AL-TP-1991-0034

ARMSTRONG LABORATORY

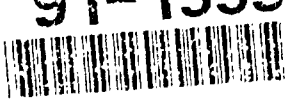
COMPARISON OF EXPERIMENTAL U.S. AIR FORCE AND EURO-NATO PILOT CANDIDATE SELECTION TEST BATTERIES

Thomas R. Carretta

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
Brooks Air Force Base, TX 78235-5000

October 1991

Interim Technical Paper for Period 21 May 1991 - 4 June 1991



AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5000

NOTICES

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.

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

This paper has been reviewed and is approved for publication.

Thomas R. Carretta

THOMAS R. CARRETTA
Project Scientist

William E. Alley

WILLIAM E. ALLEY, Technical Director
Manpower and Personnel Research Division

Roger W. Alford

ROGER W. ALFORD, Lt Col, USAF
Chief, Manpower and Personnel Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1991	Interim - 21 May 91 - 4 Jun 91

4. TITLE AND SUBTITLE
Comparison of Experimental U.S. Air Force and Euro-NATO Pilot Candidate Selection Test Batteries

5. FUNDING NUMBERS
PE - 62703F
PR - 7719
TA - 18
WU - 45

6. AUTHOR(S)
Thomas R. Carretta

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Armstrong Laboratory
Human Resources Directorate
Manpower and Personnel Research Division
Brooks Air Force Base, TX 78235-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
AL-TP-1991-0034

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Armstrong Laboratory Technical Monitor: Dr. Thomas R. Carretta, (512) 536-3943.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Air Force personnel from several countries currently are evaluating computerized test batteries as an adjunct to current pilot candidate selection procedures. This paper describes and compares proposed U.S. Air Force (USAF) and Euro-NATO Aircrew Selection Working Group (ACSWG) pilot candidate selection test batteries. A validation strategy for the ACSWG test battery is described that focuses on test item analyses (e.g., evaluating internal consistency), evaluation of test scoring procedures, evaluation of test battery factor structure, and evaluation of test scores against training performance criteria.

14. SUBJECT TERMS		15. NUMBER OF PAGES
Computerized ENJJPT Pilot	Selection testing Undergraduate Pilot Training	30
		16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

CONTENTS

LIST OF TABLES

PREFACE

This project was completed under work unit 77191845 in support of RPR-78-11, Selection for Undergraduate Pilot Training, issued by Air Training Command. This paper documents a joint USAF/NATO project to develop and validate experimental pilot candidate selection methodologies.

Appreciation is extended to Mr Robert Picascio and Mr Gene Ligon for administrative support and to Maj David Perry, Dr Joseph L. Weeks, and Dr William E. Alley for their technical comments and support during this project.

COMPARISON OF EXPERIMENTAL U.S. AIR FORCE AND EURO-NATO PILOT CANDIDATE SELECTION TEST BATTERIES

SUMMARY

Experimental computerized pilot selection test batteries currently are being evaluated by the United States Air Force (USAF) and by several North Atlantic Treaty Organization (NATO) member countries. This paper describes and compares proposed USAF and Euro-NATO Aircrew Selection Working Group (ACSWG) test batteries in terms of the types of attributes measured and test scoring procedures. A research plan for developing test scoring algorithms and for validating the ACSWG test battery is briefly discussed.

INTRODUCTION

Background

Between 1 November 1990 and 31 October 1993, USAF pilot training applicants from several North Atlantic Treaty Organization (NATO) countries will be tested on an experimental pilot candidate selection test battery developed by the Armstrong Laboratory Human Resources Directorate under the direction of the Euro-NATO Aircrew Selection Working Group (Euro-NATO ACSWG, 1990). The test battery was developed to facilitate a multinational research project regarding the validation of pilot candidate selection instruments.

The test battery is hosted on 2 modified Portable Basic Attributes Test (Porta-BAT systems with rudder pedals) Systems and 2 Determinations Gerat (DTG) test systems. Due to the limited number of test systems, a test schedule has been adopted that will enable participants from several countries to be tested. The test battery will be administered to pilot training applicants targeted for either national pilot training programs or for Euro-NATO Joint Jet Pilot Training (ENJJPT) at Sheppard Air Force Base (AFB). The testing schedule adopted by the ACSWG gives priority to nations who send pilot candidates to ENJJPT. The primary goal of the project is to enable the ACSWG to recommend a test battery to the NATO Air Force Subgroup (AFSG), which oversees ENJJPT. Table 1 summarizes the number of pilot candidates from each country that are expected to participate in this study.

TABLE 1. NUMBER OF SUBJECTS PROJECTED TO BE TESTED ON THE ACSWG TEST BATTERY

Country	Preflying Candidates	ENJJPT
1. Denmark	300	30
2. Italy	50	20
3. Netherlands	40	20
4. Norway	40	20
5. Portugal	50	0
6. Spain	90	0
Total	570	90

Note: A separate validation study is planned for pilot candidates who test on the ACSWG battery and attend their national flying training programs.

The USAF sends the largest number of pilot candidates to ENJJPT (about 150 per year). However, due to administrative reasons no USAF pilot candidates will be tested on the ACSWG test battery.

The primary obstacle to USAF participation is the availability of the test equipment. The only test systems that can support all of the ACSWG tests are being used in Europe (i.e., 2 Porta-BAT systems with rudder pedals and 2 DTG test systems). By the time these test systems are available to test USAF pilot candidates (November 1992), they will not be available for experimental testing. Starting in May 1992, all USAF pilot training applicants will be tested on a computerized test battery hosted on the BAT system (Air Training Command, 1988, Carretta, 1990).

Purpose

As shown in Table 1, only 90 out of 570 of the pilot training applicants to be tested on the ACSWG battery are expected to attend ENJJPT. The remaining 480 applicants either will attend a national pilot training program or will not be selected for training.

Validation of any test instrument involves several activities including: (1) test item analyses (e.g., evaluating internal consistency), (2) evaluation of test scoring procedures, (3) evaluation of test battery factor structure, and (4) validation of test scores against training performance criteria. The projected sample size for this study (N = 570) is large enough to perform the first 3 activities, but is insufficient to perform the validation (N = 90) for the ENJJPT sample.

Preliminary scoring procedures for the ACSWG test battery may yield more than 25 summary scores (see Appendix B). We generally recommend that there be at least 10 times as many observations (i.e., test subjects) as predictors (i.e., test scores) in a regression equation. If the ratio of observations to predictors is too small, the regression weights will not be stable.

The purpose of this study is to consider an alternate validation strategy that would use data from USAF pilot candidates tested on the BAT to bolster the European sample. The USAF and ACSWG test batteries will be compared in terms of the types of attributes being measured. If the USAF test battery adequately measures the attributes measured by the ACSWG test battery, it may be possible to adopt a "domain sampling" approach and combine data from the USAF and European samples in the validation study.

In addition to assisting in the validation study, a comparison between the two test batteries may help identify (1) characteristics not adequately measured by the test batteries, and (2) possible replacement tests for future versions of the test batteries (i.e., to avoid test compromise).

USAF AND ACSWG TEST SYSTEMS

The ACSWG test battery was designed to assess 12 abilities judged to be critical to performance in fighter type aircraft (Rodgers, Carretta, & Hansen, under review; Rodgers & Sage, 1986). The 12 critical abilities and their definitions are presented in Table 2. In the following section, the USAF tests are described in terms of these 12 abilities (where appropriate) to facilitate comparison with the ACSWG test battery.

USAF Test System

Test battery. The USAF test battery consists of 13 subtests that require about 4 h to complete. The test battery was designed to complement the Air Force Officer Qualifying Test

(AFOQT). The BAT battery assesses individual differences in hand-eye coordination, information processing ability, personality, and attitudes. The types of scores generated from these tests include tracking error/tracking difficulty, response time, response outcome (i.e., correct or incorrect), and response choice (i.e., choice between two alternatives). A summary of the USAF test battery is provided in Table 3. The table includes the test name, length, type of scores generated, and attributes measured by each test. Individual test summaries are presented in Appendix A and detailed scoring procedures are presented in Appendix B.

TABLE 2. DEFINITIONS OF ABILITY TERMS USED BY ACSWG

1. *Situational Awareness* - the state of constant mental readiness in order to respond to situational changes.

2. *Memorization* - the ability to remember information (e.g., numbers, words, procedures). Bits of information can be remembered by themselves or with other information.

3. *Reasoning* - the ability to combine separate bits of information and to apply general rules in order to derive logical answers or form conclusions.

4. *Perceptual Speed* - the ability to quickly and accurately perceive small details in patterns and configurations.

5. *Time Sharing* - the ability to (a) observe several sources of information, actions, or tasks at the same time, to (b) combine them and allot task priorities, and to (c) integrate them into actions that have to be performed.

6. *Selective Attention* - the ability to concentrate on the task one is doing (while avoiding distractions).

7. *Response Orientation* - the ability to choose between two or more actions when more than one signal (e.g., lights, sounds, pictures) is presented.

8. *Spatial Orientation* - the ability to tell where you are in relation to an object.

9. *Divided Attention* - the ability to shift back and forth between two or more sources of information.

10. *Psychomotor Coordination* - the ability to coordinate the movements of two or more limbs (i.e., as in moving equipment controls).

11. *Control Precision* - the ability to move the controls of a machine or vehicle. This involves the degree to which these controls can be moved quickly and repeatedly to exact positions.

12. *Visualization* - the ability to imagine the movement of objects in three-dimensional space.

TABLE 3. USAF BASIC ATTRIBUTES TEST BATTERY SUMMARY

	Test Name	Length (mins)	Attributes Measured	Types of Scores
1.	Test Battery Introduction	10	Biographical information	Age, gender, previous flying experience, etc.
2.	Two-Hand Coordination	10	Psychomotor coordination (rotary pursuit)	Tracking error
3.	Complex Coordination	10	Time-sharing psychomotor coordination (compensatory tracking)	Tracking error
4.	Encoding Speed	20	Reasoning (verbal)	Response time, response accuracy
5.	Mental Rotation	25	Visualization	Response time, response accuracy
6.	Item Recognition	20	Memorization (short-term)	Response time, response accuracy
7.	Time-Sharing	30	Time-sharing	Tracking difficulty, response time
8.	Self-Crediting Word Knowledge	10	Verbal ability, self-confidence, self-assessment	Response time, response accuracy, bet
9.	Activities Interest Inventory	10	Attitudes toward risk-taking	Response time, number of high risk choices
10.	Aircrew Personality Profiler	20	Extraversion, agreeableness neuroticism, openness	Response time, response choice
11.	ABCD Working Memory	20	Memorization (working memory), reasoning (verbal), self-confidence	Response time, response accuracy, confidence rating
12.	Anticipation	5	Visualization (dynamic spatial ability)	Tracking error
13.	Pattern Recognition	10	Perceptual speed	Response time, response accuracy
14.	Scanning and Allocating	15	Situational awareness, time-sharing, divided attention, control precision	Tracking error, number of switches made

Note: To facilitate comparison between the USAF BAT and ACSWG test batteries, the ability terms used by the ACSWG (see Table 2) are used to describe the attributes measured by the BAT subtests.

Apparatus. The USAF test battery is hosted on the BAT system. The BAT apparatus consists of a microcomputer built into a ruggedized chassis with glare shield and side panels designed to eliminate distractions. A test subject responds to test stimuli by using individually, and in

combination, a 2-axis joystick on the right side of the apparatus, a single-axis joystick on the left side, and a keypad in the center of the unit. The keypad includes the numbers 0 to 9, an ENABLE key in the center, and a bottom row with YES and NO keys, and 2 others labeled S/L (for same/left responses) and D/R (for different/right responses).

ACSWG Test System

Test battery. The ACSWG test battery consists of 8 subtests, and requires about 2.5 h to complete. The test battery was designed to assess individual differences in 12 psychomotor and cognitive/perceptual abilities, but does not include measures of personality or attitudes. The types of scores generated include tracking error, response time, and response choice (correct or incorrect). The ACSWG test battery is summarized in Table 4. Individual test summaries are presented in Appendix A and detailed scoring procedures are presented in Appendix B.

TABLE 4. EURO-NATO AIRCREW SELECTION WORKING GROUP TEST BATTERY SUMMARY

	Test Name	Length (mins)	Attributes Measured	Types of Scores
1.	Test Battery Introduction	10	Biographical information	Age, gender, previous flying experience
2.	Vigilance	10	Situational awareness, time-sharing, divided attention	Number of routine tasks, number of priority tasks, response time on priority tasks
3.	Matrices	15	Reasoning (nonverbal)	Response time, response accuracy
4.	Digit Recall	5	Memorization	Response time, response accuracy, weighted accuracy score
5.	Complex Coordination	10	Time-sharing, psychomotor coordination (compensatory tracking)	Tracking error
6.	Instrument Comprehension	20	Reasoning, visualization	Response time, response accuracy
7.	Time-Sharing 2	15	Time-Sharing, divided attention, control precision	Tracking difficulty, response time, response accuracy
8.	Scheduling 2	6	Situational awareness, time sharing, divided attention	Number of points accumulated
9.	DTG	20	Selective attention, response orientation	Response accuracy

Apparatus. The first 7 ACSWG tests are hosted on a modified Porta-BAT, where the left joystick is disconnected and replaced by 2 rudder pedals that are bolted to the floor of the test unit.

The DTG test device is used to administer a test measuring response orientation and selective attention (which cannot be administered by the Porta-BAT). The DTG is able to present both visual (colored lights) and auditory (high or low tone over headphones) stimuli. A subject responds by pressing buttons and/or foot pedals. The types of scores generated are based on response accuracy.

DISCUSSION

The USAF test battery provides measures of 9 of the 12 fighter-type pilot characteristics discussed in Table 1. This test battery does not include measures of selective attention, response orientation, or spatial orientation. The USAF test battery includes 2 or more measures of several characteristics including memorization, reasoning, time-sharing, psychomotor coordination, and visualization. Furthermore, the USAF test battery contains measures of several personality constructs not measured by the ACSWG test battery (i.e., self-confidence, extraversion, agreeableness, neuroticism, and openness).

The ACSWG test battery includes 2 or more measures of 4 of the 12 crucial fighter-type pilot characteristics (i.e., situational awareness, reasoning, time-sharing, and divided attention) and single measures of 6 others (i.e., memorization, selective attention, response orientation, psychomotor coordination, control precision, and visualization). Neither perceptual speed nor spatial orientation are measured. Both of these constructs were thought to be measured by the Instrument Comprehension test (Rodgers, Carretta, & Hansen, under review). The Instrument Comprehension test requires both reasoning and visualization skills to interpret the instruments and mentally rotate the planes' image in 3-dimensional space. The test does not require a point-by-point comparison of small details in patterns or configurations (i.e., perceptual speed) or the ability to tell where you are in relation to some object (i.e., spatial orientation).

The USAF and ACSWG test batteries appear to overlap enough, in terms of the characteristics being measured, in order to perform a meta-analysis.

One approach would be to combine test scores within the 2 batteries to form ability composites (i.e., combine scores hypothesized to be measuring the same characteristics or combine similar types of scores). The composite scores then could be used in place of raw test summary scores when developing regression equations to predict pilot training performance.

These composites could be constructed using either an empirical (i.e., data driven) or judgment (i.e., expert opinion) approach. The empirical approach could be accomplished through separate factor analyses of the summary scores from the 2 test batteries. The intercorrelations among the actual test scores would determine the dimensionality of the test batteries. Separate analysis will need to be performed for the 2 batteries as no subjects will be tested on both the USAF and ACSWG test batteries. USAF pilot applicants will not be able to test on the ACSWG battery due to time limitations (the USAF battery requires 4 h and the ACSWG battery 2 h) and the limited availability of the DTG. European pilot candidates cannot test on both batteries due to time limitations and the nonavailability of the USAF test battery for non-USAF applicants (i.e., the USAF test battery is being treated as controlled material).

In the judgment approach, subject matter experts (i.e., psychologists) would be required to make judgments about which scores are related to each other. Those scores judged to be measuring the same characteristic could then be combined to form unit-weighted composites. The judgment approach should be considered as a fall-back alternative, if the empirical approach

is not successful (i.e., the factor analyses do not provide interpretable solutions or do not produce similar factors for the 2 test batteries).

Once meaningful summary scores have been developed, attention should focus on interpreting these scores in the context of each participating nation is operational pilot candidate selection procedures. For instance, pilot candidates from several NATO countries are operationally tested on some version of the Instrument Comprehension Test. Range restrictions will occur on the abilities measured by this test, which will reduce the magnitude of the relationship between Instrument Comprehension test scores from the ACSWG battery and flying training performance. To correct for range restrictions due to preselection, detailed information is needed regarding the selection procedures used by each nation participating in this project (e.g., abilities measured by operational selection instruments, pilot applicant normative data, selection rate).

Additional details regarding the data analysis plan are under review by the ACSWG. A detailed report is expected by June 1992.

REFERENCES

Air Training Command (1988). *Statement of operational need (SON) for the pilot selection and classification system (PSACS)*, ATC 003-88.

Carretta, T.R. (1990). *Proposed USAF Pilot Selection and Classification Systems*. Invited address to the Eighth Annual International Conference on Aviation Physiology and Training, Southampton, PA.

Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

Euro-NATO Aircrew Selection Working Group (1990). *Proceedings of the 14th Meeting of the Euro-NATO Aircrew Selection Working Group*, Florence, Italy.

Rodgers, M.N., Carretta, T.R., & Hansen, I. (under review). *Identifying Ability Requirements and Tests for (NATO) Fast-Jet Pilot Selection*.

Rodgers, M.N., & Sage, P.N. (1986). *Identification of Ability Requirements for NATO Fast-Jet Pilot Selection*. Willowdale, Ontario. Canadian Forces Applied Research Unit.

Siem, F.M. (1990). *Predictive validity of an automated personality inventory for air force pilot selection* (AFHRL-TP-90-55, AD-A226 731). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

APPENDIX A
USAF AND ACSWG TEST BATTERY DESCRIPTIONS

USAF BATTERY DESCRIPTIONS

As with the USAF and ACSWG test battery summaries presented in Tables 3 and 4, the tests are described in terms of the 12 critical abilities from Table 2 (where appropriate) in order to facilitate comparison between the USAF and ACSWG test batteries.

Test Battery Introduction

This interactive subprogram prompts the subject to provide background information (e.g., identity, age, gender) as well as personal history and attitudes related to flying.

Two-Hand Coordination

The Two-Hand Coordination test is a variation of a rotary pursuit task. The airplane (target) moves in an elliptical path on the screen at a rate of 20 cycles per minute. The rate of movement of the airplane within each cycle varies in a fixed sinusoidal pattern. The subject controls the vertical and horizontal movement of a small "gunsight" using a left and right joystick. The left-hand joystick controls the vertical movement of the gunsight while the right-hand joystick controls the horizontal movement of the gunsight. The subject's task is to keep the gunsight on the airplane as it moves. After receiving instructions, the subject completes a 3-min practice session and a 5-min test.

The measures of interest are horizontal and vertical tracking error scores. The psychological factor for this test is psychomotor coordination (i.e., low to moderate order tracking and time-sharing ability in pursuit). This test requires about 10 min to complete.

Complex Coordination

Complex Coordination uses a dual-axis joystick (right hand) to control the horizontal and vertical movement of a cursor. The left-hand joystick controls the left-right movement of a vertical "rudder bar" of light at the base of the screen. The subject's task is to maintain the cursor (against a constant horizontal and vertical rate bias) centered on a large cross fixed at the center of the screen while simultaneously centering the rudder bar at the base of the screen (also against a constant rate bias). The instruction, practice, testing, and scoring are as in the Two-Hand Coordination test. The Complex Coordination text assesses psychomotor coordination and time-sharing ability (i.e., compensatory tracking ability involving multiple axes). This test requires about 10 min to complete.

Encoding Speed

The subject is presented simultaneously with 2 letters and is required to make a same-different judgment about the letter pair. The judgment may be based on Physical identity (AA vs. Aa), Name identity (AA vs. AH), or Category identity (vowels vs. consonants - AE vs. AH). The latency of the encoding judgment provides a measure of the speed of the cognitive encoding process.

Response time and accuracy (correct/incorrect) are recorded on each of the 96 items (32 items in each condition). The psychological factor involved in this test is verbal reasoning at several levels of cognitive operation. Test administration time is about 20 min.

Mental Rotation

The subject is presented sequentially with a pair of letters and is required to make a same-different judgment. The letter pair may be either identical or mirror images, and the pair may be either in the same orientation or rotated in space with respect to each other. A correct "different" judgment is associated with a mirror image pair and is not dependent on the relative rotation of the 2 letters.

To perform the task, the subject must form a mental image of the first letter (no longer displayed) and perform a point-by-point comparison with the second letter (which remains on the screen). In addition, when the letters are rotated with respect to each other, the subject must mentally rotate the mental image of one letter into congruence with the other before making the comparison.

Speed and accuracy of response are recorded on each of the 72 items. The psychological factor assessed by this test is visualization ability. The Mental Rotation test requires about 25 min to complete.

Item Recognition

The subject is instructed to remember a string of digits. After the digit string has been removed, the subject is shown a single digit. The subject's task is to decide if the single digit was one of those presented in the initial string. The subject is instructed to respond by pressing a keypad button marked YES if the single digit was in the string or another marked NO if it was not. The instructions inform the subject to work as quickly and accurately as possible. Speed and accuracy of response are recorded on each of the 48 items.

Short term memory storage, search, and comparison operations are the underlying psychological factors for this test. This test requires about 20 min to complete.

Time-Sharing

During a series of ten 1-min trials, the subject is required to learn a compensatory tracking task. To perform this task, the subject must anticipate the movement of a marker on a screen and operate a control stick to counteract that movement in order to keep the marker aligned with a fixed central point. Task difficulty is adjusted throughout the test depending on the subject's performance. The control dynamics are a combination of rate and acceleration components. The "disturbance" factor is a quasi-random summed sinusoidal forcing function.

After ten 1-min "tracking only" trials, the subject is required to track while canceling digits that appear at random intervals and locations on the screen (six, 1-min trials). The digits are canceled when the subject presses the corresponding key on the response keypad. A "cross-adaptive" logic forces the subject to respond to digits within 4 s after their appearance. If the subject fails to respond within 4 s, he/she loses control of the gunsight until a response is made. These dual-task trials occur in two 3-min blocks. The information processing load gradually increases during these trials. The Time-Sharing test ends with a final 3-min block of "tracking only" trials.

The effects of the secondary task loads are reflected in the pattern of level of difficulty changes caused by the adaptive logic that holds tracking error constant. The measure of interest for this test is the level of difficulty that the subject can perform at consistently. This test assesses time-sharing ability as a function of differential task load and learning rate on the tracking task. This test requires about 30 min to administer.

Self-Crediting Word Knowledge

This test is essentially a vocabulary test where the subject is presented with a "target" word and 5 other words from which its closest synonym must be chosen. There are 3 blocks of 10 questions each. The target words become increasingly difficult with each successive block. The subject is informed of this increasing difficulty and is required to make a bet before each block which reflects how well the subject expects to do. Response time and accuracy are recorded on each of the 30 items. This test assesses verbal ability, self-assessment ability, and self-confidence. The test requires about 10 min to complete.

Activities Interest Inventory

This test is designed to determine the subject's interest in various activities. The subject is presented with 81 pairs of activities and is asked to choose between them. The subject is told to assume that he/she has the necessary ability to perform each activity. The activity pairs force the subject to choose between tasks that differ on threat to physical survival, sometimes subtly, sometimes not. The measures of interest are the number of high-risk activities chosen and the amount of time required to choose between pairs of activities. The psychological factor assessed by this test is attitudes toward risk. Test administration time is about 10 min.

Aircrew Personality Profiler

This questionnaire examines the subject's attitudes and interests. The subject is given 202 items, each requiring a choice between 2 alternatives. The subject is instructed not to spend time pondering responses, but to give the first natural answer as it comes. The questionnaire is a traditionally formulated personality inventory specially compiled in cooperation with the Armstrong Laboratory and targeted for aircrew selection and classification. The questionnaire requires about 20 min to complete and assesses the psychological constructs of extraversion, agreeableness, neuroticism, and openness.

ABCD Working Memory

The subject is presented with a series of rules regarding 2 sets of letters. Set 1 always contains A and B; and set 2 always contains C and D. The subjects' task is to apply the rules to determine the order of the letters. For example: (Frame 1) "C" is not followed by D, (Frame 2) "B is preceded by A," (Frame 3) "Set 1 is not preceded by Set 2." After reading each rule the subject presses the ENABLE key to display the next rule (the previous rule is removed from the screen). Once all 3 rules have been reviewed the subject presses the ENABLE key again to display the answer alternatives as shown below:

1. ABCD
2. ABDC
3. BACD
4. BADC

5. CDAB
6. CDBA
7. DCAB
8. DCBA

The subject must respond by entering the number of the correct letter sequence using the data entry keypad. Response time and accuracy (correct/incorrect) are recorded on each of the 48 items. Response time is defined as the total amount of time spent reviewing the sequencing rules.

A self-confidence scale also is scored on each item. After completing each item the subject is required to indicate his/her level of confidence that the item was answered correctly on a 9-point scale ranging from (1) not at all confident to (9) extremely confident. Item difficulty is varied by making the rules either active ("A precedes B") or passive ("A is followed by B"); and either positive ("B follows A") or negative ("B is not followed by A"). This test assesses memorization (working memory), verbal reasoning, and self-confidence. This test requires about 20 min to complete.

Anticipation

The psychological factor assessed by this test is visualization (i.e., dynamic spatial ability). A target moves from left to right on the screen. When the target reaches point "A" on the screen it disappears, but continues to move to the right. The subjects' task is to estimate when the target reaches point "B," (located to the right of point "A"). When the subject thinks the target has reached point "B" he/she presses the ENABLE key to record the target's actual position. The target then reappears at the point where the subject stopped its movement. Target movement rate is constant within a trial but varies quasi-randomly between trials. The distance of the target from point "B" is recorded on each of the 50 trials (5 different points where the target disappears x 10 replications). This test requires about 5 min to complete.

Pattern Recognition

This test assesses perceptual speed. A 5 by 5 grid of black and white squares is displayed for 2 s. The pattern is then removed and followed, after a brief delay by 5 similar patterns. The subjects task is to indicate which of the 5 patterns is identical to the original pattern. There are 3 blocks of 10 items. Average pattern difficulty increases from the first to the third block. Response time and response accuracy are recorded on each item. This test requires about 10 min to complete.

Scanning and Allocating

The subject is presented with a box with a cross within it in the upper left hand corner of the screen. During the test, the cross will move left or right away from its vertical alignment. The subject's task is to maintain the vertical alignment of the cross using the right hand joystick. Moving the joystick to the right moves the cross to the right while moving the joystick to the left moves the cross to the left. An alignment mark at the top of the box is provided to serve as a reference point. After 1 min, and each minute thereafter, an additional box will appear in one of the remaining corners until there are a total of 4 boxes with crosses. The subject is instructed to maintain simultaneous alignment when 2 or more boxes appear on the screen, even though only 1 box can be controlled (active) at a time. To activate a particular box, the subject must press the number of that box (1-4) on the keypad. The active box is identified

by the blinking box number on the screen. The task is administered 3 times with a 1-min break between trials.

The measures of interest for this test are the amount of alignment error under different levels of workload and reaction time. Scanning and Allocating provides a psychological measure of situational awareness, time-sharing, divided attention, and control precision. The test requires about 15 min to complete.

ACSWG BATTERY DESCRIPTIONS

Test Battery Introduction

This interactive subprogram prompts the subject to provide background information (e.g., identity, age, gender) as well as previous flying experience.

Vigilance

A 9 block by 9 block grid appears on the screen. The numbers along the left side of the grid are the "A" coordinates and correspond to the rows. The numbers along the top of the grid are the "B" coordinates and correspond to the columns. Each block in the grid can be identified by its row and column coordinates.

During the test, asterisks "*" appear within the blocks of the grid. The subject's "routine task" is to cancel these asterisks as quickly as possible. An asterisk is canceled (i.e., erased) by entering its row and column coordinates on the keypad.

In addition to the asterisks, arrows "^" may appear in the blocks of the grid. The arrows represent an "emergency task." The subject is instructed to respond as quickly as possible when an arrow (i.e., emergency) appears. This function is done by pressing the ENABLE key, then entering the row and column coordinates of the arrow (i.e., emergency). The subject is instructed that the arrows must be canceled in their order of appearance. When no arrows are present (i.e., there are no emergencies) the subject is instructed to resume performing the routine task (i.e., canceling asterisks).

The measures of interest for this test are the number of routine and emergency tasks successfully completed and the average response time required to complete emergency tasks.

The psychological factors assessed by this test include situational awareness, time-sharing, and divided attention. The test requires about 10 min to complete.

Matrices

In the Matrices test, a picture of an incomplete geometric pattern appears on the screen. (The lower right-hand corner of the pattern is missing). The subject's task is to choose from several alternatives, which would correctly complete the pattern. The subject indicates his/her choice by entering the number of the chosen alternative on the response keypad.

The first 6 items are for practice only. The remaining 30 items are test items and are scored. In this test and in the Instrument Comprehension test (see below), the subject may go backwards and forwards through the test. The subject has the option to skip items, to review items previously answered or skipped, and to change answers. This test has a 10-min

time limit that begins when the subject starts the first test item (i.e., review of instructions and practice items are not timed). Response speed and accuracy are recorded on each item. The Matrices test assesses non-verbal reasoning and requires about 15 min to complete.

Digit Recall

In this test, a number string appears on the screen. After a few seconds, the number string is removed from the screen and is replaced by a string of empty boxes. (The number of boxes is equal to the number of numbers in the string). The subject's task is to enter the number string into the boxes.

The length of the strings vary from 7 to 12 numbers. Response time and response accuracy are recorded on each of the 30 test items. Response accuracy is calculated in 3 ways for each item: (1) correct or incorrect, (2) number of numbers placed in their correct position in the string, and (3) a weighted scoring algorithm that gives partial credit for numbers placed out of sequence.

This test measures short-term memory and requires about 10 min to complete.

Complex Coordination

The ACSWG Complex Coordination test is a variation of the USAF test with the same name. In this test, a dual-axis (right hand) joystick is used to control the horizontal and vertical movement of a cursor. Rudder pedals (instead of the left-hand joystick in the USAF version) are used to control the left-right movement of a vertical "rudder bar" of light at the base of the screen. The subject's task is to maintain the cursor (against a constant horizontal and vertical rate bias) centered on a large cross fixed at the center of the screen while simultaneously centering the rudder bar at the base of the screen (also against a constant rate bias). After receiving instructions, the subject completes a 3-min practice session and a 5-min test. The Complex Coordination test assesses psychomotor coordination and time sharing ability (i.e., compensatory tracking ability involving multiple axes) and requires about 10 min to complete.

Instrument Comprehension

In this test, an illustration of an airplane in 5 different positions is shown on the screen. An artificial horizon indicator and a compass are displayed above these aircraft. The subject's task is to determine which of the aircraft agrees with the readings on the artificial horizon indicator and compass. The subject indicates his/her choice on each item by pressing the numbered key that corresponds to the chosen alternative.

The test items begin after the subject has completed the instructions and practice items. As with the Matrices test (see above), the subject may go backwards and forwards through the test, and may skip items, review items previously answered or skipped, or change answers.

This test has 60 test items and a 15-min time limit that begins when the subject starts the first test item (i.e., review of instructions and practice items are not timed).

Response speed and response accuracy are recorded on each item. The Instrument Comprehension test measures reasoning and visualization. Total test administration time, including instructions, practice, and test items is about 20 min.

Time-Sharing 2

Two distinctly different kinds of tasks are involved in this test. The first is a measure of hand-eye coordination and the second is a measure of attention.

The first three 1-min trials involve tracking only to provide a pure estimate of the subject's psychomotor coordination. During these trials, a "stationary image" of an aircraft and a "gunsight" that move to the left or right are displayed on the screen. The subject must maneuver the right-hand joystick to keep the gunsight centered on the airplane.

The next two 1-min trials involve detecting and responding to missing numbers. The numbers appear one at a time in sequence on the lower part of the screen. The number sequence is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, 1, 2, etc. Occasionally, a number will be missing from the sequence (e.g., 0, 1, 2, 3, 4, 6, 7, 8 .... [5 is missing]). The subject's task is to type the missing number on the response keypad when it does not appear in the sequence. Subjects are scored on both response speed and accuracy.

The final 5 1-min trials involve both tracking and missing digits. While the subject is maneuvering the right-hand joystick to keep the gunsight on the airplane, he/she also must monitor the number counter in order to be able to detect the missing numbers.

This test assesses the psychological factors of time-sharing, divided attention, and control precision. Test administration time is about 15 min.

Scheduling 2

In this test, the subject is presented with 5 horizontal scales that can range in value between 0 and 10 points. Each scale increases at a unique, constant rate. Each scale appears on a separate screen and may be viewed by entering the scale number on the response keypad (1, 2, 3, 4, or 5). The subject "scores" points equal to the current value of the displayed scale by pressing the ENABLE key. When the ENABLE key is pressed, the current value of the displayed scale is added to the subject's total score and the scale is reset to 0 (where it will start incrementing again).

If the value of a scale reaches its upper limit, and the subject has not responded by pressing the ENABLE key, the value of the scale will return to 0 without the subject receiving any points for that scale.

The Scheduling 2 test measures the psychological constructs of situational awareness, time-sharing, and divided attention. This test requires about 6 min to complete.

Determinations Gerat

This test is hosted on the DTG test device rather than the Porta-BAT system. The subject's task is to respond as quickly and accurately as possible to auditory (i.e., high and low tones) and visual signals (i.e., colored lights) by pressing foot pedals and/or buttons. Subjects are presented with detailed instructions as to how to respond to different visual and auditory signals.

The test begins with several practice items that are neither timed nor scored. The purpose of the practice items is for the subject to learn the appropriate response to each stimulus.

After reviewing the instructions and completing the practice items, the subject is presented with a series of test items. The test consists of 5 groups of items. There are 150 items in Group 1 and 75 items in Groups 2 through 5 (150 + 4 (75) = 450 items). As described earlier, during the practice items there is no time limit imposed on the subject. During the actual test, the subject must respond to items within a fixed limit. Failure to respond in time results in an item being scored as incorrect. The interstimulus interval is decreased on each successive group of test items, thus increasing task difficulty.

The scoring procedure for this test produces several response accuracy scores including the number of correct responses, number of nonresponses, number of delayed correct responses, and number of incorrect responses. The DTG test is designed to measure selective attention and response orientation. The instructions, practice, and test items require about 20 min to complete.

APPENDIX B
SCORING PROCEDURES FOR THE USAF AND ACSWG TEST BATTERIES

USAF Test Battery

Scoring procedures for the USAF Basic Attributes Test (BAT) battery rely on a combination of tracking error and difficulty, response speed, response accuracy and response choice, and in some instances, specially derived scores (e.g., personality scale scores). Details regarding scoring procedures for the 13 USAF pilot candidate selection and classification tests are provided below.

Two-Hand Coordination

X1 Tracking Error (PS2X1). Cumulative tracking error for the X1 axis (horizontal displacement of the cross from the target) for the final 2 min of the test period.

Y1 Tracking Error (PS2Y1). Cumulative tracking error for the Y1 axis (vertical displacement of the cross from the target) for the final 2 min of the test period.

Complex Coordination

X2 Tracking Error (PS2X2). Cumulative tracking error for the X2 axis (horizontal displacement of the cross from the center of the screen) for the final 2 min of the test period.

Y2 Tracking Error (PS2Y2). Cumulative tracking error for the Y2 axis (vertical displacement of the cross from the center of the screen) for the final 2 min of the test period.

Z2 Tracking Error (PS2Z2). Cumulative tracking error for the Z2 axis (displacement of the rudder bar from the center point at the bottom of the screen) for the final 2 min of the test period.

Encoding Speed

Several of the tests in the USAF battery rely on response latencies (in milliseconds) as an indicator of test performance. The standard scoring technique for tests of this type uses data only from items that were answered correctly when computing summary scores. For the USAF battery, this procedure includes the following tests: Encoding Speed, Mental Rotation, Item Recognition, Self-Crediting Word Knowledge, ABCD Working Memory, and Pattern Recognition.

The following scores are used to evaluate performance on the Encoding Speed Test:

Average Response Time (ENCRT). Average response time in milliseconds based on all items answered correctly.

Percent Correct (ENCPER). Percent correct.

Mental Rotation

Average Response Time (MRTRT). Average response time in milliseconds, based on all items answered correctly.

Percent Correct (MRTPER). Percent correct.

Item Recognition

Average Response Time (ITMRT). Average response time in milliseconds, based on all items answered correctly.

Percent Correct (ITMPER). Percent correct.

Time-Sharing

Average Tracking Difficulty (TMSDIF). Average tracking difficulty achieved during the dual-task trials (minutes 11-16).

Average Response Time (TMSRT). Average response time in milliseconds to cancel the digits that appear during the dual-task trials (minutes 11-16).

Self-Crediting Word Knowledge

Average Response Time (WKART). Average response time in milliseconds based on all items answered correctly.

Percent Correct (WKAPER). Percent correct.

Average Bet (WKABET). Average bet across the three blocks of items.

Activities Interest Inventory

Number of High-Risk Choices (AIAHIR). This score indicates the number of high-risk choices made by the subject.

Average Response Time (AIART). Average response time across all 81 items (in milliseconds).

Aircrew Personality Profiler

The scoring procedure for this test has not yet been decided on. To date, most research with this test has focused on creating several scales (e.g., Hostility, Self-Confidence, Values Flexibility) by combining the 202 items on this test (Siem, 1990). Current plans are to reanalyze this test in terms of the Five Factor Model of Personality (Digman, 1990).

ABCD Working Memory

Average Response Time. Average response time in milliseconds based on all items answered correctly.

Percent Correct (ABCPER). Percent correct.

Confidence Rating (ABCCON). Average self-confidence score across all items. This score may range between 1 and 9.

Anticipation

Average Tracking Error (ANTAAE). This is the average of the absolute value of the tracking error score across all 50 items.

Pattern Recognition

Average Response Time (PATRT). Average response time in milliseconds based on all items answered correctly.

Percent Correct (PATPER). Percent correct.

Scanning and Allocating

Average Tracking Error (SAATE3). Average cumulative tracking error (vertical displacement of the 4 crosses from their alignment markers) for the final 4 min of the test.

ACSWG TEST BATTERY

Vigilance

Total Number of Routine Tasks (VIGNRT). Total number of routine tasks completed (i.e., number of asterisks canceled) during the 5 min test.

Total Number of Priority Tasks (VIGNPT). Total number of priority tasks completed (i.e., number of arrows canceled) during the 5 min test.

Average Response Time (VIGRT). Average response time to perform the priority tasks (i.e., average response time to cancel arrows).

Matrices

Average Response Time (MTXRT). Average response time in milliseconds based on all items answered correctly.

Number Correct (MTXNCOR). Number of items answered correctly.

Number Completed (MTXNCOM). Number of items completed (out of 30).

Percent Correct, Completed (MTXPERC). MTXNCOR divided by MTXNCOM.

Percent Correct, All (MTXPERA). MTXNCOR divided by 30.

Digit Recall

Average Response Time (DRERT). Average response time in milliseconds based on all items answered correctly (i.e., with a Perfect score equal to 1, see below).

Perfect Score (DREPS). The total number of test items where the subject entered all numbers in a string in the correct order.

Normal Score (DRENS). This scoring procedure awards the subject one point for each number that is placed in its correct position in the number string. Thus, the maximum number of points possible for any string is the number of digits in the string. DRENS is the average Normal score across all items.

Masters Score (DREMS). This scoring procedure awards 2 points for each number placed in its correct position in the string and 1 point for a number that was in the string but placed in the wrong position. DREMS is the average Masters score across all items.

Percent Correct (DREPER). Percent correct (i.e., percentage of items with a DREPS score equal to 1).

Complex Coordination

X2 Tracking Error (PS9X2). Cumulative tracking error for the X2 axis (horizontal displacement of the cross from the center of the screen) for the final 2 min of the test period.

Y2 Tracking Error (PS9Y2). Cumulative tracking error for the Y2 axis (vertical displacement of the cursor from the center of the screen) for the final 2 min of the test period.

Z2 Tracking Error (PS9Z2). Cumulative tracking error for the Z2 axis (displacement of the rudder bar from the center point at the bottom of the screen) for the final 2 min of the test period.

Instrument Comprehension

Average Response Time (ICMRT). Average response time in milliseconds based on all items answered correctly.

Number Correct (ICMNCOR). Number of items answered correctly.

Number Completed (ICMNCOM). Number of items completed (out of 60).

Percent Correct, Completed (ICMPERC). ICMNCOR divided by ICMNCOM.

Percent Correct, All (ICMPERA). ICMNCOR divided by 60.

Time Sharing 2

Average Tracking Error (TM2TED). Average cumulative tracking error per minute, dual-task trials only (minutes 6-10).

Average Response Time (TM2RT). Average response time to respond to missing numbers, dual-task trials (minutes 6-10).

Percent Correct (TM2PER). Percent correct, dual-task trials (minutes 6-10).

Scheduling 2

Total Number of Points (SD2TPA). Total number of points accumulated (minutes 1-5).

Determinations Gerat

Number Correct (DTGNC). Total number of correct responses.

Number Nonresponses (DTGNNR). Total number of items not responded to.

Number Delayed Correct (DTGNDC). Total number of items responded to correctly *after* the response interval had expired.

Number Incorrect (DTGNI). Total number of incorrect responses.