DOCUMENT RESUME

ED 166 241 TM 008 254

AUTHOR Gould, R. Bruce
TITLE Air Force Officer Qualifying Test Form N: Development and Standardization. Final Report for Period March 1974 - March 1978.
INSTITUTION Air Force Human Resources Lab., Brooks AFB, Texas.
REPORT NO AFHRL-TR-78-43
PUB DATE Aug 78
NOTE 24p.
AVAILABLE FROM Superintendent of Documents, United States Government Printing Office, Washington, D.C. 20402 (Stock Number 771-122/57)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Aircraft Pilots; Item Analysis; *Norms; *Occupational Tests; *Officer Personnel; Personnel Selection; Scores; *Test Construction; Test Reliability; Test Validity
IDENTIFIERS *Air Force Officer Qualifying Test

ABSTRACT

The construction and norming of Form N of the Air Force Officer Qualifying Test (AFOQT) is described. The new form serves the same purpose as its predecessor and possesses basically the same characteristics. References are made to the research which provided the basis for most of the changes. Other changes were made because of the admission of women to the Air Force. The test battery yields five composite scores: Pilot, Navigator-Technical, Officer Quality, Verbal, and Quantitative. Two sets of conversion tables are provided for examinees' scores according to educational level--two or more years of college, and less than two years of college. Standardization was accomplished by administering the test to samples of examinees from all major sources for commissioned personnel in the Air Force as well as to a sample of second lieutenants. Conversion tables for each raw score composite to percentile ranks are given, as are score distribution summary data and distribution curves. Summaries of item difficulty data and test reliabilities are given. (Author/CTM)

AFHRL-TR-78-43

AIR FORCE HUMAN RESOURCES LABORATORY

AIR FORCE OFFICER QUALIFYING TEST FORM N: DEVELOPMENT AND STANDARDIZATION

By

R. Bruce Gould

PERSONNEL RESEARCH DIVISION
Brooks Air Force Base, Texas 78235

August 1978
Final Report for Period March 1974 – March 1978

Approved for public release; distribution unlimited.

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When U.S. 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, 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 for publication.

LELAND D. BROKAW, Technical Director
Personnel Research Division

RONALD W. TERRY, 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-78-43	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AIR FORCE OFFICER QUALIFYING TEST FORM N: DEVELOPMENT AND STANDARDIZATION		5. TYPE OF REPORT & PERIOD COVERED Final March 1974 March 1978 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R. Bruce Gould		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Personnel Research Division Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62703F 77191212
11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235		12. REPORT DATE August 1978 13. NUMBER OF PAGES 22
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

SM Study Nrs. 6250 and 6817

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

Air Force Officer Qualifying Test	item difficulty	test construction
classification tests	officer selection and classification	test standardization
educational effects	reliability	
intercorrelations	score distributions	
internal consistency	selection tests	

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

Air Force Officer Qualifying Test (AFOQT) Form N was constructed as a replacement for AFOQT Form M in Fiscal Year 1978. The new form serves the same purpose as its predecessor and possesses basically the same characteristics. It yields Pilot, Navigator-Technical, Officer Quality, Verbal, and Quantitative composite scores. Two sets of conversion tables are provided for examinees' scores according to educational level. Standardization was accomplished by test administration to samples of examinees from all major sources for commissions in the Air Force and development of percentile conversion tables. Basic airmen with aptitude at or above the 50th percentile of the general population, Officer Training School candidates, Air Force Reserve Officers Training Corps students, and Air Force Academy Cadets composed the majority of the 2,681 cases in the standardization sample. Some 287

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:

second lieutenants were also included because of the substantial number from this population who also take the AFOQT when applying for admission to special programs such as pilot or navigator training.

PREFACE

Replacement forms of the Air Force Officer Qualifying Test are produced on a triennial cycle. The latest form of this test was produced under Project 7719, Selection and Classification Technology; Task 771912, Selection and Classification Instruments for Officer Personnel Programs.

Development of AFOQT Form N was begun by the late Dr. Robert E. Miller and completed with the assistance of Mrs. Nancy Thompson and Mr. Cecil Cannon. Senior Airmen Stan Prescott and Wayne Flikke of the Computational Sciences Division provided highly competent computer programming support for the project.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

List of Tables *(Continued)*

AIR FORCE OFFICER QUALIFYING TEST FORM N: DEVELOPMENT AND STANDARDIZATION

I. INTRODUCTION

In 1951, a selected group of paper-and-pencil subtests from the World War II aircrew classification batteries were combined with an aptitude test called the Aviation-Cadet Officer-Candidate Qualifying Test. The result was a new operational instrument known as the Air Force Officer Qualifying Test (AFOQT). In 1953, the USAF Officer Activity Inventory, the Attitude Survey, and the Information Inventory were added to the AFOQT. Five aptitude composites (Officer Quality, Observer-Technical, Pilot, Verbal, and Quantitative) were derived from the battery in much the same fashion as the current test. This test has remained the basis of the Air Force officer selection and classification testing program down to the present. During its 27 years of use, 15 different forms of the test were constructed, and from time to time, other tests were derived from it to meet special needs. The entire history of this effort has been documented (Miller, 1966, 1968, 1970, 1972, 1974; Miller & Valentine, 1964; Valentine & Creager, 1961). Extensive technical data pertaining to the AFOQT have been summarized in reports on interpretation and use of AFOQT scores (Hunter & Thompson, 1978; Mathews, 1977; Miller, 1969; Valentine, 1977).

The AFOQT is used to select candidates for most programs leading to a line officer commission, with the Air Force Academy (AFA) the only major exception. It is also used to select candidates for undergraduate pilot and navigator training and to assist in assigning nonflying officers entering their initial tour of active duty. Under current production schedules, each sequential form of the AFOQT serves these functions for the Air Force throughout a 3-year cycle. In accordance with this cycle, AFOQT Form N was scheduled for introduction in the Air Force Reserve Officer's Training Corps (AFROTC) commissioning program on 1 July 1978, approximately coinciding with the beginning of a new academic year, and in all other programs on 1 April 1978.

II. GENERAL CHARACTERISTICS

AFOQT Form N was constructed according to the same general plan as all its recent predecessors. It consists of 606 test items organized into 18 subtests from which five composite scores are derived. These are the Pilot, Navigator-Technical, Officer Quality, Verbal, and Quantitative composites. These composites are used individually in various officer personnel programs. The composition of the test is shown in Table 1.

Form N consists of four test booklets accompanied by administrative, scoring, and interpretive manuals, a set of 10 hand-scoring keys, and three special Digitek answer sheets. The answer sheets and interpretive manual are specific to Form N. The scoring manual contains two sets of tables for converting raw scores to percentiles according to the educational level of the examinee. The educational level in the various programs where the test is used varies from college freshman to college graduate. The use of separate conversion tables for different levels is supported by two studies (Gregg, 1968; Tupes & Miller, 1969) and is retested using the normative data for Form N which provide quantitative evaluation of the elevating effect of education on AFOQT scores.

Recent predecessors of Form N consisted of 13 subtests. One subtest, Officer Biographical Inventory, had been administered only to males; therefore, separate male and female composite percentile conversion tables were used. The admission of females to traditional male career areas, including pilot and navigator specialties, and the emphasis on equal treatment of males and females mandated the removal of sexist orientations in AFOQT tests. The availability of data from on-going pilot and navigator selection validation studies provided the means for modification of subtests, and items within subtests, comprising the five major composite scores.

Table 1. **Content and Organization of AFOQT Form N[a]**

Booklet and Subtest	Answer Form AFPT No.	No. of Items	Composites				
			Pilot	Nav-Tech	Officer	Verb	Quant
Booklet 1[b] (AFPT 982)	987						
Part 1 Arithmetic Reasoning		25		X	X		X
Part 2 Math Knowledge		25		X	X		X
Part 3 Data Interpretation		25		X	X		X
Booklet 2[b] (AFPT 983)	987						
Part 1 Word Knowledge		25			X	X	
Part 2 Reading Comprehension		25			X	X	
Part 3 Background for Current Events		25			X	X	
Part 4 Verbal Analogies		25	X		X	X	
Booklet 3 (AFPT 984)	988						
Part 1 Table Reading[c]		50	X	X			
Part 2 Electrical Maze[c]		30	X	X			
Part 3 Block Counting[c]		80	X	X			
Part 4 Scale Reading[c]		48	X	X			
Part 5 Tools		25	X	X			
Part 6 Mechanical Comprehension		24	X	X			
Booklet 4 (AFPT 985)	989						
Part 1 Rotated Blocks		20		X			
Part 2 Aerial Landmarks[c]		40		X			
Part 3 General Science		24		X			
Part 4 Instrument Comprehension[c]		24	X				
Part 5 Pilot Biographic and Attitude Scale		66	X				
Total		606					

[a]Associated administrative and scoring manuals are AFPT 980 and 981, respectively. Associated answer sheets are AFPT 987–989. Special answer forms (AFPT 990–992) are used in the AFROTC program. A full list of AFPT numbers for AFOQT Form N materials is included in Appendix B. Instrument comprehension is scored R-W/3 and remaining speeded subtests are scored R-W/4.

[b]Booklets 1 and 2 use the same answer form.

[c]Speeded subtests.

III. TEST MODIFICATION OVER RECENT FORMS

Test modifications of Form N over previous forms are summarized as follows: Five tests were removed and seven new tests were added. Officer Biographical Inventory, Pilot Biographical Inventory, Aviation Information, Visualization of Maneuvers, and Stick and Rudder Orientation subtests were removed, while Background for Current Events, Pilot Biographic and Attitude, Table Reading, Electrical Maze, Block Counting, Tools, and Rotated Blocks subtests were added. Two composites were further subdivided into subtests. The Quantitative Aptitude scale was subdivided into Arithmetic Reasoning, Math Knowledge, and Data Interpretation, while the Verbal Aptitude scale was subdivided into Word Knowledge, Reading Comprehension, Background for Current Events, and Verbal Analogies. Total administration time was increased from 5 hours 51 minutes to 6 hours 14 minutes. A common

male/female conversion table was developed, and the number of education-level-specific conversion tables was reduced from three to two with the combination of the "2 or more years of college but not graduates" and the "college graduate" education categories. The procedure for establishing the conversion tables was also altered as will be explained in the standardization discussion. Finally, the number of test items was increased from 522 to 606, and the test was placed in four instead of five booklets. Table 2 summarizes the content and organization of AFOQT Form M which was replaced by Form N. Appendix A summarizes the differences between Form M and N. Appendix B lists the materials associated with Form N including the Air Force Personnel Test (AFPT) numbers of the materials.

Independent studies by Valentine (1977) and Hunter and Thompson (1978) evaluated large numbers of experimental cognitive and noncognitive tests for their ability to aid prediction of navigator and pilot training success, respectively. Valentine (1977) evaluated the predictive contribution of 45 noncognitive, 17 experimental cognitive tests, and the original Form M Navigator-Technical composite subtests to performance prediction for 507 Officer Training School (OTS) students who subsequently entered Undergraduate Navigator Training (UNT). Valentine found that by retaining all but Mechanical Information and adding five new subtests, the predictive validity of the composite could be increased over the validity of Form M. Quantitative composite subscales and Scale Reading, Aerial Landmarks, General Science,

Table 2. **Content and Organization of AFOQT Form M**

Booklet and Subtest	No. of Items	Composites				
		Pilot	Nav-Tech	Officer	Verb	Quant
Booklet 1 (AFPT 972)						
Quantitative Aptitude	60		X	X		X
Booklet 2 (AFPT 973)						
Verbal Aptitude	60			X	X	
Officer Biographical Inventory[a]	96			X		
Booklet 3 (AFPT 974)						
Scale Reading[b]	48		X			
Aerial Landmarks[b]	40		X			
General Science	24		X			
Booklet 4 (AFPT 975)						
Mechanical Information	24	X	X			
Mechanical Principles	24	X	X			
Booklet 5 (AFPT 976)						
Pilot Biographical Inventory	50	X				
Aviation Information	24	X				
Visualization of Maneuvers[b]	24	X				
Instrument Comprehension[b]	24	X				
Stick and Rudder Orientation[b]	24	X				
Total	522					

[a]Not administered to female applicants.

[b]Speeded subtests.

and Mechanical Principles (now called Mechanical Comprehension) were recommended for retention, Table Reading, Electrical Maze, Block Counting, Tools, and Rotated Blocks were recommended additions. Use of the proposed Navigator-Technical composite provided a biserial validity of .64 for predicting UNT performance. See Valentine (1977) for specific subtest validities.

As part of an 8-year series of studies to develop and revise procedures for selecting undergraduate pilot trainees, Hunter and Thompson (1978) evaluated the experimental cognitive tests studied by Valentine (1977) and most of the original Form M Pilot composite scales. Data were collected on some 800 officers and officer trainees from AFROTC and OTS. Instrument Comprehension and the Pilot Biographical and Attitude Scales subtests were not part of the data reported by Hunter and Thompson (1978). Hunter and Thompson did suggest inclusion of the original Mechanical Principles (Comprehension) subtest and adding six new subtests to the Pilot Composite of the AFOQT. Verbal Analogies, Table Reading, Electrical Maze, Block Counting, Scale Reading, and Tools were suggested for inclusion in the Pilot composite because of significant validities with Undergraduate Pilot Training (UPT) performance criteria. A subsequent reanalysis of available data on the Instrument Comprehension and Pilot Biographical and Attitude Scales resulted in estimated predictive validities for the new nine-subtest Pilot composite ranging from r = .26 to r = .39—a substantial improvement in validity from that obtained by Form M (r = .19) on the same cases.

The types of test items and subtests comprising the Verbal, Quantitative, and Officer Quality Composites have not been changed except for the removal of the Officer Biographical Inventory (OBI) from the Verbal and Officer Quality Composites. The OBI was previously taken only by males since it was composed of activities associated with males. Unable to remove sex bias from the items, the subtest was eliminated from the revised AFOQT. In the other subtests, several dozen female airmen reviewed and commented on each item to aid the test developers in identifying language to be modified or items to be replaced if not modifiable.

IV. ITEM SELECTION

Each form of the AFOQT is designed to have the same difficulty as the preceding form. The selection of items is guided by the principle that the item of median difficulty in each test should be answered correctly by 50% of the examinees for whom the test is appropriate, with the other items in the test having a considerable range of difficulty about the median. The only exceptions are the biographical scale and speeded tests for which the concept of difficulty has a somewhat different meaning. Biographical items in a sense have no right or wrong answers, but responses are considered positively or negatively in terms of their relationship to an empirical criterion. Speeded subtests are constructed so that few examinees reach the final items yet most get the initial items correct so that computed difficulties represent products other than those computed where each individual has a chance to try each item. For these reasons, no difficulty levels or internal consistencies were computed for the biographical or speeded tests.

The median difficulty and range of difficulty of items in Form N are shown in Table 3. Difficulties in the table are expressed as percentages of examinees who answered the items correctly. Thus, the higher values represent the easier items. The desired median difficulty is closely approximated in each subtest.

Table 3 also presents internal consistency data for Form N. Internal consistency refers to the correlation between the correct response to an item and the total score of the subtest of which the item is a part. Again the biographical subtest and speeded subtests are special cases; low internal consistency is to be expected of them. In other subtests, it is desired that the internal consistency be high, but it is not possible to have uniformly high internal consistency in items having the desired distribution of difficulty. The range and median of the internal consistency distributions for Form N are similar to those for other forms of the AFOQT. No items having positive internal consistency coefficients for any incorrect response were included in the test. Some anchor items which appeared in previous forms were included.

Table 3. **Item Difficulty Levels and Internal Consistency of AFOQT Form N[a]**

Subtest	Difficulty Range	Level Median	Internal Range	Consistency Median
Arithmetic Reasoning	.17–.88	.53	.42–.89	.68
Math Knowledge	.45–.86	.57	.35–.99	.79
Data Interptetation	.14–.90	.53	.17–.70	.59
Word Knowledge	.19–.93	.53	.17–.77	.65
Reading Comprehension	.49–.91	.61	.49–.93	.79
Background for Current Events	.17–.87	.54	.25–.88	.61
Verbal Analogies	.23–.90	.56	.36–.81	.57
Tools	.23–.99	.61	.39–.82	.61
Mechanical Comprehension	.14–.76	.46	.17–.68	.54
Rotated Blocks	.24–.92	.53	.28–.72	.60
General Science	.20–.88	.50	.28–.77	.62

[a]Speeded tests and biographical test are not included. Internal consistency values are biserial correlations.

V. RELIABILITY, INTERCORRELATIONS, AND VALIDITY

Though various forms of the AFOQT have been used consecutively, they have in effect had the properties of alternate forms. It has therefore been assumed that such technical data as reliability, validity, and intercorrelations of composites for a new form are similar to the corresponding data for preceding forms. The assumption was warranted since only specific items were removed and similar items substituted. The assumption may still be essentially valid for the Verbal, Quantitative, and Officer Quality composites which have undergone little change in substance in the current revision. Intercorrelations between the Verbal, Quantitative, and Officer Quality composites for Forms M and N were .81, .80, and .77, respectively. Mean composite scores were not significantly different between Forms M and N. The test-retest means and correlations over a 12- to 24-month period for 266 AFROTC students attest to both the stability of the AFOQT and the equivalence of the three composites between Forms M and N. The 266 cases were from the standardization sample and represent the 3rd and 4th year AFROTC cadets for whom AFOQT Form M scores could be located. AFROTC candidates for the Professional Officers Course (POC) take the AFOQT early in their second year. Therefore, the test-retest interval ranged from 12 to 24 months. Extensive revision of the Pilot and Navigator-Technical composite subtests leaves little justification for relying on past results for these composites. Therefore, a stratified random sample of 1,000 cases was selected from the 2,681 cases in the standardization sample for analysis.

Reliability and intercorrelation data for the composites are presented in Tables 4 and 5. The reliability data are determined from the formula for the reliability of a composite (Wherry & Gaylord, 1943), which in turn is based on test-retest or Kuder-Richardson Formula 20 data for the subtests. The biographical subtest is omitted.

A convenient summary of validity data for the Officer Quality composite is contained in a technical report on interpretation and use of AFOQT scores (Miller, 1969) and in recent studies of success in predicting performance in AFROTC (Alley & Gibson, 1977) and OTS (Mathews, 1977).

Table 4. **Estimated Reliability of Composites, AFOQT Form N**

Composite	Reliability
Pilot	.91
Navigator-Technical	.95
Officer Quality	.94
Verbal	.89
Quantitative	.93

Table 5. **Intercorrelation of Composites, AFOQT Form N**

Composite	Quantitative	Verbal	Officer Quality	Pilot
Quantitative				
Verbal	.644			
Officer Quality	.896	.905		
Pilot	.740	.522	.688	
Navigator-Technical	.730	.531	.687	.969

VI. STANDARDIZATION

The AFOQT had traditionally been standardized on an AFA candidate group. After 1960, Academy candidates were no longer available for this purpose, but a new method was devised for indirectly relating a new AFOQT form to a prior Academy candidate group. The specific group was made up of 5,105 candidates for the class of 1964. The indirect method has been described in general (Dailey, Shaycoft, & Orr, 1962) and in its specific application to the AFOQT (Miller & Valentine, 1964). Briefly, the method consisted of equipercentile conversions from AFOQT Form G, which was administered to Academy candidates, through composites of tests from the Project TALENT battery to the new form of the AFOQT. The relationship between the TALENT composites and the new form was determined on samples of basic airmen stratified on the Armed Forces Qualification Test (AFQT) by deciles in the percentile range from 21 to 100 (Miller, 1974).

Revision of subtests making up composites in Form N made it impossible to develop such equipercentile conversions, so new normative data were collected. Whereas past AFOQT scores have related solely to Academy cadets, the new reference group has been changed to include all major sources for Air Force commissioning programs and sources for specialized training programs. Therefore, the standardization sample was selected from basic airmen; AFROTC, OTS, and AFA cadets; and junior officers.

In the fall of 1977, AFOQT Form N was administered to 2,681 cases for establishment of percentile conversion tables. Composition of the normative sample is as follows: Basic airmen who were in the upper 50 percent of examinees taking the Armed Forces Qualifying Test (AFQT $\geqslant$ 50) and in their seventh day of training were randomly selected and tested. Complete data were obtained on 707 airmen. The AFQT $\geqslant$ 50 limitation on selected airmen was established to obtain a sample of airmen falling in the range of college entrants. The ability level of the basics should approximate that of, for example, applicants for 4-year AFROTC scholarships.

Sixteen AFROTC detachments were selected for participation in the project. Schools were selected to represent national geographical, racial, and academic characteristics of AFROTC detachments. Data were obtained on 604 cases but 60 cases had only Pilot composite scores, so for most scales the effective number was 544. Random samples of 200 AFA cadets from each of the four classes, 1st through 4th years, were selected and 771 complete records were obtained. The basic, AFROTC, and AFA cases represented the non-college graduate portion of the sample. The college graduate sample was composed of OTS trainees and second lieutenants. Some 240 OTS students were tested but because of limited available test time, no individual took all subtests so the Ns vary considerably by composite. It should be noted that in the OTS population, most cases had science and engineering degrees and many had postgraduate training. To complete the sample, a continental United States (CONUS) wide sample of some 300 AFROTC and AFA source second lieutenants were selected, resulting in complete scores on 284 cases. Score distributions on the samples will be presented in the next section. In the total sample, there were 268 females (10%) and 294 Blacks (11%).

Past versions of the AFOQT have provided separate percentile conversion tables for three educational groupings: (1) less than 2 years college; (2) 2 years college or more but not college graduates; and (3) college graduate or post-graduate. The procedure was based on quantitative evaluation of the elevating effect of education on AFOQT scores as reported by Gregg (1968) and Tupes and Miller (1969). Inspection of raw composite score differences for the three education level groupings for Form N indicated there were no significant differences ($P < .05$) between the college graduate group and those with 2 years college who were not graduates. Composite means by education group are shown in Table 6. Significant differences between those with less than 2 years college and those with 2 or more years college were found for all five AFOQT composites. Therefore, two separate raw score to percentile conversion tables were developed to take into account effects of education on AFOQT performance. The score conversion charts are shown in Tables 7 and 8.

Table 6. **Raw Score Composite Means and Standard Deviations by Three Educational Groups, AFOQT Form N**

Composite		College Graduates	College Non-Grad Ed > 2 yrs.	College Non-Grad and Non-College Ed < 2 yrs.
Quantitative	N	460	657	1,373
	$\bar{X}$	47.02	51.26	39.41
	S.D.	13.48	12.90	15.61
Verbal	N	367	656	1,372
	$\bar{X}$	72.35	71.42	55.38
	S.D.	14.19	14.49	18.71
Officer Quality	N	367	656	1,371
	$\bar{X}$	119.14	122.68	94.80
	S.D.	24.40	24.82	31.92
Pilot	N	418	683	1,389
	$\bar{X}$	212.97	226.48	196.11
	S.D.	44.71	43.30	46.44
Navigator-Technical	N	386	652	1,356
	$\bar{X}$	191.34	202.38	171.84
	S.D.	47.04	45.86	47.80

Table 7. **Aptitude Conversion Tables, Less Than 2 Years College**

Percentile Score	Raw Score Cutoffs				
	Pilot Composite	Navigator Technical Composite	Officer Quality Composite	Verbal Composite	Quantitative Composite
95	269 and above	246 and above	144 and above	85 and above	65 and above
90	256–268	233–245	138–143	80–84	62–64
85	247–255	224–232	132–137	77–79	59–61
80	239–246	217–223	127–131	74–76	56–58
75	233–238	210–216	123–126	71–73	53–55
70	226–232	201–209	118–122	68–70	51–52
65	218–225	194–200	112–117	66–67	48–50
60	211–217	186–193	106–111	63–65	45–47
55	204–210	179–185	101–105	60–62	42–44
50	197–203	174–178	96–100	57–59	39–41
45	192–196	168–173	90–95	54–56	36–38
40	186–191	160–167	84–89	51–53	33–35
35	178–185	153–159	78–83	47–50	31–32
30	172–177	146–152	73–77	44–46	28–30
25	164–171	139–145	68–72	40–43	26–27
20	154–163	132–138	63–67	37–39	24–25
15	145–153	121–131	58–62	34–36	22–23
10	134–144	109–120	53–57	30–33	20–21
05	118–133	90–108	46–52	25–29	17–19
01	117 and below	89 and below	45 and below	24 and below	16 and below

Table 8. **Aptitude Conversion Tables, 2 or More Years of College Including College Graduates**

Percentile Score	Raw Score Cutoffs				
	Pilot Composite	Navigator Technical Composite	Officer Quality Composite	Verbal Composite	Quantitative Composite
95	281 and above	266 and above	155 and above	91 and above	69 and above
90	274–283	255–265	150–154	89–90	66–68
85	266–273	246–254	146–149	86–88	64–65
80	259–265	239–245	143–145	84–85	62–63
75	253–258	233–238	140–142	83–83	60–61
70	249–252	228–232	127–129	81–82	59–59
65	244–248	221–227	134–136	79–80	57–58
60	239–243	216–220	131–133	78–78	55–56
55	233–238	209–215	128–130	76–77	54–54
50	228–232	203–208	125–127	75–75	52–53
45	223–227	198–202	123–124	73–74	50–51
40	217–222	192–197	120–122	71–72	48–49
35	210–216	185–191	117–119	70–70	46–47
30	203–209	178–184	113–116	68–69	44–45
25	196–202	170–177	109–112	65–67	41–43
20	187–195	161–169	105–108	62–64	38–40
15	178–186	151–160	98–104	59–61	35–37
10	162–177	135–150	88–97	52–58	30–34
05	138–161	110–134	71–87	43–51	24–29
01	137 and below	109 and below	70 and below	42 and below	23 and below

VII. SCORE DISTRIBUTIONS

Table 9 presents the raw score means and standard deviations for the five samples in the standardization population. The comparisons represent a cross-sectional view of the samples and thus do not take into account sample academic differences. Most selection programs are concerned with establishing projections of available talent and assessing effects of adjusting cut-off scores on number and characteristics of the selectees. For these reasons, and to permit those concerned with qualitative differences between the samples, a series of frequency distributions are provided for the three composites (Officer Quality, Pilot, and Navigator-Technical) most frequently used in selection programs.

The Officer Quality (OQ) composite is a combination of the Verbal and Quantitative composites and can be described as an index of "general learning ability." The OQ is the AFOQT scale most used in selection programs. All candidates for 2- or 4-year AFROTC scholarship programs, AFROTC Professional Officer's Course (POC), and OTS are screened by their performance on the OQ composite. Figure 1 is the distribution of the total norm sample OQ raw scores and represents a cross-sectional view of general aptitude levels of Air Force personnel. The bimodal nature of the curve is produced by basic aptitude differences in the enlisted and commissioned entry-level personnel. The peaked and skewed nature of the curves are largely due to test characteristics which exist by design. The AFOQT was designed to maximize differentiation among aptitude levels in the lower ranges of accepted aptitudes for commissioned officers. This is one of the major differences between the AFOQT and general aptitude measures, such as the Scholastic Aptitude Test (SAT) or American College Test (ACT) which attempt to distinguish equally well between examinees at all levels of ability or even at the higher rather than the lower levels.

Figures 2 and 3 permit aptitude level comparisons between members of each of the five samples tested. From Figure 2, for example,

Table 9. **Raw Score Means and Standard Deviations for AFOQT Form N Composites by Standardization Group Subsample[a]**

Composite		Basic Airmen	OTS	AFROTC	2nd LT	AFA
Quantitative	N	722	162	546	287	773
	$\bar{X}$	27.87	49.68	47.41	46.08	54.51
	S.D.	10.00	13.41	13.17	13.18	9.05
Verbal	N	723	69	544	287	772
	$\bar{X}$	44.62	77.10	65.40	71.61	72.12
	S.D.	15.73	10.47	17.36	14.56	10.94
Officer Quality	N	722	69	544	287	772
	$\bar{X}$	72.51	129.17	112.84	117.69	126.62
	S.D.	22.30	18.62	27.78	24.69	16.87
Pilot	N	708	123	604	284	771
	$\bar{X}$	168.48	204.31	214.47	219.03	233.40
	S.D.	38.28	43.24	44.66	43.08	33.93
Navigator-Technical	N	707	90	540	285	772
	$\bar{X}$	143.81	187.16	189.97	194.88	210.09
	S.D.	39.11	49.37	46.46	44.64	36.17

[a]Variation is sample sizes due to missing subtest scores in some cases.

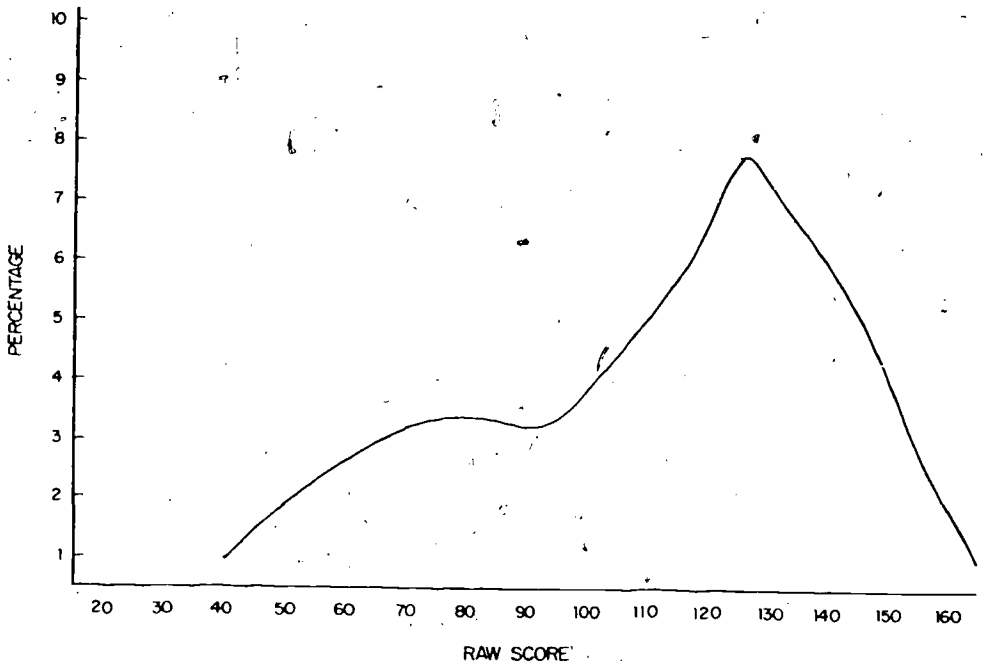

Figure 1. Total normative sample officer quality composite raw score distribution.

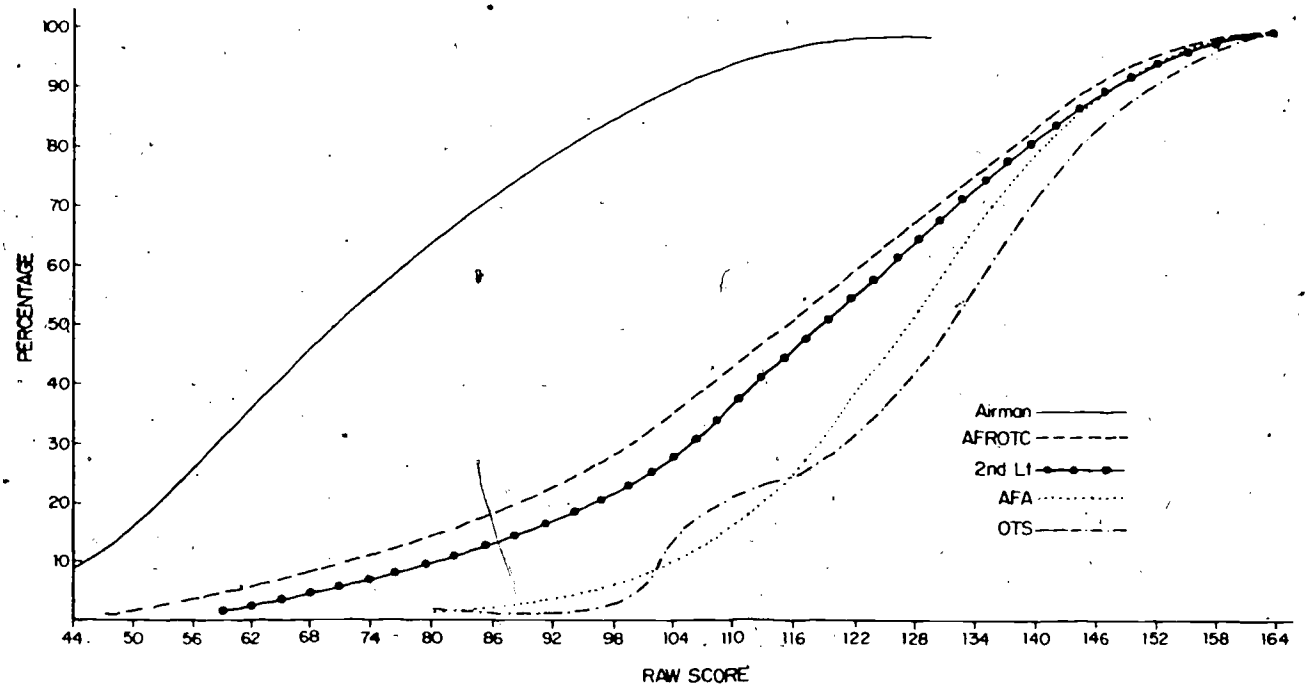

Figure 2. Cumulative frequency distribution of officer quality raw scores by sample.

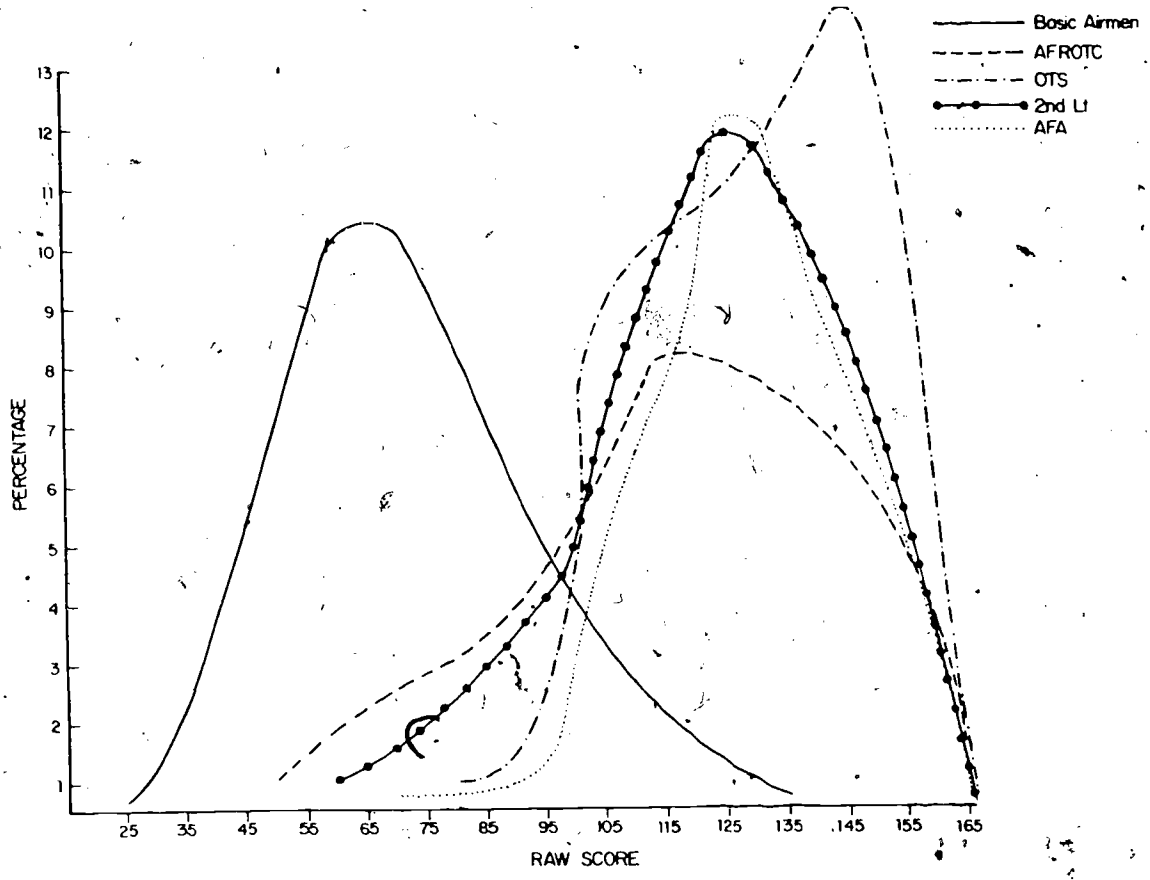

Figure 3. **Frequency distribution of officer quality raw scores by sample.**

AFROTC students scoring at the 20th percentile for that sample still scored better than 75% of the basic airmen, indicating the select nature of the AFROTC students. Another way to look at the general aptitude level of AFROTC students is to consider the full nature of the basic trainees included in the study. Only airmen with Armed Forces Qualifying Test (AFQT) $\geq$ 50 were tested; i.e., of all cases tested on the AFQT, all these basics scored better than half the norm population. Taken further, basic airmen in the top 25% of their sample have aptitudes at least two standard deviations above the mean aptitude found among the general population of 17-year-olds in the U.S. These basics obtained the same test performance as the AFROTC students scoring at the 20th percentile. Therefore, some 80% of the AFROTC students are at least two standard deviations above the mean aptitude level of the general population. Taken further, the selective nature of the OTS and AFA incumbents is apparent by noting that the lowest scoring individuals did better than 63% of the basics. From Figure 3, within each sample, the distribution of OQ raw scores tends to be approximately normally distributed with the exception of the OTS group which had a disproportionate percentage scoring in the higher score range. The OTS distribution is consistent with the high scientific and engineering background of the current OTS population. Table 10 presents selected OQ composite raw scores and relative standings within AFOQT norm samples for use if a more exact comparison of scores is desired than can be obtained from Figure 2. Figures 4 and 5 present the cumulative frequency distributions by sample for the Pilot and Navigator-Technical composites, respectively.

VIII. CONCLUSIONS AND RECOMMENDATIONS

A new form of the AFOQT (Form N) has been developed and standardized. The test item characteristics, internal consistencies of the tests and composites, validity evidence, and representativeness of the standardization sample are more than adequate. Therefore, it is recommended that Form N of the AFOQT be operationally implemented.

Table 10. Selected Officer Quality Composite Raw Score Values and Relative Standing Within AFOQT Norm Samples

Raw Score	Basics	OTS	AFROTC	2nd Lt	AFA
160	99	99	98	99	99
153	99	91	94	94	95
149	99	85	92	90	90
144	99	75	86	84	85
142	99	72	84	83	80
139	99	65	81	78	75
137	99	62	78	74	70
134	99	54	75	71	65
132	99	52	71	68	60
130	99	49	69	67	55
128	99	39	67	64	50
126	98	39	64	61	45
124	98	35	62	55	40
122	97	32	58	50	35
119	96	30	53	46	30
117	96	26	52	44	25
113	95	19	45	39	20
109	93	14	38	35	15
104	91	10	33	26	10
098	86	4	28	19	05
088	73	3	18	11	1

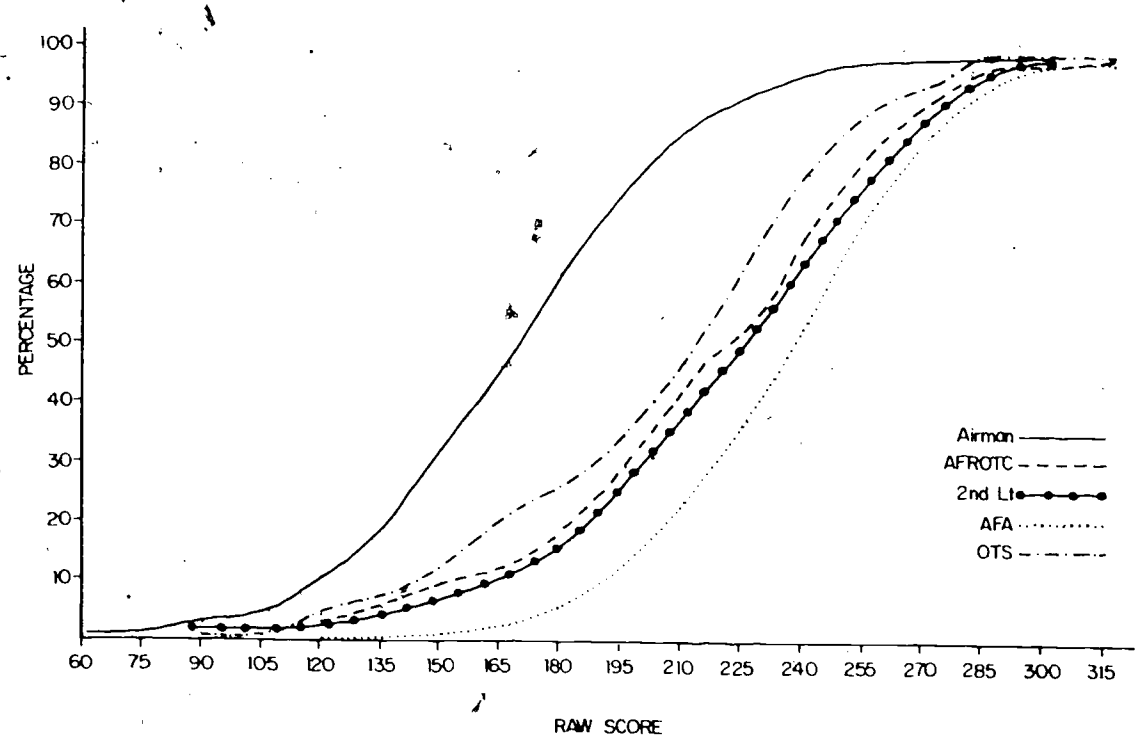

Figure 4. Cumulative frequency distribution of pilot composite raw scores by sample.

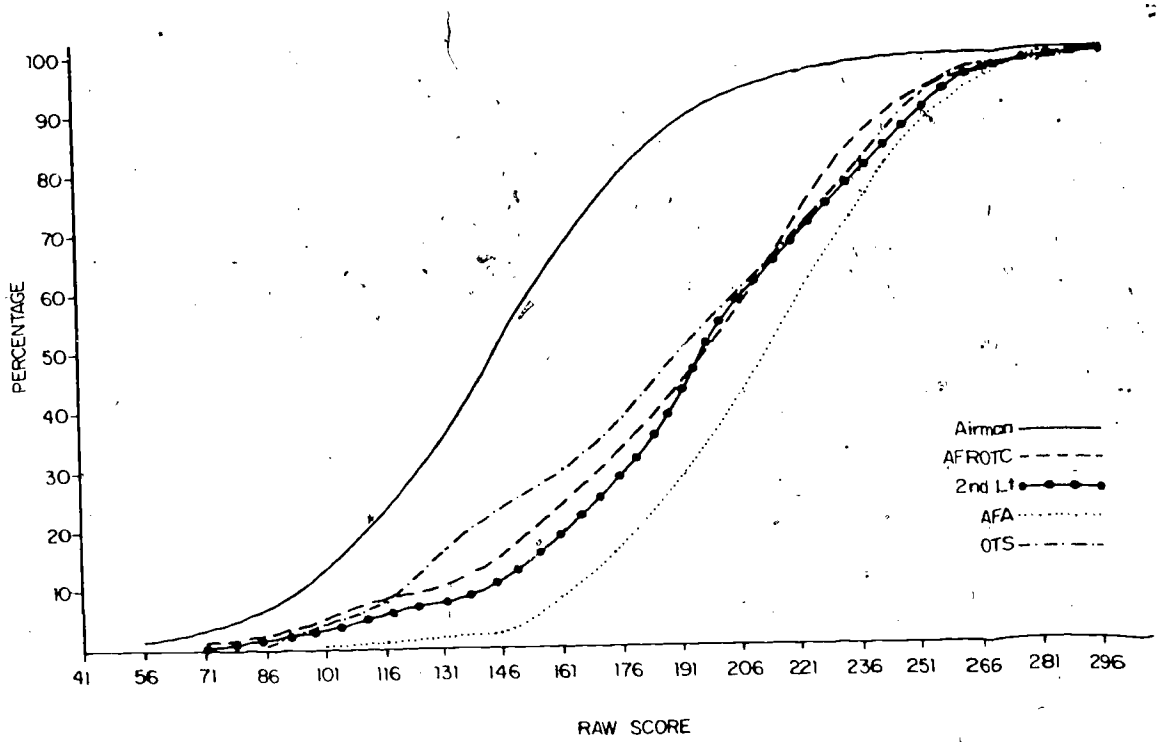

Figure 5. Cumulative frequency distribution of navigator-technical composite raw scores by sample.

REFERENCES

Alley, W.E., & Gibson, T.A. *Predicting success in the AFROTC scholarship program.* AFHRL-TR-77-11, AD-A041 132. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, April 1977.

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

Gregg, G. *The effect of maturation and educational experience on Air Force Officer Qualifying Test scores.* AFHRL-TR-68-107, AD-687 089. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, July 1968.

Hunter, D.R., & Thompson, N. *Pilot selection system development.* AFHRL-TR-78-33. Brooks AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, July 1978.

Mathews, J.J. *Racial equity in selection in Air Force officer training school and undergraduate flying training.* AFHRL-TR-77-22, AD-A043 019. Brooks AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, May 1977.

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

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

Miller, R.E. *Interpretation and utilization of scores on the Air Force Officer Qualifying Test.* AFHRL-TR-69-103, AD-691 001. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, May 1969.

Miller, R.E. *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, June 1970.

Miller, R.E. *Development and standardization of the Air Force Officer Qualifying Test Form L.* AFHRL-TR-72-47, AD-754 849. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, May 1972.

Miller, R.E. *Development and standardization of the Air Force Officer Qualifying Test Form M.* AFHRL-TR-74-16, AD-778 837. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, March 1974.

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, TX: Personnel Research Laboratory, Aerospace Medical Division, March 1964.

Tupes, E.C., & Miller, R.E. *Equivalence of AFOQT scores for different educational levels.* AFHRL-TR-69-19, AD-703 727. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, September 1969.

Valentine, L.D., Jr. *Navigator-observer selection research: Development of new Air Force officer qualifying test navigator-technical composite.* AFHRL-TR-77-36, AD-A042 689. Brooks AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory. May 1977.

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, TX: 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.

APPENDIX A. SUMMARY OF DIFFERENCES BETWEEN AFOQT FORM M AND FORM N

Form M Booklet and Subtest	No. of Items	Form N Booklet and Subtest	No. of Items
Booklet 1		**Booklet 1**	
Quantitative Aptitude	60	Part 1 –Arithmetic Reasoning	25
		Part 2 –Math Knowledge	25
		Part 3 –Data Interpretation	25
Booklet 2		**Booklet 2**	
Verbal Aptitude	60	Part 1 –Word Knowledge	25
Officer Biographical Inventory	96 omitted	Part 2 –Reading Comprehension	25
		Part 3 –Background for Current Events	25 New
		Part 4 –Verbal Analogies	25
Booklet 3		**Booklet 3**	
Scale Reading	48	Part 1 –Table Reading	50 New
Aerial Landmarks	40	Part 2 –Electrical Maze	30 New
General Science	24	Part 3 Block Counting	80 New
		Part 4 –Scale Reading	48
		Part 5 –Tools	25 New
		Part 6 –Mechanical Comprehension	24
Booklet 4		**Booklet 4**	
Mechanical Information	24	Part 1 –Rotated Blocks	20 New
Mechanical Principles	24	Part 2 –Aerial Landmarks	40
		Part 3 –General Science	24
		Part 4 –Instrument Comprehension	24
		Part 5 Pilot Biographic and Attitude	66 New
Booklet 5			
Pilot Biographical Inventory	50 omitted		
Aviation Information	24 omitted		
Visualization of Maneuvers	24 omitted		
Instrument Comprehension	24		
Stick and Rudder Orientation	24 omitted		
Total	522	Total	606

APPENDIX B. SUMMARY OF ASSOCIATED TEST MATERIALS AND AFPT NUMBERS FOR AFOQT FORM N

AFPT	Material
901	United States Air Force Officer Qualifying Test Manual for Interpretation
980	Manual for Administration (Uncontrolled)
981	Scoring Instruction (Uncontrolled)
982	Booklet 1
982a	Scoring Key, Booklet 1
983	Booklet 2
983a	Scoring Key, Booklet 2
984	Booklet 3
984a	Table Reading Chart for use with Booklet 3
984b	Scoring Key, Booklet 3 (Front-Rights)
984c	Scoring Key, Booklet 3 (Back-Rights)
984d	Scoring Key, Booklet 3 (Front-Wrongs)
984e	Scoring Key, Booklet (Back-Wrongs)
985	Booklet 4
985a	Scoring Key, Booklet 4 (Front Rights)
985b	Scoring Key, Booklet 4 (Back-Rights)
985c	Scoring Key, Booklet 4 (Front-Wrongs)
985d	Scoring Key, Booklet 4 (Back-Wrongs)
987	Answer Sheet, Booklets 1 and 2 (Uncontrolled)
988	Answer Sheet, Booklet 3 (Uncontrolled)
989	Answer Sheet, Booklet 4 (Uncontrolled)
990	Optical Scan Answer Sheet, Booklets 1 and 2 (Uncontrolled)
991	Optical Scan Answer Sheet, Booklet 3 (Uncontrolled)
992	Optical Scan Answer Sheet, Booklet 4 (Uncontrolled)

20 ☆U.S. GOVERNMENT PRINTING OFFICE: 1978 [illegible]