more divergent factors. A possible area of inquiry into the source of this disparity in test commonality is the literacy factor.

The AFQT and AQE are paper and pencil tests requiring the subject to read the instructions and test items. In administering the WAIS, however, both instructions and test items are orally presented by the psychometrist. Only the WAIS Vocabulary subtest involves the subject reading the items, and even then the administrator is pronouncing the words, so the subject is not solely dependent upon his ability to read. Thus, it appears that one approach to isolating a possible source of racial variance in test scores would be through a study of the differential literacy requirements of the two kinds of tests.

A second source of racial variance may be attributed to test administrators. The WAIS is administered individually by professionally trained psychometrists. Air Force selection tests, however, are group administered and scored by recruiting level personnel who have had minimal formal training in testing. Even in the small sample of the present study, so called recruiter effects in the distributions of scores are readily apparent. The clustering of enlistee scores at or just above AFQT and AQE cutoffs for career field entry was found in this study and others conducted by this division (Vitola & Wilbourn, 1971). For example, 37% of the black enlistees in this study fell on or just above the minimum AFQT score (31) and 21% fell at the AQE-G 40, the qualifying cutoff. Eighteen percent of the white sample fell on an AQE-M cutoff of 60.

As a final note, the results of this study have purposefully been discussed in terms of racial differences in scores rather than in terms of racial bias. Any interpretation of the cultural fairness of the AQE or AFQT based upon these data must be tempered by the fact that it is the use to which the test is put, rather than the test itself, which defines the fairness of the test (Thorndike, 1971). In other words, the fairness of a selection test is determined by its application in the selection process, not its internal characteristics. The present study only presents data in relation to the WAIS, which was utilized as a baseline for measuring general mental ability. Hopefully, some clarification of the racial differences in the AFQT and AQE will accrue from this study. But we must leave a definitive interpretation of racial bias in selection testing to other investigations.

REFERENCES

Anastasi, A. *Psychological testing.* (3rd ed.) New York: The MacMillan Company, 1968.

Brokaw, L.D. *Prediction of Air Force training and proficiency criteria from Armed Forces selection tests.* WADC-TN-59-194, AD-227 635. Lackland AFB, Tex.: Personnel Laboratory, Wright Air Development Center, August 1959.

Guinn, N., Tupes, E.C., & Alley, W.E. *Demographic differences in aptitude test performance* AFHRL-TR-70-15, AD-710 618. Lackland AFB, Tex.: Personnel Research Division, Air Force Human Resources Laboratory, May 1970. (a)

Guinn, N., Tupes, E.C., & Alley, W.E. *Cultural subgroup differences in the relationships between Air Force aptitude composites and training criteria.* AFHRL-TR-70-35, AD-715 922. Lackland AFB, Tex.: Personnel Research Division, Air Force Human Resources Laboratory, September 1970. (b)

Hays, W.L. *Statistics.* New York: Holt, Rinehart, and Winston, 1963.

Madden, H.L., & Tupes, E.C. *Estimating reading ability level from the AQE general aptitude index.* PRL-TR-66-1, AD-632 182. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, February 1966.

Madden, H.L., & Valentine, L.D., Jr. *Conversion tables for Airman Qualifying Examination and Employee Aptitude Survey scores.* PRL-TR-67-7, AD-661 998. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, August 1967.

Madden, J.H., Valentine, L.D., Jr., & Tupes, E.C. *Comparison of the Airman Qualifying Examination with the Differential Aptitude Tests.* PRL-TR-66-7, AD-639 238. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, July 1966.

Shore, C.W., & Marion, R. *Suitability of using common selection test standards for Negro and white airmen.* AFHRL-TR-72-53, AD-754 846. Lackland AFB, Tex.: Personnel Research Division, Air Force Human Resources Laboratory, May 1972.

Thompson, C.A. *Aptitude differences related to region of enlistment of basic airmen.* WADC-TN-58-65, AD-202 846. Lackland AFB, Tex.: Personnel Laboratory, Wright Air Development Center, September 1958.

Thorndike, R.L. Concepts of culture-fairness. *Journal of Educational Measurement,* 1971, **8**, 63-70.

Valentine, L.D., Jr. *Relationship between Airman Qualifying Examination and Armed Forces Qualification Test norms.* AFHRL-TR-68-106, AD-678 528. Lackland AFB, Tex.: Personnel Research Division, Air Force Human Resources Laboratory, July 1968.

Vitola, B.M., Massey, I.R., & Wilbourn, J.M. *Development and standardization of the Airman Qualifying Examination -- Form J.* AFHRL-TR-71-28, AD-730 592. Lackland AFB, Tex.: Personnel Division, Air Force Human Resources Laboratory, May 1971.

Vitola, B.M., & Wilbourn, J.M. *Comparative performance of male and female enlistees on Air Force selection measures.* AFHRL-TR-71-9, AD-726 531 Lackland AFB, Tex.: Personnel Division, Air Force Human Resources Laboratory, February 1971.