DOCUMENT RESUME

ED 274 712 TM 860 591

AUTHOR Rogers, Deborah L.; And Others
TITLE Air Force Officer Qualifying Test Form O: Development and Standardization.
INSTITUTION Air Force Human Resources Lab., Brooks AFB, Tex. Manpower and Personnel Div.
REPORT NO AFHRL-TR-86-24
PUB DATE Aug 86
NOTE 29p.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Aptitude Tests; Error of Measurement; Flight Training; Military Training; *Occupational Tests; Standardized Tests; *Test Construction; Test Items; *Test Reliability; *Test Validity; *Vocational Aptitude
IDENTIFIERS *Air Force Officer Qualifying Test

ABSTRACT

This report presents the rationale, development, and standardization of the Air Force Officer Qualifying Test (AFOQT) Form O. The test is used to select individuals for officer commissioning programs, and candidates for pilot and navigator training. Form O contains 380 items organized in 16 subtests. All items are administered in a single test booklet, accompanied by a single machine-scorable answer sheet. Five composites, expressed as percentiles, are derived from various combinations of the subtests: Pilot, Navigator-Technical, Academic Aptitude, Verbal, and Quantitative. Although Form O contains the same composites as predecessor forms of the AFOQT, this form introduced changes in content, format, administration, and scoring. Reliabilities and intercorrelations are consistent with those of previous forms. Form O was standardized to link scores to a normative group based on Form N data. The use of common items on Form N and Form O enabled a series of equipercentile equatings through the common items, resulting in the equating of the entire Form O battery to the entire Form N battery. Raw-score-to-percentile-score Form O composite conversion tables were developed from these equatings, and were implemented for operational use with the introduction of Form O in September 1981 at Officer Training School testing sites and in March 1982 at Air Force Reserve Officer Training Corps testing sites. The data indicate the test is an appropriate replacement for Form N as a selection instrument for Air Force officers. Several tables are appended. (JAZ)

AFHRL-TR-86-24

AIR FORCE HUMAN RESOURCES LABORATORY

AIR FORCE OFFICER QUALIFYING TEST FORM O: DEVELOPMENT AND STANDARDIZATION

Deborah L. Rogers, Capt, USAF
Bennie W. Roach
Toni G. Wegner, Capt, USAF

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

August 1986

Interim Report for Period April 1980 - August 1984

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

NOTICE

When U.S. Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

WILLIAM E. ALLEY, Scientific Advisor
Manpower and Personnel Division

RONALD L. KERCHNER, Colonel, USAF
Chief, Manpower and Personnel Division

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
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-TR-86-24	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Manpower and Personnel Division 6b. OFFICE SYMBOL (*If applicable*) AFHRL/MOAO	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 47

11. TITLE (*Include Security Classification*)
Air Force Officer Qualifying Test Form O: Development and Standardization

12. PERSONAL AUTHOR(S)
Rogers, Deborah L.; Roach, Bennie W.; Wegner, Toni G.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (*Year, Month, Day*)	15. PAGE COUNT
Interim	FROM Apr 80 TO Aug 84	August 1986	26

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*)
Air Force Officer Qualifying Test; item difficulty; aptitude tests; mental abilities testing; classification tests
(Continued)

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

Air Force Officer Qualifying Test (AFOQT) Form O replaced AFOQT Form N in September 1981 at Officer Training School testing sites and in March 1982 at Air Force Reserve Officer Training Corps testing sites. The new form contains 380 items organized into 16 subtests and yields a single raw score which is converted to a percentile equivalent for each of five composites: Pilot, Navigator-Technical, Academic Aptitude, Verbal, and Quantitative. Although Form O serves the same purpose as its predecessors, it differs as to content, format, administration, and scoring. Form O item statistics, reliabilities, and intercorrelations were computed and found to be similar to corresponding Form N statistics. Standardization of Form O was done through equipercentile equating of common items to the Form N normative group, and raw-score-to-percentile-score Form O composite conversion tables were developed. It was concluded that Form O is acceptable as a selection instrument for Air Force officers.

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

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

Item 18 (Concluded):

officer selection and classification
reliability
selection tests
test construction
test standardization

SUMMARY

This repo.t presents the rationale, development, and standardization of the Air Force Officer Qualifying Test (AFOQT) Form O. The AFOQT is used to select individuals for officer commissioning programs, and to select candidates for pilot and navigator training. AFOQT Form O contains 380 items (many from previous versions of the AFOQT) organized in 16 subtests. All items are administered in a single test booklet, accompanied by a single machine-scorable answer sheet. Five composites, expressed as percentiles, are derived from various combinations of the subtests: Pilot, Navigator-Technical, Academic Aptitude, Verbal, and Quantitative. There is a single raw-score-to-percentile conversion table for each composite. Although Form O contains the same composites as predecessor forms of the AFOQT, this form introduced changes in content, format, administration, and scoring. Reliabilities and intercorrelations are consistent with those of previous forms.

Form O was standardized to link scores to a normative group based on Form N data. The use of common items on Form N and Form O enabled a series of equipercentile equatings through the common items, resulting in the equating of the entire Form O battery to the entire Form N battery. Raw-score-to-percentile-score Form O composite conversion tables were developed from these equatings, and were implemented for operational use with the introduction of Form O in September 1981 at Officer Training School testing sites and in March 1982 at Air Force Reserve Officer Training Corps testing sites. The data indicate the test is an appropriate replacement for Form N as a selection instrument for Air Force officers.

PREFACE

This work was accomplished under work unit number 77191847, "Development and Validation of Civilian and Non-Rated Officer Selection Methodologies." This work unit was established in response to Air Force Regulation (AFR) 35-8, Air Force Military Testing Program.

This effort could not have been accomplished without the cooperation and dedication of several individuals at the Air Force Human Resources Laboratory. Col Ronald Kerchner, Dr. Nancy Guinn Vitola, and Dr. William Alley of the Manpower and Personnel Division provided helpful guidance during the design and editing phases of this effort. Dr. Lonnie Valentine, Lt Col Lawrence Short, Dr. Malcolm Ree, Ms. Jacobina Skinner, 1Lt Thomas Arth, Mr. Douglas Cowan, and Dr. Thomas Watson all made significant technical contributions. Special gratitude is also expressed to Technical Services Division personnel who devoted long hours to provide rapid computer programming and processing. In particular, Ms. Doris Black, Mr. James Friemann, Mr. William Glasscock, AlC Steven Hoffer, AlC Kevin Smith, and SrA Dave LeBrun put forth exceptional efforts which greatly enhanced the product. Finally, Ms. Sandy Stringfellow deserves special thanks for her quick and careful typing of this manuscript.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Tables (Concluded)

AIR FORCE OFFICER QUALIFYING TEST FORM O: DEVELOPMENT AND STANDARDIZATION

I. BACKGROUND

The Air Force Officer Qualifying Test (AFOQT) is used to select individuals for officer commissioning programs, such as Officer Training School (OTS) and Air Force Reserve Officer Training Corps (AFROTC). It is also used to select candidates for specific training programs such as pilot and navigator training. New forms of the AFOQT are developed on a periodic basis, usually a 3-year cycle, to update test information, to modify test content as needed, and to reduce the potential for test compromise. This report describes the development and standardization of AFOQT Form O, which replaced its predecessor (AFOQT Form N) in September 1981 at OTS testing sites and in March 1982 at AFROTC testing sites.

The precursor of the AFOQT was the Aircrew Classification Battery (AC8), a test developed at the onset of World War II to select and classify individuals into aircrew training. In 1951, selected subtests of the ACB were combined with an experimental aptitude test called the Aviation Cadet--Officer Candidate--Qualifying Test. The resulting combination was the AFOQT Form A. The composition and use of the AFOQT have changed over the 15 different forms of the test; however, the test has remained the basis of the Air Force officer selection testing program to the present. Five aptitude composites have been used throughout the history of the test: Pilot, Navigator-Technical (formerly Observer-Technical), Academic Aptitude (formerly Officer Quality), Verbal, and Quantitative. Although the aptitudes being measured by the composites have remained constant, the subtests making up the composites have changed over the years. These changes are documented in papers describing the history of the AFOQT and the development of earlier forms (Gould, 1978; Miller, 1966, 1968, 1970, 1972, 1974; Miller & Valentine, 1964; Rogers, Roach, & Short, 1986; Valentine & Creager, 1961).

Prior to 1960, the AFOQT was the selection test used for all commissioning sources. Since then, Scholastic Aptitude Test (SAT) scores have been the primary aptitude selection tool for the United States Air Force Academy (USAFA). The AFOQT is still used to select candidates (usually college graduates or prior enlistees) for OTS, to select AFROTC cadets for scholarships or into the Professional Officers Course, and to select candidates for Undergraduate Pilot Training (UPT) and Undergraduate Navigator Training (UNT).

AFOQT results are reported in terms of composite percentile scores. Conversion tables are used to convert raw scores on the composites to percentiles. For some previous forms, different conversion tables have been used for males and females and for those of different educational backgrounds (see Rogers et al., 1986). With the implementation of Form O in 1981, a single conversion table was developed for each composite.

II. GENERAL CHARACTERISTICS

AFOQT Form O consists of 380 items organized into 16 subtests. These subtests are used in one or more of five aptitude composites: Pilot, Navigator-Technical, Academic Aptitude, Verbal, and Quantitative. Table 1 shows the subtests and how they are organized into composites. The test requires approximately 4.5 hours to administer. Examinees are required to take all parts of the AFOQT. One conversion table is used for each of the five composites to convert all applicants' scores to percentiles.

Table 1. Content and Organization of AFOQT Form O

Subtests	No. of Items	Composites[a] P	N-T	AA	V	Q
Verbal Analogies	25	X		X	X	
Arithmetic Reasoning	25		X	X		X
Reading Comprehension	25			X	X	
Data Interpretation	25		X	X		X
Word Knowledge	25			X	X	
Math Knowledge	25		X	X		X
Mechanical Comprehension	20	X	X			
Electrical Maze	20	X	X			
Scale Reading	40	X	X			
Instrument Comprehension	20	X				
Block Counting	20	X	X			
Table Reading	40	X	X			
Aviation Information	20	X				
Rotated Blocks	15		X			
General Science	20		X			
Hidden Figures	15		X			
Total	380	205	265	150	75	75

[a]P = Pilot; N-T = Navigator-Technical; AA = Academic Aptitude; V = Verbal; and Q = Quantitative.

Form O is printed as a single test booklet, Air Force Personnel Test (AFPT) 982, with an accompanying machine-scannable answer sheet. Two sites are used to computer-score answer sheets--the Air Force Human Resources Laboratory at Brooks AFB, Texas (to score answer sheet AFPT 987), and Headquarters AFROTC at Maxwell AFB, Alabama (to score answer sheet AFROTC PTF 987). Answer sheets are scored at Maxwell AFB for individuals who test at AFROTC sites. Most, but not all, of these examinees are applying for a commission through AFROTC. Answer sheets are scored at Brooks AFB for individuals tested at all non-AFROTC sites. A majority of these examinees are applying for a commission through OTS. The administration manual is AFPT 980 for non-AFROTC sites and AFROTC PT 980 for AFROTC sites. On all subtests, only correct answers are scored. All subtests have five item-options, except for Instrument Comprehension which has four.

III. TEST MODIFICATIONS OVER RECENT FORMS

Form O differs from its most recent predecessor, Form N, in content, format, administration, and scoring. Form N was composed of 606 items divided into 18 subtests. For Form O, the number of items was reduced to 380, and the number of subtests was reduced to 16. To create Form O, four subtests (Pilot Biographic and Attitude Scale, Aerial Landmarks, Tools, and Background for Current Events) that were included in Form N were removed; one subtest (Aviation Information) that had been used prior to Form N was reinstated; and one new subtest (Hidden Figures) was added. The 16 subtests in Form O make up the same five composites derived from Form N; however, the Officer Quality composite was renamed Academic Aptitude. This was done to prevent misinterpretation of what the composite is intended to measure. Table 2 shows the subtests in Form N and how they were organized into composites. Six of the subtests in Form N were

designated as speeded subtests. For Form 0, none of the subtests is specifically designated as speeded since all subtests contain elements of both power and speed. Refer to Section IV, Item Selection, for more information on how speededness affected item analysis.

Table 2. Content and Organization of AFOQT Form N

Subtests	Items	Composites[a]				
		P	N-T	OQ	V	Q
Booklet 1						
Arithmetic Reasoning	25		X	X		X
Math Knowledge	25		X	X		X
Data Interpretation	25		X	X		X
Booklet 2						
Word Knowledge	25			X	X	
Reading Comprehension	25			X	X	
Background for Current Events	25			X	X	
Verbal Analogies	25	X		X	X	
Booklet 3						
Table Reading[b]	50	X	X			
Electrical Maze[b]	30	X	X			
Block Counting[b]	80	X	X			
Scale Reading[b]	48	X	X			
Tools	25	X	X			
Mechanical Comprehension	24	X	X			
Booklet 4						
Rotated Blocks	20		X			
Aerial Landmarks[b]	40		X			
General Science	24		X			
Instrument Comprehension[b]	24	X				
Pilot Biographic and Attitude Scale	66	X				
Total	606	372	416	175	100	75

[a]P = Pilot; N-T = Navigator-Technical; OQ = Officer Quality; V = Verbal; and Q = Quantitative.

[b]Speeded Subtests.

Form N subtests were printed in four test booklets, accompanied by three double-sided answer sheets. The answer sheets were hand-scored using 10 scoring keys. Form 0 differs in format, in that all 16 subtests are contained in a single test booklet, which is accompanied by a single machine-scannable answer sheet. The use of a single answer sheet was designed to reduce the number of incomplete test scores due to missing or lost answer sheets and to eliminate answer sheet matching problems resulting from transcription errors in social security account numbers.

Test administration differences between Form 0 and its predecessors resulted directly from the reduction in the length of the test battery. The use of a less complex set of testing materials shortened the total time for test administration from about 7 hours to about 4.5 hours. This enabled applicants to take all 16 Form 0 subtests. For previous forms, time

restrictions made it necessary to require only applicants who were applying for UPT or UNT to take the subtests unique to the Pilot and Navigator-Technical composites. Unlike previous forms then, all composite scores are available for applicants testing on Form O.

Between Form N and Form O, scoring procedures changed in three ways. First, because of the use of a machine-scannable answer sheet for Form O, computer-based scoring was initiated as a replacement for the previous inefficient, error-prone hand-scoring method used with Form N. Second, speeded subtests in Form N were corrected for guessing. The Instrument Comprehension subtest was scored as the number of right answers minus one-third of the wrong answers, and other speeded subtests were scored as the number of right answers minus one-fourth of the wrong answers. All subtests in Form O are scored as number right only. Finally, Form N used multiple tables for converting each composite's raw scores into percentile scores. These conversion tables were based on level of education, to take into account effects of education on AFOQT performance. The tables were established for two categories of applicants: those with less than 2 years of college and those with 2 or more years of college. A decision was made with the implementation of Form O to put all scores in a single metric (see Roach, 1986). This reduced the number of conversion tables to five (one for each composite).

IV. ITEM SELECTION

Over 90% of the items selected for Form O were obtained from previous versions of the AFOQT. The sources of the items in Form O are provided in Table 3. Form N items were used extensively in order to link Form O to the normative group (refer to Section VI--Standardization).

Table 3. Source of Items for AFOQT Form O

		Items selected from AFOQT Form:							
AFOQT Form O		L Only	N Only	L&M	M&N	L&N	L&M&N	New	Total
Total	n	58	162	15	102	1	9	33	380
(Nonoverlapping)	%	15.3	42.6	3.9	26.8	0.0	2.4	8.7	
Pilot	n	29	93	13	65	0	0	5	205
Composite	%	14.1	45.4	6.3	31.7	0.0	0.0	2.4	
Navigator-Technical	n	22	127	7	73	0	7	29	265
Composite	%	8.3	47.9	2.6	27.5	0.0	2.6	10.9	
Academic Aptitude	n	29	63	4	32	1	8	13	150
Composite	%	19.3	42.0	2.7	21.3	0.7	5.3	8.7	
Verbal	n	17	32	2	17	1	2	4	75
Composite	%	22.7	42.7	2.7	22.7	1.3	2.7	5.3	
Quantitative	n	12	31	2	15	0	6	9	75
Composite	%	16.0	41.3	2.7	20.0	0.0	8.0	12.0	

Each form of the AFOQT is designed to have a level of difficulty similar to that of the preceding form. Difficulty levels and biserial correlations were computed for items in each of the subtests, using a sample of 37,409 applicants (6,944 from AFROTC sites, and 30,465 from non-AFROTC sites) who were tested on Form O. The sample contained approximately 88% males and 12% females. By race, 79% were White, 13% were Black, and 5% were Hispanic. Applicants from

AFROTC testing sites have typically completed 12 or 13 years of school and are 18 or 19 years old. Applicants from non-AFROTC testing sites are usually 22- to 25-year-old college graduates. The difficulty levels were computed as the proportion of examinees who answered the item correctly. Thus, the more difficult the item, the lower the value. In a very few cases, item difficulties are below chance. There are not enough of these to negatively affect subtest scores. Biserial correlations represent the relationship between the correct response on an item and the total score on the subtest in which the item is included. These correlations were not corrected for item-test overlap.

Due to testing time limits established for the subtests, a large number of examinees do not complete all of the items on a majority of subtests. This means that most of the subtests are, to some extent, speeded in nature. The degree of speededness varies from subtest to subtest and ranges from speed of performance having a slight impact to a heavy impact on subtest score. None of the subtests, however, can be considered totally speeded. That is, even on the highly speeded subtests, power plays a part because all applicants who attempt each item do not necessarily answer it correctly. For this reason, none of the Form 0 subtests has been specifically designated as speeded. Although the same speededness considerations have similarly applied to previous forms of the AFOQT, this change is being made to better represent the actual situation.

To get the maximum amount of information about each subtest, difficulty levels and biserial correlations were computed for each item, based only on applicants who reached that item. The range and median difficulty levels and biserial correlations for each subtest are presented in Table 4. In addition, to indicate the relative speededness of each subtest, the number of items not reached by 5% and 20% of the applicants is tabled. From these data, it may be seen that subtests such as Table Reading, Electrical Maze, and Scale Reading are relatively more speeded, whereas subtests such as Mechanical Comprehension and Rotated Blocks are relatively less speeded.

Table 4. Item Difficulty Levels and Biserial Correlations for AFOQT Form 0

Subtest	Difficulty level		Biserial correlation		No. of items not reached by % applicants[a]	
	Range	Median	Range	Median	5%	20%
Verbal Analogies	.29-.91	.65	.42-.75	.58	6	2
Arithmetic Reasoning	.24-.81	.57	.38-.68	.61	9	4
Reading Comprehension	.47-.82	.69	.45-.80	.67	10	4
Data Interpretation	.14-.89	.49	.31-.60	.45	12	7
Word Knowledge	.24-.83	.56	.44-.78	.65	10	0
Math Knowledge	.37-.81	.54	.43-.81	.58	13	4
Mechanical Comprehension	.31-.83	.47	.24-.66	.53	0	0
Electrical Maze	.18-.67	.41	.37-.76	.59	15	10
Scale Reading	.29-.92	.53	.27-.60	.46	19	13
Instrument Comprehension	.34-.65	.50	.43-.76	.59	13	8
Block Counting	.30-.90	.61	.43-.73	.62	12	8
Table Reading	.22-.94	.84	.42-.86	.70	23	16
Aviation Information	.23-.72	.47	.38-.72	.56	5	0
Rotated Blocks	.25-.85	.43	.57-.73	.65	0	0
General Sciences	.10-.78	.44	.29-.71	.49	3	0
Hidden Figures	.34-.92	.66	.49-.67	.60	9	2

[a]Refer to Table 1 for the number of items in each subtest.

V. RELIABILITY, INTERCORRELATIONS, AND VALIDITY

Subtest and composite reliabilities and intercorrelations were computed on Form O data, using the combined AFROTC (n = 6,944) and non-AFROTC (n = 30,465) sample of 37,409 cases. As would be expected of two similar versions of the same test, the results are similar to the data obtained on the predecessor test, Form N (Gould, 1978). Internal consistency reliability coefficients were derived for the subtests using the Kuder-Richardson Formula 20. The subtests were then combined to form the five composites, and the reliabilities of the composites were based on a formula developed by Wherry and Gaylord (1943). Tables 5 and 6 present the reliabilities and standard errors of measurement of Form O subtests and composites, respectively. Raw scores were used in all computations. It should be noted that the reliability estimates may be overestimated due to the speededness of the subtests. Internal consistency is artificially enhanced by treating items not reached by applicants as incorrect (see Table 4 for an account of how many items are affected for each subtest). Available data make it impossible to estimate lower-bound reliability estimates.

Table 5. Reliabilities and Standard Errors of Measurement (SEM) for AFOQT Form O Subtests

Subtest	Reliability	SEM
Verbal Analogies	.796	1.915
Arithmetic Reasoning	.804	1.944
Reading Comprehension	.885	2.031
Data Interpretation	.719	2.104
Word Knowledge	.882	2.013
Math Knowledge	.867	2.144
Mechanical Comprehension	.712	1.975
Electrical Maze	.809	1.822
Scale Reading	.839	2.700
Instrument Comprehension	.844	1.912
Block Counting	.837	1.793
Table Reading	.925	1.197
Aviation Information	.794	1.961
Rotated Blocks	.769	1.600
General Science	.699	1.992
Hidden Figures	.701	1.547

Table 6. Reliabilities and Standard Errors of Measurement (SEM) for AFOQT Form O Composites

Composite	Reliability	SEM
Pilot	.964	5.395
Navigator-Technical	.967	6.657
Academic Aptitude	.959	4.963
Verbal	.944	3.455
Quantitative	.919	3.575

Intercorrelation matrices based on Pearson correlation coefficients were computed for subtest raw scores (Table 7) and composite raw scores (Table 8). Although subtest intercorrelations vary considerably (from .169 to .729), composite intercorrelations are consistently high (greater than .600). High composite intercorrelations are in part traceable to

the overlap of subtests in composites, especially Verbal and Quantitative with Academic Aptitude, and Pilot with Navigator-Technical (see Table 1).

Validity information is available for recent forms of the AFOQT in studies predicting success in non-rated officer technical training courses (Arth, 1985), air weapons controller training (Finegold & Rogers, 1985), UPT (Bordelon & Kantor, 1986), and UNT (Shanahan & Kantor, 1986). Results generally show AFOQT scores to predict performance best in classroom training. Validities are expected to remain constant across forms because of the stability of the composites. Validity information obtained on AFROTC cadets who took AFOQT Form O is available showing the relationship between AFOQT composite scores and scores on the SAT, American College Test, and school grade point average (Diehl, 1986). Results are encouraging, suggesting some unique measurement properties of the AFOQT (for example, the Pilot composite correlates less than .45 with any non-AFOQT measure), but showing overlap where it would be expected. SAT-Verbal and SAT-Math, for example, have strong correlations with their corresponding AFOQT composite counterparts (Verbal and Quantitative), and much lower correlations with noncorresponding counterparts.

Table 7. Subtest Intercorrelations[a] for AFOQT Form O

	VA	AR	RC	DI	WK	MK	MC	EM	SR	IC	BC	TR	AI	RB	GS
VA[b]															
AR	.566														
RC	.729	.563													
DI	.536	.672	.557												
WK	.682	.451	.769	.462											
MK	.534	.711	.505	.603	.404										
MC	.476	.497	.465	.466	.388	.477									
EM	.265	.362	.241	.376	.169	.377	.443								
SR	.481	.681	.462	.636	.361	.624	.497	.443							
IC	.369	.406	.346	.448	.284	.379	.502	.422	.502						
BC	.425	.509	.388	.502	.297	.486	.486	.467	.612	.498					
TR	.344	.450	.355	.466	.261	.453	.313	.321	.556	.372	.519				
AI	.340	.316	.365	.359	.331	.264	.508	.283	.363	.581	.316	.242			
RB	.404	.453	.329	.408	.257	.459	.544	.412	.499	.466	.542	.347	.350		
GS	.510	.473	.536	.437	.507	.525	.570	.336	.424	.420	.365	.263	.465	.409	
HF	.363	.368	.319	.372	.262	.370	.383	.337	.443	.358	.450	.363	.255	.428	.309

[a] All intercorrelations are statistically significant at $p < .05$.

[b] VA = Verbal Analogies; AR = Arithmetic Reasoning; RC = Reading Comprehension; DI = Data Interpretation; WK = Word Knowledge; MK = Math Knowledge; MC = Mechanical Comprehension; EM = Electrical Maze; SR = Scale Reading; IC = Instrument Comprehension; BC = Block Counting; TR = Table Reading; AI = Aviation Information; RB = Rotated Blocks; GS = General Science; and HF = Hidden Figures.

Table 8. Composite Intercorrelations[a] for AFOQT Form O

Composite	P	N-T	AA	V
Pilot (P)				
Nav-Tech (N-T)	.945			
Academic Aptitude (AA)	.753	.821		
Verbal (V)	.615	.615	.917	
Quantitative (Q)	.754	.887	.886	.627

[a]All intercorrelations are statistically significant at $p < .05$.

VI. STANDARDIZATION

The normative group for AFOQT Form O scores is a sample of basic airmen, AFROTC students, OTS candidates, USAFA cadets, and junior officers, all of whom were tested on AFOQT Form N (see Gould, 1978, for more information about the sample). Form O scores were linked to Form N scores and to the normative group using equipercentile equating of forms with common items (Angoff, 1971). This was accomplished using the 274 items which overlap between Forms N and O. This is the first form of the AFOQT to be equated to an anchor test (Form N) through the use of common items.

A three-step process was used to equate Form O to Form N (see Figure 1). In the first step, Form N item responses were obtained for a sample of applicants to the Airmen Education and Commissioning Program (AECP, n = 7,047), OTS (n = 4,581), and AFROTC (n = 2,742). Of the original sample, about 6,000 completed all of the subtests; the remainder completed only the subtests required for the Officer Quality, Verbal, and Quantitative composites. These data were weighted to represent the applicant population in the late 1970s by type of application program (approximately 10% AECP, 50% OTS, and 40% AFROTC). Composite scores for this sample were computed based on all 606 items in Form N (referred to as Full Battery, or FB) and on the 274 Form N items that were used in Form O (referred to as Short Battery, or SB). Using this sample, the SB Form N scores were equated to FB Form N scores, and SB Form N raw-score-to-percentile conversion tables were developed for each composite. These tables therefore provided conversions from SB Form N raw composite scores to percentiles, based on the original Form N normative group.

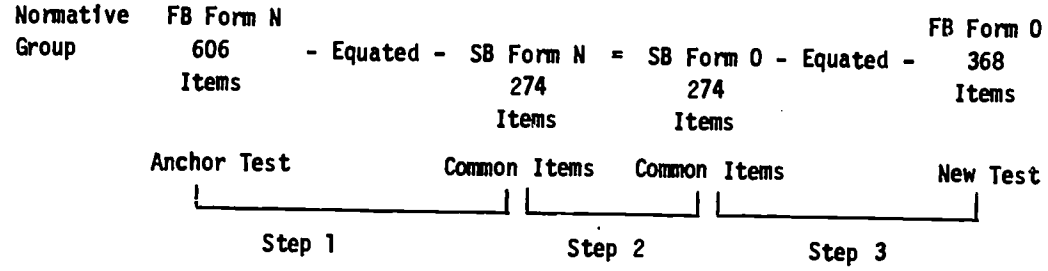

Figure 1. Equating through Common Items to Normative Group.

Because the SB Form N items were contained in Form O, the second step involved adopting the SB Form N conversion tables for Form O. Ready-made conversion tables tied to the Form N normative group were thus available, beginning with the implementation of Form O in 1981, by scoring the 274 items in Form O that were common to Form N and using the SB Form N composite conversion tables. These tables, which can also be labeled SB Form O conversion tables, are presented in Appendix A.

The third step in equating Form O to Form N used equipercentile equating of FB Form O scores[1] to SB Form O scores to link these scores to the normative group. The sample used for this equating consisted of Air Force officer applicants who tested on Form O at AFROTC (n = 6,944) and non-AFROTC (n = 30,465) test sites. This is the same sample that was used for the item analyses (Section IV) and the reliabilities and intercorrelations (Section V). Item responses were used to calculate FB and SB Form O scores for each applicant. Weights were applied prior to the equating so that applicants tested at AFROTC and non-AFROTC sites would be equally represented. The resulting FB Form O raw-score-to-percentile composite conversion tables are presented in Appendix B. These tables were implemented operationally on 24 January 1984. Prior to that date, reported AFOQT percentile scores for applicants who took Form O were based on the SB Form O composite conversion tables.

VII. CONCLUSIONS

A new form of the AFOQT (Form O) has been developed and standardized through the use of common items between Forms N and O. Tables are presented to convert Form O raw composite scores to percentiles referenced to a Form N normative group. The results from item analyses and the intercorrelations and internal consistencies of the subtests and composites indicate the test is an appropriate replacement for Form N as a selection instrument for Air Force officers.

REFERENCES

Angoff, W.H. (1971). Scales, norms, and equivalent scores. In R.L. Thorndike (Ed.), Educational measurement (2nd ed). Washington, DC: American Council on Education, 508-600.

Arth, T.O. (1985). Validation of the AFOQT for non-rated officers (AFHRL-TP-85-50, AD-A164 134). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Bordelon, V. P., & Kantor, J. E. (1986). Utilization of psychomotor screening for USAF pilot candidates: Independent and integrated selection methodologies (AFHRL-TR-86-4). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

[1]Form O contains 380 items. In the equating of the complete set of Form O items to the SB Form O items, 12 items were identified to be deleted from the complete set of Form O items. Items were deleted because they were incorrectly keyed, had incorrect item options that were bad distractors, had difficulty levels less than chance, or had low biserial correlations. Items appeared in multiple composites such that the Pilot composite contained five later-deleted items, the Navigator-Technical composite contained eight, the Academic Ability composite contained ten, the Verbal composite contained four, and the Quantitative composite contained six. Full Battery (FB) Form O scores, therefore, are based on scores from only 368 items. This is reflected in the FB Form O raw-score-to-percentile composite conversion tables.

Diehl, G.E. (1986). Correlations among SAT, ACT, AFOQT and grade point average. Unpublished manuscript, Plans and Evaluation Branch, HQ AFROTC, Maxwell AFB, AL.

Finegold, L.S., & Rogers, O.L. (1985). Relationship between Air Force Officer Qualifying Test scores and success in air weapons controller training (AFHRL-TR-85-13, AO-A158 162). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Gould, R.B. (1978). Air Force Officer Qualifying Test Form N: Development and standardization (AFHRL-TR-78-43). Brooks AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

Miller, R.E. (1966). Development of officer selection and classification tests - 1966 (PRL-TR-66-5, AO-639 237). Lackland AFB, TX: Personnel Research Laboratory, Aerospace Medical Division.

Miller, R.E. (1968). Development of officer selection and classification tests - 1968 (AFHRL-TR-68-104, AO-679 989). Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

Miller, R.E. (1970). Development and standardization of the Air Force Officer Qualifying Test Form K (AFHRL-TR-70-21, AD-710 602). Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

Miller, R.E. (1972). Development and standardization of the Air Force Officer Qualifying Test Form L (AFHRL-TR-72-47). Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

Miller, R.E. (1974). Development and standardization of the Air Force Officer Qualifying Test Form M (AFHRL-TR-74-16). Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

Miller, R.E., & Valentine, L.O., Jr. (1964). Development and standardization of the Air Force Officer Qualifying Test-64 (PRL-TDR-64-6, AO-600 782). Lackland AFB, TX: Personnel Research Laboratory, Aerospace Medical Division.

Roach, B.W. (1986). The influence of college education on standardized test performance: Should multiple conversion scales be used? In G.E. Lee (Ed.), Proceedings of the Tenth Symposium of the Psychology in the Department of Defense (USAFA-TR-86-1). U.S. Air Force Academy, Colorado Springs, CO.

Rogers, O.L., Roach, B.W., & Short, L.O. (1986). Mental ability testing in the selection of Air Force officers: A brief historical overview (AFHRL-TP-86-23). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Shanahan, F.M., & Kantor, J.E. (1986). Basic Navigator Battery: An experimental selection composite for undergraduate navigator training (AFHRL-TR-86-3, AO-A168 857). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Valentine, L.O., Jr., & [illegible] J.A. (1961). Officer selection and classification tests: Their development and use ([illegible] 61-145, AO-269 827). Lackland AFB, TX: Personnel Laboratory, Aeronautical Systems Di[illegible]

Wherry, R.J., & Gaylord, R.H. [illegible]. The concept of test and item reliability in relation to factor pattern. Psychomet[illegible] 8, 247-264.

APPENDIX A: AFOQT FORM O SHORT BATTERY CONVERSION TABLES

Table A-1. Form O Pilot Composite Short Battery Conversion Table

Raw score	Percentile	Raw score	Percentile	Raw score	Percentile
0-33	01	75	30	103	69
34-38	02	76	31	104	70
39-41	03	77	33	105	71
42-44	04	78	34	106	73
45-46	05	79	35	107	74
47-48	06	80	37	108	75
49-50	07	81	38	109	77
51	08	82	39	110	78
52-53	09	83	41	111	79
54	10	84	42	112	81
55-56	11	85	43	113	82
57	12	86	45	114	83
58-59	13	87	46	115	84
60	14	88	47	116	85
61	15	89	49	117	86
62	16	90	50	118	87
63	17	91	51	119	88
64	18	92	52	120	89
65	19	93	54	121	90
66	20	94	55	122	91
67	21	95	57	123	92
68	22	96	58	124	93
69	23	97	60	125-126	94
70	24	98	62	127-128	95
71	25	99	63	129-130	96
72	26	100	64	131-133	97
73	27	101	66	134-135	98
74	29	102	67	136-158	99

Table A-2. Form O Navigator-Technical Composite Short Battery Conversion Table

Raw score	Percentile	Raw score	Percentile	Raw score	Percentile
0-44	01	99	33	130	69
45-51	02	100	34	131	70
52-56	03	101	36	132	71
57-59	04	102	37	133	72
60-61	05	103	38	134	73
62	06	104	39	135	74
63-64	07	105	40	136	75
65-67	08	106	41	137	76
68-69	09	107	42	138	77
70-71	10	108	43	139	78
72-73	11	109	44	140	79
74	12	110	45	141	80
75-76	13	111	46	142	81
77	14	112	48	143-144	82
78-79	15	113	49	145	83
80	16	114	50	146	84
81-82	17	115	51	147	85
83	18	116	53	148	86
84	19	117	54	149-150	87
85	20	118	55	151	88
86-87	21	119	56	152-153	89
88	22	120	57	154	90
89	23	121	58	155	91
90	24	122	59	156-157	92
91	25	123	61	158-159	93
92	26	124	62	160-161	94
93	27	125	63	162-164	95
94	28	126	64	165-167	96
95	29	127	65	168-170	97
96	30	128	66	171-174	98
97	31	129	67	175-207	99
98	32				

Table A-3. Form O Academic Aptitude Composite Short Battery Conversion Table

Raw score	Percentile	Raw score	Percentile
0-20	01	60	47
21-23	02	61	49
24-25	03	62	51
26-27	04	63	52
28-29	05	64	54
30	06	65	57
31	07	66	59
32-33	08	67	61
34	09	68	63
35	10	69	65
36	11	70	67
37	12	71	69
38	13	72	70
39	14	73	72
40	16	74	74
41	17	75	76
42	18	76	78
43	19	77	80
44	20	78	81
45	21	79	83
46	23	80	84
47	25	81	86
48	26	82	87
49	28	83	89
50	29	84	90
51	31	85	91
52	33	86	92
53	34	87	93
54	36	88	94
55	38	89	95
56	40	90- 91	96
57	42	92	97
58	44	93- 94	98
59	45	95-104	99

Table A-4. Form O Verbal Composite Short Battery Conversion Table

Raw score	Percentile	Raw score	Percentile
0-08	01	30	44
09	02	31	47
10	03	32	50
11	04	33	53
12	05	34	56
13	06	35	60
14	07	36	63
15	09	37	66
16	10	38	69
17	12	39	73
18	14	40	75
19	16	41	78
20	18	42	81
21	20	43	85
22	22	44	87
23	25	45	90
24	27	46	92
25	31	47	95
26	33	48	97
27	36	49	98
28	39	50-52	99
29	41		

Table A-5. Form O Quantitative Composite Short Battery Conversion Table

Raw score	Percentile	Raw score	Percentile
0-09	01	30	54
10	02	31	58
11	03	32	61
12	04	33	65
13	05	34	68
14	07	35	71
15	09	36	75
16	10	37	77
17	13	38	80
18	15	39	83
19	18	40	85
20	21	41	88
21	24	42	90
22	27	43	92
23	30	44	93
24	33	45	95
25	37	46	96
26	40	47	97
27	44	48	98
28	47	49-52	99
29	51		

APPENDIX B: AFOQT FORM O FULL BATTERY CONVERSION TABLES

Table B-1. Form O Pilot Composite Full Battery Conversion Table

Raw score	Percentile	Raw score	Percentile	Raw score	Percentile
0-40	01	97	33	128	67
41-45	02	98	34	129	69
46-51	03	99	35	130	70
52-55	04	100	36	131	71
56-57	05	101	37	132	73
58-60	06	102	38	133	74
61-63	07	103	39	134	75
64-65	08	104	41	135	76
66	09	105	42	136	77
67-68	10	106	43	137	78
69-70	11	107	44	138	79
71-72	12	108	45	139	80
73-74	13	109	46	140	81
75	14	110	47	141	82
76	15	111	48	142	83
77	16	112	50	143-144	84
78-79	17	113	51	145	85
80	18	114	52	146-147	86
81	19	115	53	148	87
82-83	20	116	54	149	88
84	21	117	55	150	89
85	22	118	56	151	90
86	23	119	57	152	91
87-88	24	120	58	153	92
89	25	121	60	154	93
90	26	122	61	155-156	94
91	27	123	62	157-158	95
92	28	124	63	159-161	96
93	29	125	64	162-166	97
94	30	126	65	167-172	98
95	31	127	66	173-200	99
96	32				

Table B-2. Form 0 Navigator-Technical Composite Full Battery Conversion Table

Raw score	Percentile	Raw score	Percentile	Raw score	Percentile
0-58	01	127	34	163	67
59-67	02	128	35	164	68
68-72	03	129	36	165	69
73-76	04	130	37	166	70
77-79	05	131-132	38	167	71
80-81	06	133	39	168	72
82-83	07	134	40	169-170	73
84-86	08	135	41	171	74
87-89	09	136	42	172	75
90-91	10	137-138	43	173	76
92-93	11	139	44	174	77
94-95	12	140	45	175	78
96-97	13	141	46	176-177	79
98-99	14	142	47	178	80
100-101	15	143	48	179-180	81
102-103	16	144	49	181	82
104-105	17	145	50	182-183	83
106-107	18	146	51	184	84
108	19	147	52	185	85
109-110	20	148	53	186	86
111-112	21	149	54	187-188	87
113	22	150	55	189-190	88
114	23	151	56	191-192	89
115	24	152	57	193-194	90
116-117	25	153	58	195-196	91
118	26	154	59	197	92
119	27	155	60	198-199	93
120	28	156	61	200-202	94
121	29	157	62	203-205	95
122-123	30	158	63	206-209	96
124	31	159	64	210-214	97
125	32	160-161	65	215-218	98
126	33	162	66	219-257	99

Table B-3. Form O Academic Aptitude Composite Full Battery Conversion Table

Raw score	Percentile	Raw score	Percentile	Raw score	Percentile
0-25	01	70	29	98	69
26-30	02	71	31	99	70
31-34	03	72	33	100	71
35-36	04	73	34	101	72
37-39	05	74	35	102	75
40-41	06	75	36	103	76
42-43	07	76	37	104	78
44	08	77	38	105	79
45-47	09	78	40	106	80
48-49	10	79	41	107	81
50	11	80	43	108	82
51	12	81	44	109	83
52	13	82	45	110	84
53	14	83	47	111	85
54	15	84	49	112	86
55-56	16	85	50	113	87
57	17	86	51	114	88
58-59	18	87	52	115	89
60	19	88	53	116	90
61	20	89	54	117	91
62	21	90	57	118	92
63	22	91	59	119-120	93
64	23	92	61	121	94
65	24	93	62	122-123	95
66	25	94	63	124-125	96
67	26	95	65	126-127	97
68	27	96	67	128-129	98
69	28	97	68	130-140	99

* U.S. GOVERNMENT PRINTING OFFICE:1986- 659- 055/1770

Table B-4. Form O Verbal Composite Full Battery Conversion Table

Raw score	Percentile	Raw score	Percentile
0-11	01	41	41
12	02	42	44
13-14	03	43	46
15	04	44	48
16	05	45	50
17	06	46	53
18	07	47	55
19	08	48	57
20	09	49	60
21	10	50	62
22	11	51	64
23	12	52	67
24	13	53	69
25	14	54	72
26	15	55	74
27	17	56	77
28	18	57	78
29	19	58	81
30	21	59	84
31	23	60	86
32	24	61	87
33	26	62	90
34	27	63	92
35	30	64	93
36	32	65	96
37	33	66	97
38	36	67	98
39	38	68-71	99
40	40		

Table B-5. Form O Quantitative Composite Full Battery Conversion Table

Raw score	Percentile	Raw score	Percentile
0-12	01	41	54
13-14	02	42	57
15-16	03	43	59
17	04	44	61
18	05	45	64
19	06	46	66
20	08	47	69
21	09	48	71
22	10	49	75
23	11	50	76
24	14	51	78
25	15	52	80
26	17	53	82
27	19	54	85
28	21	55	86
29	24	56	88
30	26	57	90
31	28	58	91
32	31	59	92
33	33	60	93
34	34	61	94
35	38	62	95
36	41	63	96
37	43	64	97
38	45	65	98
39	48	66-69	99
40	52		