Special Report - 77-2

12

AD A048105

AVIATOR SELECTION 1919 - 1977

Robert A. North and Glenn R. Griffin

DDC
RECEIVED
JAN 4 1978
B

4 October 1977

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
PENSACOLA FLORIDA

AD NO.
DDC FILE COPY

Approved for public release; distribution unlimited;

Approved for public release; distribution unlimited.

Aviator Selection 1919-1977

Robert A. North and Glenn R. Griffin

Naval Medical Research and Development Command
MF51.524.002-5012DX5X

Approved by

Ashton Graybiel, M.D.
Assistance for Scientific Programs

Released by

Captain R. E. Mitchel, MC, USN
Commanding Officer

4 October 1977

Naval Aerospace Medical Research Laboratory
Naval Air Station
Pensacola, Florida 32508

DDC
RECEIVED
JAN 4 1978
B

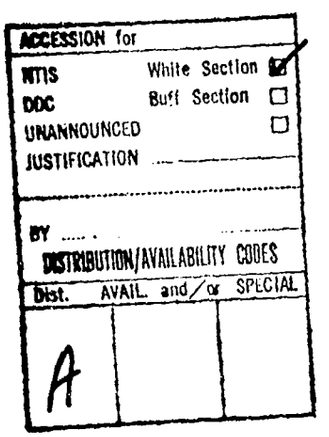
Table of Contents

		Page
1.	INTRODUCTION	1
2.	OVERVIEW OF AVIATOR SELECTION, WORLD WARS I AND II	1
	Selection Testing	1
	Early Selection Test Development: World War II Aviators	3
	Civilian Pilot Training (CPT) Selection Efforts (1939-1941)	4
	World War II Candidate Selection: Military Studies	5
	Apparatus Tests Used by the Army Air Force	7
	Overall Success of Army Air Force Tests	10
3.	POST-WAR AVIATOR SELECTION RESEARCH	11
	Psychomotor Testing	11
	Perceptual/Cognitive Paper-and-Pencil Test Selection Research	12
	Personality Test Selection Research	13
	Neurological Research	19
4.	NEW DIRECTIONS IN POST-WAR AVIATOR SELECTION RESEARCH	19
	Task and Factor Analysis Studies	19
	Isolating Behaviors for Psychological Assessment	22
	Concurrent Information Processing	24
	Decision Making Capability	24
	Residual Attention	25
	Tests of Anticipatory Behavior	25

AD

Table of Contents
(Continued)

		Page
5.	CURRENT EMPHASIS IN AVIATOR SELECTION RESEARCH . . .	25
	Criterion Measurement	25
	ACMR/ACMI the Ultimate Criterion?	28
	Synthetic Selection Research	30
	Air Combat Simulators	31
	Perceptual Psychomotor Performance	31
	Division of Attention	32
	Voice Analysis as a Measure of Psychological Stress . . .	33
	Vestibular Disorientation Research	34
	Human Anthropometry in Aircraft Assignment	34
6.	CONCLUSIONS/RECOMMENDATIONS	34
	Conclusions Concerning Selection Measures	34
7.	REFERENCES	37



SUMMARY PAGE

THE PROBLEM

This report is the second in a series of two literature reviews designed to provide an updated summary of psychological assessment research in aviator selection. The first report was specifically related to Navy aviator attrition research. One purpose of this review is to provide a wide-range description of tri-service aviator selection testing methods and assess their predictive improvement. A second purpose is to suggest methods to improve the prediction of aviator success based upon results and findings in the research literature. Hopefully, this review will act to stimulate additional relevant research and evaluation efforts which have the potential for the improved selection of aviators for initial undergraduate training and advanced performance in operational environments.

FINDINGS

The potential for increased success in predicting aviator performance is high. The fact that current selection tests normally account for less than half of the total variance associated with aviator success (in training) suggests that there are additional factors associated with aviator performance which are not now being adequately assessed. The lack of any prominent breakthrough in perceptual/cognitive paper-and-pencil testing since the war years (WW-II) suggests that non-paper-and-pencil performance tests should be investigated more fully to determine their relationship to aviator performance in both a training and operational setting.

RECOMMENDATIONS

Relating aviator performance to better and more appropriate performance measurement criteria is a continuing psychological assessment goal. New technological advancements such as the Navy and Air Force Air Combat Maneuvering Ranges have the potential to identify and reliably measure relevant physical and psychological human attributes which may provide more accurate and valid prediction of aviator operational performance.

Still, such obviously valid criteria as ACMR performance, pose an interesting assessment problem. It is unclear whether the prediction variables presently utilized in aviator selection to predict successful performance in undergraduate training, are related to successful performance in post-graduate operational environments.

It is suggested that research be oriented toward the identification of highly relevant criterion-oriented performance measures for use as criteria in the evaluation of present and new selection prediction variables and identification and development of non-paper-and-pencil performance prediction measures to

improve prediction of criterion performance in undergraduate training, and in post-graduate operational flying environments. Examples of non-paper-and-pencil performance prediction measures recommended for future study are Selective and Divided Attention, Stress and Anxiety Motivational Measurement, and Perceptual Psychomotor skill assessment.

INTRODUCTION

From the beginning of aircraft development and the subsequent integral role of pilots in military and civilian transportation systems, there have been efforts to select individuals for aviator training that possess both physical and mental attributes conducive to success in flying training. The high cost of flight training and the relative high rate of failure with its resultant loss of monetary expenditure justifies a continual selection research effort. The cost of training pilots is extremely high and continues to increase. Majesty (92) indicated that in 1975 the cost of Air Force Undergraduate Pilot Training averaged $160,000 per individual pilot, with an additional $300,000 expended in the pilot's transition to an operational F-4 aircraft. A comparable figure is expended by the Navy in its jet pilot training program (25).

There is presently a relatively sizeable attrition rate in pilot training programs. Griffin and Mosko (62) indicate that the Navy attrition rate averaged approximately 30 percent from 1962 to 1977. Schweitzer (118) indicates that the Air Force has experienced an attrition rate of from 23 to 28 percent from 1965 to 1975. The types and descriptions of aviator attrition of the two major service producers of fixed wing pilots is depicted in Table 1.

Present rates of attrition, though excessively high, are a far cry from those reported prior to the utilization of psychological testing devices. Majesty (62) states that in those early periods (pre-World War II) it was not uncommon for attrition rates to be as high as 60 percent. The perceptual/cognitive paper-and pencil and psychomotor tests which were implemented at the beginning of World War II are believed to be the major factor responsibie for the reduction in attrition in flying training to present levels. However, the present 25 or 30 percent rates of attrition mean that 1 in 3 (Navy), or 1 in 4 (USAF), fail to complete training; a rate of failure which is extremely high considering the cost of instructor training, materials, fuel, and aircraft. Thus, the elimination of potential failures prior to or very early in flight training represents a great saving in material and human resources. Identifying those candidates whose skill acquisition rate and cognitive processing will not meet the demands or the time constraints involved in flying training represents additional knowledge, which may ultimately lead to a considerable reduction in the cost of flying training.

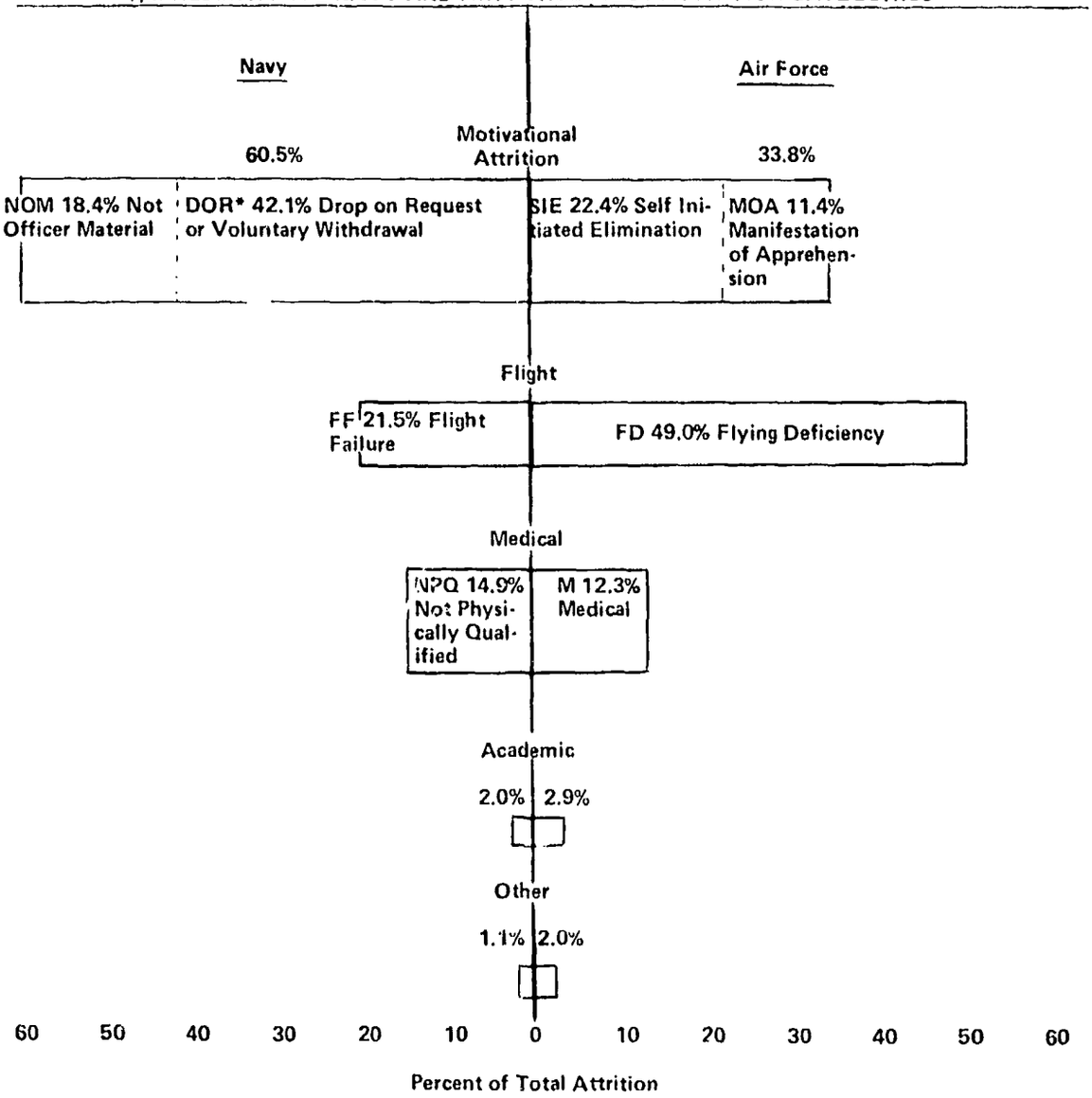
OVERVIEW OF AVIATOR SELECTION, WORLD WARS I AND II

SELECTION TESTING

Early selection tests were primarily paper-and-pencil perceptual/cognitive tasks supplemented by psychomotor devices and were able to screen pilot applicants with a fair amount of validity. By the end of World War II, and certainly by the early 1950s, the present state-of-the-art had been achieved with respect

Table I

REPRESENTATIVE NAVY AND AIR FORCE PILOT ATTRITION CATEGORIES ☆★



*Includes 3.6% Air Sickness

☆ From reference (118).

★ From reference (62).

to paper-and-pencil testing. The major advances in testing in the last twenty years have been in the area of statistical methodology rather than in test content.

Selection procedures and predictor variables should identify candidates who will be successful over long periods of time. Predictors of individual ability to adapt and cope with stressful and rapidly changing situations and accommodate rapid decision-making should be valuable in the selection of candidates for flying training. Much of the effort in the development of selection tests in the United States has been directed toward the prediction of success in undergraduate training. Few long-term prediction studies of operational performance have been made. The emphasis on the former is the result of wartime demands requiring the production of a large number of pilots in a short period of time. A continuing problem with the latter is the absence of a suitable, reliable and objective criterion. Other factors complicate the criterion problem. While the pilot profession is considered highly important, it is secondary to the role of an officer in the U. S. Armed Forces. The adage, "officer first, pilot second" emphasizes command, management, and executive responsibilities expected of military officers. As the officer gains tenure, these demands increase, so that by the time the individual is a lieutenant commander (Navy) or major (Air Force) his responsibilities may often be more management than pilot oriented. Management effectiveness as a criterion is typically difficult to measure in both the military and civilian communities.

The purpose of this review is twofold: To develop an historical background of perceptual/cognitive paper-and-pencil, psychomotor, and other selection testing methods and assess the predictive success improvement over the years. A second purpose is to suggest measures designed to improve the prediction of pilot success; based upon an analysis of the flight task, past and current research, and the opinion of successful aviators and performance assessment experts.

EARLY SELECTION TEST DEVELOPMENT: WORLD WAR I AVIATORS

The field of aviation was only fifteen years old when the first need for the aviator in combat was apparent. Hundreds of volunteers wished to fulfill their military obligations as pilots, and training centers were quickly established. As the war progressed, and data became available concerning pilot casualties, it was apparent that many accidents and failures in combat were due not to equipment or aircraft failures, but were produced by human error. After the war, several efforts were initiated to predict pilot training success. In France, measures of emotional behavior variability of pilots were tested by measuring reaction times for comparison with non-aviators. Italy became involved in selection testing efforts as early as 1919, and England stressed the measurement of physiological parameters of candidates enrolled in flight school, including the effects of high-altitude flight, pulse rate, blood pressure, and volition. The latter parameter was measured by variations in the maintenance of a column of mercury by blow-

ing into one end of a manometer fitted with a rubber tube. These examples are documented in an early review article by Dockeray (35).

In the United States, testing procedures were highly undeveloped, and aviation psychologists assigned to selection test development were not initially convinced that perceptual or sensory testing had relevance toward the prediction of flight performance. The Barany chair test, used to study disorientation and nystagmus, was found to be of little use in predicting success in flight school, although McFarland (97) indicates that some success had been observed in France. Many American psychologists seemed convinced that psychomotor aptitude tests offered more predictive validity, although attempts to implement such testing were rare before World War I.

Kelley Field Study. The first comprehensive attempt to validate tests in selecting candidates occurred in 1919 at Kelley Field, Texas. The investigation by Henmon (68) included the use of a group of predictor variables called "emotional stability" measured by hand tremors when a pistol was fired, mental alertness measured by the Thorndike Intelligence Test, and several perceptual tests such as blindfolded perception of tilt angle, and amount of "swaying" when standing blindfolded for an extended period of time. The highest predictive validity observed among these measures was the emotional stability measure and the mental alertness test ($r = .35$). No reference was made to any multiple regression technique for adding to predictive power, until the 1940s, when World War II selection tests were being developed.

Between Wars. The development of selection tests continued between the First and Second World Wars. During this period substantial effort was devoted to development of psychomotor test devices by Mashburn and colleagues (95). Several testing devices were produced including the Serial Reaction Time Apparatus (also called "complex coordination"), which was later revised and used in the Army Air Corps selection battery. This device simulated the stick and rudder movements of the airpline. Perceptual/cognitive paper-and-pencil tests continued to be developed during this period, but the predictive validity of both psychomotor and written tests was difficult to establish. A contributor to the lack of validity was the deficiency in external criteria of in-flight performance.

CIVILIAN PILOT TRAINING (CPT) SELECTION EFFORTS--(1939-1941)

The prospect of World War in 1939 encouraged the development of a civilian combat pilot force and subsequent recruitment of these aviators into the Army and Navy flight programs. The recruitment process was under the auspices of the Committee on Selection and Training of Aircraft Pilots, initiated in 1939. The evaluation of predictor variables such as biographical inventories, psychomotor ability, and other cognitive written test scores had not been overly encouraging in the course of developing the Civilian Pilot Training (CPT) program; however, the studies of predictive validity had been plagued with poor criterion measures. McFarland's review (97) of psychological factors in the

selection of pilots notes that these early criterion deficiencies were caused by inconsistent instructor ratings, failure to implement accurate recording equipment, and the low percentage of candidates that actually failed, making the pass-fail criterion virtually worthless. The pass-fail criterion was much more useful in the military studies that followed, because higher failure rates were observed.

WORLD WAR II CANDIDATE SELECTION: MILITARY STUDIES

Naval Studies. World War II produced a demand for Naval Aviators who could be trained in a very short time. The increased costs of training created the necessity of minimizing the number of candidates lost due to poor or unsatisfactory proficiency. The naval testing battery included test items previously evaluated in the Civilian Pilot Training (CPT) program and those recommended by consulting psychologists and aviators. McFarland was a participant in an extensive study of the effectiveness of this selection program for the Navy at Pensacola, Florida (98). The validity of approximately 60 psychological, physiological, and psychomotor tests were evaluated utilizing a sample population of over 900 Navy flight candidates. The criterion utilized in the study was success in flight training. This evaluation is known in the research literature as the Pensacola 1000 aviator study. Franzen and McFarland (50) indicated that certain test components had predictive validities that were sufficiently high to have successfully eliminated 44 percent of the candidates who eventually failed, and would have eliminated only 14 percent of the cadets who successfully completed flight training. More importantly, the results of this study indicated that psychological and psychomotor measures had more validity for the prediction of success in flying training than did physiological measures. Viteles (130) indicated that of the more than 21 physiological measures evaluated, none differentiated the criterion groups at a better than chance level.

The Pensacola 1000 Aviator Study led to the finalization of a Navy aviator testing program which had been previously initiated utilizing three perceptual/cognitive paper-and-pencil psychological tests as a routine part of its selection process. These included the Wonderlic Personnel Test (a test of general intelligence), Bennet Mechanical Comprehension Test (a test of mechanical interest and abilities), and the Purdue Biographical Inventory which was a measure of morale, interest, and attitudes. Viteles (130) suggests that these tests had been selected primarily on the basis of previous research conducted by the Committee on Selection and Training of Aircraft Pilots. The results of the Pensacola 1000 aviator study verified the effectiveness of these psychological instruments and in addition indicated that psychomotor tests had validity in the prediction of flight success. Still, no psychomotor tests were ever used in the Navy selection program even though Viteles (130) states that several psychomotor tests (Two-hand Coordination, Mashburn Serial Reaction, and Eye-hand Coordination) had predictive utility. It was Navy policy that test devices which could not be easily and inexpensively administered at decentralized test stations would be excluded from its selection program. Apparently, Navy testing personnel were already aware of the problems of unreliability associated with the psychomotor tests which eventually

resulted in the Air Force's decision to eliminate them from their selection program ten years later.

Fiske (43) reviewed the results of the Navy selection studies and reported that multiple regression techniques were used in the development of a test composite entitled the Flight Aptitude Rating, which became known simply as FAR. When the Biographical Inventory and Mechanical Comprehension Test scores were combined in the regression equation the multiple regression coefficient for the prediction of success in training was .41. Reports of other testing efforts indicated validities in the .50 to .60 range. Channell (24) utilized a psychomotor test (serial reaction time) to supplement the written tests and observed a multiple regression coefficient of .61.

Although the primary purpose of these efforts was to predict success in basic flying training, other Navy selection research was being conducted to determine desirable pilot qualities in combat environments. For example, Jenkins (73) asked a group of experienced combat flyers what attributes they placed above others in deciding the question "with whom would you most like to fly?' Jenkins, Ewart & Carroll (74) utilized a peer ranking (sociogram) technique to identify poor and good combat flyers. Their research suggested almost no ralationship between aptitude variables (Personnel Test, Flight Aptitude Rating, and Biographical Inventory) and the peer ranking criterion. A similar study by Bair (16) indicated that the opinions of combat and highly successful pilot trainees were quite similar in describing the attributes of good aviators. Similar to this approach was the substantial amount of research being conducted by the Navy between 1940 and 1960 attempting to deal with the problem of stress (or anxiety, which was the popular term in the Navy research literature) and its relation to aviator performance. A review of Navy research studies related to stress can be found in Griffin & Mosko (62). Although several research studies indicated a relationship between stress or anxiety and training failure in the Navy flying training program there was not satisfactory reliability, validity or confidence in the anxiety predictor variables to warrant their use for the selection of aviation candidates.

Other Navy research efforts were concerned with job sample or flying training tasks to determine their relation to aviator success. Evaluations of the link trainer by Page & Lyon (108), and Poe & Lyon (112), an approach landing trainer by Creelman (30), and an aircraft trimming device by Johnson (75) proved unsuccessful as predictors of success or failure in flying training. The authors of these studies noted the possibility of inadequate measurement technology in the development of training evaluation criteria.

Concurrently, substantial research was being conducted related to physiology oriented human performance research and sense reaction in the aircraft environment. Examples of these research efforts are the studies concerned with sound localization by Clark & Graybiel (26); Graybiel & Niven (59), vertigo and spatial disorientation in flight by Graybiel (58), Page (107), and Vinacke (130),

high altitude environmental effects by Houston (70), Miller (101) and Schaefer (119), vestibular system functioning by Guedry (63), Mann (93), and Mann & Dauterive (94), and speech intelligibility research performed by Atkinson (11), Camp (23) and Peters (110).

Army Air Force Studies. Efforts to develop selection criteria for Army aviators were the responsibility of a program headed by Colonel J. C. Flanagan from 1941-1946 (44). In 1942, the Army Air Force implemented the first edition of the Army Air Force Qualifying Examination (AAFQE). Davis (34) indicated that the test was used for initial selection of aircrew personnel, and consisted of General Vocabulary, Reading Comprehension, Math, Mechanical and Contemporary Affairs items. Prior to the development of the AAFQE, aircrew personnel were initially selected on the basis of successful completion of two year's work at a recognized college, in addition to a standard interview and medical examination. Aircrew personnel selected on the basis of the AAFQE were further evaluated and classified into pilot, navigator, bombardier, and gunner positions, on the basis of service needs and the individual's performance on the aircrew classification test battery administered at the Aviation Cadet classification centers. The Aircrew classification battery consisted of fourteen more elaborate and time-consuming tests (in comparison to the AAFQE) of general intelligence, mechanical comprehension, perception, vocabulary, reading comprehension, and a number of psychomotor, or - as they were then called - apparatus tests. Melton (100) indicates that eleven different psychomotor tests were used in the Aircrew classification battery between 1942 and 1945.

Cronbach (31) noted that Army test selection research resulted in the development of a standard score procedure called stanine, with 5 representing the mean, 7 being one standard deviation above the mean, and 3 being one standard deviation below the mean. The individual was placed in the training program reflecting his highest stanine score if his score was above the criterion for the specific training program. Each year the stanine score procedure and the weights of the battery tests were revalidated. Appparatus tests were weighted heavily in the Army Air Force classification process. Complex Coordination was highly related to fighter pilot success in training until late in 1944 when Rudder Control was given slightly more weight. Complex Coordination continued to have more weight for the classification of bomber pilots through the end of the war. Melton (100) indicates that Finger Dexterity, and Discrimination Reaction Time were used to predict bombardier success and Two-Hand Coordination and Discrimination Reaction Time were predictive of navigator success. The predictive validity of the individual tests of the battery were, of course, based on success in primary training, rather than operational performance in combat.

APPARATUS TESTS USED BY THE ARMY AIR FORCE

Melton (100) provides a thorough discussion of the apparatus tests used by the Army Air Force. The final version of the aircrew classification battery (1945) included the Complex Coordination Test (bombardiers and pilots), Discri-

mination Reaction Time (bombardiers, navigators, fighter pilots), Finger Dexterity (bombardier), Rotary Pursuit with Divided Attention (bomber pilots), Rudder Control (bomber and fighter pilots), and Two-Hand Pursuit (navigators and fighter pilots). These tests are briefly described in the following text. Typical apparatus test validity coefficients for pilots, navigators and bombardiers are presented in Table 2.

Complex Coordination. Attempts to construct a device measuring flight aptitude in the 1930s eventually led to the development of such devices as the Complex Coordination Test. The original prototype was named the Serial Reaction Apparatus and was developed by Mashburn (95). It had been used in the testing of cadets at Randolph Field, Texas, in 1931, and proved successful in predicting elimination of candidates. Its predictive validity was sufficient to warrant its inclusion in the Aircrew Classification Battery. Cronbach (31) indicated that the complex coordination test (also called Mashburn serial reaction time) was the most useful and most highly weighted test used in the World War II Army Air Force selection battery. McGrevy and Valentine (99) report that the Complex Coordination Test was used by the Air Force until 1951 when its use was discontinued for administrative rather than validity-related reasons.

The Complex Coordination Task requires the candidate to make simple controlled movements of a stick and rudder in response to patterns of visual stimuli. The stick and rudder were used to match the position of a target light and a follower when the target moved to a new position. A panel of vertical lights, horizontal lights and a curved panel of lights represented forward-backward stick, left-right stick, and left-right rudder movements, respectively. Each was a double panel with one panel for the stimulus and one for the response. The measure of performance was reaction time to match discrete changes in the positions of the three stimulus lights.

Rotary Pursuit With Divided Attention. This test was an adaptation of the Koerth (83) pursuit rotor task. Gilliland (53) found it to be appreciably correlated with flight proficiency in the Civilian Pilot Training (CPT) program. Rotary pursuit with divided attention consists of a pursuit rotor task with a side task requiring one of four lights to be extinguished by pressing a telegraph key. The side task requiring divided attention was added by Army researchers as a result of findings by the Aviation Psychology Program. Melton (100) indicates that the basis of its inclusion was the belief that a measure of divided attention would be a valid predictor of pilot success. Time on target was the performance measure (contact by applicant's stylus on a metal disk on the circumference of the rotor) associated with the divided attention task.

Discrimination Reaction Time. This test was included to assess all-or-none type manual responses to visual signals. The test required the applicant to push one of four toggle switches in response to certain lighting configurations of a red and green signal lamp The time taken to operate correct switch sequences was the performance measure.

Table II

VALIDITY OF APPARATUS TESTS ☆

Complex Coordination	Pilots 1943, 3,154 subjects r = .33, r^c = .40	Navigators 1942, 1,022 subjects r = .17, r^c = .24	Bombardiers 1943, 1,829 subjects r = .10, r^c = .13
Rotary Pursuit with Divided Attention	1943, 3,146 subjects r = .14, r^c = .22		
Finger Dexterity	1943, 4,779 subjects r = .07, r^c = .10	1943, 1,021 subjects r = .10, r^c = .13	1943, 1,828 subjects r = .13, r^c = .15
Discrimination Reaction Time	1943, 4,779 subjects r = .25, r^c = .28	1942, 1,022 subjects r = .27, r^c = .35	1942, 1,829 subjects r = .22, r^c = .25
Rudder Control Test	1943, 3,146 subjects r = .22, r^c = .30		
Two Hand Pursuit	1943, 1,385 subjects r = .27 (average)		1943, 421 subjects r = .20
Two Hand Coordination	1943, 4,779 subjects r = .31, r^c = .35	1942, 1,022 subjects r = .26, r^c = .29	1943, 1,828 subjects r = .09, r^c = .12

r = validity coefficient based on dichotomous pass/fail training criterion.
r^c = validity coefficient corrected for restriction in range.

☆ From reference (100).

Finger Dexterity. The Santa Ana Finger Dexterity Test (named after the Santa Ana Army Air Base where it was devised) was designed to involve precision and speed of movement in withdrawing, inverting, and replacing pegs in a form board. The finger dexterity test was used primarily in the classification of bombardiers in the Aircrew classification battery.

Rudder Control Test. The Rudder Control Test was designed to measure fine control sensitivity and psychomotor coordination. According to Melton (100) its primary use had been as a training tool for early flight training to teach rudder and braking movements. The test required the applicant to control the movements of a chair, such that a light bar mounted in front of the chair was always pointed toward a target. The requirement of the movements of the rudder were directed toward appropriate adjustments to keep the chair in an upright position at all times. The measure of proficiency was the accuracy maintained in keeping the chair light bar pointed within 2.5 degrees of the target.

Two-Hand Pursuit Test. The two-hand pursuit test was used to assess the candidate's spatial relations ability and coordination of both hands to control the movement of a target-follower in response to a visual target moving on an irregular path. The Two-Hand Pursuit Test replaced the Two-Hand Coordination Test with which it had been highly correlated (r = .60).

OVERALL SUCCESS OF ARMY AIR FORCE TESTS

Guilford and Lacey (64) pointed out that the primary objective of the selection program had been one of quick selection and classification of potential aviation candidates to meet training requirements and become available for duty in a short period of time. A large portion of candidates who would have failed in training or would have required extra training were undoubtedly identified before acceptance into training programs, resulting in the savings of considerable material and instructor time.

The success of the Army testing program is summarized by Davis (34). "For every 100 graduates from advanced pilot training . . . desired . . . in the summer of 1943, it was necessary to start 397 men in pilot preflight school . . . When the men were selected by both the AAFQE and Aircrew Classification battery (using a stanine score of 7 . . .) only 155 men were required to obtain 100 graduates."

The problem of longer range validity of these measures was still unsolved after the war. The task required the development of objective and reliable ratings in performance over a long period of time on such dimensions as promotions, type of duty assignment, and success in combat missions. With this goal in mind, Army Air Force and Navy psychologists joined efforts in a Pilot Candidate Selection Research program in 1947, to improve selection measures for pilot candidates. A total of 35 paper-and-pencil tests and 20 psychomotor apparatus tests were given to entering aviation students.

McFarland (98) indicated that these efforts were useful with regard to predicting preliminary training success but were not productive in the prediction of advanced operational training and combat proficiency. Flannigan (44) indicated that attempts to correlate the tests with other criteria such as promotions, awards, and combat duty effectiveness, were unsuccessful. Sheeley (126) suggests that perhaps the major accomplishment of the joint Air Force/Navy effort was the incorporation of an Air Force paper-and-pencil spatial test into the Navy Selection Test Battery.

POST-WAR AVIATOR SELECTION RESEARCH

PSYCHOMOTOR TESTING

Psychomotor Tests remained a part of the USAF selection battery until 1951 when their use was discontinued. Guilford (64) states that there was no doubt that the psychomotor tests contributed to the effective selection of pilots, even though there was considerable overlap between the psychomotor test (complex coordination) and paper-and-pencil spatial and mechanical ability tests. Cronbach (31) notes that the unique contribution to validity of the psychomotor test (the multi-limb coordination factor) could not be provided by any paper-and-pencil tests. In spite of this fact, there were substantial administrative, reliability and quality control problems associated with the psychomotor tests. As a consequence, the Air Force gave up use of the psychomotor tests in selection in spite of their unique contribution to selection validity. There were continued attempts to validate psychomotor tests for predicting pilot candidate success in the post war years, as evidenced by the research of Dailey & Gragg (32), Leiman & Friedman, (87), Fleischman, (45), and Creager, (28). Most validities obtained were similar to those found previously, and fell in the .40 to .50 region.

The first attempt toward revision of the psychomotor tests to eliminate the administrative and reliability problems was undertaken by Adams (2) at Lackland Air Force Base. The approach was to fabricate a series of very simple motor tasks which required virtually no hardware for testing the same abilities as the more complex apparatus tests previously used. Such tasks as placing marbles in holes, drawing dots in circles, and gross muscular tests such as chin-ups or push-ups were used. The results of this study indicated that the use of simple motor skill tests offer little predictive value for flight school success. The Navy reported similar results in research of the predictive value of gross muscular tasks by Creelman (29), and Schwarts and Lowe (117). However, study efforts by Hutchins and Pomarolli (72), and Willingham (143) indicated some relationship between more finely coordinated muscular skills (gymnastic and swimming performance, for example) and success in pilot training.

Cronbach (31) reviewed attempts to develop paper-and-pencil measures of motor performance and concluded that apparatus tests and paper/pencil tests of motor ability represent different factors. Fleishman and Ellison (46) indicate

that printed tests should not be used to measure more complex dexterity and coordination skills.

Adams (2) suggested that the failure of simple motor tasks to demonstrate predictive validity was due to: (1) the unreliability of both predictor and criterion scores; (2) the use of inappropriate motor tests for the particular criteria being evaluated; and (3) the task of flying an airplane is too complex for simple motor skill tests to be of substantial benefit. This latter point suggests that the more complex psychomotor tests which were used through the war years were more appropriate for assessment of the motor skills required of the pilot. The fact that both psychomotor and perceptual/cognitive paper-and-pencil tests could not account for a substantial number of failures implies that other complex abilities are necessary for successful pilot performance. The goal of selection research must be to define these abilities, quantify the behaviors making up the abilities, and devise tests to measure them accurately.

Passey and McClaurin (109) provide a comprehensive review of psychomotor selection testing over the years, and review studies measuring complex behaviors for the purpose of developing selection tests. Their review covers factor analytic techniques which were used to analyze aviator tasks for the purpose of developing proficiency tests, the use of the light plane as a selection device, and the development of a rationale behind new types of ability tests.

The Air Force continues to utilize the light plane as a preliminary selection tool. Majesty (92) reported a validity coefficient of .07 between the light plane screening program (T-41 final grade) and the criterion (pass/fail) in Air Force Undergraduate Pilot Training (UPT). The Navy utilized the light plane selection concept in the late 1950s and into the 1960s in its ROTC program, but discontinued the effort because of excessive cost. However, the Navy still uses the light plane in a familiarization role at the recruiting level. Prospective aviation officer candidates without previous flight experience are taken up in a light plane and are allowed a minimal amount of experience handling the controls. In the Navy, then, the light plane serves essentially as a self-selection device in that the flight experience may allow the potential aviator candidate to determine if he is really interested in actively pursuing a flying career.

PERCEPTUAL/COGNITIVE PAPER-AND-PENCIL TEST SELECTION RESEARCH

Efforts have continued to refine and develop new perceptual/cognitive paper-and-pencil predictors of aviator performance. Although thirty years have passed since the conclusion of World War II, perceptual/cognitive paper-and-pencil test predictors of aviator performance have changed very little despite advances in test technology. Thus, the U. S. military service paper-and-pencil selection test batteries consist of: (1) a general intelligence component composed of verbal and quantitative items; (2) mechanical comprehension (usually an adaptation of the Bennett Mechanical Aptitude series); (3) a spatial component (usually adapted from the Air Force's spatial aptitude series); and (4) a back-

ground or biographical inventory composed of a miscellaneous subset of items usually of an historical nature known to relate to aviator success.

A description of the selection test batteries of the three military services, together with validity coefficients, are provided in Tables 3, 4 and 5. These test battery descriptions and validity coefficients have appeared previously in separate service research by Miller (103), Doll (36) and Kaplan (79). The description of Navy tests are from the Navy Examiners Manual and Scoring Instructions (41).

Research conducted on the use of perceptual/cognitive paper-and-pencil tests has led to the general consensus that the state-of-the-art has been obtained in the use of such tests for predicting success in undergraduate flying training. An extensive review of this research may be found in psychological testing texts such as those prepared by Guilford and Lacey (64), or Cronbach (31). As a result, current research utilizing perceptual/cognitive paper-and-pencil tests often has diverse objectives. For example, Ambler and Smith (9) recently evaluated perceptual/cognitive paper-and-pencil tests to determine their potential for selecting students for assignment to flying training pipelines. Egan (38) recently studied a perceptual/cognitive paper-and-pencil test to determine if question response times (latency) are related to training performance in a Navy undergraduate flying training environment.

Perceptual/cognitive paper-and-pencil tests have recently been the subject of considerable scrutiny to determine if they might be biased against certain cultural or ethnic population sub-groups. Typically, the results of such research have indicated that certain population subgroups have both lower test and criterion scores. Recent Air Force work by Mathews (96) suggest that perceptual/cognitive selection tests tend to overestimate the later performance of non-white groups in flying training. Similar results were found by Guinn, Tupes and Alley (65) for non-white groups in non-flying training. Similar research is presently being conducted by Navy representatives to determine the fairness of Navy aviation selection tests to minority population sub-groups, and potential women naval aviators.

PERSONALITY TEST SELECTION RESEARCH

A great deal of research effort since the 1930s has been devoted to the investigation of paper-and-pencil and projective personality inventories to determine their usefulness in predicting motivational categories of attrition in aviator training programs. These motivational categories of attrition are Drop on Request (DOR) (also called voluntary withdrawal) and Not Officer Material (NOM) in the Navy, Self Initiated Elimination (SIE) and Manifestation of Apprehension (MOA) in the Air Force (see Table 1).

Richardson and Rusis (114) indicated that personality factors associated with success in various occupations have been the subject of literally thousands

Table III

U. S. NAVY AVIATION SELECTION TEST BATTERY
U. S. NAVY AVIATION SELECTION TEST BATTERY

Composition and Validity Coefficients

Subtest☆	Validity, Pass/Fail Criterion Uncorrected★		Corrected ♦
Academic Qualification	.12		.40
Mechanical Comprehension	.19		
Spatial Apperception	.11	FAR	
Biographical Inventory	.19		
FAR	.23		.63

Subtest Description

Academic Qualification Test (AQT)
This is a test of general intelligence. Research has shown that this test is particularly adapted to the prediction of ground school performance. Individuals who score low tend to have difficulty in the academic portions of training.

Mechanical Comprehension Test (MCT)
This is a test dealing with ability to perceive physical relationships and handle familiar concepts of everyday mechanics rather than with technical subject matter found in textbooks.

Spatial Apperception Test (SAT)
This is a test of ability to orient in space or, specifically, to visualize the relationship between the attitude of a plane and the territory over which it flies.

Biographical Inventory (BI)
This is a questionnaire containing elements of personal history, expressions of interest and attitudes, and selected information items. No single item is heavily scored or significant in itself, but certain total patterns have been found to differentiate between successful and unsuccessful flight students.

Flight Aptitude Rating (FAR)
Scores made on the MCT, SAT and BI are combined into a single index called the Flight Aptitude Rating or FAR. The FAR, expressed in terms of a numerical grade, indicates the applicant's measured probability of success or failure in the flight training program.

☆ From reference (41).

★ Based on 1973 Pilot Input 2,109 subjects, NAMRL Computer Analysis.

♦ From reference (36) Validity coefficients are corrected for restriction in range of subjects and attenuation in the criterion.

Table IV

AIR FORCE OFFICER QUALIFYING TEST

Composite Composition ☆

Subtest	Aptitude Composite		
	Pilot	Nav-Tech.	Off.Qual.
Quantitative Aptitude		x	x
Verbal Aptitude			x
Officer Biographical Inventory			x
Scale Reading		x	
Aerial Landmarks		x	
General Science		x	
Mechanical Information	x	x	
Mechanical Principles	x	x	
Pilot Biographical Inventory	x		
Aviation Information	x		
Visualization of Maneuvers	x		
Instrument Comprehension	x		
Stick and Rudder Orientation	x		

Validity Coefficients ☆

Composite	Criterion (pass/fail in training ★)		
	Uncorrected		Corrected
	Pilots	Navigators	Pilots
Pilot	.26	.07	.40 ♦
Nav-Tech	.18	.02	
Officer Quality	.12	.04	

☆ From reference (103).

★ Coefficients based on 1500 students in Undergraduate Pilot Training and 2132 students in Undergraduate Navigator Training.

♦ Pilot correlation corrected for restriction of subject range.

Table IV (Con't)

Air Force Officer Qualifying Test (AFOQT)

Subtests☆

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

Verbal Aptitude consists of items pertaining to vocabulary, verbal analogies, reading comprehension, and understanding of the background for world events.

Officer Biographical Inventory consists of items pertaining to past experiences, preferences, and personality characteristics known to be related to success in officer training.

Scale Reading consists of items in which readings are taken of various printed dials and gauges. Many of the items require fine discriminations on nonlinear scales.

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

General Science consists of items related to the basic principles of physical science. The emphasis is on physics, but other sciences are also represented.

Mechanical Information consists of items pertaining to the construction, use, and maintenance of machinery. Some of the items are concerned with the use of tools.

Mechanical Principles consists of diagrams of complex apparatus. Understanding of how the apparatus operates or the consequences of operating it in a specified manner is required.

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

Aviation Information consists of semi-technical items related to various types of aircraft, components of aircraft, and operations involving aircraft.

Visualization of Maneuvers consists of items requiring identification of the silhouette which expresses the attitude of an aircraft in flight after executing a verbally described maneuver.

Instrument Comprenhension consists of items similar to those in Visualization of Maneuvers except that the maneuvers are indicated by readings of a compass and artificial horizon.

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

The subtests are organized into several composite scores used for different selection purposes. For example, the Officer Quality Composite consisting of Biographical Inventory, Verbal and Quanitiative sub-tests is typically used for the selection of nonflying officers. The Pilot and Officer Quality Composites are used in the selection of pilots. The Navigation/Technical and Officer Quality Composites are used in the selection of navigators.

☆ From reference (103).

Table V
Army Flight Aptitude Selection Tests (FAST)

FAST Composition☆ Subtests	and Validity Coefficients[1] Officer		Warrent Officer	
	Rotary	Fixed	Rotary	Fixed
Biographical Information	.18	.18		
Mechanical Principles		.24		
Flight Orientation		.32		
Aviation Information - Fixed Wing		.27		.270
Aviation Information - Rotary Wing	.243		.243	
Mechanical Information		.24		.240
Mechanical Functions	.299		.299	
Visualization of Maneuvers	.277	.28	.277	.285
Instrument Comprehension		.21		.210
Complex Movements	.342		.342	
Stick and Rudder Orientation	.279		.279	
Self Description			.361	.361
Composite Validity Coefficients	.424	.390	.478	.457

Subtest Description[1]

Biographical Information - Items of this inventory relate to the individuals family, education, hobbies, etc. and contains personality orientated self description items and self estimates of ability.

Mechanical Principles - This test requires the examinee to solve problems on the basis of principles of mechanics.

Flight Orientation - This test is a measure of ability to visualize the relationship between an airplane and the territory over which it flies.

Aviation Information - Fixed Wing - The items of this test related to general and technical aspects of fixed-wing aviation, e.g., flying terminology, specific maneuvers, use of controls, etc.

Aviation Information - Rotary Wing - This test relates to the flying, uses, terminology, and theory of the helicopter.

Mechanical Information - This test is a measure of knowledge about general mechanics and tool functions.

Mechanical Functions - This test is a measure of ability to understand general mechanical principles. Pictures are shown and questions are asked on the mechanical principles illustrated. The pictures are of practical real life situations.

Visualization of Maneuvers - This test is a measure of ability to visualize airplane maneuvers.

Instrument Comprehension - In this test, each item consists of pictures of two instruments, an artificial horizon and a compass, followed by pictures of 5 planes. The problem is to determine which of the 5 planes has a position and direction consistent with the instrument readings.

Complex Movements - This test, previously named Coordinate Movements Test, requires the examinee to judge distances and visualize movements quickly and relate these distances and movements to a set of symbols.

Stick and Rudder Orientation - This test presents the examinee with three photographs taken from the cockpit of a plane doing simple maneuvers (banking, turning, climbing, and diving) or combinations of maneuvers (turning while climbing, for example). The examinee is required to relate the maneuvers shown to stick and to rudder positions on the answer sheet.

Self Description - This is a personality oriented test in which the individual selects phrases which are least and most descriptive of himself.

☆ From reference (79).

of studies by many competent investigators using various techniques and instruments. Griffin and Mosko (62) in a recent review of Navy selection research indicate that approximately 40 different personality paper-and-pencil test devices have been evaluated from 1950 to 1976 for pilot selection without any appreciable impact on the selection of aviator candidates.

The problem with the utilization of the personality paper-and-pencil test devices and projective tests is their reliance on the individual to provide an honest and objective evaluation of himself even though such an evaluation has the potential to prohibit the individual's entry, or continuation, in aviator training. Obviously, such behavior is rarely exhibited, resulting in relatively low or non-significant correlations with the criterion - aviator success in training.

Hathaway and McKinley (67) suggest that the personality inventory may have some validity for separating normal and abnormal individuals in society. However, Cronbach (31) concludes that personality inventories are apparently poor predictors of occupational performance. Freeburg (51) reaches a similar conclusion concerning academic performance. Still, on those occasions when personality tests are administered under a no-threat, no-consequence condition (i.e., when subjects are told, "Your performance on these tests will in no way affect your continuation in flying training"), or after attrition has occurred, small relationships with motivational criteria in military settings occasionally occur.

However, when the tests are applied "for real," the relationship typically disappears, or becomes so small and variable that its usefulness is severely limited. This occurs as a direct result of subjects' ability to select the test item response which is more socially acceptable or more congruent with success in aviation training. Cronbach (31) indicates that this phenomena is commonly known as "faking the test," or test response bias. Bucky (21), Bucky, Spielberger & Bale (22), Jones (76), Voas (131-133), Wallon & Webb (134), (135), and Waters (136) have noted the susceptibility of personality inventories to faking and response bias in Navy studies. When one considers the quality of the aviator trainee population--practically all have college degrees, are above average in intelligence, and have taken literally hundreds of tests during their academic careers--it is not surprising that highly motivated potential aviators can readily determine appropriate and inappropriate responses for selection to aviator training.

In spite of the discouraging results reported by both the Navy and the Air Force in the use of personality devices for selection, the Army has apparently had sufficient success with the use of personality measures to include their use in the Army Fixed and Rotary Wing selection battery. The Army reports validity coefficients of .18 for officers, and .36 for enlisted personnel. Kaplan (79) reports that the tests are most useful in predicting training failure in preflight rather than in actual flying training.

NEUROLOGICAL RESEARCH

EEG (Electroencephalogram) recordings have been studied repeatedly to determine their relationship to aviator performance and as a possible predictor of aircraft accidents. Lennox-Buchthal, Buththal and Rosenfalck (89) indicated that individuals with abnormal EEG recordings have an accident involvement more than 3 times higher than controls. These findings have resulted in the Danish Air Force's use of EEG recordings in the selection of pilot candidates.

Sem-Jacobsen and Sem-Jacobsen (121) have investigated the relationship of EEG recordings to inflight stress or G forces on Norwegian and USAF pilots in aircraft flight. These findings suggest agreement between the clinical appearance of the pilots experiencing G forces and EEG abnormal recordings; and indicated a relationship between inflight G force stress and pilot error accidents.

Ades (4) investigated the relationship between EEG recordings and altered consciousness during flight. His findings suggested a positive relationship between the two. Ades speculated that a substantial number of accidents per year may be attributable to altered consciousness. As a direct result of Ades research, the Navy implemented a program of EEG recordings for student naval aviators in 1961, which continues to the present day. Evidence of an abnormal EEG may be sufficient to prohibit prospective student naval aviators and student naval flight officers from continuing in naval aviation training.

Despite the studies which suggest that EEG recordings may have potential in the selection of aviators to reduce the pilot accident potential, and its use in Navy aviator secondary selection, there is skepticism of the value of EEG recordings for pilot selection by the scientific community. The skepticism apparently is a result of a variety of studies which have shown no relationship, or an extremely low relationship between EEG recordings and pilot performance, as reported by Forbes, Davis & Davis (49), Franzen & McFarland (50), Gastant, Lee & Labourer (52), Kennard (82), McFarland & Franzen (99), Mundy-Castle (105), and Picard, Labourer, & Navarronne (111).

NEW DIRECTIONS IN POST-WAR AVIATOR SELECTION RESEARCH

TASK AND FACTOR ANALYSIS STUDIES

A necessary prerequisite for developing valid aviator selection tests is a more complete understanding of the flying tasks. Because the aviator acts as a complex integrator involved in sorting out appropriate behaviors to fit the particular demands of the moment, one cannot hope to predict performance from only one or two tests. The proficiency of the performance measurement process must be taken into account also, and objective reliable testing of in-flight performance must be obtained to provide selection tests the opportunity to achieve high validity. The first step in this approach is a task analysis of the aviator's job. (A detailed task analysis can be exceedingly complex. For example, Shannon, Waag

& Long (125) identified twenty-four discrete and sequential task activities associated with a basic - but complex - spin maneuver in Naval Aviation Flying training in an effort to isolate recurring student pilot errors in primary flight training.)

Initial task analytic efforts were directed at analyzing criterion measures. Gordon (57) used a questionnaire approach to specify the abilities of the successful airline pilot, Miller (102) attempted to identify reasons for failure in both training and combat, and Ericksen (40) analyzed the comments of flight instructors in student logbooks during flight training. The latter study revealed that instructors most often commented on factors relating to motivation, attitude, aggressiveness, planning, judgment, and division of attention.

Fleischman and Ornstein (48) presented factor analytic data on 24 flight maneuvers as scored by trained observers. The six factors determined were labeled as:

a. Control Precision: fine control sensitivity.

b. Spatial Orientation: judgment of position in three dimensional space.

c. Multilimb Coordination: performance of simultaneous tasks with hand or feet.

d. Response Orientation: rapid response to changing stimulus conditions.

e. Rate Control: responses in anticipation of velocity or rate changes.

f. Kinesthetic Discriminations: reactions to slow movements of the aircraft, as in stalls.

A previous factor analysis by Fleischman and Hempel (47) on psychomotor and written tests had revealed factors related to the first five mentioned above, indicating that these tests had indeed been measuring the responses that they had intended to measure. A number of factor analytic efforts have been applied to selection and training performance variables in naval aviation training. Bair, Lockman & Martoccia (17) identified four factors accounting for 51 percent of the total predicted variance in their study of selection predictors and training performance criteria in Naval Aviation "Basic Stage" Training. The four factors were labeled: (1) perceptual analysis involving visualization of symbols; (2) academic potential; (3) comprehension of relationships involving the understanding of written and oral instructions; and (4) applied spatial relations or the relationship of objects in three dimensions. Waters and Wherry (137) applied factor analysis techniques to selection test and performance measures in preflight and conducted factor analysis studies of primary and basic stages of training performance for both jet (138) and multi-engine (139) student aviators in naval avia-

tion training. More recently, Booth & Berkshire (18) analyzed the factor structure of naval aviation training measures and the performance of Marine fixed and rotary wing pilots in operational squadrons. Academic ability, flying skill and systems comprehension factors were identified for both jet and helicopter pilots. Bale, Smith and Ambler (15) conducted the most recent and extensive Navy factor analytic study involving the study of Navy training performance measures leading to Naval Aviator "designation," and included post-graduate performance in the Replacement Air Group (RAG). (In Navy aviation, RAG or RTS[1] training is used to transition newly designated naval aviators to high performance aircraft used in operational squadrons.) This comprehensive study identified nine factors accounting for 45 percent of the total variance in Navy undergraduate and post-graduate training. The factors identified are labeled and described below.

I. Basic Flight Capacity, a factor associated with flight skills and aeronautical adaptability in combination with inflight mechanical operation skills.

II. Operational Flying Indoctrination, associated with precision flying and combat tactical skills in the military use of aircraft.

III. Academic Capacity, associated with the motivation to acquire knowledge together with verbal and mathematical cognitive skills.

IV. Advanced Military Flying Capacity, similar to Factor II. This factor appeared to be oriented toward high level tactical flying ability and motivational aspects of skill application in a military setting.

V. Instrument Flying Indoctrination, associated with aircraft instrument flying skills.

VI. Instrument Flight Skill, similar to factor V. This factor is associated with the intellectual ability to understand the theory of instrument flight and its operational application to new situations.

VII. RAG Operational Flying Skill, representing operational combat flying required in the fleet.

VIII. Day Carrier Landing Skill, includes skills associated with the capacity and ability to maneuver a high performance aircraft onto a moving landing platform.

IX. Night Carrier Landing Skill, related to Factor VIII, but performed in a darkened environment.

---

[1]Replacement Air Group training is now designated as Readiness Training Squadron (RTS) training.

Perhaps the major finding of this research was the distinction between aviator skills required in DAY and NIGHT Carrier Landings. Also, the fact that carrier landing skills loaded significantly as separate factors suggested that this aspect of naval flying is separate and distinct from other service operational flying skills. The authors conclude that there does not appear to be a single "Flight Training" factor and that independent skills appear to be taught in each phase (basic, advanced, RAG) of naval aviation training. As a result, in Navy flying training the student aviator must apparently be required to acquire new skills in new phases of training.

Relatively little is known concerning the analysis of Socialistic countries regarding individual abilities required for successful flight performance. This is partially a result of their ideological and political philosophy which emphasizes the basic equality of individuals, while deemphasizing the importance of special individual abilities and skills.

Lin (90), of the Peoples Republic of China, identifies the following as important psychological attributes closely associated with flying.

- quick and accurate perceptual skills.
- good distribution and shifting attention abilities.
- coordination of hand and foot movements.
- good simulation ability.
- good memory.
- reaction sensitivity (vestibular system oriented).
- stable emotions, and strong will.

Passey and McClaurin (109) summarize ability domains believed to be germane to successful aircrew performance. These abilities are (1) adaptability to changing surroundings, (2) capacity for integrating and processing information, (3) storage reorganization, (4) comparison and combination of data inputs, and (5) endurance under demanding situations. The task confronting the developer of selection tests is to isolate certain specific behaviors and abilities which comprise these rather broad domains.

ISOLATING BEHAVIORS FOR PSYCHOLOGICAL ASSESSMENT

Several behaviors have been measured using simple testing techniques, but they have not typically been used in any attempt to predict aircrew performance. The functions will be mentioned briefly here and will provide the basis for an expanded look at several potential selection test measures.

Behavioral Functions[1]

a. Motor skill acquisition rate: the ability to learn certain skills quickly, requiring modification of rate of improvement to fit training constraints.

b. Automatization of response: the integration of sense modalities, making automatic reactions to specified stimuli.

c. Resistance to prolonged operation: ability to perform long-term missions with a minimum in performance decrement.

d. Performance consistency: low variability in performance, reliability of responses over long time period.

e. Resistance to distraction: ignoring irrelevant stimuli; extracting relevant stimuli.

f. Adaptive capacity: stress capacity; alerting responses under changing conditions, resisting emotional distractions.

g. Kinesthetic discrimination: the ability to use kinesthetic and proprioceptive cues.

h. Concurrent information processing: measuring reserve capacity, organizing and performing simultaneous tasks.

i. Anticipatory behavior: predicting future stimuli, predicting eventual states from existing information; prediction of a rapidly deteriorating condition.

j. Behavioral rigidity: failure to respond to changes in the environment.

k. Short-term memory: kinesthetic feedback and specific movement retention in storage for the purpose of recall for the same situations.

l. Perceptual speed: recognizing or comparing rapidly.

m. Attention: including attention span, duration of attention, fixation on a particular input source.

n. Estimation: of time; velocity, extent, direction of certain events.

o. Discrimination reaction time: differential response speed, use of visual or auditory input to signal a response.

[1]From reference (109).

p. Visualization: or the ability to manipulate objects in spatially related matters.

The value of such a listing of behavioral functions is that it can serve as a directive for developing test selection measures. Although the measurement of all of the above behaviors would certainly prove beneficial to the test developer, it is impractical to consider a battery of tests which measures every item individually. This practical constraint does not limit us to selection of only two or three of these functions to measure, but does limit the overall size of the battery. An alternative to limiting the number of abilities assessed is to incorporate several required abilities into one test, providing parallel assessment. This also limits the amount of hardware implementation needed. In order to provide an understanding of the type of tasks which have been used to assess some of the above behavioral functions, several studies will be reviewed concerning the topics of concurrent information processing, decision-making, attention, and anticipatory behaviors.

CONCURRENT INFORMATION PROCESSING

Research efforts concerning concurrent information processing have not been numerous, but several studies involving dual-task performance have been conducted. Brown (20) suggested the use of dual tasks which overload the individual enough to study performance deficiency under a variety of conditions. This technique had been successful in studies by Griew (61) and Kalsbeek (78). Griew used a mixed mode approach in a task involving an auditory input and a continuous pursuit tracking task. He observed that performance on the tasks performed singly was superior to the simultaneous performance. Kalsbeek studied the deterioration in performance caused by distraction stress and used a primary-secondary task approach. Both of his tasks required choices among alternative actions. The results of simultaneous choice making led Kalsbeek to conclude that when the subject is confronted with concurrent choice making, the choices will be made successively rather than simultaneously. This suggests a "single channel" monitoring hypothesis when choices about movements must be made.

DECISION MAKING CAPABILITY

Adiseshiah (5) used a rapid decision making task to study decisions of pilot candidates under stress. Stress was manipulated by varying the time available to make the decisions. The task was comparing two stimulus cards with aircraft symbols and reporting the number of symbols in common to the two cards. The time to make this decision was varied from 1 to 20 decisions per minute. Three levels of pilot experience were used including student pilots, instructors, and experienced airline pilots. The results indicated that experience was related to ability to handle increased speed demands in making decisions. Student pilots had the sharpest decline in performance, followed by instructor and airline pilots.

RESIDUAL ATTENTION

Several attempts have been made to study the extra attention capacity which the pilot has in addition to the workload of performing routine flight tasks. Most of the studies of residual attention have used a dual task approach in which one task is defined as the primary task and one as the secondary. Ekstrom (39) was fairly successful in quantifying the reserve capacity for pilots flying X-15 missions. Slocum, Williges, and Roscoe (127) used the residual attention approach to study reserve capacity when the primary task is altered. The primary task involved controlling common aircraft functions with a series of rotary switch knobs. In one condition, these knobs were coded in a meaningful fashion, while two other conditions provided no coding and arbitrary coding respectively. Performance on the secondary loading task revealed that the meaningfully coded condition produced the highest scores on the secondary task.

A problem with using the primary-secondary task approach has been the lack of control over difficulty levels of the two tasks. This has been facilitated by the use of adaptive techniques devised by Kelley and Prosin, (81), and Kelley and Kelley, (80). This approach allows the subject to perform the secondary task at his own difficulty level as long as his performance on the primary task(s) is within some error limit. This technique was recently used by both Damos (33), and North and Gopher (106) in studies concerned with the prediction of pilot performance in an introductory flight course.

Measuring residual attention of prospective candidates may have utility because it is often the reserve capacity that is used to deal with deteriorating situations in flight. This ability may be extremely important in handling emergency procedures smoothly.

TESTS OF ANTICIPATORY BEHAVIOR

Although this ability has been studied by many investigators in psychology interested in cognitive processes, such tests have not been used to predict pilot success. Adams and Chambers (3) asked subjects in these experiments to anticipate a sequence of lights or other events by selecting the next event. Usually, a stochastic rule governs the event sequence, and after a learning period, the subject is asked to predict future events. These tests should be evaluated for predictive validity, as the ability to anticipate future events is important in pilot performance.

CURRENT EMPHASIS IN AVIATOR SELECTION RESEARCH

CRITERION MEASUREMENT

Development of objective and reliable criterion measures is extremely important in obtaining high validity of prediction. Efforts to develop objective rating schemes of pilot performance began in the Civilian Pilot Training (CPT)

program in 1939, and continued through World War II, as evidenced by the work of Jenkins (74). However, these and other similar attempts failed to produce an objective set of measures for in-flight performance. Other efforts have been made by Edgerton and Walker (37) and Miller (102) as part of the postwar Civil Aeronautics Administration program and the Army Air Force program, respectively. Each of these studies produced rating procedures which were costly and time consuming to administer, and subsequently their use proved limited. An alternative approach is the development of the automated recording of pilot performance. This strategy is discussed by Connelly, Schuler, and Knoop (27) in a USAF study. The objective of this research was development of a pilot assessment measure for training.

Recently, several researchers have developed a flight performance rating scale from the Federal Aviation Administration's "Private Pilot's Test Guide". The "Illinois Private Pilot Flight Performance Scale" was developed for grading student performance on the required maneuvers for pilot certification. Povenmire, Alvares, and Damos (113) report the initial implementation of this rating scale in terms of observer-observer reliability. Reliability indexes were quite high, indicating that a relatively simple rating procedure could yield consistent results across performance raters. Later checks on reliability of this performance scale were conducted by Selzer, Hulin, Alvares, Swartzendruber, and Roscoe (120) and the same result of high observer-observer reliability was found. For a more detailed discussion of the problem of criterion measurement, the reader should consult a recent review of the literature in the development and use of synthetic flight training devices (Williges, Roscoe, and Williges (142).

Navy research personnel have indicated a strong interest in the identification and development of advanced criterion-oriented operational performance measures to serve as more "valid" criteria for the selection of student naval aviators. Rickus and Berkshire (115) investigated the use of flight surgeon ratings of aviators as a combat criterion. Unsatisfactory aviators were identified from performance descriptions such as "turned in wings," "had wings taken away, transferred due to poor performance," or were identified as ". . . men others refuse to fly with." Men thus identified had poorer preflight, basic flight, and advanced flight grades. The authors suggested that peer ratings obtained during the eighth week of pre-flight training had the potential to predict unsatisfactory aviator performance in the fleet. Bale, Rickus and Ambler (13) utilized Replacement Air Group (RAG) performance measures as advanced criteria. A number of undergraduate training grades were predictive of RAG performance, as were two initial selection variables, the Mechanical Comprehension Test (MCT), and Biographical Inventory (BI), of the Navy Flight Aptitude Rating. The undergraduate performance variables most predictive of RAG performance were tactical weapons grades and instrument grades in advanced training. The MCT and BI carried significant but low weights in the multiple prediction formula. Interestingly, the MCT prediction weight was negative. The Spatial Apperception Test (SAT) was negatively related to the advanced per-

formance criteria, as were certain undergraduate training grades, (Presolo, Engineering, Transition, Basic and Advanced Ground Grades.)

Brictson, Burger and Gallagher (19) utilized Initial Carrier Landing Performance as advanced criteria for the prediction of F-4 pilot performance in an operational environment. Selection tests, Basic and Advanced Flight grades, and Replacement Air Group Grades resulted in a multiple correlation of .72 using a composite night landing score as the criterion. This relationship accounted for 50 percent of the variance associated with the criteria. Selection tests accounted for six percent of the total variance. The Aviation Qualifying Test (AQT) was negatively related to the criterion, as were a number of undergraduate training parameters (Presolo, Precision, Instruments, Night Familiarity, Radio Instruments, Carrier Qualification, Flight Grade, and Conventional Weapons Delivery Grades). Shannon, Waag & Ferguson (124), and Shannon & Waag (122) conducted analyses of critical skills in a Replacement Air Group (RAG) training environment in an effort to develop advanced training criteria. These studies demonstrated that a small subset of critical performance skills are predictive of overall RAG performance. A follow-on study by Shannon & Waag (123) utilized the previously developed RAG criteria (RAG final grade), ratings completed by squadron commanders (fleet evaluations), and critical incidents, as the basis for the prediction of F-4 Pilot Performance. The RAG criteria included pilot selection test scores and undergraduate flight grades. Regression analysis results indicated that eight variables predicted final RAG grade, yielding a multiple correlation of .51. Of all variables, experience level (time in service after designation as a Naval Aviator), was most predictive of pilot performance. Flight Aptitude Rating and the Aviation Qualification Test scores entered the prediction formula; however, these variables carried a negative weight. Five variables produced a multiple correlation of .40 in the prediction of Fleet Evaluations. Again, experience level was highly related to Fleet Evaluation ratings completed by squadron commanders. The Flight Aptitude Rating entered the prediction formula, again with negative weight.

Bale, Rickus, and Ambler (14) utilized a success/failure criterion in RAG to determine the relationship of selection and undergraduate training performance to later performance in this near operational environment. A multiple regression analysis indicated that 15 variables were suitable predictors of the criterion. (R = .43). A cross-validation effort resulted in a reduced multiple correlation of .36. Advanced tactical training skills accounted for the greatest proportion of the explained variance. Advanced, basic, and primary flight grades contributed to prediction in that order. Certain undergraduate performance was negatively related to the criterion. (Presolo, Basic Instruments, Basic Final, and Advanced Basic Instrument Grades). Selection tests which entered the prediction formula were the Mechanical Comprehension Test (negative weight) and the Biographical Inventory. Selection tests accounted for 6 percent of the explained variance in the prediction of success or failure in the RAG.

The most recent naval efforts to produce an advanced criteria are those associated with the development and evaluation of an "operational rating" of pilot effectiveness across aircraft and aircraft squadrons. This work is still in the developmental stage, although preliminary results have been documented by Lane and Ambler (85), and Ashburn (10).

The Navy research presents an assessment dilemma. It is assumed that performance in undergraduate training is predictive of future performance in an operational setting. The factor analytic research results discussed above tends to confirm this assumption. The research data suggest a close relationship between advanced undergraduate training performance measures, and RAG and Fleet Performance criteria. The same data also indicates little relationship or even a negative relationship with very early basic and presolo undergraduate training performance parameters. Most surprisingly, however, is the indication that certain selection variables utilized to initially select personnel into aviation training have little, no relationship, or even a negative relationship with the advanced performance criteria.

In summary, these data suggest that certain selection variables utilized to predict success in undergraduate training, and early aviation training performance measures may not be related, or may be inversely related; to performance in an advanced operational environment.

ACMR/ACMI, THE ULTIMATE CRITERION?

The Navy Air Combat Maneuvering Range (ACMR), and the Air Force counterpart, the Air Combat Maneuvering Instrumentation Facility (ACMI), represent a high fidelity criterion for the assessment of fighter pilot performance short of actual air combat in a wartime environment. These facilities function to allow multiple fighter aircraft to engage and maneuver in a tactical environment, allowing the simulated employment of air-to-air missiles in ACMR, missiles and guns in ACMI, as a means of providing training in fighter aircraft tactical skills, weapon systems capabilities and weapon envelope recognition. The Air Combat Maneuvering Ranges provide training in conditions highly similar to combat; however, the high psychological stress levels associated with air combat with its capability to produce aviator injury or death, may be partially absent from these air combat engagement simulations. The adjective "partially" is used, since those who have experienced high fidelity combat simulation environments verify that these engagements evoke an amount of psychological excitement similar to that of actual combat.

The Air Combat Maneuvering Ranges are highly advanced engineering systems which allow the development of tactical skills in real-time in an environment where both the "victor" and "loser" adversary may subsequently confront each other and discuss the tactical maneuvers and skill execution which resulted in the final engagement outcome. These simulations of combat may enable the necessary psychometric control not previously available to permit the identifi-

cation and measurement of critical fighter pilot physical and psychological attributes. It must be cautioned, however, that performance in this environment should not be considered the only criterion. This is particularly true in Navy aviation in that the ability to land an aircraft on a carrier deck for refueling and rearmament may be just as important as the capability to effectively utilize the weapon platform and associated weapons systems in an air combat tactical environment. A more thorough description of the ACMR and ACMI is provided, respectively, by Lau (86), and the USAF ACEVAL-AIMVAL Test Plan (1).

A number of assessment problems must be solved before ACMR performance can be effectively utilized as a criterion. For example, in a one-on-one tactical encounter, individual performance is dependent on that of the adversary to such an extent that performance outcome (victory) may be either the result of superior performance by one individual in maneuvering and utilizing his weapons systems to advantage, or simply very poor adversary effectiveness. This problem quickly compounds itself in unit actions; i.e., 2 on 1, or 2 vs 2, engagements. For example, in a unit context, success in combat may conceivably be the result of a previously developed plan of tactical engagement developed by a unit individual who is never in position to deliver the products of his weapon systems on adversary aircraft in simulated combat.

Participants, training managers and aviation psychologists must be aware that mission success in an environment such as ACMR is not necessarily always the result of good tactical planning and maneuvering execution. Likewise, failure is not always the result of poor tactical planning and maneuvering. It is possible for the results of a given combat simulation to be attributed to either good execution by one adversary, poor execution by the other, or a combination of the two. The effectiveness of ACMR training and the use of ACMR facilities in the development of selection predictor variables depends in part on recognizing this distinction. Mission accomplishment, therefore, is an imperfect criterion for the evaluation of tactical decisions and flying performance skill. Even so, fighter pilot performance over time in an ACMR environment may well be one of the best criteria available. Finally, the ability to control many variables in these high fidelity simulated environments gives ACMR an advantage over actual combat in measuring potential combat performance effectiveness.

The aspects of performance assessment in ACMR environments (noted above) suggest that highly controlled experimental procedures must be utilized in the identification of critical skills and attributes associated with tactical combat performance. Some will suggest that a highly controlled experimental procedure is inconsistent with actual combat. This is because, by its very nature, each aircraft engagement in combat is different. These individuals may argue that the use of a highly controlled experimental procedure in a simulated combat environment is a classic example of a measuring instrument biasing what is measured. Despite the potential problems suggested by such an approach, it is essential to isolate specific variables associated with success in combat. While it is true that no two situations are alike in combat, it is just as true that the

flight training experience is never completely identical for any two individuals undergoing training; and yet considerable gains have been made in isolating factors associated with student success in undergraduate aviation training.

ACMR Performance as Selection Criteria

The ACMR facilities are so new that resultant performance in them has yet to be effectively utilized as criteria in the prediction of aviator performance. Also, there is some controversy concerning which specific ACMR performance parameters should be utilized as criteria. (Time in weapon envelope, kill probability, etc.). Very few of these ACMR assessment problems are insurmountable, and plans are underway to utilize the ACMR facilities to allow a better understanding of the aviator skills and attributes which appear to enhance successful performance in ACMR. Navy Fighter Pilots, for example, have indicated that the individual who sees the other first in an intercept encounter achieves a tactical advantage. As a result, human visual acquisition ability is presently being studied in the ACMR environment, with the idea that certain aviator visual skills may be related to success in a combat environment. Ferguson & Goodson (42) have described the air-to-air visual acquisition task. Jones and Doll (77) in preliminary research suggest that peer rankings are potential predictors of air-to-air visual acquisition capability; and Hutchins and Jones (71) have identified altitude separation as a critical variable in visual target acquisition. In addition, plans are underway to utilize ACMR performance measures as criteria for present selection variables, and for new proposed selection research.

SYNTHETIC SELECTION RESEARCH

There is renewed interest in the use of Flight Simulators for selection purposes as evidenced by current work being sponsored by the Air Force. This renewed interest is in part due to advancing technology; i.e., capability to automate flight simulator performance measures, as evidenced by the work of Hill & Gobel (69), and research which has established a positive relationship between ground based simulator performance and instructor evaluations of student performance in actual flight, Gobel, Baum & Hagin (54); and time to complete training, Woodruff & Smith (144).

LeMaster and Gray (88) evaluated the use of the T-40 instrument trainer as a selection device for the identification of flying abilities possessed by Air Force Undergraduate Pilot Training candidates. Their research indicated that performance in the T-40 Instrument Trainer was predictive of pilot flying performance based on the overall T-37 phase grade, but was not useful in the prediction of ultimate success or failure in undergraduate pilot training. The author notes that these findings are inconsistent and suggests that the bulk of attrition in Air Force UPT results from motivational rather than from flying skill factors.

More recent synthetic selection research conducted by Long & Varney (91) consisted of an evaluation of a reconfigured General Aviation Trainer (GAT-1)

for the selection of pilots in UPT. The GAT-1 performance measurement system is a five-hour learning sample of flight tasks. The automated GAT-1 measurement system, which both administers and scores performance, is called the Automated Pilot Aptitude Measurement System (APAMS). Majesty (92) reports that a preliminary validation study of the Automated GAT-1 System resulted in a correlation of .58, using the criterion pass/fail in Air Force UPT. The system is now undergoing an extensive validation process as one part of an Air Force effort to develop more effective predicters of pilot success, with the ultimate goal of reducing pilot attrition from the present level (25 percent) to 10 percent (104).

AIR COMBAT SIMULATORS

Though not currently being utilized in selection, a number of Air Combat Simulators are available which provide training in many of the performance skills associated with air-to-air combat. While it is the popular consensus that these simulations are less valid than the actual use of aircraft on an ACMR facility, air combat simulators often provide training in skill areas not adaptable to an actual training engagement simulation facility because of technological and/or safety considerations. Additionally, these devices have the potential to expedite aviator acquisition (learning) of combat tactical skills on a more cost effective basis prior to their utilization and execution on the high cost ACMR ranges. These air combat simulations may also serve as more cost effective intermediate selection criteria assuming a positive relationship can be demonstrated to exist between performance in the computerized air combat simulator and ACMR environments. A detailed description of computerized air combat simulator systems[1] with a summary of results associated with attempts at their validation have been documented in a feasibility study to predict combat fighter pilot effectiveness by Youngling, Levine, Mocharnuk and Weston (145).

PERCEPTUAL PSYCHOMOTOR PERFORMANCE

Since psychomotor testing has been known to be related to aviator performance since World War II, why are psychomotor tests no longer used?

Factor Analysis of the complex coordination test indicated the major reason for its predictive goodness. Cronbach (31) suggested that it measures an appropriate amount of cognitive, spatial and mechanical comprehension abilities in addition to the unique contribution of a psychomotor or multilimb coordination factor which no paper-and-pencil tests have yet measured. Psychologists realized that paper-and-pencil tests available to measure non-psychomotor skills were much more economical and easy to administer than the hardware oriented psychomotor tests. Additionally, there was the great problem of unreliability

[1] McDonnell-Douglas Manned Air Combat Simulator, St. Louis, Mo. The Differential Maneuvering Simulator at NASA-Longley, Virginia, and the Simulator for Air-to-Air Combat - Luke AFB, Arizona.

with the psychomotor tests. In fact, the unreliability of these devices became such a problem that the Air Force gave up the use of its psychomotor selection tests in the early 1950s. McGrevy and Valentine (99) indicate that the rationale behind this decision was that the extra amount of predictive variance accounted for by the psychomotor tests was not worth the extensive device upkeep, and maintenance and calibration effort.

With recent technological advances there has been a revival of interest in perceptual/psychomotor assessment. A recent USAF contractual effort resulted in the development of two solid state perceptual psychomotor tests based in part on the old two-hand coordination and complex coordination (stick and rudder test) of World War II fame. Sanders, Valentine and McGrevy (116) report that both tests were transfigured into a solid-state independent testing apparatus of high reliability. Subsequent validation of the tests indicated that complex coordination was a reliable and valid predictor of success vs failure (graduation) and flight training deficiency (similar to the Navy term Flight Failure) in Undergraduate Pilot Training (UPT). McGrevy and Valentine (99) report that the perceptual psychomotor complex coordination test made a unique contribution to the prediction of graduation from Air Force UPT above and beyond that provided by the Air Force paper-and-pencil test selection instrument, the AFOQT. The Air Force is now completing a relatively large scale validation of the AFOQT, GAT-1, and the perceptual psychomotor tests. Discussion with an Air Force laboratory representative[1] indicates that the perceptual psychomotor test (complex coordination) continues to provide additional and unique variance. Additionally, the complex coordination test is highly related to GAT-1 performance. Since the perceptual psychomotor test is less costly, takes less time to complete, and is easier to administer, it is probable that the perceptual psychomotor performance measure may be used in place of the GAT-1 as a predictor variable in USAF Pilot Selection.

DIVISION OF ATTENTION

Recently, several efforts have shown predictive success with tests measuring the ability to perform more than one task simultaneously. Trankell (128) reported selection test efforts conducted by the Scandinavian Air Lines System on a Simultaneous Capacity Test which combined a problem solving task and a simple motor task consisting of rhythmic tapping. A biserial correlation of .42 was observed for predicting training success.

Divided attention during the performance of simultaneous tasks was used successfully by Damos (33) to predict success in introductory pilot training. Subjects performed a one-dimensional tracking task while cancelling lights

[1]Personal communication (22 March 1977) with Dr. David Hunter, Research Psychologist, Selection and Classification Branch, Air Force Personnel Research Division, Lackland AFB, Texas.

appearing on an adjacent display with the opposite hand. Tracking under divided attention was used to predict check flight scores and produced validities in the .50 to .60 range.

A refined technique for measuring divided attention was recently offered by North and Gopher (106). The technique provides several desirable methodological controls: (1) measuring the candidate's capacity on the tasks performed separately using adaptive logic to selectively adjust task difficulty, and (2) controlling and adjusting the priorities between tasks during concurrent performance. Scores on time-shared performances combining tracking with digit processing performance were predictive of performance of students in introductory pilot training and differentiated between experienced instructor pilots and flight-naive subjects. Single-task performance was not predictive of student success, lending further support that multi-task skills rather than single-task skills have potential for predictive validity.

A selective attention test using a dichotic listening technique has been investigated by the Israeli Air Force and was successful in predicting training success in high-performance jet aircraft. In dichotic listening tests the subject is told to ignore one message while listening for relevant words or items in the designated channel. Gopher and Kahenman (56) report validities in the range .30 to .40 using 100 Israeli flight candidates. More recently, Gopher (55) presents new data on the dichotic listening test based on a population of 200 individuals finishing training. (The final population group will consist of approximately 2,000 subjects). His initial data indicate a low but significant relationship with success in jet training ($r = .18$). Although the correlation is low, it has virtually no relationship with other predictor measures, thus it offers a new and welcome dimension to the prediction of pilot success.

The previous examples demonstrate the predictive utility of measures designed to test the time-sharing capabilities of the aviation candidate. More data with larger samples are needed to further assess this utility. A large scale study using several divided and selective attention tests is desirable.

VOICE ANALYSIS AS A MEASURE OF PSYCHOLOGICAL STRESS

There is current interest in the evaluation of individual speech characteristics under stress to determine their relation to aviator performance, especially the specific motivational components of flying, including stress and fear. Williams and Stevens (140, 141) analyzed vocal recordings of pilots in aircraft mishaps and report that acoustic analysis of speech samples may reveal the underlying emotional condition of a speaker under extreme conditions of stress or anxiety. New voice analysis techniques are currently available as a result of minicomputer hardware and software development which are an improvement over analysis methods previously used. Application of these new techniques and developments may provide the means for an improved understanding of the effect of stress/anxiety on human flying performance. Apparently, voice analysis as a

technique of anxiety measurement has potential as a research tool worthy of further evaluation to determine its ability to objectively identify anxiety prone individuals and to determine the relationship of anxiety to motivational attrition in aviator training programs. This area of research has the potential to result in effective prediction measures of motivational attrition.

VESTIBULAR DISORIENTATION RESEARCH

A relatively recent Navy development, though related research has been conducted for years; concerns a vestibular disorientation procedure which appears effective in the prediction of aviator motion, or flying sickness. The procedure consists of a rotating chair and a series of head movements which measures the potential aviator's reaction to mild rotating dynamic forces. Ambler and Guedry (6-8) and Harris, Ambler and Guedry (66) report that the Pensacola Brief Vestibular Disorientation Test (BVDT) is related to both airsickness and anxiety with correlations in the .4 and .2 range, respectively. The higher relationship between aviator performance in the rotating chair and airsickness seems to indicate a primary motion sickness relationship. The BVDT is scheduled to become a Navy secondary selection device in FY-78 (October 1977). Majesty (92) indicates that the Vestibular Disorientaton Procedure is currently undergoing evaluation by the USAF for aviator selection purposes.

HUMAN ANTHROPOMETRY IN AIRCRAFT ASSIGNMENT

Recent Naval Aerospace Medical Research Laboratory (NAMRL) research by Gregoire (60) to improve pilot aircraft pertormance and reduce the potential for accidents, involved the measurement of human physical dimensions (foot, leg, arm/hand functional reach, and eye height relative to a sitting position, etc.) in relation to the cockpit work space requirements of Navy operational aircraft. Essentially, this very practical effort represents an attempt to eliminate the practice of placing individuals in aircraft in which they are physically unsuited to operate one or a number of controls. This study effort resulted in a procedure - now being implemented - requiring the measurement of the physical dimensions of each aviator to identify those aircraft to which the individual should not be assigned. Preliminary research has been conducted by Baisden (12) and Lane (84) at NAMRL to delineate the anthropometry characteristics of potential female naval aviators.

CONCLUSIONS/RECOMMENDATIONS

CONCLUSIONS CONCERNING SELECTION MEASURES

The decision to develop new prediction measures of aviator success by assessing a wide variety of abilities is not a recent one. Aviation experts have known that the aviator's task is such that no one ability can provide all the necessary behaviors to become a highly successful flyer. Initial attempts to develop effective predictive test batteries were plagued with time constraints and objec-

tive limits because of the demand for the quick screening of applicants for training. Today's objectives are different, however, as the demand now is for predictors which will: (1) reduce attrition in Undergraduate Aviator Training and lead to important cost savings; and (2) be effective in predicting pilot performance from 4 to 8 years into the officer's career.

The potential for success in predicting aviator performance is just as bright today as it was in the 1940s. Typically, test batteries utilized to select aviators into undergraduate training account for approximately 25-40 percent of the variance associated with aviator success. The lack of any prominent breakthrough in perceptual/cognitive paper-and-pencil testing since World War II years suggests that non-paper-and-pencil performance tests should be investigated more fully to determine their relationship to aviator performance. This review of aviation selection research has attempted to provide insight into the historical development of past and current selection research as a basis for the development of future research efforts. It has been written in an attempt to motivate research personnel to more fully investigate those research areas which appear to have potential for the future selection of aviators for initial undergraduate training programs and for the prediction of mid-term aviator success (4-8 years after designation). The outlook for long-term prediction is unclear. Typically, the criteria utilized in operational flying validation studies continue to be poorly defined, or have questionable reliability and objectivity. Additionally, long-term longitudinal validation studies are typically hard to validate because of the loss of cases due to leaving the service, and changes in assignment. More importantly, as the officer/aviator gains tenure in his respective service, he gains increasing rank and with it new (management) responsibilities. In fact, by the time the officer has spent 10-12 years in the service, his management responsibilities may be considered more important than his flying duties. Until more reliable and objective management and performance evaluation techniques become available and are effectively used, there will continue to be the problem of relating individual performance skills and abilities to poorly defined and poorly measured management and operational flying performance criteria.

New technological advancements, such as those resulting in aircraft engagement simulation combat environments (ACMR facilities and computer simulations of air combat) may provide the means to identify, and reliably measure, relevant physical and psychological attributes and performance skills to enable the more valid selection of aviator trainees. It is unclear, however, whether the factors utilized in aviator selection to predict success in undergraduate training will be related to successful performance in post-graduate operational environments.

Several goals appear important for the test developer. One should be the identification of highly relevant performance measures for use as criteria in test prediction. A second goal is the identification and development of non paper-and-pencil performance measures to better predict criterion performance in

undergraduate training, and in post-graduate operational flying environments. The most encouraging types of non paper-and-pencil performance prediction measures worthy of investigation appear to be selective and divided attention capabilities, stress and anxiety motivational measurement, and perceptual-psychomotor skill assessment.

REFERENCES

1. ACEVAL-AIMVAL Final Test Plan. Nellis Air Force Base, Nevada. AIMVAL Joint Test Force, August 1976.

2. Adams, J. A. An evaluation of test items measuring motor abilities. USAF Personnel and Training Research Center, Research Report No. AFPTRC-TN-56-55, Lackland AFB, Texas, 1956.

3. Adams, J. A., and Chambers, R. W. Response to simultaneous stimulation of two sense modalities. Journal of Experimental Psychology, 63, 198-206, 1962.

4. Ades, H. W. Electroencephalographic findings in Relation to episodes of altered consciousness in aviators, NSAM 182, Naval School of Aviation Medicine, Pensacola, Florida, 1961.

5. Adiseshiah, W. T. V. Speed in decision making under single channel display conditions. Indian Journal of Psychology, 32, 105-108, 1957.

6. Ambler, R. K., Guedry, F. E. The validity of a brief vestibular disorientation test in screening pilot trainees. NAMI-947. Naval Aerospace Medical Institute, Pensacola, Florida, 1965.

7. Ambler, R. K., Guedry, F. E. Cross-validation of a brief vestibular orientation test administered by a variety of personnel. NAMI 1009. Naval Aerospace Medical Institute, Pensacola, Florida, 1967.

8. Ambler, R. K., Guedry, F. E. Reliability and validity of the brief-vestibular disorientation test compared under 10 and 15 RPM conditions. NAMRL 1115. Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1970.

9. Ambler, R. K., & Smith, M. J. Differentiating aptitude factors among current aviation specialities, NAMRL 1207, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1974.

10. Ashburn, J. H. Resolution of overall aviator proficiency rating. Paper presented at the 1977 Annual Convention, Aerospace Medical Association, Las Vegas, Nevada, May 1977.

11. Atkinson, C. J. Intelligibility of speech heard at high altitude and sea level, NSAM 592, Naval Aerospace Medical Institute, Pensacola, Florida, 1952.

12. Baisden, A. G. Research on the aeronautical adaptability of women. Paper presented at the 18th Annual Conference of the Military Testing Association, Gulf Shores, Alabama, 18-22 October 1976.

13. Bale, R. M., Rickus, G. M., & Ambler, R. K. Replacement Air Group, performance as a criterion for naval aviation training, NAMRL 1126, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1970.

14. Bale, R. M., Rickus, G. M., & Ambler, R. K. Prediction of advanced level aviation performance criteria from early training and selection variables. Journal of Applied Psychology, 58, (3), 347-350, 1973.

15. Bale, R. M., Smith, M. J., & Ambler, R. K. Factor analysis of undergraduate and postgraduate flight training grades. Aerospace Medicine, 43, 373-375, 1972.

16. Bair, J. T. The characteristics of the wanted and unwanted pilot-in-training and in combat. Morale Report #2. NSAM 135. Naval School of Aviation Medicine, Pensacola, Florida, 1962.

17. Bair, J. T., Lockman, R. F., & Martoccia, C. T. Validity and factor analyses of naval air training predictors and criterion measures. Journal of Applied Psychology, 40: 213-219, 1956.

18. Booth, R. F., Berkshire, J. R. Factor analysis of aviation training measures and post-training performance evaluations. NAMI 1050, Naval Aerospace Medical Institute, Pensacola, Florida, 1968.

19. Brictson, C. A., Burger, W. J., & Gallagher, T. J. Prediction of pilot performance during initial carrier landing qualifications. Aerospace Medicine, 43: 483-487, 1972.

20. Brown, I. D. The measurement of perceptual load and reserve capacity. The Transactions of Industrial Medical Officers, 14, 44-49, 1964.

21. Bucky, S. F. The California Psychological Inventory given to incoming AOCs and DORs with normal and ideal instructions. NAMRL 1127. Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1971.

22. Bucky, S. F., Spielberger, C. D., & Bale, R. M. State and trait anxiety in student naval aviators. NAMRL 1125, Naval Aerosapce Medical Research Laboratory, Pensacola, Florida, 1970.

23. Camp, R. T. The effect of noise environment upon speaker intelligibility. Report #63, NSAM 641. Naval School of Aviation Medicine, Pensacola, Florida, 1956.

24. Channel, R. C. Psychomotor Tests as Predictors of Success in Flight Training, Naval Reserve Air Base, Squantum, Mass., 1942.

25. Chief of Naval Education and Training FY-75 Pilot Cost Figures. Naval Air Station, Pensacola, Florida, 1975.

26. Clark, B., & Graybiel, A. The effect of angular acceleration on sound localization. The Audiogyrol illusion. Report #21. NSAM 422, Naval School of Aviation Medicine, Pensacola, Florida, 1948.

27. Connelly, E. M., Schuler, A. R., & Knoop, P. A. Study of adaptive mathematical models for deriving automated pilot performance measurement techniques. Vol. I and II, AFHRL TR-69-7, Human Resources Laboratory, Wright-Patterson AFB, Ohio, 1969.

28. Creager, J. A. Validation of the February 1947 Aircrew Classification Battery for the 1950 pilot training classes. Technical Memorandum 57-8, USAF Personnel and Training Research Center, Lackland AFB, Texas, 1957.

29. Creelman, J. A. An Analysis of the Physical Fitness Index in relation to training criteria in naval air training. NSAM Report 180. Naval School of Aviation Medicine, Pensacola, Florida, 1954.

30. Creelman, J. A. Evaluation of approach training procedures. NSAM 329, Naval School of Aviation Medicine, Pensacola, Florida, 1955.

31. Cronbach, L. J. Essentials of Psychological Testing, 3rd Edition, Harper and Row, Inc., New York, 1970, pp. 97.

32. Dailey, J. T., & Gragg, D. B. Postwar research on the classification of aircrew, Research Bulletin 49-2, USAF Human Resources Research Center, Lackland AFB, Texas, 1940.

33. Damos, D. L. Cross-adaptive measurement of residual attention to predict pilot performance. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, Technical Report ARL-72-25/AFOSR-72-12, 1972.

34. Davis, F. B. (Editor) The AAF Qualifying Examination, Army Air Forces - Report #6, Washington, D.C., 1947 (AD 651 782).

35. Dockery, F. C., & Isaacs, S. Psychological research in aviation in Italy, France, England, and the A.E.F. Journal of Comparative Psychology, 1: 115-148, 1921.

36. Doll, R. E. Estimating the "true validity" of the naval aviation selection tests. Paper presented at the 1977 Annual Convention, Aerospace Medical Association, Las Vegas, Nevada, May 1977.

37. Edgerton, H., & Walker, R. H. History and development of the Ohio State Flight Inventory. Part I: Early versions and basic research. Washington, D.C.: Civil Aeronautics Administration, Division of Research, Report 47, 1945.

38. Egan, D. E. Accuracy and latency scores as measures of spatial information processing, NAMRL 1224, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1976.

39. Ekstrom, P. Analysis of pilot workloads in flight control systems with different degrees of automation. Paper presented at the Institute of Radio Engineers International Congress on Human Factors in Electronics, Long Beach, Calif., May 3-4, 1962.

40. Ericksen, S. C. Analysis of basic grade slip folders. Research Note PILOT: 52-5, Wright-Patterson AFB, Ohio, 1952.

41. Examiners Manual and Scoring Instructions, U. S. Navy and Marine Corps Aviation Selection Tests, NAVMED P-5098 (1971), Aerospace Operational Psychology Branch, Bureau of Medicine and Surgery, Navy Department, Washington, D.C.

42. Ferguson, J. C., & Goodson, J. E. A description of the Air-to-Air Visual Acquisition Task. Aerospace Psychology Technical Memorandum 72-2, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1972.

43. Fiske, D. W. Validation of naval aviation cadet selection tests against training criteria. Journal of Applied Psychology, 31(6), 601-614, 1947.

44. Flanagan, J. C. (Ed.) The Aviation Psychology Program in the Army Air Forces, Army Air Forces, Aviation Psychology Program, Research Report 1, Government Printing Office, Washington, D.C., 1948.

45. Fleishman, E. A. A factorial study of psychomotor skill abilities. USAF Personnel and Training Research Center, Research Bulletin, 54-15, 1954.

46. Fleishman, E. A., & Ellison, G. D. A Factor analysis of five manipulative tests, Journal of Applied Psychology, 46, 96-105, 1962.

47. Fleishman, E. A., & Hempel, W. E., Jr. Factorial analysis of complex psychomotor performance and related skills. Journal of Applied Psychology, 40, 96-104, 1956.

48. Fleishman, E. A., & Ornstein, G. N. An analysis of pilot flying performance in terms of component abilities. Journal of Applied Psychology, 44, 146-155, 1960.

49. Forbes, A., Davis, H., & Davis, P. A. Electroencephalography of Naval Aviators, Washington, D.C. CAA Division of Research, Report #13, 1943.

50. Franzen, R., & McFarland, R. A. Detailed statistical analysis of data obtained in the Pensacola study of naval aviators. Rept. 41, Division of Research, Civil Aeronautics Administration, Washington, D.C., 1945.

51. Freeberg, N. E. The Biographical Interest Blank as a predictor of student achievement: A review. Psychological Report, 20: 911-925, 1967.

52. Gastaut, H., Lee, M. C., & Labourer, P. Comparative EEG and psychometric data for 825 French naval pilots and 511 control subjects of the same age. Aerospace Med., 31: 547-552, 1960.

53. Gilliland, A. R., Studies in pilot selection at Northwestern University. Progress Report (to NRC Committee on Selection and Training of Aircraft Pilots), 1940.

54. Gobel, R. A., Baum, D. R., & Hagin, W. V. Using a ground trainer in a job sample approach to predicting pilot performance. AFHRL TR-71-50, Flying Training Division, Williams AFB, Arizona, 1971.

55. Gopher, D. Letter of Correspondence, Subject: Major initial results of the use of the dichotic listening selective attention test in a large scale validation evaluation effort in the Israeli Air Force, 1976.

56. Gopher, D., & Kahneman, D. Individual differences in attention and the prediction of flight criteria. Perceptual and Motor Skills, 33, 1335-1342, 1971.

57. Gordon, T. A. The airline pilot: A survey of the critical requirements of his job and of pilot evaluation and selection procedures (Rep. No. 73). Washington, D.C.: Civil Aeronautics Administration, Division of Research, 1947.

58. Graybiel, A. Spatial disorientation in flight. Report #24, NSAM 425, Naval School of Aviation Medicine, Pensacola, Florida, 1950.

59. Graybiel, A., & Niven, J. I. The effect of a change in direction of resultant force on sound localization: The audiogravic illusion. Report #23, NSAM 424, Naval School of Aviation Medicine, Pensacola, Florida, 1950.

60. Gregoire, H. G. Failure in cockpit geometry. Paper presented to the Society of Flight Test Engineers, Flight Test Technology Symposium, Washington, D.C., 1977.

61. Griew, S. Set to respond and interrupting signals upon tracking performance. Journal of Experimental Psychology, 57, 333-337, 1959.

62. Griffin, G. R., & Mosko, J. D. A review of naval aviation attrition research 1950-1976: A base for the development of future research and evaluation. NAMRL 1237, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1977.

63. Guedry, F. E. Age as a variable in post-rotational phenomena. Report #19, NSAM 478, Naval School of Aviation Medicine, Pensacola, Florida, 1950.

64. Guilford, J. P., & Lacey, J. E. (Ed.) Printer Classification Tests, Parts I and II. Army Air Forces, Aviation Psychology Program, Research Rept. 5. Government Printing Office, Washington, D.C., 1947.

65. Guinn, N., Tupes, E. C., & Alley, W. E. Cultural subgroup differences in the relationships between Air Force Aptitude composites and training criteria, AFHRL-TR-70-35. Personnel Research Division, Lackland AFB, Texas, 1970.

66. Harris, C. S., Ambler, R. K., & Guedry, F. E. A brief vestibular disorientation test. NSAM 856, Report #82, Naval School of Aviation Medicine, Pensacola, Florida, 1963.

67. Hathaway, S. R., & McKinley, J. C. A multiphasic personality schedule (Minnesota): III. The measurement of symptomic depression. Journal of Psychology, 14, 73-84, 1942.

68. Henmon, V. A. C. Air service tests of aptitude for flying. Journal of Applied Psychology, 3: 103-109, 1919.

69. Hill, J. W., & Gobel, R. A. Development of automated GAT-1 performance measures AFHRL-TR-71-18, Flying Training Division, Williams AFB, Arizona, 1971.

70. Houston, C. A. A Study of the physiological changes which occur during acclimatization to high altitude. Occurrence of aero-embolism at moderate altitudes. Report #2, NSAM 83, Naval School of Aviation Medicine, Pensacola, Florida, 1946.

71. Hutchins, C. E., Jones, T. N. An initial investigation of those ACMR parameters related to initial air-to-air visual acquisition. Aerospace Psychology Technical Memorandum 75-2, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1975.

72. Hutchins, C. W., & Pomarolli, R. S. Performance in the preflight water survival course as a predictor of success in flight training. SR 65-3, Naval School of Aviation Medicine, Pensacola, Florida, 1965.

73. Jenkins, J. G. The nominating technique as a method of evaluating air group morale. Journal of Aviation Medicine, 19(1), 12-19, 1948.

74. Jenkins, J. G., Ewart, E. S., & Carroll, J. B. The combat criterion in naval aviation (Report #6), Washington, D.C. National Council Committee on Aviation Psychology, 1950.

75. Johnson, J. H. An evaluation of a device designed to teach the principles of trimming an aircraft. SR 61-1, Naval School of Aviation Medicine, Pensacola, Florida, 1961.

76. Jones, M. B. Aspects of the autonomous personality - VII. The Pensacola Z Survey. NSAM 250, Naval School of Aviation Medicine, Pensacola, Florida, 1958.

77. Jones, T. N., Doll, R. E. Peer ranking as criterion measure for initial acquisition of targets. Aerospace Psychology Technical Memorandum 74-2, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1974.

78. Kalsbeek, J. W. H. On the measurement of deterioration in performance caused by distraction stress. Ergonomics, 7, 187-195, 1964.

79. Kaplan, H. Prediction of success in army aviation training. Technical Research Report 1142, U.S. Army Personnel Research Office, OCRD, 1965.

80. Kelley, C. R., & Kelley, E. J. A manual of adaptive techniques. Santa Monica, Calif.: Dunlap and Associates, Final Report on Contract Nonr-4986 (00), 1970.

81. Kelley, C. R., & Prosin, D. J. Adaptive performance measurement. Santa Monica, Calif.: Dunlap & Associates, Final Report on Contract Nonr-4986(00), 1968.

82. Kennard, M. A. The electroencephalogram in psychological disorders. Psychosom. Med., 15: 95-115, 1953.

83. Koerth, W. A pursuit apparatus. Psychological Monographs, 31, 228-292, 1922.

84. Lane, N. E. Female anthropometry and cockpit geometrics. Paper presented at Women in Aviation Symposium, American Psychological Convention, New Orleans, La., 29 Aug - 2 Sep. 1974.

85. Lane, N. E., Ambler, R. K. Operational criteria of pilot performance: Development of a postgraduate rating form. Paper presented at the Aerospace Medical Association 46th Annual Meeting, San Francisco, Calif., April 27-2 May 1975.

86. Lau, C. Air Combat Maneuvering Range (ACMR) system accuracy as it relates to pilot performance measures to air-to-air combat. Aerospace Psychology Technical Memorandum 76-2, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1976.

87. Leiman, J. M., & Friedman, G. Validation of aircrew classification battery against advanced flying training - single engine jet - criterion. USAF Human Resources Research Center, Research Note PERS: 52-2, Lackland AFB, Texas, 1952.

88. LeMaster, W. D., & Gray, T. H. Ground training devices in job sample approach to UPT selection and screening. AFHRL TR-74-86, Flying Training Division, Williams AFB, Arizona, December 1974.

89. Lennox-Buchthal, M., Buchthal, F., & Rosenfalck, P. Correlation of electroencephalographic findings with crash rate of military jet pilots. Copenhager, Institute of Neurophysiology, University of Copenhagen, Tech. Note No. 2, AD 253 774. 1960.

90. Lin, C. What psychological qualities should be cultivated in preparation for flight training? Hang K'ung Chih Shih (Aeronautical Knowledge No. 3, 1960, pp. 24-25. (Translation by Foreign Technology Division, Wright Patterson AFB, Ohio (AD 693 456), 1969.

91. Long, G. E., & Varney, N. C. Automated Pilot Aptitude Measurement System, AFHRL-TR-75-58, Personnal Research Division, Lackland AFB, Texas, 1975.

92. Majesty, M. S. The making of an aircraft pilot selection system. Officer Training School, Air Training Command, Lackland AFB, Texas, 1976.

93. Mann, C. W. The effects of auditory-vestibular nerve pathology on space perception. Report #22. NSAM 481, Naval School of Aviation Medicine, Pensacola, Florida, 1951.

94. Mann, C. W., & Dauterive, H. J. The perception of the postural vertical. 1. The modification on non-labyrinthine cues. Report #4, NSAM 463, Naval School of Aviation Medicine, Pensacola, Florida, 1949.

95. Mashburn, N. C. Mashburn automatic serial action apparatus for detecting flying aptitude. Journal of Aviation Medicine, 5, 155-160, 1934.

96. Mathews, J. J. Racial equity in selection in air force officer training school and undergraduate flying training. AFHRL-TR-77-22, Personnel Research Division, Lackland AFB, Texas, 1977.

97. McFarland, R. A. Human Factors in Air Transportation. New York, McGraw-Hill Book Co., 1953.

98. McFarland, R. A., & Franzen, R. The Pensacola study of naval aviators--final summary report. Rept. 38. Division of Research, Civil Aeronautics Administration, Washington, D.C., November 1944.

99. McGrevy, D. F., Valentine, L. D. Validation of two aircrew psychomotor tests, AFHRL TR-74-4, Personnel Research Division, Lackland AFB, Texas, 1974.

100. Melton, A. W. (Ed.) Apparatus Tests. Army Air Forces, Aviation Psychology Program, Research Report 4. Government Printing Office, Washington, D.C., 1947.

101. Miller, E. F. Effect of breathing 100 per cent xoygen at atmospheric pressure upon the visual field and visual acuity. Report #1, NSAM 51, Naval School of Aviation Medicine, Pensacola, Florida, 1958.

102. Miller, N. E. (Ed.) Psychological research on pilot training. Army Air Forces Aviation Psychology Program Research Report No. 8 Washington: U.S. Government Printing Office, 1947.

103. Miller, R. E. Interpretation and utilization of scores on the AFOQT-AFHRL-TR-69-103. Personnel Research Laboratory, Lackland AFB, Texas, 1969.

104. Mission Analysis on Future Undergraduate Pilot Training: 1975 through 1990. TR-72-001, USAF Mission Analysis Study Group, AFSC, Wright-Patterson AFB, Ohio, 1972.

105. Mundy-Castle, A. C. An appraisal of electroencephalography in relation to psychology. J. Nat'l Inst. Pers. Res. (Johannesburg, South Africa), Monogr. Suppl. No. 2, 1958.

106. North, R. A., & Gopher, D. Measures of attention as predictor training performance. Human Factors, 18(1), 1-14, 1976.

107. Page, H. E. Vertigo incidence among naval aviators. Report #16. NSAM 475, Naval School of Aviation Medicine, Pensacola, Florida, 1950.

108. Page, H. E., & Lyon, V. W. Preliminary report on the performance in the SNJ cycloramic link and its relationship to subsequent performance in Stage A of flight training. NSAM 351, Naval School of Aviation Medicine, Pensacola, Florida, 1950.

109. Passey, G. E., & McClaurin, W. A. Perceptual-Psychomotor Tests in Aircrew Selection: Historical Review and Advanced Concepts. PRL-TR-66-4, Personnel Research Laboratory, Lackland AFB, Texas, 1966.

110. Peters, R. W. The effect of acoustic environment upon speaker intelligibility Report #26, NSAM 604, Naval School of Aviation Medicine, Pensacola, Florida, 1954.

111. Picard, P., Labourer, P., & Navarronne, P. Examen electroencephalographique du personnel volant. Rev. Med. Aeronautique, Paris, p. 53, 1955.

112. Poe, A. C., & Lyon V. W. The effectiveness of the cycloramic link trainer in the U.S. Naval School, Preflight. NSAM 513, Naval School of Aviation Medicine, Pensacola, Florida, 1952.

113. Povenmire, H. K., Alvares, K. M., & Damos, D. L. Observer-observer flight check reliability. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, Technical Report LF 20-2, 1970.

114. Richardson, W. G., & Rusis, G. Future undergraduate pilot training system study - Appendix VII FUPTS Student Selection and Career Considerations. Aeronautical Systems Division, AFSC Wright Patterson AFB, Ohio, 1971.

115. Rickus, G. M., & Berkshire, J. R. Development of an aviation combat criterion - Preliminary Report NAMI 1047, Naval Aerospace Medical Institute, Pensacola, Florida, 1968.

116. Sanders, J. H., Valentine, L. D., & McGrevey, D. F. The development of equipment for psychomotor assessment. AFHRL-TR-71-40, Lackland AFB, Texas, Personnel Division, Lackland AFB, Texas, 1971.

117. Schwarts, M., & Lowe, B. H. Gymnastics grades as predictors to attrition to flight training. SR #58-15, Naval School of Aviation Medicine, Pensacola, Florida, 1958.

118. Schweitzer, J. J. Reducing attrition rate and training costs in undergraduate pilot training through improved screening and motivational methods. Air University Research Study, Air University, Maxwell AFB, Alabama, 1976.

119. Schaefer, H. J. Evaluation of present day knowledge of COSMIC radiation at extreme altitude in terms of the hazard to health. Report #1, NSAM 20, Naval School of Aviation Medicine, Pensacola, Florida, 1950.

120. Selzer, U., Hulin, C. L., Alvares, K. M., Swartzendruber, L. E., & Roscoe, S. N. Predictive validity of ground based flight checks in Private Pilot training and certification (Tech. Rep. ARL-72-19/AFOSR-72-8). Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, 1972.

121. Sem-Jacobsen, C. W., & Sem-Jacobsen, Ingebjorg, E. Selection and Evaluation of Pilots for High-Performance Aircraft and Spacecraft by Inflight EEG Study of Stress Tolerance. Aerospace Medicine, 34, 605-609, 1963.

122. Shannon, R. H., & Waag, W. L. Toward the development of a criterion for Fleet Effectiveness in the F-4 Fighter Community. NAMRL 1173, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1972.

123. Shannon, R. H., & Waag, W. L. The Prediction of pilot performance in the F-4 aircraft. NAMRL 1186. Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1973.

124. Shannon, R. H., Waag, W. L., & Ferguson, J. C. A new approach to criterion development in the replacement air group. NAMRL 1158. Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1972.

125. Shannon, R. H., Waag, W. L., & Long, G. M. Isolation of recurring student pilot errors during primary flight training: I. Spin Maneuver. NAMRL 1189, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1973.

126. Sheeley, J. Psychological selection and testing of naval aviators and flight officers: An historical sketch, NAMRL Preliminary Draft Report, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, 1977.

127. Slocum, G. K., Williges, B. H., & Roscoe, S. N. Meaningful shape coding for aircraft switch knobs. Aviation Research Monographs, 1 (3), 27-40, 1971.

128. Trankell, A. The psychologist as an instrument of prediction. Journal of Applied Psychology, 43, 170-175, 1959.

129. Vinacke, W. E. The concept of aviators' "vertigo." Report #7, NSAM 408, Naval School of Aviation Medicine, Pensacola, Florida, 1946.

130. Viteles, M. S. The aircraft pilot: Five years of research, a summary of outcomes. Committee on Selection and Training of Aircraft Pilots, Civil Aeronautics Administration Report #46, Washington, D.C., 1945.

131. Voas, R. B. Anxiety, intelligence and distortion toward social favorability. NSAM 297, Naval School of Aviation Medicine, Pensacola, Florida, 1956.

132. Voas, R. B. A procedure for reducing the effects of slanting questionnaire responses toward social acceptability. NSAM 215, Naval School of Aviation Medicine, Pensacola, Florida, 1956.

133. Voas, R. B. Relationships among three types of response sets. NSAM 302, Naval School of Aviation Medicine, Pensacola, Florida, 1958.

134. Wallon, E. J., & Webb, W. B. The effect of varying degrees of projection on test scores, NSAM 210, Naval School of Aviation Medicine, Pensacola, Florida, 1956.

135. Wallon, E. J., & Webb, W. B. A note on the effect of test set on the Rosenzweig Picture Frustration Test. NSAM 217, Naval School of Aviation Medicine, Pensacola, Florida, 1956.

136. Waters, L. K. Relationships among inventoried self-rated and peer-rated interpersonal values. NAMI 1015, Naval Aerospace Medical Institute, Pensacola, Florida, 1967.

137. Waters, L. K., & Wherry, R. J., Jr. Factor analysis of selection tests and performance measures in U.S. Naval School, Pre-Flight, NSAM 252, Naval School of Aviation Medicine, Pensacola, Florida, 1959.

138. Wherry, R. J., Jr., & Waters, L. K. Factor analysis of primary and basic stages of flight training: Advanced jet pipeline students. NSAM-254, Naval School of Aviation Medicine, Pensacola, Florida, 1960.

139. Wherry, R. J., Jr., & Waters, L. K. Factor analysis of primary and basic stages of flight training: Advanced multi-engine pipeline students. NSAM-256, Naval School of Aviation Medicine, Pensacola, Florida, 1960.

140. Williams, C. E., & Stevens, K. N. On determining the emotional state of pilots during flight. An exploratory study. Aerospace Medicine, 40 (12), 1369-1372, 1969.

141. Williams, C. E., & Stevens, K. N. Emotions and speech: some acoustal correlates. Journal of the Acoustical Society of America, 52, #4 (Part 2), 1238-1250, 1972.

142. Williges, B. H., Roscoe, S. N., & Williges, R. C. Synthetic flight training revisited. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, (Tech. Rep. ARL-72-21/AFOSR-72-10), 1972.

143. Willingham, W. W. A note on the predictive validity of physical training phase grade. NSAM 220, Naval School of Aviation Medicine, Pensacola, Florida, 1957.

144. Woodruff, R. R., & Smith, J. F. T-4G simulator and T-4 ground training devices in USAF undergraduate pilot training AFHRL TR 74-78. Williams AFB, Arizona, Flying Training Division, 1974.

145. Youngling, E. W., Levine, S. H., Mocharnuk, J. E., & Weston, L. M. Feasibility study to predict combat effectiveness for selected military roles: Fighter pilot effectiveness. McDonnell-Douglas Astronautics Company - Contract No. MDA903 76C 0169 (ARPA), 1977.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(14) REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NAMRL-SR-77-2 · SPECIAL REPORT 77-2	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) Aviator Selection 1919-1977	5. TYPE OF REPORT & PERIOD COVERED (9) INTERIM rept.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Robert A. North and Glenn R. Griffin	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Aerospace Medical Research Laboratory Naval Air Station, Pensacola, Florida 32508	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS MF51.525.002-5012DX5X
11. CONTROLLING OFFICE NAME AND ADDRESS (11)	12. REPORT DATE 4 October 1977
	13. NUMBER OF PAGES 57
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (15) 57p.	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(16) F51525

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(17) X F51525 001

18. SUPPLEMENTARY NOTES

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

Aviation selection	Drop on request	Manifestation of apprehension
Aviation personnel	Voluntary withdrawal	
Aviator selection tests	Flight failure	Flying deficiency
Motivation	Self initiated elimination	

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

The potential for increased success in predicting aviator performance is high. The fact that current selection tests normally account for less than half of the total variance associated with aviator success (in training) suggests that there are additional factors associated with aviator performance which are not now being adequately assessed. The lack of any prominent breakthrough in perceptual/cognitive paper-and-pencil testing since the war years (WW-II) suggests that non-paper-and-pencil performance tests should be investigated more

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

406 061

B

fully to determine their relationship to aviator performance in both a training and operational setting.

Relating aviator performance to better and more appropriate performance measurement criteria is a continuing psychological assessment goal. New technological advancements such as the Navy and Air Force Air Combat Maneuvering Ranges have the potential to identify and reliably measure relevant physical and psychological human attributes which may provide more accurate and valid prediction of aviator operational performance.

Still, such obviously valid criteria as ACMR performance pose an interesting assessment problem. It is unclear whether the prediction variables presently utilized in aviation selection to predict successful performance in undergraduate training are related to successful performance in post-graduate operational environments.

It is suggested that research be oriented toward the identification of highly relevant criterion-oriented performance measures for use as criteria in the evaluation of present and new selection prediction variables and identification and development of non-paper-and-pencil performance prediction measures to improve prediction of criterion performance in undergraduate training, and in post-graduate operational flying environments. Examples of non-paper-and-pencil performance prediction measures recommended for future study are Selective and Divided Attention, Stress and Anxiety Motivational Measurement, and Perceptual Psychomotor skill assessment.