AFHRL-TR-68-104

July 1968

Development of Officer Selection

and Classification Tests--1968

AD 679939

By
Robert E. Miller

D D C

JAN 6 1969

B

PERSONNEL RESEARCH DIVISION
AIR FORCE HUMAN RESOURCES LABORATORY
AIR FORCE SYSTEMS COMMAND
Lackland Air Force Base, Texas

CLEARINGHOUSE

23

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.

ABSTRACT

In accordance with the normal two-year replacement cycle, new forms of the Air Force Officer Qualifying Test (AFOQT) and the AFROTC Pre-Enrollment Test (PET) were constructed for implementation in Fiscal Year 1968. Both tests are designated by their fiscal year of implementation. AFOQT-68 closely resembles the previous form in type of content, organization, and norming strategy. It yields Pilot, Navigator-Technical, Officer Quality, Verbal, and Quantitative composite scores. Standardization was accomplished with reference to the Project TALENT battery in a way which permits relating AFOQT scores to performance of Air Force Academy candidates and 12th grade males. A new feature of AFOQT-68 is the provision of separate norms for AFROTC and other use. These norms take into account the effects of differences in level of formal education at the time of testing in various commissioning programs. Differences in educational level are also provided for in the norms of PET-68. To facilitate test administration, this test is considerably shorter than the previous form but otherwise resembles it. PET-68 yields a total score based on verbal and quantitative items. It is intended as a screening device for AFROTC candidates.

DEVELOPMENT OF OFFICER SELECTION
AND CLASSIFICATION TESTS--1968

By
Robert E. Miller

FOREWORD

Officer selection and classification tests developed for implementation in Fiscal Year 1968 are the Air Force Officer Qualifying Test–68 and the AFROTC Pre-Enrollment Test–68. Development of these tests was accomplished under Project 7717, Selection, Classification, and Evaluation Procedures for Air Force Personnel; Task 771706, Selection and Classification Instruments for Officer Personnel Programs.

This report has been reviewed and is approved.

F. L. McLanathan, Lt Col, USAF
Chief, Personnel Research Division

TABLE OF CONTENTS

LIST OF TABLES

DEVELOPMENT OF OFFICER SELECTION AND CLASSIFICATION TESTS-1968

I. AFOQT-68

Background

The United States Air Force has relied on the Air Force Officer Qualifying Test (AFOQT) as the principal instrument for the selection and classification of officer personnel since 1953. A description of early forms of the AFOQT and several related tests has been presented elsewhere, along with a discussion of the origin and history of this testing program (Valentine & Creager, 1961). Recent forms of officer tests have been described in a series of reports (Miller & Valentine, 1964; Miller, 1966). The present report describes the officer selection and classification tests developed for implementation in Fiscal Year 1968.

The principal test for officer selection and classification continues to be the AFOQT. A new form of this test, known as AFOQT-68, became operational on 1 September 1967 in the AFROTC program and on 1 January 1968 in other Air Force programs. Current schedules call for development of new forms of this test on a two-year cycle.

Description

In type of content and organization, AFOQT-68 is identical with its immediate predecessor, AFOQT-66. The operations involved in construction and standardization of the new form are essentially those used for both AFOQT-64 and AFOQT-66. However, AFOQT-68 differs from previous forms in that two sets of normative data were provided for it. One set, published in the scoring manual, is appropriate for use with Officer Training School (OTS) applicants and all others who have completed or nearly completed college. The other set is for use in the AFROTC program, where testing ordinarily is accomplished in the freshman or sophomore year. Since it is known that formal education influences scores obtained on the AFOQT, the use of the appropriate norms which take this effect into account gives greater comparability of meaning to AFOQT percentile scores in the various commissioning programs.

The AFOQT-68 subtests and their organization into five operational aptitude composites are shown in Table 1. Following is a brief description of each subtest:

Quantitative Aptitude consists of items involving general mathematics, arithmetic reasoning, and interpretation of data presented in tables and graphs.

Verbal Aptitude consists of items involving verbal analogies, vocabulary, reading comprehension, and understanding of the background of current events.

Officer Biographical Inventory consists of items pertaining to experiences, preferences, and personality characteristics related to measures of officer effectiveness.

Scale Reading consists of items in which readings are to be taken of various scales and gauges, many of them calling for very fine discriminations.

Aerial Landmarks consists of pairs of photographs showing terrain as seen from different positions of an aircraft in flight. Landmarks indicated on one photograph are to be located on the other.

General Science consists of items pertaining to the basic principles of the physical sciences, with emphasis on physics.

Mechanical Information consists of items related to understanding of mechanics and knowledge of the function or operation of mechanical devices.

Mechanical Principles consists of drawings of complex apparatus and requires ability to determine the effects of prescribed operations of the apparatus.

Pilot Biographical Inventory consists of items pertaining to background experiences and interests related to measures of success in pilot training.

Aviation Information consists of semi-technical items concerned with types of aircraft, components of aircraft, and operation of aircraft.

Visualization of Maneuvers consists of pictorial items calling for identification of the attitude of an aircraft in flight after executing a verbally specified maneuver.

Instrument Comprehension consists of items similar to Visualization of Maneuvers except that the maneuvers are specified by readings of a compass and artificial horizon.

Stick and Rudder Orientation consists of sets of photographs of terrain as seen from an aircraft executing a maneuver. The proper movements of the control stick and rudder bar to accomplish this maneuver are to be indicated.

Table 1. Content and Organization of AFOQT-68[a]

		Composites				
Subtest	No. of Items	Pilot	Nav-Tech	Officer Quality	Verbal	Quantitative
Booklet 1 (AFPT 941)						
Quantitative Aptitude	60		X	X		X
Booklet 2 (AFPT 942)						
Verbal Aptitude	60			X	X	
Officer Biographical Inventory[b]	100			X		
Booklet 3 (AFPT 943)						
Scale Reading[c]	48		X			
Aerial Landmarks[c]	40		X			
General Science	24		X			
Booklet 4 (AFPT 944)						
Mechanical Information	24	X	X			
Mechanical Principles	24	X	X			
Booklet 5 (AFPT 945)						
Pilot Biographical Inventory	50	X				
Aviation Information	24	X				
Visualization of Maneuvers[c]	24	X				
Instrument Comprehension[c]	24	X				
Stick and Rudder Orientation[c]	24	X				
Total	526					

[a]Associated manuals are AFPT 939 and 940. Associated answer sheets are PRT 87, AFPT 946, and AFPT 947. Special manuals and answer forms are used in the AFROTC program. Scale Reading and Aerial Landmarks are scored R-W/4; Visualization of Maneuvers and Instrument Comprehension are scored R-W/3. Other subtests are scored rights only.

[b]Not administered to female applicants.

[c]Speeded subtests.

Item Selection

Data on the difficulty level and internal consistency of items selected for each subtest are presented in Table 2. The difficulty levels are proportions of an experimental sample choosing the correct response. The internal consistency data are phi coefficients based on the upper and lower 27 per cent of the sample. Item data for the biographical subtests are not included in the table. Analysis of their content suggests that their internal consistency is low and that "difficulty" does not have the same meaning for items of this type as for aptitude items. The other subtests show good internal consistency and a suitable spread of difficulty level about the desired median of .50.

Table 2. Item Difficulty Levels and Internal Consistency for AFOQT-68

Subtest	No. of Items	Difficulty Level		Internal Consistency	
		Range	Median	Range	Median
Quantitative Aptitude	60	.19 - .84	.54	.18 - .88	.54
Verbal Aptitude	60	.18 - .85	.53	.16 - .79	.48
Scale Reading	48	.26 - 92	.58	.15 - .81	.40
Aerial Landmarks	40	.25 - .78	.52	.30 - .74	.53
General Science	24	.23 - .82	.52	.32 - .75	.55
Mechanical Information	24	.28 - .84	.52	.26 - .81	.50
Mechanical Principles	24	.20 - .83	.51	.16 - .64	.48
Aviation Information	24	.20 - .85	.48	.30 - .75	.47
Visualization of Maneuvers	24	.20 - .83	.60	.21 - .65	.40
Instrument Comprehension	24	.24 - .76	.57	.28 - .73	.59
Stick and Rudder Orientation	24	.40 - .84	.60	.10 - .68	.48

Reliability and Intercorrelations

It was necessary to estimate reliability and intercorrelations from AFOQT-66 data pending the accumulation of operational data for AFOQT-68. Normative samples for AFOQT-68 do not contain the same distribution of aptitudes as the population for which the test was intended and hence will give misleading results. The estimated data for AFOQT-68 are presented in Tables 3 and 4. While they are best regarded as approximations, the data suggest a high degree of stability and consistency for AFOQT and reasonable intercorrelations in view of the partially overlapping content of the composites. The subtest reliabilities were obtained by Kuder-Richardson Formula 20 or test-retest, depending on whether the subtest was speeded. Composite reliabilities were computed from subtest reliabilities by application of the formula for the reliability of a composite (Wherry & Gaylord, 1943, p. 250).

Table 3. **Reliability Estimates for AFOQT-68**

Subtest or Composite	Reliability	Subtest or Composite	Reliability
Quantitative Aptitude	.93[a]	Visualization of Maneuvers	.77[b]
Verbal Aptitude	.89[a]	Instrument Comprehension	.65[b]
Scale Reading	.85[b]	Stick and Rudder Orientation	.73[b]
Aerial Landmarks	.68[b]	Pilot Composite (without Pilot Biographical Inventory)	.91[c]
General Science	.81[a]	Navigator-Technical Composite	.95[c]
Mechanical Information	.75[a]	Officer Quality Composite (without Officer Biographical Inventory)	.94[c]
Mechanical Principles	.75[a]	Verbal Composite	.89[a]
Aviation Information	.78[a]	Quantitative Composite	93[a]

[a]Determined by Kuder-Richardson Formula 20.

[b]Determined by test-retest.

[c]Determined by Wherry and Gaylord formula.

Table 4. **Estimated Intercorrelation of Subtests and Composites for AFOQT-68, Based on AFOQT-66**

(N = 450)

Subtest or Composite	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Quantitative Aptitude															
2. Verbal Aptitude	.55														
3. Officer Biographical Inventory	.12	.11													
4. Scale Reading	.65	.40	.15												
5. Aerial Landmarks	.38	.27	.06	.42											
6. General Science	.64	.59	.06	.47	.29										
7. Mechanical Information	.39	.39	.04	.32	.29	.57									
8. Mechanical Principles	.52	.38	.01	.42	.36	.58	.62								
9. Pilot Biographical Inventory	.04	-.05	.17	.05	.17	.14	.36	.24							
10. Aviation Information	.20	.41	.10	.23	.16	.46	.47	.37	.29						
11. Visualization of Maneuvers	.41	.41	.15	.42	.45	.45	.39	.41	.22	.35					
12. Instrument Comprehension	.39	.25	.04	.38	.41	.34	.31	.38	.23	.34	.51				
13. Stick and Rudder Orientation	.42	.35	.02	.39	.44	.44	.38	.39	.16	.31	.45	.53			
14. Pilot Composite	.50	.43	.09	.46	.50	.59	.65	.63	.48	.58	.66	.71	.85		
15. Navigator-Technical Composite	.87	.57	.11	.76	.63	.75	.62	.71	.18	.36	.55	.50	.55	.70	
16. Officer Quality Composite	.85	.80	.47	.60	.35	.64	.41	.47	.06	.33	.47	.35	.40	.50	.79

Standardization

The norming strategy for AFOQT-68 was generally the same as that used for AFOQT-64 and AFOQT-66. It is based on the previous administration of an old form of the AFOQT and tests of the Project TALENT battery to large groups of basic airmen who had been stratified by deciles on the Armed Forces Qualifying Test (AFQT). This operation permitted development of weighted composites of TALENT tests for the optimal prediction of each AFOQT composite. Thereafter, the process of standardizing a new form of the AFOQT has consisted of administering the composites of the new form and the appropriate TALENT composites to basic airmen stratified as in the samples from which the TALENT composites were developed. Equipercentile conversions from percentile norms of the old AFOQT form through the old and new TALENT composite distributions to the composite distributions of the new AFOQT form completes the process. This approach has been described in detail elsewhere (Dailey, Shaycoft, & Orr, 1962), and its application to AFOQT-64 and AFOQT-66 has been reported (Miller & Valentine, 1964; Miller, 1966).

Because of this complex strategy, the markedly different distributions of scores for airmen and officers on officer tests does not lead to inappropriate norms. Each new form is still indirectly tied to an old form which was normed on an Air Force Academy candidate population. The strategy also permits a constancy of meaning of percentile scores for a given AFOQT composite across the various forms of the test which were standardized in this way. In addition, it is possible to relate AFOQT scores to the 12th grade male sample tested in the original Project TALENT study of over 400,000 high school students.

The above procedure was applied in a straightforward way to the norming of each AFOQT-68 composite with the exception of the Officer Quality composite. This composite has tended to present special problems because of its large biographical component. In AFOQT-64 and AFOQT-66 this component was omitted from the norming process until, as a final step, its mean in an AFROTC sample was added as a constant to results obtained through a modification of the TALENT Officer Quality composite. For AFOQT-68 a variant of this procedure was used in which the modified TALENT Officer Quality composite was replaced by the so-called TALENT Academic composite. This composite was developed in a supplement to the study in which the other TALENT composites were defined. The TALENT Academic composite is intended as an optimal predictor of the AFOQT Officer Quality composite without the Officer Biographical Inventory. By using the TALENT Academic composite in the same manner as the other TALENT composites, an AFOQT-68 Academic composite distribution was obtained and partitioned by percentile values. This constituted the Female Officer Quality composite. The addition to this distribution of a constant representing the mean Officer Biographical Inventory score in an AFROTC sample yielded the Male Officer Quality composite. The underlying assumption that the Officer Biographical Inventory correlates zero with the Officer Quality composite minus the Officer Biographical Inventory appeared to be well met. The actual correlation for AFOQT-66 in a sample of 440 Officer Training School students was .03.

The composition of the several TALENT composites used in standardizing AFOQT-68 is shown in Table 5. The integral weights applied to each component test to yield the optimal prediction of the corresponding AFOQT composites are also shown. These weights approximate the raw score regression weights computed during the development of the TALENT composites. The component tests to which the weights are applied are all from the original Project TALENT battery.

Norm tables constructed in the manner described above yield percentile scores having the same meaning as the corresponding scores on earlier forms of the AFOQT. However, it is known that formal education has an effect on percentile scores such that the meaning of a given score may be different in the several commissioning programs where the testing is done at different educational levels. Thus, AFROTC cadets are usually tested relatively early in college, while OTS students are tested upon graduation from college or just prior to graduation. One result is a largely spurious difference in score distributions for the two programs. It was not feasible with previous forms of the AFOQT to take educational effects on test scores into account in the norming process because no adequate sample was available in which these effects could be measured.

Table 5. Composition of TALENT Composites Corresponding to AFOQT-68 Composites[a]

Composite	TALENT Component	Weight
Pilot	110 Aeronautics and Space (Info)	3
	111 Electricity and Electronics (Info)	1
	112 Mechanics (Info)	3
	270 Mechanical Reasoning	3
	281 Visualization in Two Dimensions	1
	282 Visualization in Three Dimensions	2
	333 Mathematics III: Advanced	2
Navigator-Technical	106 Mathematics (Info)	3
	111 Electricity and Electronics (Info)	2
	270 Mechanical Reasoning	3
	282 Visualization in Three Dimensions	3
	312 Mathematics II: Introductory	3
Academic	106 Mathematics (Info)	2
	110 Aeronautics and Space (Info)	2
	250 Reading Comprehension	1
	312 Mathematics II: Indtroductory	2
	333 Mathematics III: Advanced	3
Verbal	102 Vocabulary (Info)	2
	103 Literature (Info)	2
	106 Mathematics (Info)	2
	110 Aeronautics and Space (Info)	3
	250 Reading Comprehension	1
Quantitative	106 Mathematics (Info)	2
	312 Mathematics II: Introductory	2
	333 Mathematics III: Advanced	3

[a]Data assembled from Dailey et al., 1962, and unpublished supplement.

During the construction of AFOQT-68, data which permitted this measurement became available in a usable form through the Department of Defense Officer Record Examination program. In this program, junior officers in each military service are tested for research purposes with an instrument having verbal and quantitative items essentially similar to those of the AFOQT. By using Officer Record Examination scores in effect as control variables, AFOQT scores in AFROTC and OTS samples could be equated in terms of actual level of aptitude. As a result, it became possible to construct two sets of conversion tables for AFOQT-68. The set constructed without adjustment for educational effects was considered most suitable for the AFROTC and AECP programs, where educational effects are minimal because of the educational level at which testing generally occurs. This set of tables was provided to Headquarters AFROTC for use with their centralized scoring operations. An additional set, containing a correction for educational effects, was published in the AFOQT-68 *Scoring Instructions*, AFPT 940. This is considered the appropriate set for OTS and most other programs where the AFOQT is used.

The values of the raw TALENT composites corresponding to AFOQT percentiles in the AFROTC conversion tables are shown in Table 6. These raw distributions were derived from the study initially defining the TALENT composites and have been used in norming subsequent AFOQT forms. The distributions of TALENT raw composite scores corresponding to AFOQT percentiles in the conversion tables where educational effects are taken into account are shown in Table 7. These were obtained by equipercentile conversion to the TALENT raw score distributions from AFOQT Form G norm tables into which the educational correction had been incorporated. Data in Tables 6 and 7 have been smoothed slightly to eliminate irregularities peculiar to the samples on which they were developed. These tables can be used directly in norming future forms of the AFOQT without recourse to Officer Record Examination scores.

In standardizing any form of the AFOQT it would be desirable to use a single sample of approximately 1,000 examinees from which all data for each composite could be obtained. However, this would impose a prohibitively lengthy testing schedule on the examinees. As a feasible alternative, three separate samples, each containing about 1,000 cases, were used for AFOQT-68. One sample was used for norming the Pilot composite, one for the Navigator-Technical composite, and one for the remaining three composites. Though each sample was stratified by AFQT decile, some differences among the three AFQT distributions existed. To insure that these differences did not change the meaning of a given AFOQT percentile score from one composite to another in the same form of the test, the AFQT distributions of the three samples were tested against each other by chi-square. The results are shown in Table 8. None of the chi-square values are significant. The conclusion is that the three samples contain the same distributions of aptitudes measured by the AFQT.

Table 9 shows the cumulative percentage distriubutions of the principal TALENT composites in AFOQT-68 samples for the AFROTC norms and in the sample from which the TALENT composites were originally derived. Considerable similarity of the distributions for a given composite persists, although the AFOQT-68 samples tend to show somewhat higher aptitudes on each composite than the original sample. The small percentage of cases near the top of the distributions results in some instability in the placement of the upper partition values of the percentile scale in the final conversion tables. The percentages at levels where minimum qualifying scores are set, however, are large enough for reasonable stability. No current minimum for any program is higher than the 60th percentile; most are lower.

The marked skewness seen in all distributions in Table 9 does not produce distortions in standardization because these distributions are used only indirectly in norming the AFOQT. Raw score means and standard deviations, however, will differ considerably in these samples from the corresponding statistics in operational samples. The differences are illustrated in Table 10, which shows raw score means and standard deviations of AFOQT-68 composites for the normative samples, a 12th grade male sample, and two populations of student officers for which AFOQT-68 is an appropriate test. Except for the normative samples, all the data are estimates from AFOQT-68 conversion tables or unpublished tables developed as part of the study defining the TALENT composites. Differences between the four groups are generally in the direction expected on the basis of the probable distribution of aptitudes in such groups.

Table 6. TALENT Composite Raw Scores Corresponding to AFOQT-68 Percentiles in AFROTC Conversion Tables

AFOQT Percentile	TALENT Composites				
	Pilot	Navigator-Technical	Academic	Verbal	Quantitative
95	223 and above	270 and above	194 and above	205 and above	129 and above
90	218-222	266-269	188-193	198-204	125-128
85	214-217	262-265	185-187	194-197	123-124
80	210-213	258-261	182-184	190-193	121-122
75	206-209	254-257	179-181	186-189	119-120
70	202-205	250-253	176-178	182-185	117-118
65	198-201	246-249	173-175	178-181	115-116
60	194-197	242-245	170-172	176-177	113-114
55	192-193	239-241	167-169	174-175	111-112
50	190-191	236-238	164-166	172-173	109-110
45	186-189	232-235	161-163	169-171	106-108
40	182-185	227-231	157-160	165-168	102-105
35	178-181	221-226	153-156	161-164	98-101
30	174-177	214-220	149-152	157-160	94-97
25	170-173	207-213	143-148	153-156	89-93
20	165-169	198-206	136-142	149-152	84-88
15	158-164	189-197	127-135	143-148	79-83
10	150-157	176-188	117-126	134-142	74-78
05	135-149	156-175	101-116	119-133	58-73
01	134 and below	155 and below	100 and below	118 and below	57 and below

Table 7. TALENT Composite Raw Scores Corresponding to AFOQT-68 Percentiles in Conversion Tables for OTS and Other Programs

AFOQT Percentile	TALENT Composites				
	Pilot	Navigator-Technical	Academic	Verbal	Quantitative
95	247 and above	277 and above	196 and above	210 and above	132 and above
90	243-246	271-276	193-195	206-209	129-131
85	239-242	269-270	191-192	202-205	126-128
80	235-238	267-268	189-190	198-201	125
75	231-234	265-266	187-188	195-197	124
70	227-230	263-264	185-186	193-194	123
65	223-226	261-262	183-184	191-192	122
60	219-222	259-260	181-182	189-190	121
55	215-218	257-258	179-180	187-188	120
50	211-214	255-256	178	185-186	119
45	207-210	253-254	177	183-184	118
40	204-206	250-252	175-176	181-182	117
35	201-203	247-249	173-174	179-180	116
30	198-200	243-246	170-172	177-178	115
25	194-197	239-242	167-169	174-176	111-114
20	190-193	234-238	163-166	170-173	107-110
15	186-189	225-233	155-162	166-169	102-106
10	182-185	212-224	146-154	158-165	94-101
05	178-181	197-211	136-145	149-157	86-93
01	177 and below	196 and below	135 and below	148 and below	85 and below

Table 8. Homogeneity of AFOQT-68 Normative Samples
with Respect to AFQT Deciles

Samples Compared	Chi-Square	df	P
Pilot and Navigator-Technical	8.21	7	>.30
Pilot and Officer Quality	5.50	7	>.50
Navigator-Technical and Officer Qaulity	4.11	7	>.70

Table 9. Cumulative Percentage Distributions for TALENT Composites in Original
Air Force TALENT Sample and AFOQT-68 Normative Samples

*(AF TALENT Sample N = 2,489; Pilot normative Sample N = 1,041; Navigator-Technical
normative Sample N = 910; Officer Quality normative Sample N = 998)*

AFOQT Percentile	TALENT Composites					
	Pilot		Navigator-Technical		Academic	
	AF TALENT	Normative	AF TALENT	Normative	AF TALENT	Normative
95	0.8	1.0	0.6	0.8	0.1	0.3
90	1.6	2.1	0.9	1.0	0.4	0.5
85	2.5	3.1	1.3	2.1	0.5	0.8
80	3.3	5.2	1.6	2.6	0.6	1.0
75	4.4	6.5	2.0	3.6	0.9	1.3
70	6.1	8.3	2.7	4.2	1.3	2.1
65	7.4	10.7	3.2	4.4	1.7	2.6
60	9.2	12.8	3.8	5.2	2.2	3.6
55	10.4	13.3	4.5	6.0	2.7	4.9
50	11.6	15.1	5.3	6.8	3.4	5.5
45	13.3	16.7	6.2	7.8	3.8	5.7
40	15.5	19.3	7.3	9.1	4.5	6.2
35	17.7	21.6	8.3	10.2	5.4	8.1
30	21.2	25.0	10.2	12.0	6.6	9.9
25	25.1	27.3	12.4	14.1	8.2	13.3
20	29.3	32.8	15.2	16.4	10.5	15.9
15	34.8	39.6	18.6	19.5	13.6	20.3
10	42.5	47.9	23.5	27.6	18.4	26.0
05	56.6	62.8	34.4	40.9	29.5	41.1
01	100.0	100.0	100.0	100.0	100.0	100.0

Table 10. Raw Score Means and Standard Deviations pf AFOQT-68
Composites for Four Groups

Composite	Normative Samples [a]		12th Grade Male Sample [b]		AFROTC Population [c]		OTS Population [c]	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Pilot	72.2	24.9	64.9	28.1	95.5	24.8	115.5	22.4
Navigator-Technical	66.9	21.9	67.7	25.6	101.5	22.0	115.5	20.4
Officer Quality	69.6	15.3	72.0	23.3	103.5	15.4	114.5	13.6
Verbal	21.8	9.3	19.6	13.1	34.5	11.4	40.5	8.6
Quantitative	17.3	7.9	18.2	9.5	33.5	9.6	39.5	6.8

[a]Ns vary from 910 to 1,041.

[b]Data estimated from unpublished tables by Dailey et al., based on 4% subsample of 12th grade males in original Project TALENT study.

[c]Data estimated from AFOQT-68 conversion tables.

II. AFROTC PET-68

Background

In 1965 a test known as the AFROTC Pre-Enrollment Test (PET) was developed and implemented in compliance with a requirement from Headquarters USAF. This test was intended as a short screening device for freshman applicants for the AFROTC program. The initial form was an adaptation of the earlier and no longer operational Air Force Precommission Screening Test-62. Both the earlier test and the adaptation have been described elsewhere (Valentine, 1961; Miller, 1966). The adaptation was designated as the AFROTC Pre-Enrollment Test-66.

A two-year revision cycle was established to coincide with that of the AFOQT. Consequently, PET-66 was replaced during the summer of 1967 by a new form known as the AFROTC PET-68. The official implementation date for the new form was 1 September 1967, but the test was administered in a few AFROTC detachments beginning in July.

Description

PET-68 consists of Verbal and Quantitative subtests, each containing items of the same types as those in the corresponding AFOQTsubtests. A single total score, based on right responses only, is obtained. The test manual and booklet bear AFPT numbers 650 and 651, respectively, and both are dated 1 September 1967. A standard PRT 81 answer sheet is used.

Because of scheduling difficulties encountered in the AFROTC detachments with PET-66, it was decided to shorten PET-68 sufficiently to permit the test to be administered within a normal class period of 50 minutes. To accomplish this, the number of items was reduced to 20 verbal and 20 quantitative.

Two sets of norms were developed and submitted to Headquarters AFROTC. Since the total score on the PET is in effect a predicted score on the AFOQT Officer Quality composite, it was desirable to take into account the effects of formal education on AFOQT scores in norming PET. One set of norms is for use in detachments where the testing schedule calls for administration of the two tests within a period of less than six months. The other set is for use where the tests are administered more than six months apart and significant additional effects of education can therefore be anticipated. PET-68 is normally administered at the beginning of the freshman year. Most detachments adminster the AFOQT later in the same year, but no uniform schedule is imposed on the detachments.

Item Selection

Criteria for item selection were the same as for AFOQT-68. The item difficulties and internal consistency, defined as in Table 2. are presented for PET-68 in Table 11. The data are highly similar to those for the verbal and quantitative items of AFOQT-68. Again they represent a satisfactory degree of internal consistency and a sufficient spread of difficulty about the desired median of .50.

Reliability and Intercorrelations

Reliability data for the PET-68 subtests and total score were estimated from the PET-66 reliability data by use of the Spearman-Brown formula. Results are shown in Table 12. The level of reliability is satisfactory but somewhat lower than that for PET-66 because of the reduction in test length. The estimated correlations of PET-68 scores with relevant composites of AFOQT-68 are shown in Table 13. These correlations also are slightly lowered as a result of shortening the PET. They may be interpreted as concurrent validity coefficients.

Standardization

The AFROTC PET-68 was standardized by approximately the same process used for the Officer Quality composite without the Officer Biographical Inventory in AFOQT-68. Percentile ranges in the TALENT raw Academic composite were established for the PET-68 normative sample with reference to their location in the TALENT Academic composite for AFOQT in the case without educational effects. By equipercentile conversion, the corresponding percentiles in the distribution of PET raw total scores were located. This process yielded percentiles for PET-68 having the same meaning as those for the Officer Quality composite of AFOQT-68. Since no educational effects enter the standardization process by this method, the resulting conversion table was considered suitable for the case where a relatively short time intervenes between administration of the PET and the AFOQT.

A separate norm table was provided for use where a longer period intervenes between test administrations. To construct this second table, medians of the Officer Quality distributions by month of testing in the AFROTC program were examined separately and in various combinations to observe the amount of increase as a result of educational effects. It was assumed that increases are due only to educational effects and not to selection policies or other characteristics of colleges hosting AFROTC detachments. The large number of detachments contributing to score distributions for most months lends credibility to this otherwise dubious assumption.

It appeared that inversions could be eliminated and an increase of about five points on the percentile scale shown if the second set of norms were made to apply where six or more months intervene between PET and AFOQT testing dates. Utilizing this known increase in Officer Quality scores after six or more months of college, special Officer Quality percentile norms were constructed from the operational AFROTC norms by linear interpolation. These special norms were studied in relation to the operational Officer Quality norms for the AFROTC program and the PET-68 raw scores in the normative sample. The relationships were such that a PET-68 raw score corresponding to a given raw score in these special norms is 1 to 2 points lower than in the norms without educational effects. The initial set of PET-68 norms was adjusted accordingly. The results agree with the expectation that the minimum qualifying raw score on the PET should be lower when the interval between tests is longer in order that educational effects can operate significantly to increase Officer Quality aptitude before the AFOQT is administered.

The cumulative percentage distributions for the TALENT Academic composite in the stratified basic airman norming sample for PET-68 and the original Air Force TALENT sample are shown in Table 14. These data correspond to those for AFOQT-68 in Table 9. They indicate a somewhat higher aptitude level in the PET-68 sample than in the original Air Force TALENT sample despite the stratification of both samples. However, the aptitude levels of the PET-68 sample and the Officer Quality sample in Table 9 are very similar by inspection. To resolve the question of possible significant differences, the distributions of AFQT scores in the two samples were compared by chi-square in the manner used in comparing AFOQT 68 norming samples. The results are presented in Table 15. No significant difference was found.

11

Raw total score means and standard deviations for the PET-68 normative sample, a 12th grade male sample, and a student officer population are shown in Table 16. Differences in the distribution statistics for these groups are generally in the expected direction. Data from the normative sample were computed. The other data were estimated from the PET-68 norm table for the interval of less than six months or from unpublished tables generated in the original Air Force TALENT study. Early results of using the PET-68 in the AFROTC program suggest that the actual AFROTC mean is somewhat lower than the estimated mean.

Table 11. Item Difficulty Levels and Internal Consistency for PET-68

Subtest	No. of Items	Difficulty Level		Internal Consistency	
		Range	Median	Range	Median
Verbal	20	.20 - .84	.52	.20 - .70	.48
Quantitative	20	.24 - .84	.52	.12 - .81	.56

Table 12. Reliability Estimates for PET-68

Subtest or Composite	Reliability
Verbal	.73[a]
Quantitative	.82[a]
Total	.82[a]

[a]Estimated by Spearman-Brown formula.

Table 13. Estimated Correlation between PET-68 and AFOQT-68
(N = 387)

Subtest or composite	1	2	3	4	5
1. PET-68 Verbal					
2. PET-68 Quantitative	.26				
3. PET-68 Total	.72	.78			
4. AFOQT-68 Quantitative Composite	.28	.82	.72		
5. AFOQT-68 Verbal Composite	.73	.29	.68	.32	
6. AFOQT-68 Quantitative plus Verbal	.63	.73	.82	.85	.77

12

Table 14. Cumulative Percentage Distributions for TALENT Academic Composite in Original Air Force TALENT Sample and PET-68 Normative Sample

(AF TALENT sample N = 2,489; PET normative sample N = 1,024)

AFOQT Percentile	TALENT Academic Composite	
	AF TALENT	Normative
95	0.1	03
90	0.4	0.8
85	0.5	1.3
80	0.6	1.6
75	0.9	1.8
70	1.3	2.1
65	1.7	2.3
60	2.2	3.4
55	2.7	4.2
50	3.4	4.7
45	3.8	6.0
40	4.5	7.6
35	5.4	9.1
30	6.6	10.9
25	8.2	13.0
20	10.5	15.1
15	13.(	22.9
10	18.4	29.4
05	29.5	40.4
01	100.0	100.0

Table 15. Homogeneity of Officer Quality and PET-68 Normative Samples with Respect to AFQT Deciles

Samples Compared	Chi-Square	df	P
PET-68 and AFOQT-68 Officer Quality	4.51	7	>.70

Table 16. Raw Score Means and Standard Deviations of PET-68 for Three Groups

Score	Normative Sample[a]		12th Grade Male Sample[b]		AFROTC Population[c]	
	Mean	SD	Mean	SD	Mean	SD
Total	15.1	5.6	14.0	7.9	25.5	6.8

[a]N = 1,024

[b]Data estimated from unpublished tables by Dailey et al., based on 4% subsample of 12th grade males in original Project TALENT study.

[c]Data estimated from PET-68 conversion tables. Early empirical results suggest that the mean is somewhat overestimated.

III. SUMMARY

New replacement forms of the AFOQT and AFROTC PET were constructed for implementation in Fiscal Year 1968. These tests are designated as the AFOQT-68 and AFROTC PET-68. Both are similar to their immediate predecessors in terms of type of content and organization and in norming strategy. The principal new feature of AFOQT-68 is the provision of separate conversion tables for programs where different amounts of formal education at the time of testing can be anticipated. PET-68 is provided with separate conversion tables for examinees differing in the amount of formal education received between administration of the PET and the AFOQT. To make test administration more feasible in the AFROTC detachments, PET-68 was considerably shortened from the previous form.

REFERENCES

Dailey, J.T., Shaycoft, Marion F., & Orr. D.B. *Calibration of Air Force Selection tests to Project TALENT norms.* PRL-TDR-62-6, AD-285 185. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, May 1962.

Miller, R.E. *Development of officer selection and classification tests-1966.* PRL-TR-66-5, AD-639 237. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, June 1966.

Miller, R.E. & Valentine, L.D., Jr. *Development and standardization of the Air Force Officer Qualifying Test-64.* PRL-TDR-64-6, AD-600 782. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, March 1964.

Valentine, L.D., Jr. *Development of the Air Force Precommission Screening Test-62.* ASD-TN-61-146, AD-269 527. Lackland AFB, Tex.. Personnel Laboratory, Aeronautical Systems Division, October 1961.

Valentine, L.D., Jr. & Creager, J.A. *Officer selection and classification tests. their development and use.* ASD-TN-61-145, AD-269 827. Lackland AFB, Tex.: Personnel Laboratory, Aeronautical Systems Division, October 1961.

Wherry, R.J. & Gaylord, R.H. The concept of test and item reliability in relation to factor pattern. *Psychometrika,* 1943, 8, 247-264.

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified

1 ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Personnel Research Division Lackland Air Force Base, Texas 78236	2b GROUP

3 REPORT TITLE

DEVELOPMENT OF OFFICER SELECTION AND CLASSIFICATION TESTS—1968

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

5 AUTHOR(S) (First name, middle initial, last name)

Miller, R.E.

6 REPORT DATE	7a. TOTAL NO OF PAGES	7b NO OF REFS
July 1968	14	6

8a CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO 7717	AFHRL-TR-68-104
c Task 771706 d.	9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10 DISTRIBUTION STATEMENT

This document has been approved for public release and sale, its distribution is unlimited.

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY
	Personnel Research Division Lackland AFB, Texas 78236

13 ABSTRACT

In accordance with the normal two year replacement cycle, new forms of the Air Force Officer Qualifying Test (AFOQT) and the AFROTC Pre enrollment Test (PET) were constructed for implementation in Fiscal Year 1968. Both tests are designated by their fiscal year of implementation. AFOQT-68 closely resembles the previous form in type of content, organization, and norming strategy. It yields Pilot, Navigator-Technical, Officer Quality, Verbal, and Quantitative composite scores. Standardization was accomplished with reference to the Project TALENT battery in a way which permits relating AFOQT scores to performance of Air Force Academy candidates and 12th grade males. A new feature of AFOQT-68 is the provision of separate norms for AFROTC and other use. These norms take into account the effects of differences in level of formal education at the time of testing in various commissioning programs. Differences in educational level are also provided for in the norms of PET-68. To facilitate test administration, this test is considerably shorter than the previous form but otherwise resembles it. PET-68 yields a total score based on verbal and quantitative items. It is intended as a screening device for AFROTC candidates.

DD FORM 1 NOV 65 1473

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
selection tests						
classification tests						
screening tests						
officer selection						
test construction						
test standardization						
Air Force Officer Qualifying Test						
AFROTC Pre-Enrollment Test						
project TALENT						