

AFCAPS-TR-2012-0003

**Self Description Inventory
Plus Initiative: Assault on
Occam's Razor**

IMTA Conference
October 2007



Johnny J. Weissmuller
Kenneth L. Schwartz

AFPC/Strategic Research and Assessment
Branch



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

Available 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)		2. REPORT TYPE Final		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Self-Description Inventory Plus Initiative: Assault on Occam's Razor				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Johnny J. Weissmuller and Kenneth L. Schwartz				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Personnel Center Strategic Research and Assessment Branch 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-TR-2012-0003	
12. DISTRIBUTION / AVAILABILITY STATEMENT Available for public release. Distribution Unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Purpose of the paper is to outline military testing and selection and focus on the need to develop secondary measures, like the Self Description Inventory (SDI+), in order to increase efficiency of person-job match and develop realistic profiles of applicants and military positions beyond the typical job-performance measures. Covers the history of military selection using measures like the Armed Services Vocational Aptitude Battery (ASVAB) and Air Force Officer Qualifying Test (AFOQT), along with the history of the SDI stemming from the Learning Abilities Measurement Project (LAMP). Paper discusses various problems of traditional military selection procedures based on changing population of applicants, and the idea that Occam's Razor may not be the solution to future Air Force personnel selection and classification issues. Suggested solution of this paper is the development of personality measures to capture and form a "profile" of applicants before classification and assignment are done to ensure better assignment, satisfaction, and retention.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT U	18. NUMBER OF PAGES 55	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)

Self-Description Inventory Plus Initiative: Assault on Occam's Razor

Johnny J. Weissmuller and
Kenneth L. Schwartz
HQ Air Force Personnel Center
Force Management Liaison Office
Randolph AFB, Texas

Abstract

In the late 1950's the Air Force conducted a factor-analytic meta-analysis of existing data sets for personality models. What they identified have become known as the "Big Five" (BF) factors of personality (BF acronym - OCEAN=Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism). The Self-Description Inventory Plus (SDI+) is an expanded personality assessment instrument designed to *cover the waterfront* of normal personalities for use in today's Air Force personnel systems. The "Plus" adds factors for "Service" and "Team" orientations yielding the new acronym of OCEAN ST.

One of the unique challenges in giving military recruits a choice of careers is that many lack a realistic knowledge of what those jobs entail. Recruits choose a career field from a list of available Air Force Specialty Codes (AFSCs). Each AFSC, however, represents an entire career field and may be suitable to many different kinds of personalities. Hence, rather than trying to find "the perfect profile" for each AFSC, the goal of the SDI+ Initiative is to *identify all profiles* which show promise for an Air Force career choice. A *profile with promise* is one with a positive self-reported job satisfaction from current valued workers already on the job. With this information the Air Force can advise new recruits on the best initial career choices and improve long-term retention by assisting reenlisting service members in planning strategic career directions.

Last minute update...

What follows is not the paper I envisioned presenting today. Operational demands rightly take precedence over research. In April, I requested that the 12,000 answer sheets with the Self-Description Inventory Plus (SDI+) response data be scanned for this analysis. The contractor hired to do the work officially began on 1 Oct – three days before my departure to this, the 49th Annual Conference here in Australia. For this reason, rather than reporting on the results of a relatively unique analysis of what is now 20,000 "Big Five" personality inventories on the shelf, I am going to discuss the structure of our current selection and classification system and the challenges ahead in integrating use of personality data by the United States Air Force. This paper is intended as a concept paper to encourage discussion and feedback, not to announce a well-formulated plan under way.

Informal Opening Remarks

As I approach age 60, I am a relatively “old-timer” with International the Military Testing Association (IMTA). My first attendance was 34 years ago at the 1973 Conference in San Antonio, Texas – of course, my role there was as a sergeant flipping transparencies for the real attendees. In all those years I’ve not only had access to full IMTA proceedings dating back to 1959, I’ve also had the chance to meet and work with some of the best minds in our business. I have been personally mentored by a few in the United States Air Force. Indulge me, if you will, with a more personal and conversational style of presentation that was common in the early 1960s from whence my primary sources come.

In today’s presentation, I arrive with less current data than I desire, but with a historical perspective and a clear vision for direction I believe we need to consider for a more mission-ready military in the future. Because of the unique United States Air Force personnel conditions ahead, these comments are applicable for the next 4-10 years.

Being that this is a symposium on “Personality,” I thought I would open with a few anecdotes which prove I have very little personality of my own. I poured over my extensive lists of funny stories – a whole handful – if you don’t count the thumb... And, to my surprise, I found two which are relevant to my presentation today and reflect the fact that my focus really is on “test and measurement” issues.

Abraham Lincoln and the Dog’s Tail

The first story comes from Abraham Lincoln, our 16th US President, but from early days as a lawyer and an invited speaker for many occasions. In the presentation at hand, Honest Abe, as he was known, asked a question of the audience... “Assume you have a faithful pet dog who is in perfect physical condition...Do you have that animal in mind?”...There were some murmurs in the audience from cat lovers, but he continued.. “Now, if one were to say, for the sake of argument, that the dog’s tail was a “leg,” then how many “legs” does this dog have?”

Abe stepped back from the podium and immediately side arguments broke out in the audience. It was clear that were two camps of thought and it had nothing to do with cat- or dog- preferences. From where he stood, Abe could overhear the arguments on both sides. One side said, “Well, if you call a tail a leg, then the dog has FIVE legs.” People on the other side said, “No. A tail is not like a leg. If you call the tail a “leg” then the dog only has ONE leg and we must rename the old legs to be called something else.” When the arguments became overheated, Abe stepped back up to the podium and raised his arms to quiet the discussions. Abe said, “The point here is that we believe too much in the reality of the labels we use. In point of fact, the dog still has only FOUR legs. *Calling* it a “leg” doesn’t MAKE it a leg.”

Remember this when we get to the discussion about CONSTRUCT VALIDITY later on.

The Lamp Post and the Circle of Light

The second story supposedly took place some eighty years ago in the United States. Although as the story was told to me it took place in Boston outside of a tavern, I've replaced the stereotypical characters by "just plain folks" for presentation today....

Visualize that it's a clear starry night at 2 o'clock in the morning on what is now Saturday for two hours. We're just off the city square and a helmeted police officer with a fine, full mustache is walking his beat. His billy club (night stick) in hand, he slowly taps his leg in sync with his casual stroll. His progress is steady in and out of the yellow circles illuminated on the ground from the city lamp posts.

As he rounds the corner, the officer is startled to see a person on his hands and knees two light circles away he runs to offer aid. Upon arriving in the circle of light, the officer offers to help the citizen to his feet. The citizen replies, "No. no I'm not injured. I just dropped my keys and I'm on my hands and knees just trying to find them..."

Being very observant, the officer quickly scans the well-lit circle and informs the citizen, "Sir, I can see that your keys are not here. Perhaps you dropped them before you came into this circle from which direction did you come?" The citizen looked at the police officer dourly and pointed far off to a dark alley across Main Street from the city square. The police officer inquired further, "Well, sir, exactly where were you standing when you dropped your keys?"

The citizen replied, "I just told you... over there, in the alley. I was trying to get my key into the lock of my car in the alley, and they fell from my hand." The officer thoughtfully tapped the tip of his billy club to his chin, carefully considering his next question... "Sir, if you dropped your keys across the street in the dark alley, why are you looking for them here?"

The citizen looked the police officer directly in the eye, as if to lecture a young student... "Why here? Why here? I'd think that would be obvious..." The citizen waited for the officer to respond, which he did not. Finally, the citizen gleefully finished his argument... "Why here?...obviously because the light is so much better over here!"

Remember this when we get to the discussion on THE CRITERION PROBLEM.

Military Selection Testing and Validity Issues

Aptitude Testing and the U.S. Military

Now, to return to the present day United States... To qualify applicants for enlisted military service, our Department of Defense (DoD) uses a standardized testing system known as the Armed Services Vocational Aptitude Battery (ASVAB). After talking with a school counselor or recruiter, potential applicants report to a DoD managed Military Entrance Processing Station (MEPS) for evaluation. The MEPS administers the ASVAB and performs medical, physical, and general psychological screening.

By law, entrance into the military is controlled by an applicant's score on a given subset of ASVAB tests called the Armed Forces Qualifying Test (AFQT). The AFQT is composed only of math and verbal tests. In addition, each service branch (Army, Navy, Air Force, Marine Corps) uses service-specific combinations (composites) of the individual ASVAB test scores to aid in classifying the successful applicants into job families within that service.

In the validation phase of any professionally responsible testing system, test developers ensure that individual tests are appropriate to the pool of applicants to be assessed and that the tests produce scores appropriate to the intended purpose. The purpose of any test is to rank and stack individuals based on how well they perform on the test with the expectation that the score represents a reasonable estimate of probable future success in the designated area.

Based on numerous studies, it has been shown that the ASVAB composites do predict success in the U.S. military. Within the Air Force the standard measure of success has been pass/fail or course scores for completion of technical training. In the late 1980s, the US Congress mandated that the services validate these test scores against measures of on-the-job performance. As a contractor at the time, I was involved in working out these Job Performance Measurement (JPM) protocols first for the US Navy and later for the US Air Force. While the project was very expensive, the relationships between tests scores and job performance were at best, moderate.

Since my return to civil service four years ago I've worked in enlisted promotion testing. While waiting for the SDI+ data to become available I completed a comprehensive study which documents that US Air Force classification composites are positively related to promotions in mid-career grade levels (E-5 through E-7). These are people being promoted by a constrained compensatory model. It's a relatively objective promotion system in which promotions occur from 4 to 14 years after the member took the ASVAB test. The strength of the ASVAB and promotion relationship drops off as the grade level increases – a natural effect once everyone remaining has passed the objective promotion system multiple times. As discussed later, less than acceptable performers aren't allowed to reenlist or test for promotion and hence "refining the force" means restricting the range is the active goal of the personnel system.

Linkage of score performance for classification purposes is an on-going validation process. In the Air Force, for example, we are just concluding a contract to evaluate whether or not our four present composites should be replaced by ten new composites. Our four long-standing composites measure aptitude in Mechanical, Administrative, General, and Electronic areas. Each of our 150 job families has established minimum scores for entrance on one or more of these four composites. The basis of the current evaluation is increase in the pass rate expected in technical training. We will be briefed next month.

Testing and Population Changes

Over time, however, both the ability levels of the applicant pool change and the job demands in the military evolve. ASVAB scores are reported as percentile standing in a reference or norm population sample. The ASVAB norms are based on a large-scale measurement of the ability levels (norms) in the Population of American Youth (PAY) which were conducted in 1946, 1980, and, most recently in 1997.

At the same time that the ASVAB developers sensed the need for recalibration of the abilities in American youth in the late 1970s, the Air Force began to address a related issue – test score validity for applicants with diverse educational or cultural backgrounds. Aptitude testing (for verbal and numerical ability, at least) is predicated on an assumption of a common history of access to traditional educational sequencing and an equally supportive culture in the home, school, and community – things that are less and less true each year in our evolving multi-cultural society.

The Learning Abilities Measurement Program (Project LAMP)

In the early 1980s, U.S. Air Force interest was rekindled in alternative measures for applicants from other than the mainstream background. The Learning Abilities Measurement Program (Project LAMP – 1981-1998) set the goal to identify and assess qualified applicants who failed to meet minimum cutoffs on standardized tests. Project LAMP brought in world-class researchers and established liaisons with psychological research centers and universities around the world.

The LAMP Premise

The Air Force training program for the enlisted force takes high school graduates and transforms them into specialists in one of many highly technical fields. While individuals with high ASVAB scores do tend to do better in training programs, this does not necessarily imply that low-scoring individuals are doomed to fail. This was not a new concept. Tri-Service Conferences sponsored by the Office of Naval Research (ONR) in the early in the early 1960s stated we were getting all the unique variance possible from measuring general intelligence (“g”) with our then current aptitude tests and additional variance should be sought elsewhere such as in non-cognitive assessments.

A goal of Project LAMP was to identify people for whom traditional academic testing did not adequately project their full potential value to the Air Force. Hence, in the modern, diverse applicant pool, training and future success in the Air Force might be better predicted, for some, by their ability to learn. If this can be demonstrated, then it becomes a policy question as to whether a “learning ability” measure should be used as a modifier of traditional aptitude scores or, in some cases, as a valid alternative path to accession.

Project LAMP’s Evolving Strategy

Project LAMP began by studying low levels of cognitive processing -- the “snappiness of neurons” in the words of the program creator, Dr. Raymond Christal. This is the same “Christal” of Tupes and Christal who did pioneering work on the Big Five Factors of personality in the late 1950s. It is not surprising then that Project LAMP expanded from measures of learning ability into assessments in the non-cognitive domains of personality.



Dr. Christal became my mentor upon my arrival at the Air Force Human Resources Laboratory (AFHRL) in August 1971 when I took the cross-base bus directly from USAF Basic Training to AFHRL. He was the Chief of the Occupation and Manpower Research Division and I (eventually) was his primary liaison to the Computation Sciences Division – advisor on data in Master Files, Occupational Analysis databases, statistical software, and designer of new software as needed.

The focus of our jobs at AFHRL was on Occupational Analysis – developing reliable ways to describe and measure Air Force jobs to inform policy makers for recruiting, classification, training, promotion, retention and other personnel programs.

During his “mentoring” time, Dr. Christal would reminisce about his work on personality back in the 50’s and 60’s and how that offered the brightest hope for adding reliable information for improving “person-job-match (PJM).” PJM was the Holy Grail that he sought because he believed if the service got the right person in the right classification, the right training program, and the right job, both the person and the Air Force could expect a promising career and a greater return on investment (ROI).

Personality Comes of Age

During those early 70s, we worked on software to capture, report, and apply Occupational Reinforcer Pattern (ORPs) profiles for various occupations from the Industrial Relations Center at the University of Minnesota (IRC Bulletin 48, 1968). I frequently told Dr. Christal that I couldn’t imagine a scenario which would persuade the Air Force to use personality measures in operational personnel programs. Twenty years later (1994), just months before he passed away, I was able to report to Dr. Christal results of a Critical Incident Study of sorties in the first Gulf War --solid Air Force data that documented the very scenario that I believed couldn’t be found – that personality issues affected the delivery of ordnance during times of armed conflict.

Project LAMP and the Self-Description Inventory Plus (SDI+)

By the mid 1990s Project LAMP had begun to collect experimental SDI data on enlisted personnel in Basic Military Training (BMT). These data were collected using evolving forms of the Air Force's SDI personality inventory. The current SDI+ is the result of over ten years (1994-2004) of further research to create a Big Five Model (BFM) instrument with two new synthetic factors ("Service" and "Team" Orientation), producing an expanded acronym: OCEAN ST. The SDI+ went "operational" in July 2005 with its inclusion in the new Air Force Officer Qualifying Test (AFOQT, Form "S"). Although the SDI+ is part of the AFOQT, it is not scored nor used in any selection or classification decision making at present. Its intended use now is envisioned more as a career counseling tool as explained below.

Changes in the Environment

The General U.S. Social Climate

In the fifty years of SDI+ development (1957-2007), the social climate has also continued to change. Had all the recommendations from the early 1960s been implemented, they would have been shortly overturned by evolving US legislation which prohibits federal agencies from establishing differing norms for gender or race. Personality inventories at the time were showing differences along those lines and the standard psychological practice is to establish separate norms for each identified group. In addition, limitations now exist on the kind of questions that may or may not be asked of any job applicant. These restrictions flow down from statutes to regulations by various agencies and include such examples as the Americans with Disabilities Act (ADA), the Equal Employment Opportunity Commission (EEOC), and the Merit Labor Relations Board (MLRB). As reported as recently as 2005 (Murphy, 2005), personality measures are still considered risky for inclusion in formal employment selection systems – although convinced of their value, private sector firms are moving more and more in that direction.

Changes in the U.S. Military Environment

For many years, all "personality" measures were considered too controversial for use in the US military. Prior to 1974, the selective service system brought unwilling people into the military and "personality measures" were only considered in the negative sense -- for screening out undesirables and, at the same time, encouraging "faking" to avoid the draft.

In today's all volunteer force, there is a suspicion that centralized data bases of personality profiles may be used to adversely affect a person's career, whether it be a restriction on the opportunity to enter a particular career field or later on, to deny, without apparent reason, reenlistment. In the past, the concern was about "faking" in order to look bad to be *selected-out*. Today's concern is about "gaming the system" to give the "desired" answers in order to be *selected-in*. The bottom line – if systems can be

manipulated, they will be – by all parties. The goal then is to design a system in which each party pursuing its own long-term best interest works to everyone’s benefit – the Air Force Self-Description Plus (SDI+) Initiative is that kind of design.

Changes in the U.S. Military Applicant Pool

Cross-Generational Differences – an EOD Example

Differences between generations have been noted in the popular literature on both ability and areas of interest. Let me give one military example I observed first hand. As part of a US DoD inter-service team evaluating possible ways to improve the ASVAB, we (Navy and Air Force) toured the Naval Explosive Ordnance Disposal (EOD) school in Pensacola, Florida. Each service trained there (Army, Navy, Marines, and Air Force) has a different procedure for identifying and qualifying students to attend the school. In the US Air Force procedure, potential students must have high scores on both the Mechanical and Electronic aptitude composites based on the ASVAB. This dual requirement for high scores on BOTH tests highly restricts the percent of Air Force recruits who can qualify for these training slots.

One of the truly valuable “best practices” observed at the EOD School that the instructor corps is primarily composed of former students who served successfully in the field and have returned to inculcate the next generation. New students were being taught by people who had made the same career choice and racked up both experience and stories which they relayed as the training progressed. Traditionally, the “high mechanical” aptitude students tend to be less verbal. Trading stories from different instructors encouraged communication and helped to bridge motivation across the generational gap.

The point of this story, however, is an incident which occurred as our team was departing. I noted quite a number of highly polished, immaculately maintained large motorcycles in the parting lot. I remarked that a lot of people here must like motorcycles. One of the instructors replied, “Well, we all do. We’re mechanically inclined and it builds team spirit and a little competition to work on them at the same time.” I asked, “How many of these are student motorcycles and how many belong to instructors?” The instructor looked at me oddly, “These all belong to the instructors.”

I asked further, “Are students not allowed to bring bikes?”

The instructor continued, “Sure they’re allowed to bring bikes or other transportation. Back in my days as a student, the hottest competition was between instructor and student machines.”

I pursued the issue, “So what happened to that tradition?”

He responded, “Now they bring cars they don’t service themselves. As soon as class lets outside, they’re all back in the dorms playing Internet computer games.”

“Ok” I replied – perhaps something has been lost across these generations.

In the late 1990s, the US DoD revised the ASVAB and dropped two speeded tests, Numerical Operations (NO) and Coding Speed (CS), from the test battery. The Air Force composites for both Mechanical and Administrative aptitudes had to be reconfigured.

The new Air Force Mechanical composite increased the relative importance of verbal aptitudes. While I thought this may have contributed to a reduced interest in mechanics by new Air Force EOD students, the change in free-time hobbies was prevalent across all the services.

Broad Generational Characteristics and Time Frames

These generational groups are typically defined in the popular literature by the years spanned at their birth. Wikipedia currently identifies the following “main” generations: 1911-1924 (the Greatest Generation), 1925-1945 (the Silent Generation), 1946-1964 (the Boomers – my generation), 1965-1979 (Generation X), and 1980-Present (Generation Y, the Millennials, or Internet Generation). Our interest, of course, lies in their ages at the time of initial eligibility for military enlistment. Note that the US DoD ASVAB norms were developed on 18 year-olds from samplings in 1946, 1980, and most recently 1997 which correspond to birth years of 1928, 1962, and 1980.

Today’s pool of potential military applicants lies squarely in the Millennial or Internet generation. Studies have shown a marked shift from older patterns and this is not tied to the broadening cultural diversity which is taking place at the same time. The characteristics attributed to Millennials sounds exactly like the military’s description of the future infantryman – fiercely mission-oriented, self-motivated, net-centric, auto-organizing, team oriented, intelligent and situationally aware. The only disconnect is that Millennials tend to distrust authority figures and information providers until they prove their integrity.

Many of the assumptions we make about the validity of our selection tests need to be reexamined in light of these changes in the applicant pool. The US DoD and our services are addressing these issues with a number of studies and initiatives

The Path to Proficiency

As noted as early as 1980 in Project LAMP, high test scores still provide valid predictions about success in training in the military, but it is the source of low scores which should demand our attention. The Millennial or Internet generation is the first to break the fundamental assumption underlying arguments for the face validity of our aptitude tests.

Before the Internet Generation, other than public libraries, access to educational resources remained, for the most part, a de facto monopoly of traditional public and private schools. Success in those institutions confirmed not only achievement but also exposure to fairly standardized curriculum sequencing and conformance with social norms. Failure in those institutions (expulsions or voluntary withdrawals) indicated either inability to master material, abide by the rules, or parental desire for specialized instructional curricula.

Based on empirical grounds, the US DoD recognizes three “levels” or “tiers” for acceptance into the military. Tier 1 is a bona fide high school graduate with diploma. Tier 2 is a non-high school graduate but with a General Educational Development (GED) certificate or certified home-schooled graduate. Tier 3 is a person without a high school education. Standards for acceptance into the military vary based upon the Tier of the applicant. The US Air Force, for example, does not accept Tier 3 applicants.

The Fork in the Road

When the first motivated student logged onto the Internet and found something interesting to read and study, the educational institution’s single path to proficiency encountered a fork in the road. All of a sudden, the brightest and best students (with access) could explore topics of interest to them, when they pleased, and as deeply as they cared to pursue it. When search engines like Google, Ask, Lycos, or Copernic came on the scene, the student’s ability to locate and explore far surpassed the ways and means of the traditional educational systems and their planned curricula.

Finally, “communities of practice” evolved out of “chat room” technologies. At this stage, ordinary students (like my 16 year old daughter five years ago), could have her original stories read, evaluated and critiqued by professional editors who are looking to mentor emerging talent. These interactions do not show up on any school transcript – but her later awards for writing in college attest to its value.

Some believe that all this time alone on the Internet is detrimental to a student’s social development, but that isn’t necessarily so. I was fascinated to hear my daughter describe to me the nuances in how these communities of practice support responsible members and shun disruptive or pompous non-productive participants, no matter what their “real world” credentials may be.

Measuring Progress on the Alternate Path

While progressive educational institutions have moved to incorporate these capabilities into their curricula, their ability to “lock-step” or limit the student’s exposure to new ideas has been compromised. The situation here is very much like the issue with high test scores – high scores predict success and graduation from traditional school settings ALSO predict success. The problem is not assessing the ability of those who have the traditional background, but rather in assessing the ability of those who have not taken the traditional path, and can in fact profit from training and be successful on the job.

Consider, for a moment, a true story which highlights this problem. There was a young lady who had been home-schooled. She later applied to a local junior college and she was given a placement test in mathematics. Based upon the test results she was told she needed to take remedial math courses. This was quite a surprise since she was very smart

in math and could take the most complex word problems with many extraneous facts, and still produce the correct answer. What had gone wrong here?

The problem, of course, was the placement test assumed that traditional educational sequencing had been followed. In other words, one first memorized the addition tables, then the subtraction tables, then the multiplication tables and finally, the division tables. Only after all this memorization work would a student branch off into word problems and algorithms for applications. Our young lady in question, however, had been taught from the outset how to identify the essential elements in any context, to think out real world problems, and to use a calculator to quickly arrive at the correct answer.

The question arises, and it is a policy question, can an individual function in today's world of work without all that memorization effort? How many people today do "math in their head" on the job? There is an argument that memorization is good for the discipline it teaches, but the same argument was made for learning Latin not too many generations ago. Being a math major myself and having been raised in the era of slide rules, I understand the value of being able to estimate, at least, the order of magnitude of the desired answer. On the other hand, all other things being equal, a person who can perform calculations in their head is a more valuable human asset, if only marginally so.

Questioning the Evaluation Algorithm, Not the Score

The underlying assumption used in this placement test recommendation, however, was that if you couldn't do math in your head, your development must have stopped at that level. This is a mistake for which future test developers should be vigilant when dealing with Millennial or Internet Generation students – assumptions about sequencing of educational content are no longer a "given." This means that, in some cases, test developers may have a harder time writing items. What we call "aptitudes," because they are thought to represent basic skills, may not be the cornerstone building blocks in alternate educational paths. What is the solution? Perhaps items need to be closer to the styles and modalities required on the job. Work sample tests, however, tend to assume prior knowledge of job concepts and this greatly limits the range of applicants to which the test can justifiably be administered. Alternatively, the algorithm to recommend placements should take into account that traditional precursors may play a part in a comprehensive evaluation, but that demonstration of higher level proficiency should perhaps carry more weight.

Assault on Occam's Razor I: Aptitude Testing

"Occam's Razor" is a well-accepted (though often misquoted) tenet in science. The common translation is something like "The simplest answer is the RIGHT answer." This translation deviates from the spirit of the original which is more closely captured by the phrase: "Do not generate more assumptions than are actually necessary to arrive at a valid answer."

The bottom line in this train of thought is that there is a fork in road to proficiency. Trying to find the ONE best approach (i.e. Occam's Razor) may be doomed to failure. Instead, perhaps, we need to be developing multiple criteria and multiple evaluation methods. Once each method is tested and found valid, then there remains a policy decision to be made on how to use two (or more) systems concurrently. For the reasons identified above, whatever methods are validated they should NOT be considered sequential hurdles. More appropriate models would include a compensatory system (full or constrained) or as totally independent ways to predict success. The DoD system for identifying the three educational tiers for evaluating applicants is just one example.

This section addressed issues with the “well-behaved” area of aptitude testing. Now, let's turn our attention back to the less well-defined domain of non-cognitive measurement.

Challenges to Operationalizing SDI+

ASVAB Correlations: The Curse of Success

The plan is for the Self-Description Inventory Plus to be an augmentation, in some sense, to the traditional aptitude testing model which has worked so well for so many years. In this sense, it is important to realize that the goal is to find and use incremental validity or unique variance over and above that already provided by the ASVAB and the individual service composites. Sometimes these new “personality” initiatives get mistaken as proposed replacements for ASVAB and that needs to be clarified before proceeding.

Stakeholder Lamp Posts – The Criterion Problem I

Every stakeholder, however, has their own personal lamp post and circle of light. Despite the desire to find a criterion where the light is good, the actual key may lie in darker places. Recruiters, for example, evaluate proposed systems in terms of whether or not it increases or restricts the flow of qualified candidates through their offices. Trainers evaluate models based on how well the actual process reduces attrition or improves the average final course grade. Field supervisors measure success in terms of perceived job performance of technical school graduates.

Because of so many confounding factors, measures in these areas may not yield the stable criterion measures desired. Counting the flow of applicants into recruiting stations is subject to fluctuations due to world events, the wavering influence of family and friends, as well as the prevailing economic environment.

While some, such as our US Congress, rightly hold that “job performance” should be our gold standard for measuring success, that perspective, too, has its drawbacks in the military setting. A set of finely tuned selection, classification, and training systems is designed to reduce the variance in the quality of personnel who are actually deployed into the field. First we filter civilian applicants on general intelligence (aptitude) and on broad medical and mental standards such as general physical health, upper body strength, lower body strength, hearing, eyesight, and general psychological well-being. When an applicant is found to meet all these standards, he or she becomes a recruit who must be classified into a job family. In the US Air Force today, each of our approximately 150 job families uses these same data to establish a profile of minimum requirements for assignment into that career field. Next, these pre-screened recruits are sent to technical school to be trained in a relative narrow specialization. The net effect of the training is two-fold – to weed out those who cannot learn to perform the job requirements to standards and to give special attention to those who are marginal performers in order to bring them up to performance standards. Once again, the goal is to reduce performance variance of graduates who are turned over to field supervisors. Because supervisors see only a limited number of subordinates, on-the-job training by these supervisors tend to be non-comparable across settings and is affected by traditional issues such as halo effect, invisibility, or outright personality conflicts.

Historically, for the United States Air Force at least, the traditional measure of “ASVAB” success had been training outcome. This is the criterion of choice because it is quantifiable, available, and supposedly objective.

Recently, the Air Force has begun to question the value of the Armed Services Vocational Aptitude Battery (ASVAB). Analyses show lower than desired correlations between ASVAB scores and attrition, pass/fail rates and final course grades. Why is this? There are two answers.

Classification Done Right

Oddly, the recent internal criticism of the ASVAB is due to its own success. The existing selection, classification, and training system is doing its job quite well. The military classification system properly uses the ASVAB scores for initial person-career field match based on recruits meeting or exceeding published minimum requirements. The higher the aptitude minimum for any occupational group, the more restricted the bandwidth of abilities. Because of *restriction of range effects*, it is not surprising that the same entrance test score does not predict the rank order of final course grades nor job performance within these narrowly defined groups. A few Air Force career fields, like Security Police, have a relatively low entrance minimum profile and a burgeoning demand in the field. These select fields have a range of talent comparable to the full Air Force and the ASVAB composites demonstrate the strength of their predictive power.

Crafted Training Courses

The second answer is relatively subtle. When trainers run statistics on class performance, they assume the current course is a “given.” The course curriculum, however, was designed against a field-validated need for airmen of a pre-selected ability level. Course material, presentation styles and instructional design are all crafted to achieve a target level of proficiency on items negotiated between the school house and the career field managers at a Utilization and Training Workshop (U&TW). The length of the course is established assuming an expected ability level of students, availability of resources (classrooms, instructors, simulators, non-durable materials, etc. and the graduate production requirements of the field.

The course is crafted to produce a given number of qualified graduates per year which is programmed well in advance. Attrition happens. When conditions change, for whatever reason, the number of graduates may fall off and attrition goes up. If the increased attrition is attributed to individual student quality, the normal response within the training community is to increase remedial training and/or set the student back in the course curriculum (washback). Although it is rare, natural disasters have destroyed classrooms and rendered essential equipment and material unusable – all with the end result of denying students the same access to instructional materials as previous classes and causing marginal students to fail at a greater rate. When attrition becomes a concern of higher headquarters, more stringent actions are employed and attrition will go down. All the elements in crafting the original model of the course can be re-negotiated. In other words, the *objective measure* of attrition is a value sometimes affected by policy and is not just a product of student quality. For this reason, the correlation of ASVAB (student quality input) to attrition can appear weaker than it should.

Basically, if you separate the qualified from the non-qualified before they enter the course, reason for failures by the qualified personnel is not a question of qualifications – but rather it’s a result of things not yet being directly measuring such as their *ability to cope, conscientiousness, or motivation*.

Demonstrating Incremental Validity – Criterion Problem II

USAF – The Occeumetrics Laboratory

As noted in recent literature, where and how measures are validated (especially personality measures) affects the strength of the relationship you find (Barrick and Mount, 2005). While traditional academics tended to discount military studies as a “special case,” in some ways military studies offer almost laboratory conditions for evaluating personnel system models. With a broad range of occupational domains with varying technical demands and ability level entrance requirements, the military systems, especially the US Air Force, provide controls and criterion measures nearly unattainable in the private sector. Studies conducted under these unique set of conditions have been called (by those outside the military) the science of Occeumetrics (awk U Metrics) – for “occupationally good (eu) measures.”

Implications of the Military Pipeline

With few notable officer exceptions (doctors, lawyers, etc) all military positions are filled from the next lower rank, officer O-1 through O-7 and enlisted E-1 through E-9. This military pipeline concept makes all US military services very cautious about new enlisted personnel programs. This motivates them to study personnel policy proposals very carefully. Consider, for example, if a personnel policy was implemented which was not well received by enlisted Air Force personnel at the paygrade of E-6, Technical Sergeant. These are people with about 7-12 years service. If these people, in large numbers, chose to not reenlist at their next four year enlistment window, it would create a vacuum at the critical point in the full force structure. While this would create more promotion opportunities for the E-5 airmen, the ripple effect would not end there. There would be an increased need to fill positions at the E-5 level from the E-4 population.

This example gets into branch of service promotion policy differences and the following only applies to the US Air Force promotion system. Promotions in the Air Force from paygrades E-1 through to E-4 are basically fixed time intervals given normal progress on self-study upgrade coursework. This means that E-4s eligible for promotion to E-5 can advance in this ripple. Personnel in paygrades E-1 through E-3 and those E-4s without sufficient prerequisite time-in-service or time-in-grade minimums cannot be promoted ahead of schedule to fill the gap. To restore the manpower levels, additional high school students must be recruited and trained beyond the already programmed limits and it may be years before the force returns to its former state of military readiness.

--All this trouble in response to one bad personnel policy decision. It is no wonder the Air Force is cautious and elevates this force management concern to the level of a science.

The Refined Enlisted Force

Through their careers, the military pipeline process in the US Air Force flows airmen through successive 4-year reenlistments and paygrade promotion filters. Those who remain are those who have been invited to reenlist because they've demonstrated sufficient proficiency and potential to remain "on track" with promotions. The US Air Force enforces a "High Year of Tenure" (HYT) concept. For each paygrade there is a maximum number of years in service you can serve. Hence, if you are at a given grade, say E-5, and you've not been promoted to E-6 by your high year of tenure mark, you are not eligible to reenlist.

This refining process continues with each paygrade promotion. The range on variables important for promotion from lower levels gets smaller. The Weighted Airman Promotion System (WAPS) controls promotions into paygrades E-5 through E-7. The Senior Non-Commissioned Officer Promotion Program (SNCOPP) controls promotions to the paygrades of E-8 and E-9. These promotion systems use a combination of objective and subjective measures. At the lower ranks, WAPS uses six factors: Specialty

(Job) Knowledge Test (SKT) score, Promotion Fitness Exam (PFE) score, Points for Decorations, Time-in-Service points, Time-in-Paygrade points and a 10-year history of the supervisor's Enlisted Performance Report ratings. At the higher ranks, SNCOPP uses similar items except the SKT score is replaced by a heavily weighted Board Score -- a consensus rating by a panel of three senior members reviewing each candidate's personnel file against all other current eligibles.

Measures, Measures Everywhere – What to do?

In the enlisted ranks in the US Air Force, several member “quality” measures are developed over a person's career. Because of the WAPS and SNCOPP promotion programs, many measures are the basis for promotion and therefore experience restriction of range over time in the service. There are three interesting measures which are not so constrained. The three measures of interest for validating a “successful” recruit/service member include: ASVAB scores, the act of reenlistment, and self-reported job satisfaction. This section addresses only the first, the relevance of the ASVAB to personality assessment.

ASVAB and USAF Promotion Success

The ASVAB scores (or more specifically, the Air Force composites based on ASVAB) are used to stratify personnel and select them INTO career fields. Once in the career field, however, the ASVAB scores play no direct role in promotion. ASVAB will only occasionally affect a person's ability to reenlist -- only if that person is changing career fields – either voluntarily or involuntarily. Voluntary changes occur for career broadening, stepping up to cover critically under staffed areas or to take advantage of variable enlistment bonuses or improved promotion potential. Involuntary changes occur due to Air Force reorganizations or downsizing and reduced demands for the previous specialty.

While minimum aptitude levels are established for entrance into individual Air Force career fields, these standards are not formally revalidated on a regular basis. Requests for changes in minimums by career field managers are handled on a case-by-case basis and the current standards are issued four times per year – recently increased from a traditional twice-a-year update. It is much rarer to have an across-the-board, Air Force-wide, assessment of the linkage between minimum entrance requirements and a selected criterion. One such study is currently under way.

In an evaluation of the quality of Air Force promotion tests, I reviewed overall promotion statistics for the past six years (2001-2006). There are currently five promotion/testing cycles per year in the US Air Force. Each corresponds to promotion to paygrades E-5 through E-9 respectively. The focus of our study was on the technical mid-level paygrades covered by WAPS: E-5 through E-7. For these grades, four of the six promotion factors are data-based and pulled directly from the master personnel file. The

other two factors are tests – the SKT and PFE. The PFE is an Air Force-wide test assessing knowledge of broad military topics. The SKT is specific to each job family and grade level. It is important to note there are around 400 *different* SKTs every year.

Personality-Aptitude Interaction Hypotheses

As an ancillary note to this review of promotion tests, the behavior of the incidental selection effect on the ASVAB score was documented. This was the source of the earlier observation that Air Force Specialties (AFSs) with high entrance minimums show little promotion impact on incidental selection in ASVAB scores while broader AFSs, like Security Police demonstrate a strong relationship between ASVAB and success.

When moving into the domain of personality measures, the key point of this discussion is that it will be necessary to control for the amount of “success” variance that the ASVAB has already captured. My working hypothesis is that personality measures are probably best used as moderators to the aptitude/ability measures which should do the “heavy lifting” in selection and classification. Hence, my expectation is that personality measures will help explain the rank order of final score grades when the ASVAB, due to restriction in range cannot. Conversely, in career fields with a wide span of job demands low entrance thresholds and surging personnel demands, I anticipate that in both the school house and on-the-job the ability differences will be the high-drivers in differentiating between sufficient and successful service members.

Implications of a Second-Order Effect

In the United States, we feel the need, professionally, economically, legally, and/or morally to document the magnitude of this contribution in an environment of many constraints. If carefully executed, I believe the benefit of personality measures added to the ASVAB can be demonstrated using insights from the ongoing USAF occupational analysis program described below. The restriction in range of ASVAB works to accentuate the effect of personality and I anticipate this will primarily affect the hardest to fill career fields. The dollar savings will not be as great if the effect could be shown across the board, but there will be benefits in strategically vital areas. Even if the magnitude of the contribution can only be documented in limited environments, there will be larger benefits in broader and less measurable arenas that I discuss below.

Wrapping Up -- The SDI+ Initiative

The challenge is to not only validate that personality measures can inform us on a person’s temperament, but also to devise a way to integrate this additional knowledge in into a no-nonsense, actionable system that benefits the Air Force. The constraints surrounding personality measures have been reviewed. The AFSs where validity can be documented have been characterized. So, what does the Self-Description Inventory Plus (SDI+) Initiative propose?

USAF Personnel System – New(?) Stakeholder

The stakeholders discussed thus far are the recruiters, the trainers, and the field supervisors. Here is a proposal for a fourth stakeholder – the Air Force personnel system with its own lamp post and circle of light. The first three stakeholders are responsible for acquiring, training, and utilizing service members. The mission of the personnel system is to manage those developed human resources and strive to get the right qualified person into the right job at the right time. For every loss of a qualified airman who does not reenlist, there is a requirement to promote from within and bring in a new recruit to complete the ripple effect. On the other hand, increasing retention maintains a more mission-ready force. In addition to a mission-ready force, additional costs are avoided in recruiting and in school-house technical training. Not all cost savings are visible or apparent. As noted earlier, the school house only teaches what was negotiated at the U&TW. As tech school graduates arrive at the job site, they undergo local on-the-job training (OJT). OJT requires manpower that could otherwise be spent in direct mission support. Hence, when a person in direct mission support fails to reenlist, we must burden the remaining staff to absorb not only an additional OJT requirement, but also to make up for work formerly done by the individual who was lost.

New Lamp Post, New Circle

With new weapons systems (F-22, F-35, etc) coming online in a big way early in the next decade, major shifts are planned in the US Air Force structure. Older weapon system platforms are being retired and replaced with multi-function systems. Naturally, the US Air Force classification system will change to accommodate these new requirements. Another fact of the future resource-constrained environment has arisen – paying for these new weapon systems. “Transformation” is the keyword of the day. Being proactive in response to the programmed costs, strategic efforts are underway to trim future costs through technology insertion today. By taking this opportunity to automate formerly person-intensive support programs, savings in future personnel costs can not only defray the cost of modern weapon systems but also improve the speed and quality of future personnel support functions.

SDI+ - The Goal: Retention The Key: Job Satisfaction

The impact of “transformation” on personnel is downsizing and the reallocation of current personnel to different or highly modified career fields. To prepare for a future, smaller, smarter force, the goal must be set to retain the best during these stressful times. Not only will retained quality airmen maintain mission readiness during this transition, but they will also be the senior on-the-job trainers of the future should a surge in the force be required.

If the goal is retention, can anything be done other than awarding bonuses and putting forth sound personnel policies? The key is to measure and increase job satisfaction. There is strong evidence of this linkage as shown in Figure 1 -- these two measures are strongly linked in many AFSs. Many factors other than job satisfaction will affect actual

reenlistments, but many are outside the control of the personnel system. Of all the factors which really impinge on reenlistment decisions, other than bonuses, increasing job satisfaction is something the personnel system can measure and take action on directly.

Classification Structure: The Backbone

It is a little noted fact that the Air Force Personnel Center (AFPC) “manages” the Air Force military classification system. AFPC can amend it or, in conjunction with Air Staff, restructure it to align with emerging Air Force-wide strategic visions.

Consolidating AFSs can ease the constraints on the assignment system or conversely, sub-dividing AFSs can eliminate wasted training, better control highly specialized talent, and provide an agile response to emerging changes in the way the work is really accomplished in the field. The military classification system is the backbone on which recruiting, training, and the field build the body of the force. The classification system is also the “single sheet of music” from which the field defines their personnel needs, the school house manages their training pipeline, and the recruiters receive and manage their quotas for new recruits.

Because of the central importance of the military classification system to mission readiness, timely feedback from the field is essential. The Air Force, under the direction of Dr. Raymond E. Christal researched occupational analysis methods of various large agencies and developed over a ten year period (1957-1967) a comprehensive methodology to capture essential data on rapidly changing, real-world Air Force jobs.

AFOMS: Measuring the Pulse in the Field

After 10 years of R&D, an operational unit – the Air Force Occupational Measurement Center (OMC) was launched to help support this analysis as an ongoing mission. This function continues today in the US Air Force Air Education and Training Command’s (AETC) Air Force Occupational Measurement Squadron (AFOMS). Because of the strong ties between current field feedback and promotion tests that are valid and up-to-date, from the personnel system’s perspective as a client, AFOMS has two primary missions – one for occupational analysis and one for development of enlisted promotion tests. At present, enlisted occupational areas are surveyed about once every three years and promotion tests (over 400) are revised annually.

The methodology developed for capturing, organizing, and reporting detailed job information specified two common elements in each USAF Job Inventory survey booklet – a “background section” and a “task list section.” The task list section typically contains 700-1200 task statements recently validated for the AFS being analyzed. Current information is critical for aligning the AFS structure with emerging needs in the field at the U&TW mentioned above. Changes from the U&TW imply new curriculum for the school house and AFOMS provides tailored training extracts to the school house for basic reference material in course development and revision.

The background section of the USAF Job Inventory booklet collects demographic information, job satisfaction items, and career field specific information. The demographic section includes detailed identification information such as name, social security number, phone number, email address, organizational position, classification of the job, and traditional demographics. The job satisfaction items include intent to reenlist (along with a checklist of reasons for reenlisting or not reenlisting), specific questions about job satisfaction and the degree to which talents and training are utilized in the current position. AFOMS collects and maintains these data and archives them on over 100 current and historical Air Force occupational areas.

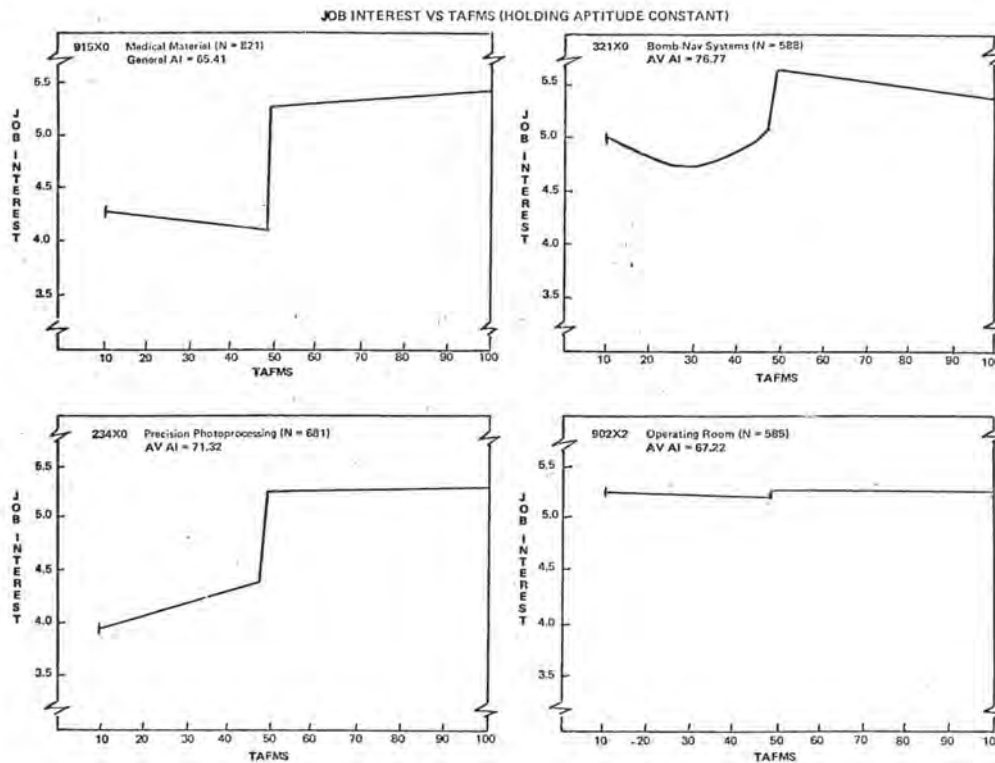
“Encouraging Constellations” – The Virtual Mentors

Over and above the traditional job satisfaction data, AFOMS has an Air Force-owned Internet-based tool they use to survey the entire enlisted force. One of the thrusts of the SDI+ Initiative is to tie SDI+ based job satisfaction constellations to AFSs and more specifically to the job type or job title within an AFS. This more molar linkage would ensure the robustness of this approach under any future Air Force restructuring efforts decisions. The AFOMS Internet survey tool not only makes this possible, it makes it easy and nearly automatic if added to their already published survey schedule. It is important to note that in order for this to work the SDI+ must be co-administered with the standard occupational survey. The only challenge here is getting authorization for the extra 40 minutes of incumbent/respondent time needed to administer the SDI+ as part of the standard USAF Job Inventory administration.

Once collected, the SDI+ profiles for job incumbents along with their standard self-reported job satisfaction will be treated as a statistical item constellation. Those constellations associated with incumbents reporting high job satisfaction will be deemed “encouraging constellations.” Constellations associated with incumbents reporting low job satisfaction will be identified as “discouraging constellations” While there are plans to recode, cluster, and factor analyze these data for statistical analysis and explanatory purposes, the goal is to operationally use the constellations in their most raw form and characterize these constellations (both pro and con) as “virtual mentors.”

The notional charts below in Figure 1 show differential effect by AFS on job interest when 4-year enlistees pass the critical decision point (Christal, 1974, Gould, 1972). Note that the figure represents a “full AFS” snapshot at one given point in time and the plot shows the average job satisfaction rating for each month-cohort group across a 100-month time frame. Further analysis is planned to compare and contrast constellations pre- and pro- reenlistment to capture and explain the yes/no “intent to reenlist” and the reasons associated with each group.

Figure 1: Reported Job Satisfaction for Air Force Career Fields in Occupational Surveys
 Note: TAMFS = Total Active Military Federal Service (in months) 1 enlistment=48 months



Giving Millennials what they Need

Fidelity - Making it Happen

In contrast to most other approaches, the SDI+ Initiative will focus ONLY on the constellation patterns of people with high job satisfaction following their first reenlistment (49+ months of service.) In other words, rather than focusing in on reasons for failure with further research needed to study proposed “corrective measures,” the SDI+ Initiative proposes to accentuate the positive and focus in on success.

There are several reasons for this decision. First, the act of reenlistment means certain things in unequivocal terms. Most importantly, since reenlistment is not a “right,” but an invitation, the act of reenlisting means that both the incumbent was willing AND the Air Force deemed the person a valuable asset. In the coming years of the draw down of forces, these invitations will be harder and harder to obtain and will better distinguish between higher and higher levels of successful Airmen as valued by the Air Force. Just as “Transformation” was an opportunity to benefit from technology insertion during stressful times, data collection for “virtual mentoring” is now entering a critical window of opportunity.

Operationalizing an SDI+ benefit to the Air Force can be obtained with a fairly reasonable investment of time by job incumbents who understand the vision. Surveying the entire enlisted force would be accomplished using current AFOMS schedules and would be completed in about three years. As soon as the surveys are completed for each AFS, key constellations can be identified and codified for rapid look-up and retrieval. This establishes the baseline of “virtual mentor” constellations tied to both AFSs and job types within those AFSs.

Because this approach focuses on “winning constellations,” if the Air Force decides to restructure multiple career ladders into a new larger AFS, then the winning profiles are simply pooled and relabeled to reflect the consolidation. If AFSs are sub-divided the constellations associated with the specific job types can easily be relabeled accordingly.

Pulling the Dog’s “Leg” – Beware of our labels

Another reason for using direct positive indicators of job satisfaction is that mediated interpretations are not required. Each time that data are summarized and translated into labels, we run the risk of believing our labels rather than the respondent’s data. In his final progress report on Project LAMP in November 1994, Dr. Christal reported success and a partial failure in a construct validation study of SDI (Christal 1994). This study compared the interpreted “Big Five” dimensions and reported interest areas using the Air Force Vocational Interest – Career Examination (VOICE). As he related it:

Although the relationships of rated personality traits with rated interests are only modest, they are highly significant and produce meaningful and sensible patterns. (see Appendix 1).

Appendix 1: Correlations of Interest Composites with Personality Traits

One of the goals of the work task was to evaluate the relationship between personality and interests. The Air Force Vocational Interest – Career Examination (VOICE) was reprogrammed so that it could be administered by computer... This computerized VOICE was administered along with 205 personality trait inventory to 363 airmen in Basic Military Training. Eighteen interest composites were computed for each of these subjects according to instructions presented in the technical report by Alley, Berberich and Wilborne in 1977 (AFHRL-TR-76-88). These composites were in turn correlated with each of the 205 personality trait self ratings collected in a 363 case sample. A correlation of .17 is significant at the .001 level. Only correlations above .20 are presented below, all of which are well beyond the .001 level of significance.

(A sample of these items include)

Electronics –

Industrious	.22
Inventive	.20

Science –

Inventive	.33
Philosophical	.30
Innovative	.28
Inquisitive	.27

Outdoors

Adventurous	.48
Brave	.34
Daring	.32
Active	.27

He continued with the less than desirable results:

Two of the interest areas are strange indeed. The first is food service. Individuals rating higher interest in this area rate themselves higher than other individuals on being absent minded, sluggish, gullible, lazy and insecure. The second unusual interest area is craftsman. Those showing interest in the craftsman area rated themselves higher than other individuals on being inefficient, absent-minded, forgetful, dull, scatterbrain, ignorant, and not proud. Individuals showing a strong interest in these two areas certainly seem to have a poor self image. All in all, the relationships between rated interests and rated personality traits tends to add creditability to both measures.

Two points are important here. First, can you imagine showing this description (gullible, lazy, etc.) to a job applicant and then indicating the Air Force job recommendation based on this description? Providing factor or facet descriptions just hide the basis of the match one or two levels.

The SDI+ Initiative proposes NOT to give “factor interpretations” to applicants, but rather to provide only an ordered list of Air Force job recommendations in descending order of the “goodness of fit” metric.

Second, are these dismal descriptions really accurate? From my experience with these career fields, I totally understand these self-descriptions coming from Basic Trainees. These two cases are caused by different dynamics. Traditionally, the food service AFS had very low entrance standards. Recruiters would “sell” these jobs to people with aptitudes just high enough to get into the US Air Force. It is easy to imagine these people having been told they were losers for much of their life and so getting an Air Force job, any job, offered a step up and out of poor circumstances.

The issue with the craftsman area is different. I know. My father was a carpenter. Growing up in his house, I was always being told I was handling the tool wrong – not as an empty criticism, but rather as an instruction of how to improve with more exacting internal standards. I assume that other craftsman areas generated offspring with similar high standards for self-evaluation.

The point here is that it is important to understand what we are really measuring and how it can be used most effectively. Responses on the Self-Description Inventory are verbal BEHAVIOR, not a measure of ability. Similarly the use of “interest” inventories of past experience have limited power to direct applicants to critical and uniquely military fields like air traffic control or explosive ordnance disposal.

Because of my independent personality I would not have responded with the typical “son of a carpenter” pattern. I am not really “Mechanical” craftsman material. My “Mechanical” aptitude score on the Airman Qualifying Exam (AQE, ASVAB’s precursor) was at the very top of the chart. This also was misleading. That aptitude test was based on identifying the names of particular tools – assuming that if you knew the tool’s purpose or name, you had experience and you were probably more mechanically inclined than those who didn’t know about the tools.

With me this was not true. While I did watch my father work with these tools, I never had any interest in following that line of work. My dad would have me hold tools while he worked and hand them to him when he needed them. Between these interchanges, I would disappear and then return shortly. Finally, my dad had to ask, “Where are you going every five minutes?” I replied that his tools were dirty so I had to go and wash my hands. Later in college I went into mathematics because it was clean and my dad had no interest in the area. Now, as an adult, when forced into home repairs, I judge the size of the project by the number of injuries I expect to inflict upon myself.

Because I knew all the “labels” or names for the tools, I was labeled a person with high mechanical aptitude. Like Abraham Lincoln and the dog, “*calling*” me a mechanic doesn’t make me a mechanic.

Assault on Occam’s Razor II – Personality Measures

Showing an interest in improving the job satisfaction of current job incumbents may, in itself, be able to improve retention rates for the right people. The SDI+ Initiative has the more modest goal of providing personalized career counseling at the point of reenlistment/career field change decisions. Moreover, it is possible that by using the same “virtual mentor” approach on the Internet we can increase the number of qualified applicants from the general population.

As time progresses, more and more our applicant pool and service members are part of the Internet or Millennial Generation. While generational generalizations are over simplifications, there are truths that are clear about the younger population. They demand choice. They demand honesty and forthright interactions. They know how to find “secret” information and evaluate motives behind informational sources. Clever screening “tricks” on instruments from psychologists can be detected, compromised and broadcast over the Internet in a matter of hours.

Psychological instruments used for selection in a pass/fail mode are viewed as “high stakes” games and draw more attention from would-be hackers. Instruments designed to identify suitable candidates for just one job are easier to compromise than those which merely compute a “job compatibility metric” to be used in conjunction with other evaluation factors or instruments which assess a broader spectrum of jobs. I’ve reviewed two operational psychological tests over the past four years that I could “compromise” in less than two minutes of strategy coaching to potential applicants.

Because the SDI+ Initiative is being proposed as an applicant-driven self-help counseling process, the concept of “faking” would be a mere waste of time and Millennials don’t put up with things that waste THEIR time. Test compromise is not an issue because every Millennial already has his or her own lamp post, circle of light and personality which is the ultimate key to person-job-match and job satisfaction.

The Answer is “Give up the illusion of central control”

If I seemed a bit traumatized by the Millennial Generation, I assure you it is from personal experience. I have a daughter, now 21, who, as a true Millennial, was home-schooled (on the Internet) from age 13 through 17. Later she took the GED, placed at the top 2% of high school graduates for that year and started college. About 8 years ago, I remember coming home and checking on her daily work. I happened to mention that we were going to do a training-needs/design analysis of the JSF. She asked, “What’s a JSF?” And I replied “A Joint Strike Fighter – a new aircraft for the 21st century.” She said “Oh,” in a fairly disinterested tone. The next morning as I was leaving for the office she handed me about 100 sheets of paper with descriptions, specs, and comparisons of the two bidders on Air Force evaluation documents labeled “For Official Use Only.” Her idle curiosity overnight was more productive than a week’s work by my research assistant at the office.

As long as Millennials believe the information provided to them is “managed, manipulated or “slanted” they will be less inclined to act on information, no matter how valid it may be. Every time a measure is “interpreted” by a specialist, it loses something in the translation. The modern generation desires raw information in high fidelity in the same manner as the previous generation embraced “organic produce.” No amount of technical assurances of fidelity will remove the stain of manipulated information. When they get information they question in the least, it’s off to the Internet for different perspectives. The Millennial generation’s outlook on life can be summarized by two mottos from the X-Files TV program, “Trust No One” and “The Truth is Out There.” A strategic study in South Texas investigated how today’s Hispanic high school students gather information about possible jobs. Even among this group, searching the Internet far surpasses other sources such as school counselors, friends, etc.

There are nearly insurmountable obstacles for the US Air Force to centrally manage an effective person-job-match system. Most troublesome is the legal prohibition against separate norms in recognizing protected group differences. There is the logistical issue of administering the SDI+ early enough in the recruitment process to make a difference in the Person-Job-Match decision. The DoD MEPS have made it clear that there are no additional testing resources available for a new single service “special test.” Moreover, testing at the MEPS may, in fact, be too late in the process. Sixty percent of the recruits coming to Basic Military Training have guaranteed jobs negotiated with the recruiter who has quota-driven motives.

In an “all-volunteer force,” it should be clear that “volunteer” comes before “force.” We’ve heard through out this IMTA conference that “two-way trust” is essential for command and control of the “strategic private” in the modern theater. To recruit Millennials, we need to give them what they want, raw information in actionable format. Trust them to make the best decision. While one of the strongest motivators for joining the all-volunteer force is still patriotism, it doesn’t hurt to structure our job offers in terms of “what’s in for me” so that both the service and the individual can look forward to a promising career.

Summary

The theme of this year's IMTA Conference was "Applying the Science." The SDI+ Initiative is a perfect model of that goal in an increasingly complex military, social, and economic environment. In normal terms, "applying the science" translates into forging scientific truths into practical technologies. Technology means gaining the benefits of scientific investigations after the scientists have departed.

This is the same goal the United States Air Force is pursuing in its proactive response to future demands by automating "Personnel Services Delivery" (PSD). The SDI+ instrument forged and validated in the Big Five model, will operationally be "the little 220," by proposing "field" to "applicant" Person-Job-Match counseling without intervening interpretations by psychological professionals. Because of its purpose, recruiting and retention, the thorny issues associated with selection are avoided as part of the technological implementation strategy.

Self-Description Inventory Plus (SDI+) Initiative is to use self-help access for the prospect or service member to find possible US Air Force careers with which they are unfamiliar, but where "people with a temperament like theirs" have found high job satisfaction. This approach addresses basic force management issues from two perspectives. First, it is aimed at increasing the SUPPLY of potential applicants by exposing "prospects" to a wider CHOICE of possibilities than they may have imagined. Second, it is the ultimate approach to maximizing DIVERSITY. Each individual gains insight on his or her own personal, not group, characteristics and the advice is directly from those who know, young, virtual mentors known to be valuable assets in today's mission-ready force. The goal is to provide insight for both potential recruits and re-enlisting members on the kinds of jobs that today's US Air Force has for them – with a future both will find rewarding.

References

- Christal, R.E. *The United States Air Force Occupational Research Project*. AFHRL-TR-73-74., AD-774 574. Lackland AFB, TX: Occupational Research Division, Air Force Human Resources Laboratory, Jan 1974
- Christal, R.E., *R&D Summary Report, Contract F33615-91-D-0010, Feb 1993-Nov 1994*.
- Gould, R.B. *Reported job interest and perceived utilization of talents and training by airmen in 97 career ladders*. AFHRL-TR-72-7, AD-745 099. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, January 1972 .
- Murphy, K.R. & Dzieweczynski, J.L., *Why Don't Measures of Broad Dimensions of Personality Perform Better As Predictors of Job Performance?* Human Performance, 18(4), 343-357.
- Barrick, M.R. & Mount, M.K. Yes, *Personality Matters: Moving on to More Important Matters*. Human Performance, 18(4), 359-372.

Table of Contents

Abstract.....	1
Informal Opening Remarks.....	2
Abraham Lincoln and the Dog’s Tail.....	2
The Lamp Post and the Circle of Light.....	3
Military Selection Testing and Validity Issues.....	4
Aptitude Testing and the U.S. Military	4
Testing and Population Changes	5
The Learning Abilities Measurement Program (Project LAMP).....	5
The LAMP Premise	5
Project LAMP’s Evolving Strategy	6
Personality Comes of Age.....	6
Project LAMP and the Self-Description Inventory Plus (SDI+).....	7
Changes in the Environment.....	7
The General U.S. Social Climate	7
Changes in the U.S. Military Environment	7
Changes in the U.S. Military Applicant Pool.....	8
Cross-Generational Differences – an EOD Example.....	8
Broad Generational Characteristics and Time Frames.....	9
The Path to Proficiency.....	9
The Fork in the Road	10
Measuring Progress on the Alternate Path	10
Questioning the Evaluation Algorithm, Not the Score.....	11
Assault on Occam’s Razor I: Aptitude Testing.....	11
Challenges to Operationalizing SDI+	12
ASVAB Correlations: The Curse of Success	12
Stakeholder Lamp Posts – The Criterion Problem I.....	12
Classification Done Right.....	13
Crafted Training Courses.....	14
Demonstrating Incremental Validity – Criterion Problem II.....	14
USAF – The Occumetrics Laboratory.....	14
Implications of the Military Pipeline	15
The Refined Enlisted Force.....	15
Measures, Measures Everywhere – What to do?	16
ASVAB and USAF Promotion Success.....	16
Personality-Aptitude Interaction Hypotheses.....	17
Implications of a Second-Order Effect	17
Wrapping Up -- The SDI+ Initiative	17
USAF Personnel System – New(?) Stakeholder	18
New Lamp Post, New Circle.....	18
SDI+ - The Goal: Retention The Key: Job Satisfaction.....	18
Classification Structure: The Backbone.....	19
AFOMS: Measuring the Pulse in the Field	19
“Encouraging Constellations” – The Virtual Mentors	20

Giving Millennials what they Need.....	21
Fidelity - Making it Happen.....	21
Pulling the Dog's "Leg" – Beware of our labels.....	22
Assault on Occam's Razor II – Personality Measures	24
The Answer is "Give up the illusion of central control".....	25
Summary.....	26
References	26

Headquarters Air Force Personnel Center

Integrity - Service - Excellence



The USAF Self-Description Inventory Plus Initiative

**Johnny J. Weissmuller
AFPC/DPST
Force Mgt Liaison Office**



Advance Apology

- *Humor is hard in an international setting, but since that's all I've got, let's give it a go...*



Objectives for Today's Presentation

- **Summary of Progress to Date**
- **The Final Version of the Abstract**
- **The Recantation and Situational Update**
- **Boring High-Level Summary of the 26-page paper with no data to facilitate the after-lunch nap**
- **Funny Stories to fill the Time**



Summary of Progress to Date

- **The USAF Big Five Model celebrates its 50th anniversary of “further research is required.”**
- **Concessions to Current Operational Demands...**
 - **Data Requested in April, not yet ready**
 - **Paper, not yet cleared for full release**
- **Realized that if a valid SDI+ instrument was on the shelf today, there is no way it could be implemented in the current environment.**
- **All in All, life is good.**



The “Big Five” History (1)

Late 1950's US military services realize that they've captured all the available relevant variance in general intelligence (“g”) using current aptitude tests

AF reviews the professional literature the world of personality measurement



The “Big Five” History (2)

Using state-of-the-art electronic data processing equipment, in 1957 USAF performs factor analytic meta analysis of data sets from leading researchers under various conditions

Assigns Names to the recurrent “Big Five” factors which emerge under diverse conditions: **OCEAN**

Openness (Surgency)

Conscientiousness

Extraversion

Agreeableness

Neuroticism (Emotional Stability)



Big Five History (3)

- In May 1960, The US Navy sponsors a Tri-Service Conference on Selection to discuss THEORETICAL CONSIDERATIONS IN THE DEVELOPMENT AND USE A NON-COGNITIVE BATTERY
- All the US services (including USAF Tupes) present some findings and plans for future research
- Some thing happens --- nothing



Big Five History (4)

- **Evolving External Social Climate Pre-empt deployment options**
- **The “Big Five” Model is researched and commercialized in the private sector (1960-1980).**
- **1979 - Concern Arises that with growing cultural diversity and increased complexity in Air Force jobs, ASVAB needs augmentation**



Big Five History (5)

- **1981, USAF Launches the Learning Abilities Measurement Program (Project LAMP) headed by Dr. Raymond E. Christal (of Tupes & Christal).**
- **Project LAMP's focus grows from basic cognitive processes to learning ability measurement to assessment of non-cognitive constellations.**
- **The SDI Development and Refinement begins and data are collected from Basic Trainees**



Big Five History (6)

- **1994 A USAF multi-dimensional scaling (ALSCAL) of Critical Incidents from the first Gulf War demonstrates the relevance of personality to operational readiness**
- **1995 – Dr Christal passes away and research continues on the development of a Self-Description Inventory for operational use in the Air Force.**



Big Five History (6)

- **1995-2004 A series of contracts to operationalize SDI+ - Big Five plus “Service” & “Team” orientations – Now OCEAN ST**
- **July 2005, SDI+ becomes Part 12 of USAF Officer Qualifying Test**
- **Aug 2007, the Air Force Personnel Center re-organizes and the Force Management Liaison Office is created.**



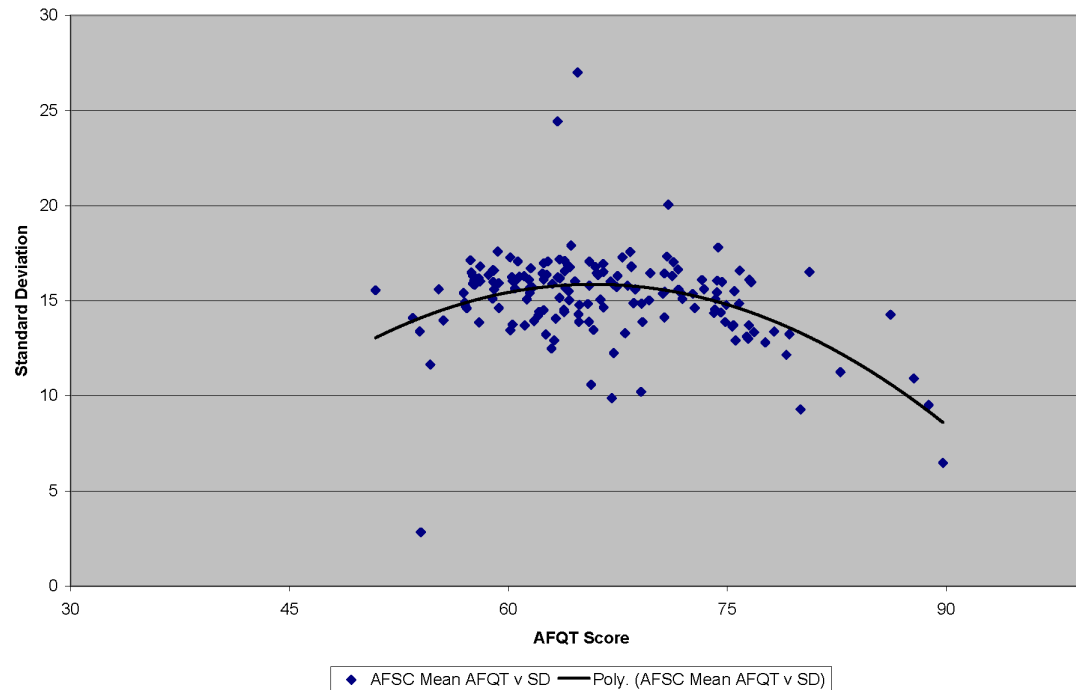
Operational Demands & Critical Air Force Specialties

- **8 Enlisted Aircrew jobs**
- **Air Traffic Controllers**
- **Explosive Ordnance Disposal**
- **Aircraft Structural Maintenance**
- **Civilian-Military Security Police**
- **Cyber Command**
- **Space Systems**



Higher Aptitude – Lower Variance

Plot of AFSC Mean AFQT versus Standard Deviation (05E7 Only)

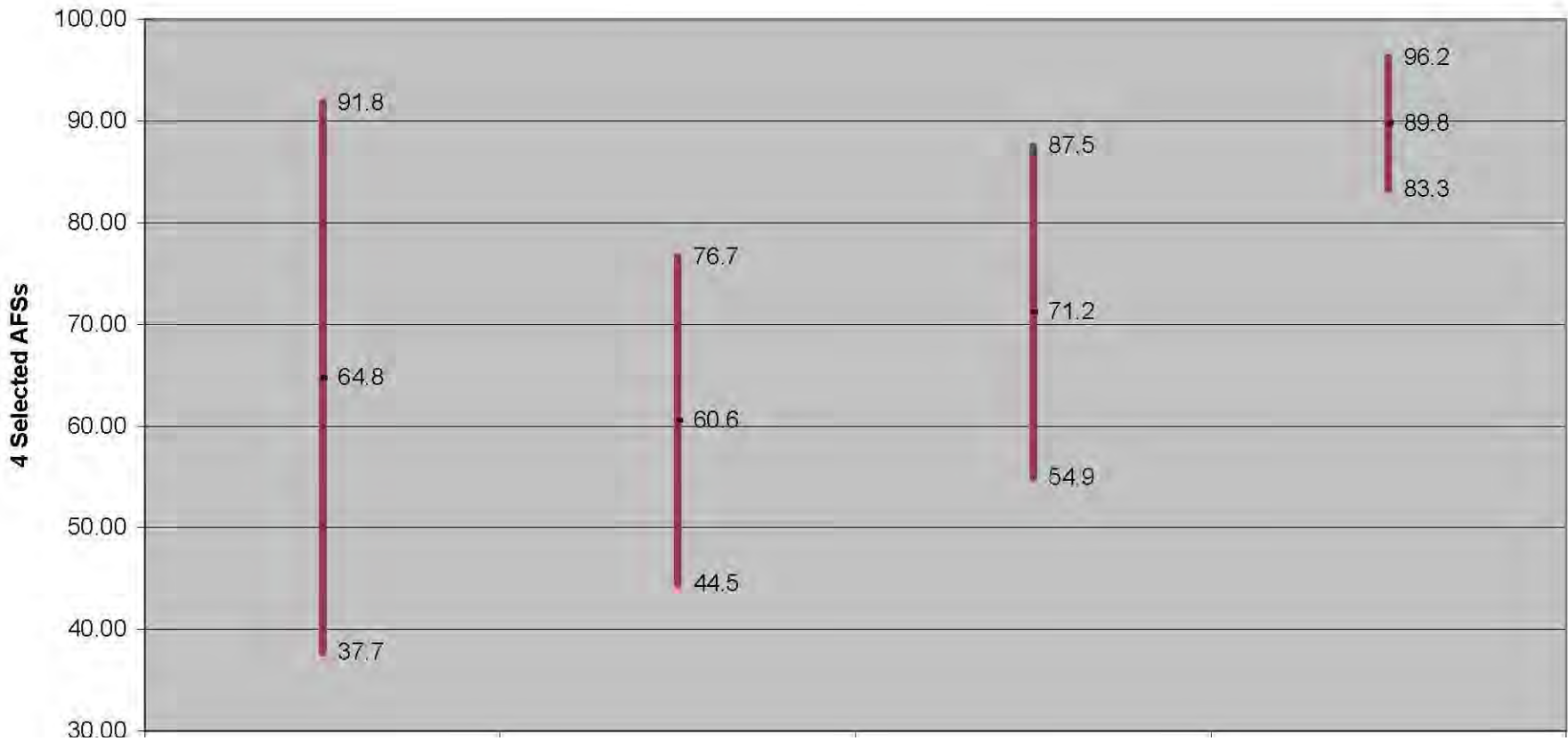




AFS AFQT Range Restrictions

2005 E7 Promotion Pool

**Expected Ability Range:
AFQT Scores (05E7 Sample)**



Integrity - Service - Excellence



Job Difficulty Across AFS

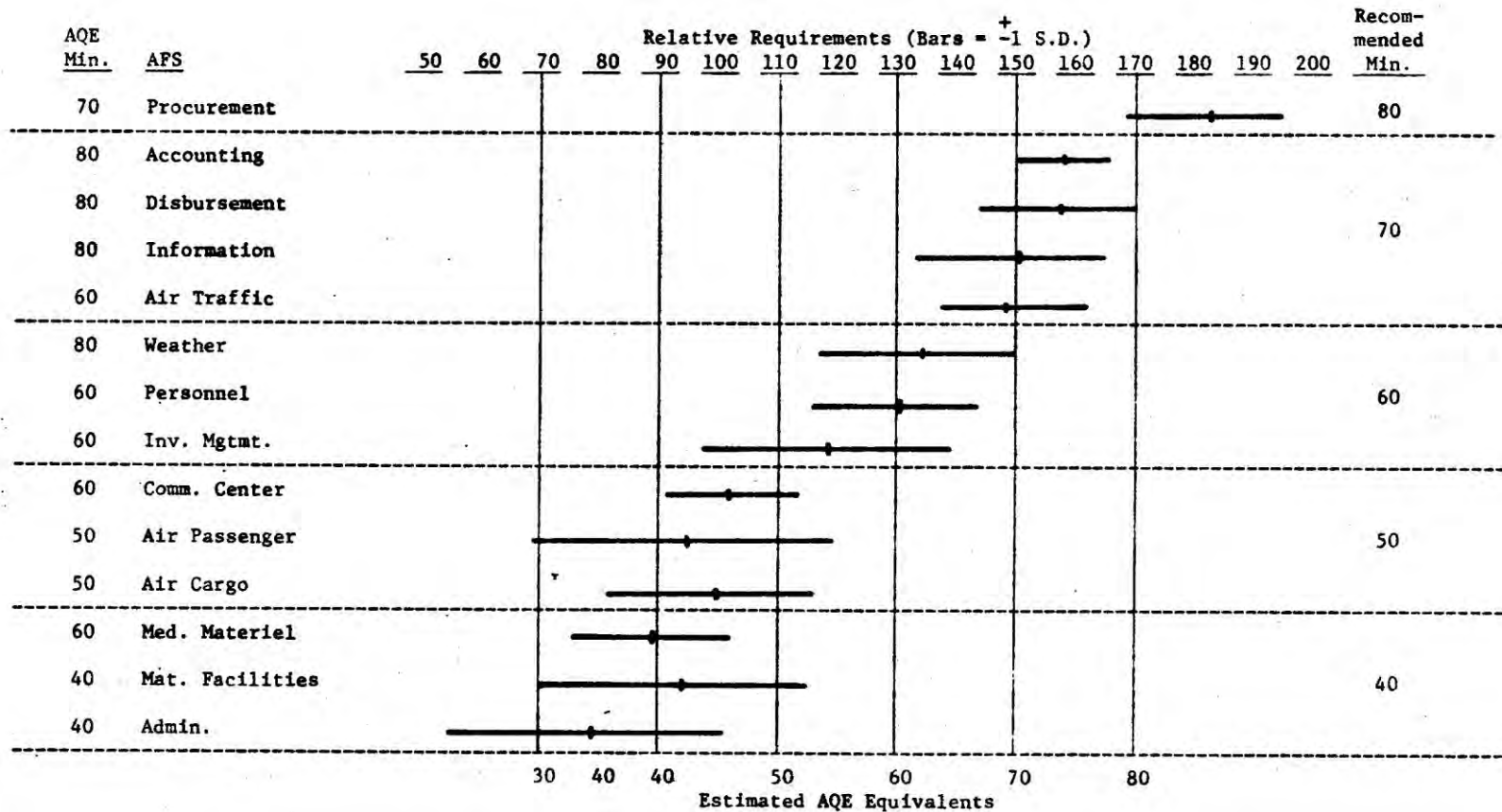


Figure 4. Relative Aptitude Requirements for 1st-Term Jobs in 14 Career Ladders.



Job Difficulty WITHIN an AFS

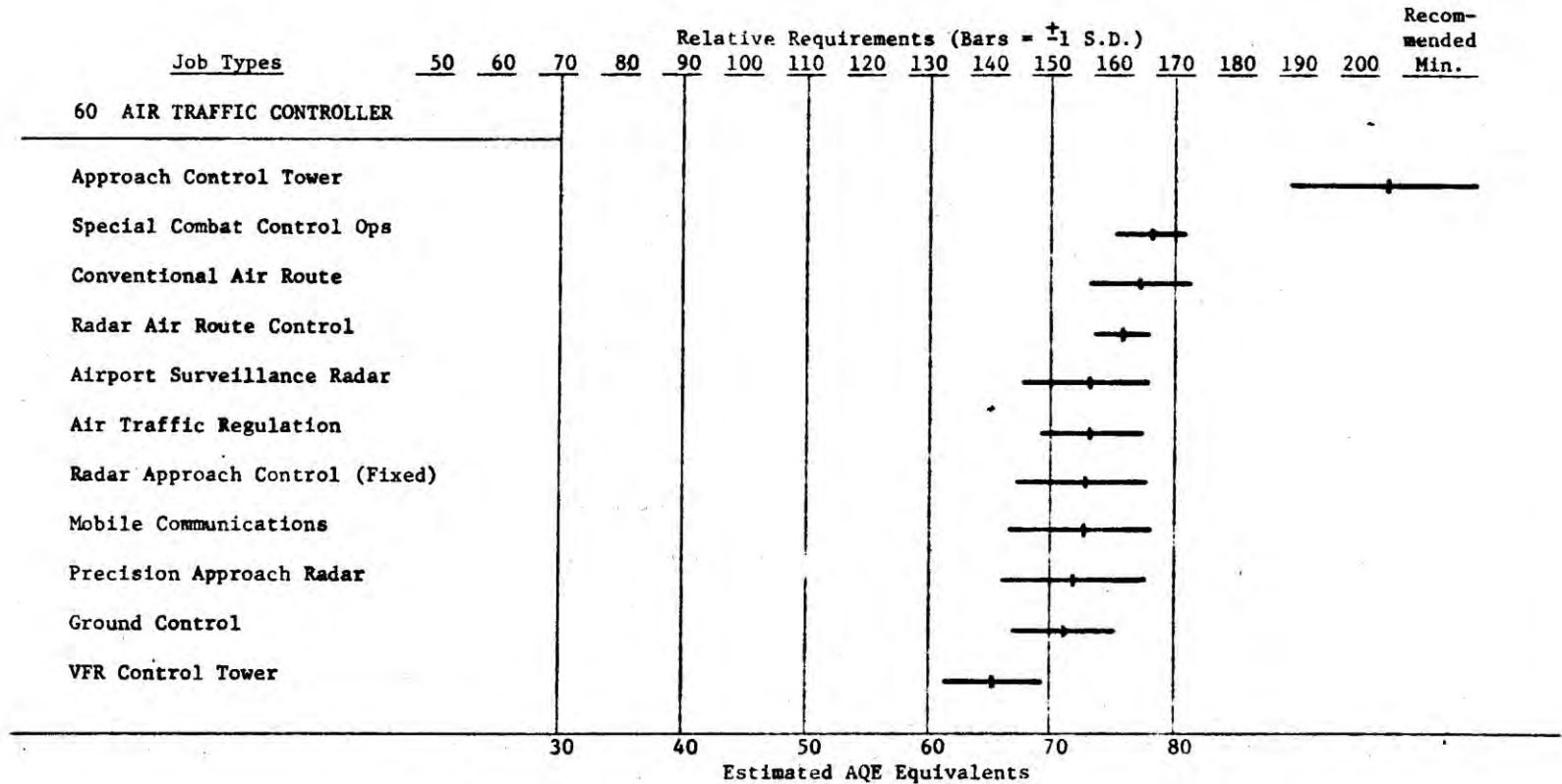


Figure 5. Relative Aptitude Requirements for 1st-Term Job Types in the Air Traffic Controller Career Ladder.



Job Satisfaction, Aptitude & Reenlistment

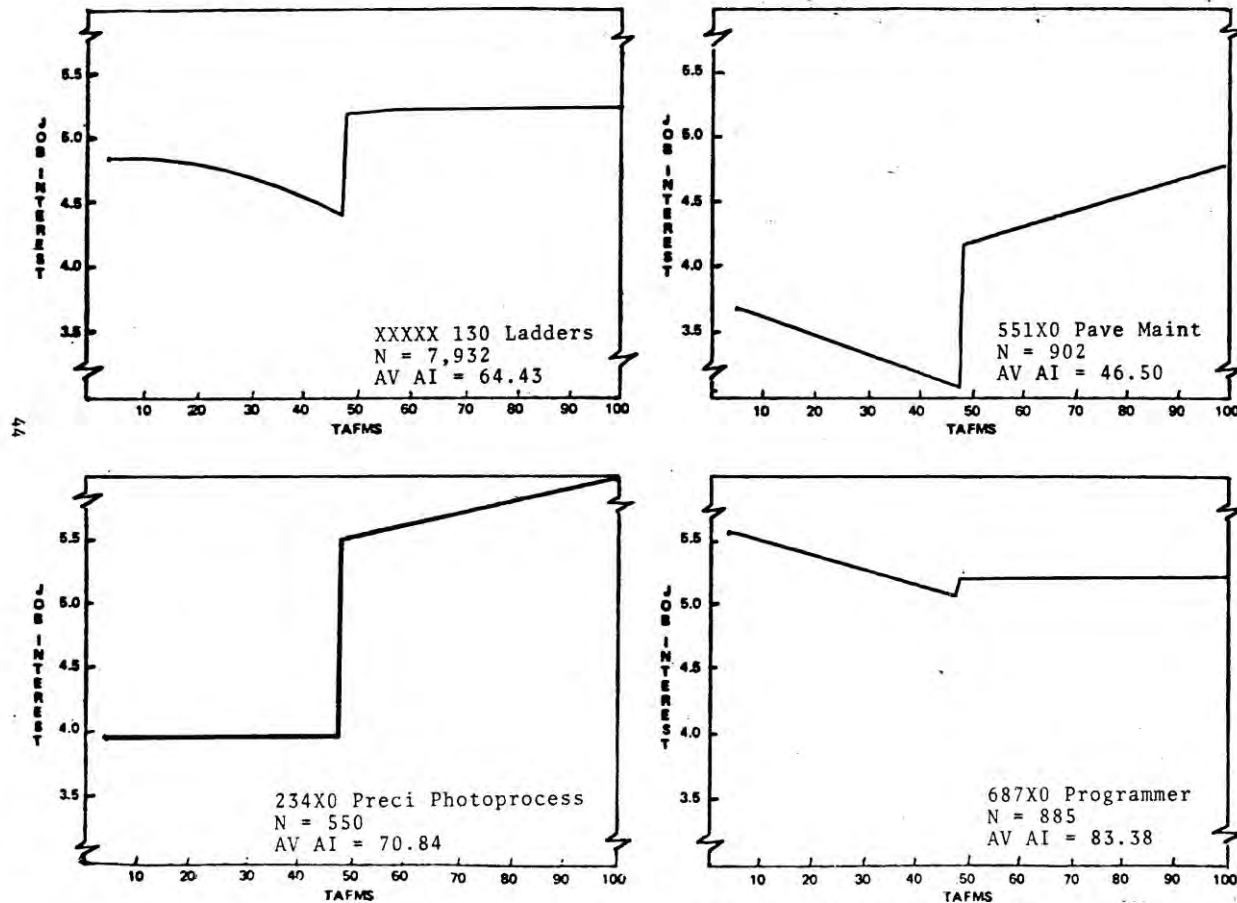


Figure 8. Job Interest versus TAFMS (Holding Aptitude Constant) for Career Fields XXXXX, 551X0, 234X0, and 687X0



SDI and Vocational Interest

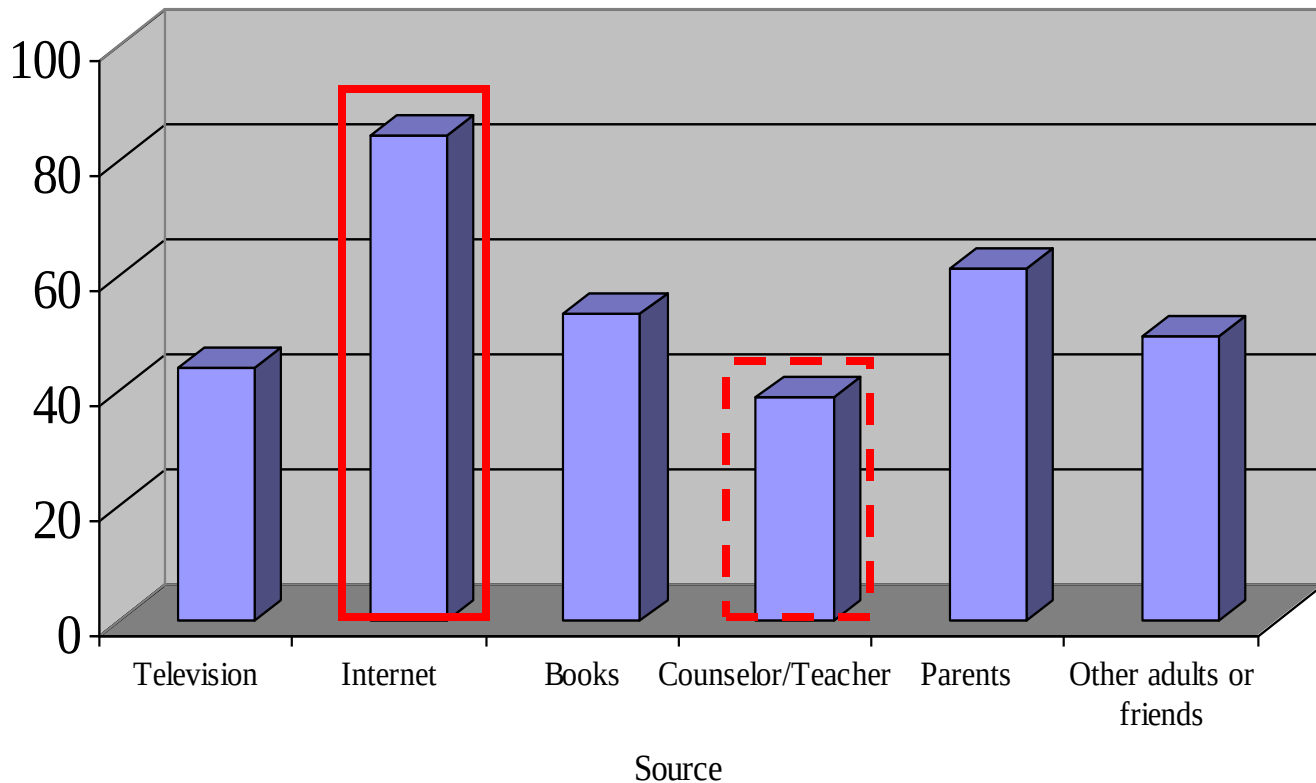
- **Electronics –**
- **Industrious .22**
- **Inventive .20**
- **Science –**
- **Inventive .33**
- **Philosophical .30**
- **Innovative .28**
- **Inquisitive .27**
- **Outdoors**
- **Adventurous .48**
- **Brave .34**
- **Daring .32**
- **Active .27**



Complications

- In the US “All volunteer force”, “volunteer” comes before “force”
- The Millennial Generation Applicant pool...
 - Doesn’t trust “manipulated” information
 - Want’s Choice
 - Searches the Internet for validation

Where would you look for information about careers that interest you?



The Internet is by far the most used source of information, followed by parents, and books. School counselors and teachers are at the bottom of the list.



The Conundrum

- **There are nearly insurmountable obstacles for the Air Force to centrally manage an effective person-job-match system in today's environment.**
 - **Legal – can't use different norms/challenges**
 - **Logistical – can't get into MEPS**
 - **60% of new accessions arrive at Basic Training with “guaranteed” jobs negotiated by the recruiter who has quotas.**



The SDI+ Initiative Proposal

- **Use, in-place Occupational Survey Program to tie personality constellations to job titles within AFSs**
- **Use self-reported Job Satisfaction to determine which response profiles “encourage” or “discourage” reenlistment**
- **Develop Person-Job-Match from job incumbents positive/negative constellations to applicant profiles for self-service Internet-based delivery.**
- **Encourage applicants to bring printed results to recruiters to match “best fit” to available jobs.**

Headquarters Air Force Personnel Center



Integrity - Service - Excellence