

AFCAPS-FR-2012-0001

**Recent Developments in
USAF Officer Testing and
Selection**

IMTA Conference
October 2004



Johnny J. Weissmuller
Kenneth L. Schwartz
Stanley D. Kenney
Dr. C. Wayne Shore
Dr. R. Bruce Gould

Air Force Personnel Center
Force Management Liaison Office
Randolph AFB TX 78150



Kenneth L. Schwartz
Johnny J. Weissmuller

**AFPC/Strategic Research and
Assessment Branch (SRAB)**

Air Force Personnel Center
Strategic Research and Assessment
HQ AFPC/DSYX
550 C Street West, Ste 45
Randolph AFB TX 78150-4747

Approved for Public Release. Distribution Unlimited

UNCLASSIFIED

NOTICE

When 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.

This report was cleared for release by HQ AFPC/DSYX Strategic Research and Assessment Branch and is releasable to the Defense Technical Information Center.

This report is published as received with minor grammatical corrections. The views expressed are those of the authors and not necessarily those of the United States Government, the United States Department of Defense, or the United States Air Force. In the interest of expediting publication of impartial statistical analysis of Air Force tests SRAB does not edit nor revise Contractor assessments appropriate to the private sector which do not apply within military context.

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct request for copies of this report to:

Defense Technical Information Center - <http://www.dtic.mil/>

Available for public release. Distribution Unlimited. Please contact AFPC/DSYX Strategic Research and Assessment with any questions or concerns with the report.

This paper has been reviewed by the Air Force Center for Applied Personnel Studies (AFCAPS) and is approved for publication. AFCAPS members include: Senior Editor Dr. Thomas Carretta AFMC 711 HPW/RHCI, Associate Editor Dr. Gregory Manley HQ AFPC/DSYX, Dr. Lisa Hughes AF/A1PF, Dr. Paul DiTullio AF/A1PF, Kenneth Schwartz HQ AFPC/DSYX, Johnny Weissmuller HQ AFPC/DSYX, Dr. Laura Barron HQ AFPC/DSYX, Dr. Mark Rose HQ AFPC/DSYX, and Brian Chasse HQ AFPC/DSYX.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 01-10-2004		2. REPORT TYPE Final		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Recent Developments in USAF Officer Testing and Selection				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Johnny J. Weissmuller, Kenneth L. Schwartz, Stanley D. Kenney, Dr. C. Wayne Shore, Dr. R. Bruce Gould				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Personnel Center Force Management Liaison Office Randolph AFB TX 78150				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Personnel Center Strategic Research and Assessment Branch Randolph AFB TX 78150				10. SPONSOR/MONITOR'S ACRONYM(S) HQ AFPC/DSYX	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFCAPS-FR-2012-0001	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release. Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Report covers the basics of military testing and some of the recent/future developments of testing within the Armed Forces. Summarizes which departments of the Air Force are responsible for supervising and creating tests and standards while also examining some of the more important military tests. The tests covered by this report were the Armed Services Vocational Aptitude Battery (ASVAB); the enlisted promotions tests which include, Specific Knowledge Test (SKT), Promotion Fitness Exam (PFE), and the USAF Supervisor Exam (USAFSE); Foreign Language Proficiency Testing (DLAB); Air Force Officer Qualifying Test (AFOQT); Test of Basic Aviation Skills (TBAS) and Basic Attributes Test (BAT). Report also covers some of the experimental tests that were being developed and recent changes to the AFOQT and Pilot-Navigator composites and testing.					
15. SUBJECT TERMS ASVAB, SKT, PFE, DLAB, TBAS, PCSM, AFOQT, BAT, Pilot, Navigator					
16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT U	18. NUMBER OF PAGES 39	19a. NAME OF RESPONSIBLE PERSON Kenneth L. Schwartz
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code) 210-565-3139

(Page Intentionally left blank)

Recent Developments in USAF Officer Testing and Selection

Johnny J. Weissmuller, Kenneth L. Schwartz, and Stanley D. Kenney

USAF Military Testing, HQ Air Force Personnel Center

Randolph AFB, Texas USA

johnny.weissmuller@randolph.af.mil

C. Wayne Shore, Ph.D. and R. Bruce Gould, Ph.D.

Operational Technologies Corporation (OpTech)

San Antonio, Texas, USA

wayne@otcorp.com

Background

Military testing in the United States Air Force (USAF) falls into several categories. Testing associated with *Education & Training* is vital for preparing a team-oriented, a mission-ready force. The other forms of testing, collectively known as *Personnel Tests*, have a direct impact on defining an *individual's* official job and level of responsibility. The goal of personnel testing is to select, promote, or certify the most qualified individuals. The Air Force Personnel Center at Randolph AFB (HQ AFPC) is charged with managing this force throughout the world to ensure the right person is at the right place at the right time. For this reason, all military personnel tests are centrally managed by the Air Force Military Testing Section of AFPC.

The Air Force Military Testing Section (AFPC/DPPPWT) is responsible for the integrity and application of personnel tests that are used across the Air Force. This section performs many functions related to testing including 1) recommending new research/validation needs, 2) evaluating emerging test technologies; 3) managing test development contracts, 4) validating proposed instruments against professional standards, 5) applying accepted psychometric procedures in response to non-standard testing situations, 6) interpreting, reviewing, and updating Air Force testing policy instructions, 7) responding to inquiries and requests for waivers, 8) coordinating the Test Control Officer network, 9) providing oversight and coordination for testing and associated study materials printing and distribution operations, 10) receiving and scanning completed test answer sheets, and 11) scoring, uploading, and answering routine inquiries on all test results. Because many of these tests will affect the quality of people in the Air Force (as well as a person's pay), AFPC/DPPPWT is a key player in combating test compromise, including support for detection, statistical analysis, and prosecution, working in close concert with the Air Force Office of Special Investigations (AFOSI) and the Test Development Flight of the Air Force Occupational Measurement Squadron (AFOMS/TE) and the AFOMS Commander.

The major categories of the personnel tests managed by DPPPWT include selection tests (for all USAF officer and enlisted personnel), enlisted promotion tests, as well as certification tests for foreign language ability level (and its associated proficiency pay). Other special use tests such as reading ability, cross-training assessment or selection into specific fields like computers or recruiting also are managed by this office. The entire range of tests managed is available through the public web site at <http://www.afpc.randolph.af.mil/testing/>.

The Armed Services Vocational Aptitude Battery (ASVAB)

To say that the Military Testing Section “manages” a test means different things depending on the test. The type of interaction with a test varies based on 1) who developed/administers the instrument, 2) the type of test, and 3) its ultimate application or use within the Air Force. The first example is the test battery used to select civilians into the United States military enlisted ranks. This battery of tests is the Armed Services Vocational Aptitude Battery (ASVAB). The ASVAB is administered under the auspices of the Department of Defense (DoD). Results of four of these tests are combined to produce a single score called the Armed Forces Qualifying Test (AFQT) score by which Congress mandates enlistment goals and guidelines. Each of the uniformed military services (Army, Air Force, Navy, and Marines) further combine these nine individual test scores into their own set of composite scores which are used within that service for selection and classification.

Whenever the ASVAB is changed or anchored to a new population norm, the Air Force Military Testing Section acts as the AFPC coordination point with the Pentagon’s (AF/DPX) office tracking the changes. While the DoD is responsible for updating, distributing, and scoring the ASVAB, the local section must ensure that all AFPC “users” are taking appropriate corrective measures which include 1) updating master personnel file tables of allowable test-IDs, 2) adjusting the minimum scores required to enter each career field, or 3) re-calibrate Air Force cross-training tests which are based on the ASVAB.

Many of the reports documenting this and related personnel research history from the Air Force Human Resources Laboratory (AFHRL) and its successors are now available online. HQ AF/DPX recently scanned in over 1,000 technical reports archived in the non-profit Innovation Center for Occupational Data Application and Practices’ (ICODAP) hardcopy Technical Report (TR) library (<http://www.icodap.org>). As of 23 September 2004, these reports are available to users from “.mil” and “.gov” sites at: <https://www.dp.hq.af.mil/dpx/techreps/>

Enlisted Promotion Testing

Enlisted personnel in the Air Force serve in the paygrade ranks of E-1 through E-9. Airmen in paygrades E-1 through E-3 are promoted on a “fully qualified” basis. Eligible airmen in paygrades E-4 through E-6, after recommendation by their commander, are selected for promotion based on their total score produced in the Weighted Airman Promotion System (WAPS). The cutoff score between “selected” and “non-selected” is driven by force shaping needs of the Air Force at the time each testing cycle ends. A person’s total score is based on several factors including Performance Reports, Decoration points, Time-in-Service, Time-in-Grade, and test scores. These airmen compete for promotion and are given two tests, one for job knowledge (the Specialty Knowledge Test – SKT) and one for knowledge of Air Force customs, procedures and practices (the Promotion Fitness Exam – PFE). Together these tests can contribute up to 200 points out of the 460 total available points in promotion evaluations. When one realizes that people are competing against only those in their own career area at the same grade level, it becomes clear that the test scores are the high drivers in this promotion equation.

Airman in the ranks of E-7 and E-8 are covered by the Senior NCO Promotion Program (SNCOPP) and use similar factors as above. This promotion program, however, uses only one

test (the USAF Supervisory Exam – USAFSE). These airmen have their records reviewed and scored by a promotion board. While the board score contributes between 270 to 450 total points (a range of 180), the USAFSE represents 100 points and remains an important factor in determining actual promotions.

The Military Testing Section recently completed a comprehensive re-validation of the formulae used to combine the promotion factors in both the WAPS and the SCNOPP. This was the first such overall evaluation in 18 years. Because of all the careful personnel research that went into the development of WAPS in the late 1960s, it was not surprising to find that the system is still selecting the best people.

Although the Military Testing Section is in charge of the policy instruction on how testing will be accomplished, the actual work of testing is carried out by others. Tests are ordered, secured, controlled, administered, and answer sheets shipped back for scoring by Test Control Officers (TCOs) in the field. The TCO Network is composed primarily of civilians in the pay grades of GS-5 through GS-7 who perform all manner of military testing (except for the initial ASVAB) and serve as the local eyes and ears for minimizing, detecting, and combating test compromise.

The test content for all the Specialty Knowledge Tests (SKTs), Promotion Fitness Exams (PFEs) and USAF Supervisory Exam (USAFSE) is created at the Test Development Flight of the Air Force Occupational Measurement Squadron using teams of subject matter experts (SMEs) under the supervision of Test Development Psychologists. Test booklets (SKTs, PFEs, USAFSEs – over 220,000 in 2004) are all printed and shipped “on demand” by the Air Force Occupational Measurement Squadron’s on-site printing capability known as the Document Automation and Production Service (DAPS).

Foreign Language Proficiency Testing

Like the ASVAB, for all Foreign Language Proficiency tests, the contents are developed and controlled by a DoD level organization, the Defense Language Institute (DLI). Unlike the ASVAB, however, the actual tests are administered by the Air Force Test Control Officers. The goal of this testing is to keep track of current language skills available in the force to quickly support ever-changing Air Force mission needs. Air Force members are encouraged to self-report these language skills even if they are not currently serving as a linguist or in a “Language Designated Position (LDP).” To encourage the widest possible response and currency of this information, once tested and annually recertified, Air Force personnel receive Foreign Language Proficiency Pay (FLPP) tied to both proficiency level and the number of distinct languages spoken by a member. Specifics of this program vary based on the needs of each service and other federal agencies involved in this program. The Military Testing Section coordinates access to controlled testing materials, manages the Air Force testing policy, audits the level of compensation, and coordinates pay-affecting adjustments with the Defense Finance Accounting System (DFAS).

The United States Air Force Officer Qualifying Test (AFOQT)

The Air Force Officer Qualifying Test (AFOQT; Carretta & Ree, 1996; Skinner & Ree, 1987) is the test battery used to screen applicants for entrance into Air Force officer corps. Except for Air Force Academy cadets, scores from this test are used by various accession sources as a key element in deciding to accept or reject applicants into their commissioning programs. These primary accession sources include the Reserve Officer Training Corps (ROTC – in colleges and universities around the country/world), the Officer Training School (OTS at Maxwell AFB in Alabama), and the Airman Education and Commissioning Program (AECF).

Historically, the AFOQT was developed and validated by the Air Force Human Resources Laboratory and its successors – Armstrong Lab and the Air Force Research Laboratory. The AFOQT is now managed by the AFPC Military Testing Section with updates and revalidations done under contracts funded by Headquarters, Air Force, Directorate of Personnel, Plans and Programs (HQ AF/DPX) - the Air Force Personnel Operations Agency (AFPOA). The current work on the AFOQT will be discussed below.

Pilot & Navigator Selection Tests

While the AFOQT is used to screen potential officers, there are specialized, computer-based tests to help select pilots. These scores along with the AFOQT and evaluation of flying history are combined to produce a Pilot Candidate Selection Method (PCSM) score used to accept individuals into flying training programs. (<http://www.aetc.randolph.af.mil/SAS/pcsm/>) The current test battery is called the BAT (Basic Attributes Test) and a replacement system, TBAS (Test of Basic Aviation Skills) is under development and evaluation. These batteries are interactive, cognitive and psychomotor tests. The management model for the BAT and TBAS are similar to the AFOQT in that they are Air Force initiated (with contractor development support) and Air Force administered and scored. They differ from the AFOQT in that new computer development work (on TBAS) is being done in-house by the Air Education & Training Command, Studies and Analysis Squadron (AETC/SAS) with validation efforts supported by contract.. Because all other personnel tests use optical scan answer sheets, the world-wide deployed BAT has the additional issue of computer hardware upgrade and maintenance. The maintenance on the BAT computers is supplied world-wide under a contract administered by AETC/SAS. (<http://www.aetc.randolph.af.mil/SAS/pcsm/BatLocations/BatLocations.htm>)

Recent Cooperation in Emerging Selection Methods

Many changes are in progress which will affect the demands on Air Force pilots and navigators – both in the weapons systems they fly and in the operational environment within which they will operate. Advance thought is being given to needed adjustments in both the selection methods as well as the qualities of the ideal aircrew. While it may be too early to institute new policies, it is the right time to begin to collect data that may become the bedrock for future selection systems. In addition, looking forward to a new, more highly integrated environment of not just multi-service but joint international operations, working groups are being formed to search for common approaches to selection and training.

This cooperation is taking many forms. In fact, the new Air Force TBAS software was built by AETC/SAS after receiving and reviewing an early Navy prototype with a similar purpose. The Air Force-developed source code for the TBAS has been provided to the Naval Operational Medicine Institute (NOMI), the organization responsible for the Navy's Aviation Selection Test Battery (ASTB) for their evaluation in experimental testing for a revision of the ASTB. Related to this effort, the Air Force Military Testing Section has had interactions with other Navy personnel and contractors supporting this effort. These other contacts include personnel from such offices as the Naval Aerospace Medical Research Laboratory (NAMRL), the NAVAIR Orlando, Training Systems Division, and the Navy Personnel Research, Studies, and Technology (NPRST) organization. In one such instance, the Air Force provided testing materials to NOMI for the current AFOQT (Form Q) so that naval cadets could be tested as a benchmark in comparing the relationship between the ASTB and the AFOQT. The Air Force scanned the answer sheets and provided the results back to the researchers. The Navy shared the results of their testing and offered to share the ASTB source code so that the Air Force can explore possible adaptation to their own emerging pilot selection practices.

In another example, a Navy representative from NAVAIR, while here for meetings on Navigator training, gave the Military Testing Section a briefing on the Navy's initial efforts to perform a job analysis of pilots and navigators in preparation for a future revision of the ASTB. Air Force personnel advised the Navy visitor that a pilot task analysis had already been done as part of a 1999 Joint Strike Fighter (JSF) training analysis project. That project did a global training needs survey with Air Force, Navy, and Marines, as well as the United Kingdom's (UK) Royal Navy (RN) and Royal Air Force (RAF) pilots. The data and final report for the JSF project were located, organized and forwarded to the Navy visitor, along with contact information for other Navy personnel who were active in that project. In addition, it was pointed out that the Air Force developed the Internet Survey tool used in that JSF project and it is the same program in operational use by the Air Force Occupational Measurement Squadron. This software (GenSurv/MTISurv) was specifically designed for large-scale occupational surveys. The Navy visitor was given contact information to obtain that government-developed software for their anticipated survey. He has since made contact and expects to use that survey tool.

This level of cooperation extends beyond the limits of pilot & navigator testing and selection. The Military Testing Section is a charter member in the new working group on Enlisted Promotions and Testing with members from not only the Army (who suggested the group), but also the Navy and Coast Guard. A hot topic here is computerized promotion testing as well as service differences on promotion policy. An active coordination web page allows quick interactions and a quick way for replacement/new personnel to "catch up."

The Military Testing Section also provides support to an AETC Memorandum of Agreement with the Federal Aviation Agency (FAA) for the development of automated selection tools for screening air traffic controllers (ATCs) as well as one with NASA to administer and score the AFOQT for candidate astronauts. In addition to attending international conferences (such as this one), the Military Testing Section has had productive contacts with the Psychology Department at UK RAF College and responded to inquiries about our normal testing procedures such as the one from the embassy of the Republic of Korea.

Form “S” - Recent Developments in the AFOQT

Along with the development of the new AFOQT Form “S” instrument, a new evaluation was conducted for using revised metrics in the standard AFOQT composites (Verbal, Quantitative, and Academic) and updated Pilot and Navigator composites.

Reduction in Administration Time

In response to the changing demands on Air Force officers in general, a new form of the AFOQT has been developed and is ready to be deployed. This version, Form “S,” was under development for four years and represents a streamlined instrument with experimental items for future implementation. While the current version is Form “Q,” an experimental research version (Form “R”) allowed sufficient analysis to demonstrate that the same factor structure and comparable reliability was achievable using only 11 of the 16 current tests. This demonstration provided the justification to produce Form “S” which only takes 3½ hours to administer, rather than the original 4 ½ hours. Testing now takes LESS THAN a half day and provides new, strategic, forward-looking measures sure to be relevant in future selection, assignment and/or career development systems.

Same AFOQT Factor Structure, Fewer Tests

Subtests included in Form Q are Verbal Analogies (VA), Arithmetic Reasoning (AR), Reading Comprehension (RC), Data Interpretation (DI), Word Knowledge (WK), Math Knowledge (MK), Mechanical Comprehension (MC), Electrical Maze (EM), Scale Reading (SR), Instrument Comprehension (IC), Block Counting (BC), Table Reading (TR), Aviation Information (AI), Rotated Blocks (RB), General Science (GS), and Hidden Figures (HF). Compare Tables 1 and 2 for test content and schedule differences.

Revised Pilot Composite

AFOQT subtests in the current pilot composite (AFOQT Form Q) are VA, IC, BC, TR, AI, MC, EM, and SR. The subtest scores are simply added (i.e., are unit weighted) to compute the pilot composite score. The last three subtests are not included on Form S.

There were data available for 216 Undergraduate Pilot Training (UPT) eliminees (“elims”), along with the data from 1,925 UPT pilot graduates and 287 navigator graduates to use in determining the effectiveness of AFOQT subtests in predicting success in graduating from UPT. Excluded from the analyses were 50 individuals who had been tested more than once.

This study recommends retaining three of the subtests, IC, TR, and AI, but replacing VA and BC with AR and MK. Notably, this reduces the importance of verbal aptitude while increasing the emphasis on quantitative skills. Later analyses dropped the BC subtest.

The validity coefficients for the recommended pilot composite are higher for Form S than those for Form Q and yet fewer subtests are used. Using the dichotomous “graduation/elimination” decision as a criterion, the uncorrected coefficients for Form S are .13, and those for Form Q are .10. Using the Total Flying Training Course Score (T-37) as a criterion, the uncorrected coefficients for Form S are .35, and those for Form Q are .33. Note that the Form Q subtests are unit weighted, while the recommendation accepted for Form S is for regression weighted subtests.

Table 1. AFOQT Forms Q1 and Q2 Testing Schedule and Structure

Subtest	Test/Activity Description	Number of Items	Testing Time ^a in Minutes	AFOQT Composites				
				Pilot	Nav-Tech	Academic	Verbal	Quantitative
	Pre-Test Activities		24					
1	VA – Verbal Analogies	25	9	X		X	X	
2	AR - Arithmetic Reasoning	25	30		X	X		X
3	RC* - Reading Comp	25	19			X	X	
4	DI* - Data Interpretation	25	25		X	X		X
5	WK - Word Knowledge	25	6			X	X	
6	MK - Math Knowledge	25	23		X	X		X
7	MC* - Mechanical Comp	20	23	X	X			
8	EM* - Electrical Maze	20	13	X	X			
9	SR* - Scale Reading	40	18	X	X			
10	IC - Instrument Comp	20	9	X				
11	BC - Block Counting	20	5	X	X			
12	TR - Table Reading	40	9	X	X			
13	AI - Aviation Information	20	9	X				
14	RB - Rotated Blocks	15	15		X			
15	GS - General Science	20	11		X			
16	Hidden Figures (HF)	15	10		X			
	*=Scheduled for removal in Form “S”							
	Collection of Materials & Break		12					
TOTAL		380	270					
	Testing Time: 4 hr 30 mins							

^aSubtest times listed include subtest directions and test performance.

Table 2. Proposed Shortened AFOQT S1 & S2 Testing Schedule and Structure

Subtest	Test/Activity Description	Number of Items	Testing Time ^a in Minutes	AFOQT Composites				
				Pilot	Nav-Tech	Academic	Verbal	Quantitative
	Pre-Test Activities		24					
1	VA – Verbal Analogies	25	9		X	X	X	
2	AR - Arithmetic Reasoning	25	30	X	X	X		X
3	WK - Word Knowledge	25	6			X	X	
4	MK - Math Knowledge	25	23	X	X	X		X
5	IC - Instrument Comp	20	9	X				
6	BC - Block Counting	20	5		X			
7	TR - Table Reading	40	9	X	X			
8	AI - Aviation Information	20	9	X				
9	GS - General Science	20	11		X			
10	RB – Rotated Blocks	15	15					
11	HF - Hidden Figures (HF)	15	10					
12	SDI - Self-Description Inventory	220	40					
	Collection of Materials & Break		12					
TOTAL		470	213					
	Testing Time: 3 hr 33 mins							

Revised Navigator-Technical Composite

AFOQT subtests in the current NAV-Tech composite (AFOQT Form Q) are AR, DI, MK, MC, EM, SR, BC, TR, RB, GS, and HF. As with the pilot composite, subtests are unit weighted to compute the composite score. Subtests DI, MC, EM, and SR are not included on Form S and VA was added to the NAV-Tech composite. RB and HF, though no longer in any composite, continue to be collected as baseline for future study like the experimental SDI. The validity coefficients for the recommended navigator composite are higher for Form S than those for Form Q and yet the number of subtests used is less. Using the Total Flying Navigator Training Course Score as a criterion, the uncorrected coefficients for Form S are .435, and those for Form Q are .33. Note that the Form Q subtests are unit weighted, while the approved method for Form S is to use regression-weighted subtests.

Experimental Tests Considered

Ten candidate tests in the cognitive domain and 10 in the non-cognitive domain were identified. These are shown in Table 3 with designations of the most relevant target populations of Air Force officers: Pilots (P), Navigators (N), or officers in general (O).

The twenty test concepts were presented to the Air Force for review. Of the original proposals, four test concepts were selected for further development. The selection was based on consideration of job requirements and attributes for successful officers, taking into account the coverage provided by the operational tests already included in the AFOQT.

The two cognitive tests selected were Figure Analogies and Abstract Reasoning. The title of Abstract Reasoning test was later changed to Missing Figures to be consistent with subtest naming conventions in the operational battery. Data were collected on experimental forms and showed promise for inclusion in future instruments. These tests are still being evaluated for future use, especially for validity, incremental validity, and fairness.

Table 3. Candidate Officer Selection Tests/Concepts

Test	Cognitive	Target Group	Test	Non-Cognitive	Target Group
1	Logic Gates	O	11	Management Decision	O
2	Figure Analogies	O	12	Conscientiousness/Integrity	O
3	Weather Comprehension	P	13	Team Functions (Team Orientation)	P
4	Chart Reading	N	14	Organizational Commitment (Service Orientation)	O
5	Abstract Reasoning	O	15	Reaction to Stress	P, N
6	Symbol Coding	N	16	Motivation to Fly	P, N
7	Deductive Reasoning	O	17	Impulsiveness	P, N
8	Spatial Assembly	P	18	Competitiveness vs. Cooperation	P,N
9	Word Discrimination	O	19	Motivation to Manage	O
10	Decoding Operations	N	20	Social Skills	O

Non-Cognitive Test History

The two non-cognitive measures selected by the Air Force for further development were labeled Service Orientation and Team Orientation. The U.S. Air Force has a history of personality test development extending back to the mid-1950s. Results of this early work led to the identification of five recurring personality factors which later became known as the Big Five (Tupes & Christal, 1958). The factors were characterized as Conscientiousness, Agreeableness, Neuroticism, Openness, and Extroversion. Later work outside the military verified the ubiquitous nature of the factors across a broad range of personality tests and subject populations

(Digman, 1989; Costa & McRae, 1980, 1985; Goldberg, 1990). Anesgart and Callister (1999) found that completion of pilot training in the Air Force was related to both Neuroticism and Extroversion. Candidates who were high on Neuroticism and low on Extroversion were more likely to self-eliminate than their counterparts.

Support for the utility of measures of personality for personnel selection is widespread. In a broad meta-analysis of selection methods for many occupations across an 85 year period, Schmidt and Hunter (1998) examined the validity of general mental ability (*g*) and 18 other selection procedures for predicting training and job performance. They concluded that the personality construct of integrity (or conscientiousness) significantly added to the predictiveness of cognitive ability for predicting training and job performance.

Christal (1994) revisited his earlier work and constructed a contemporary inventory suitable for computer administration. Working with adjective clusters defined by Goldberg (1990), Christal assembled an inventory comprised of single-word trait descriptions and more lengthy behavioral statements. Analyses showed that inventory captured the well-documented five-factor personality structure. The inventory was labeled the Self Description Inventory (SDI).

In the present effort, the Air Force sponsored work to bring a Big Five personality inventory nearer to operational implementation and to extend the traditional personality measures in new directions by measuring traits deemed relevant to officer selection. The new traits explored were Service Orientation and Team Orientation. This extended SDI is a 220 item inventory and forms the last section, part 12, of AFOQT Form "S."

The Service Orientation (SO) test grew out of a perceived need by Air Force senior leadership to assess an officer applicant's potential for organizational commitment prior to service entry. Hence, the assessment focused on measuring predispositions for recognizing that organizational goals sometimes supersede an individual's natural striving for personal benefit. People differ on their inclinations for this capacity from totally egocentric to those quite willing to invest whatever is required by the organization with little thought about the personal consequences that might entail. Descriptions of egocentric behaviors would include corporate executives who provide generously for themselves at the expense of the corporation or its stockholders. The loyalty shown to the organization would often be superficial and dependent on personal gains rather than on genuine concern for other organization members or the organization as a whole. People high in Service Orientation would typically be more inclined toward enduring self-sacrifice if the good of the organization was at stake or if self-subornation was perceived as a way to achieve organizational goals.

The Team Orientation (TO) test, which was formerly called Team Functions, was designed to assess predispositions for working comfortably in groups versus preferences for working alone. Items focused on the respondent's preferences and capacity for working in groups to achieve organizational goals. The test items asked whether group activities were a preferred mode of work or whether the respondent found working alone as generally more suitable. Prototypical "high TO" respondents would be expected to be comfortable in groups, able to lead as well as to follow, and be capable of focusing on team goals and of compromising on specific methods of achieving overall goals when group cohesion required it. People low on TO would be uncomfortable working closely with others, impatient with other team members and prefer to rely on themselves exclusively rather than on other team members.

Summary

These are exciting times in the Military Testing Section – new tests, new technology, new partners, and new visions. The deployment of AFOQT Form “S” begins a data collection effort which will last for years. While training success offers good clues, real validation must wait for actual track records of success in real officer careers.

References

- Alley, W. E. (2004) *Experimental Test Development for the AFOQT Form “S”*. Draft Report. Operational Technologies Corporation (OpTech), San Antonio, TX.
- Anesgart, M. N., & Callister, J. D. (1999). Predicting training success with the NEO: The use of logistic regression to determine the odds of completing a pilot screening program. *Proceedings of the 10th International Symposium on Aviation Psychology*. Columbus, OH: The Ohio State University, 565-570.
- Berger, F. R., Gupta, W. B., Berger, R.M., & Skinner, J. (1988). *Air Force Officer Qualifying Test (AFOQT) Form P: Test construction* (AFHRL-TR-88-30, AD-A200 678). Brooks AFB, TX: Air Force Human Resources Laboratory.
- Berger, F. R., Gupta, W. B., Berger, R. M., & Skinner, J. (1990a). *Air Force Officer Qualifying Test (AFOQT) Form P: Test Manual* (AFHRL-TR-89-56, AD-A221 004). Brooks AFB, TX: Manpower and Personnel Research Division, Air Force Human Resources Laboratory.
- Berger, F. R., Gupta, W. B., Berger, R. M., & Skinner, J. (1990b). *Prototypes of cognitive measures for Air Force officers: Test development and item banking*. (AFHRL-TP-89-73). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Carretta, T. R., & Ree, M. J. (1996). Factor structure of the Air Force Officer Qualifying Test: Analysis and comparison. *Military Psychology*, 8, 29-42.
- Christal, R. E. (1994). *R&D summary report on contract F33615-91-D-0010*. Unpublished final status report for Delivery Order 0010 for the period Feb 93 to Nov 94. Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Costa, P. T. Jr., & McCrae, R. R. (1980). Constancy of adult personality structure in males: Longitudinal, cross-sectional and times-of-measurement analyses. *Journal of Gerontology*, 35, 877-883.
- Costa, P. T., Jr. & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Digman, J. M. (1989). Five robust trait dimensions: Development, stability, and utility. *Journal of Personality*, 57, 195-214.
- Glomb, T. M., & Earles, J. A. (1997, July). *Air Force Officer Qualifying Test (AFOQT): Forms Q Development, Preliminary Equating and Operational Equating* (AL/HR-TP-1996-0036). Brooks AFB, TX: Human Resources Directorate, Armstrong Laboratory.
- Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216-1229.
- Gould, R. B. (1978). *Air Force Officer Qualifying Test Form N: Development and standardization* (AFHRL-TR-78-43, AD-A059 746). Brooks AFB, TX: Air Force Human Resources Laboratory.

- Gupta, W. B., Berger, F. R., Berger, R. M., & Skinner, J. (1989) *Air Force Officer Qualifying Test (AFOQT): Development of an item bank*. (AFHRL-TR-89-33, AD-A216 228). Brooks AFB, TX: Air Force Human Resources Laboratory.
- Ree, M. J. (2003). *Test of Basic Aviation Skills (TBAS) – Scoring the Tests and Compliance of Tests with the Standards of the American Psychological Association*. Final Report. . Operational Technologies Corporation (OpTech), San Antonio, TX.
- Rogers, D. 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, AD-A173 484). Brooks AFB, TX: Air Force Human Resources Laboratory.
- Rogers, D. L., Roach, B. W., & Wegner, T. G. (1986). *Air Force Officer Qualifying Test Form O: Development and standardization* (AFHRL-TR-86-24, AD-A172 037). Brooks AFB, TX: Air Force Human Resources Laboratory.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Shanahan, F. M., & Kantor, J. E. (1986). *Basic navigator battery: An experimental selection composite for undergraduate navigator training* (AFHRL-TR-86-3, AD-A168 857). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Shore, C.W, Gould, R.B., Ree, M.J., Alley, W.E., Skinner, J. (2002). *Reduction of AFOQT Administration Time*. Final Report. . Operational Technologies Corporation (OpTech), San Antonio, TX.
- Shore, C.W, Gould, R.B. (2003). *Developing Pilot and Navigator/Technical Composites for the Air Force Officer Qualifying Test (AFOQT) Form S*. Final Report. . Operational Technologies Corporation (OpTech), San Antonio, TX.
- Skinner, M.J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analysis of Form O* (AFHRL-TR-86-68, AD-A184 975). Brooks AFB, TX: Air Force Human Resources Laboratory.
- Steuck, K. W., Watson, T.W., & Skinner, J. (1988). *Air Force Officer Qualifying Test (AFOQT): Forms P pre-implementation analyses and equating* (AFHRL-TP-88-6, AD-A201 100). Brooks AFB, TX: Air Force Human Resources Laboratory.
- Tupes, E. C., & Christal, R. E. (1958). *Stability of personality trait ratings obtained under diverse conditions* (WADC-TN-58-61, AD-151 041). Lackland AFB, TX: Personnel Research Laboratory, Wright Air Development Center.
- Tupes, E. C. & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings*. (ASD-TR-61-97 AD-267 778). Lackland AFB, TX: Personnel Research Laboratory, Wright Air Development Division.
- Tupes, E. C. & Christal, R. E. (1992). *Recurrent personality factors based on trait ratings*. Reprint of original technical report in Special Issue on the Five-Factor Model: Issues and Applications, *Journal of Personality*, 60, 2, 225-251.
- Ziebell, T.S., (Nov 1989). Policy Issues for Air Force Personnel Testing Programs. *Proceedings of the 31st Annual Conference of the Military Testing Association*, (pp. 554-558). San Antonio, TX: U.S. Air Force.

Recent Developments in USAF Officer Testing and Selection

Johnny J. Weissmuller, Kenneth L. Schwartz, and
Stanley D. Kenney – USAF Military Testing
Randolph AFB, Texas USA

C. Wayne Shore and R. Bruce Gould – Operational Technologies
(OpTech)
San Antonio, Texas USA

Testing in the USAF

- Education & Training Applications

Personnel Selection &
Promotion Applications

To increase:

To manage:

- Individual Competency
- Operational Cooperation
- Mission Coordination

Acceptance into Service
Promotion in Pay Grade
Increase in Authority

- Operational Readiness

Personnel Management

USAF MILITARY TESTING Section

HQ USAF Personnel Center

- Manage the USAF program for personnel tests for selection and promotion:
 - Recommend test research
 - Evaluate new technologies
 - Validate new/proposed tests
 - Establish operational testing policy
 - Coordinate Test Control Officer network

USAF MILITARY TESTING Section

Types of Tests Managed

- USAF Officer Qualifying Test (AFOQT – all officer, pilots, and navigators)
- All USAF Enlisted Promotion Tests
 - **Promotion Fitness Exam (PFE)**
 - **USAF Supervisory Exam (USAFSE)**
 - **130+ Specialty Knowledge Tests (SKTs)**
- Special selection tests (programmers, etc.)
- DoD controlled tests – Armed Services Vocational Aptitude Battery (ASVAB) and Foreign Language Aptitude & Certification Tests

USAF MILITARY TESTING Section

Coordination & Cooperation

US Navy –

- AFOQT reference testing for new automated Navy pilot selection test
- Internet survey tool sharing with Navy of USAF job survey tool for use in pilot/nav job analysis

Interservice Enlisted Testing & Promotion Working Group (Army, AF, Navy, Coast Guard)

Federal Aviation Agency (FAA) – Automated Selection Tool for Air Traffic Controllers (ATCs)

NASA – AFOQT screening for all astronaut applicants

US Air Force Officer Qualifying Test (AFOQT)

- Forms of this test have been in use since the 1950s
- The current operational form, Form Q, has been in use 10 years (1994)
- A 4-year research effort just finished (using Research Form R) by OpTech
- Results – a shorter, equally powerful version S was constructed and augmented to include a “Big Five” personality inventory (SDI)

US Air Force Officer Qualifying Test (AFOQT)

- Previously maintained by the Air Force Human Resources Laboratory (AFHRL), its predecessors & successors until 1999
- New version motivated by need for faster, better, less expensive form
- Form S can be administered in 3½ hours instead of 4½ hours (under ½ day)
- Includes benchmark and research subtests for future applications
- Form S will be scored using new scanner equipment greatly increasing throughput with no loss in accuracy.

The US Air Force Officer Qualifying Test (AFOQT)

- A complete listing of the subtests in Version Q and Version S are given in the paper. The paper will be in the proceedings and is currently available at:
- <http://www.icodap.org/041025>
- In summary, 16 subtests were used in Form Q. These 16 subtests are combined to create FIVE (5) composites.

AFOQT Summary

- The five composites are:
 - Pilot
 - Navigator
 - Academic
 - Verbal
 - Quantitative

AFOQT Form S Summary

- Reduced subtests from 3 to 2 per factor dimension – examples:
- Verbal composites:
 - Form Q:
 - VA – Verbal Analogies
 - RC – Reading Comprehension
 - WK – Word Knowledge
 - Form S:
 - VA – Verbal Analogies
 - WK – Word Knowledge

AFOQT Form S Summary

- Reduced subtests from 3 to 2 per factor dimension – examples:
- Quantitative composite:
 - Form Q:
 - AR – Arithmetic Reasoning
 - DI – Data Interpretation
 - MK – Math Knowledge
 - Form S:
 - AR – Arithmetic Reasoning
 - MK – Math Knowledge

AFOQT Form S Summary

The Pilot and Navigator composites used eight (8) and eleven (11) subtests in Version Q

In Form S, the Pilot and Navigator composites will now use five (5) and six (6) subtests

AFOQT Form S Summary

After validation to training outcomes in undergraduate flying training (UFT), the pilot composite is now computed from five (5) subtests:

AR – Arithmetic Reasoning

MK – Math Knowledge

IC – Instrument Comprehension

TR – Table Reading

AI – Aviation Information

AFOQT Form S Summary

The subtests dropped from the pilot composite are:

VA – Verbal Analogies

MC – Mechanical Comprehension

EM – Electrical Maze

SR – Scale Reading

BC – Block Counting

AFOQT Form S Summary

The navigator composite is now computed from six (6) subtests:

VA – Verbal Analogies

AR – Arithmetic Reasoning

MK – Math Knowledge

BC – Block Counting

TR – Table Reading

GS – General Science

AFOQT Form S Summary

The subtests dropped from the navigator composite are:

RC – Reading Comprehension

DI – Data Interpretation

MC – Mechanical Comprehension

EM – Electrical Maze

SR – Scale Reading

RB – Rotated Blocks

HF – Hidden Figures

AFOQT Form S Summary

Experimental Tests

20 candidates – selected 3 –

Big Five factors of:

Personality

Service Orientation

Team Orientation

AFOQT Form S Summary

Big Five Factors of Personality

- Discovered by USAF in 1950s at Lackland AFB, 25 miles from HQ AFPC
- Early factor analysis studies by Tupes & Christal
- Adopted and disseminated in private sector
- In 1990s; USAF reevaluated, researched, validated on officer and enlisted force – coordinated research with the University of Plymouth, UK
- Developed the “Self-Description Inventory” or SDI for operational use in Form S

AFOQT Form S Summary

Self-Description Inventory

- 220 items include Big Five plus new dimensions for Service Orientation (SO) and Team Orientation (TO)
- Big Five contains (OCE AN) –
 - Openness
 - Conscientiousness
 - Extroversion
 - Agreeableness
 - Neuroticism

AFOQT Form S Summary

Research SubTests

- The following subtests are collected and scored, but are NOT currently used in any selection decisions:
 - Rotated Blocks
 - Hidden Figures
 - Self-Description Inventory
- Future use of these data are expected in long-term validations.

Use of Personality Tests

- The problem – test compromise –
unidimensional tests compromised by strategy
leaks in under 2 minutes
- The Big Five approach –
 - Multiple dimensions
 - “More” is NOT better (neuroticism or too much openness)
 - Multiple uses/multiple profiles

Use of Personality Tests

- Possible uses:
- Selection (obvious but unlikely)
- Assignments –
 - Best match to position for productivity
 - Lessen personal stress in jobs requiring technical development
 - Lessen organizational stress in accepting new managerial individuals
 - Best match to position for personal development or broadening
- Predicting retention

Questions?