AD-A279 981

NAMRL-1391

COMPUTER-BASED PSYCHOMOTOR TESTS IN OPTIMAL TRAINING TRACK ASSIGNMENT OF STUDENT NAVAL AVIATORS

D. R. Street, Jr., and D. L. Dolgin

Naval Aerospace Medical Research Laboratory
Naval Air Station
Pensacola, Florida 32508-5700

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
51 HOVEY ROAD, PENSACOLA, FL 32508-1046

NAMRL-1391

COMPUTER-BASED PSYCHOMOTOR TESTS IN OPTIMAL TRAINING TRACK ASSIGNMENT OF STUDENT NAVAL AVIATORS

D. R. Street, Jr., and D. L. Dolgin

Approved for public release; distribution unlimited.

Reviewed and approved 14 Feb 94

A. J. MATECZUN, CAPT, MC USN
Commanding Officer

This research was sponsored by the Naval Medical Research and Development Command under work unit 63706N M0096.002-M00960.01.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

Volunteer subjects were recruited, evaluated, and employed in accordance with the procedures specified in the Department of Defense Directive 3216.2 and Secretary of the Navy Instruction 3900.39 series. These instructions are based upon voluntary informed consent and meet or exceed the provisions of prevailing national and international guidelines.

Trade names of materials and/or products of commercial or nongovernment organizations are cited as needed for precision. These citations do not constitute official endorsement or approval of the use of such commercial materials and/or products.

ABSTRACT

The purpose of our investigation was to determine if computer-based selection tests could predict training track assignment for student naval aviators. This study evaluated the predictive efficacy of an experimental battery of computer-based pilot selection tests for training classification. Student naval aviators are currently assigned to an aircraft training track based primarily on performance in primary training. Students were tested on the experimental test battery and classified into one of three aircraft training tracks based on their test scores. The resulting classifications were compared to actual selections made as the students progressed through naval aviation training. Using a sample of 237 students, linear analyses were conducted to evaluate the efficacy of predicted decisions. The unique contribution of the experimental battery was determined by comparing scores on the experimental battery to scores on the Navy/Marine Corps Aviation Selection Test Battery, a paper-and-pencil pilot selection test used by the United States Navy and Marine Corps, and student primary flight training grades. A significant classification model including one of the experimental selection tests was derived. The model was able to significantly predict fast attack pipeline selections before flight training.

Acknowledgments

We thank Ms. Kathleen Helton and Ms. Sylvia Starling for their assistance in data collection. We also thank Ms. Kathy Mayer for her editorial support in the preparation of this report. We also wish to cite the extensive work of LCDR David Blower and Dr. Harold Delaney in the initial derivation of composite scores for the Naval Aerospace Medical Research Laboratory computer-based tests used in this investigation.

INTRODUCTION

Psychological testing has proven to be an integral component in the medical screening of potential aviators. Performance in flight training has been statistically linked to performance on a broad range of these tests (see Burke, 1993). Research has grouped these tests into five broad areas: 1) general cognitive ability, 2) information processing ability, 3) psychomotor coordination, 4) personality traits, and 5) background (Burke, 1993; Street, Helton, & Dolgin, 1992). General cognitive ability has been the most widely tested domain in actual pilot selection (Burke, 1993). For example, until 1993 the U.S. Navy (USN) and Marine Corps based the selection of physically qualified pilot candidates on the paper-and-pencil Academic Qualification Test (AQT), a test of flight-related academic abilities, and the Flight Aptitude Rating (FAR), an aptitude-related measure. A revision of the AQT/FAR, the Aviation Selection Test Battery, was implemented in 1993. Similar paper-and-pencil tests, the U.S. Army (USA) Flight Aptitude Selection Test (FAST) and the U.S. Air Force (USAF) Officer Qualifications Test (AFOQT), have been used by the Army and Air Force. These tests are generally economical and easily administered to large numbers of candidates. Even so, general cognitive ability has proven to be of somewhat more limited value than other domains in aviator selection research (Burke, 1993). Specifically, psychomotor coordination and background have historically accounted for more variance in predictions of flight training grades.

Psychomotor coordination tests have been the most robust strategies for training performance prediction. These strategies typically focus on eye-hand-foot coordination in their simplest forms, although more promising strategies have combined such skills with information processing, problem solving, and reaction time in an aircraft-like environment (Damos, 1987; Blower & Dolgin, 1991). Information processing tests measure the speed and efficiency with which an individual makes decisions about sensory information (i.e., hearing, vision, and touch) in an aircrew environment. Of particular importance are the strategies implemented by the USA and USAF. The first operational psychomotor and information-processing selection battery, the computer-based Basic Attributes Test (BAT), was implemented by the USAF in 1992 to augment selection decisions formerly based on the AFOQT. The USA implemented a similar computer-based battery to augment classification decisions student aviators selected on the basis of the FAST. These batteries and their contribution to selection and classification will be discussed in more detail later in this article.

Results have been mixed for background and personality tests, although background tests are generally considered to be the best predictor of training attrition (Hilton & Dolgin, 1991). Theoretically, background tests reflect what a person has done in the past. Presumably, tests that measure a person's knowledge and interest in aviation predict the individual's ultimate interest in an aviation career (Street & Dolgin, 1992). Research with personality tests has generally tried to identify behavioral and emotional characteristics that improve or lessen the likelihood of success in aviation. The operational contribution of personality and background tests has been questioned because of the high susceptibility of such tests to faking (Davis, 1989; Retzlaff & Gibertini, 1987). Nevertheless, the USN AQT/FAR and its revision, the ASTB, include a background questionnaire that is used to predict early training attrition. The USA and USAF no longer include background questionnaires in testing of pilot candidates.

During the late 1980s, Damos (1987) argued that pilot-selection decisions could be improved by testing a broader spectrum of the five major domains. This idea had been alluded to as early as WWI (see Parsons, 1918), although the complexity and expense of such broad-spectrum batteries was prohibitive until the advent of high-speed microcomputers. Street, Chapman, and Helton (1993) found that broad-spectrum batteries explain more variance in training outcome predictions than narrow-spectrum batteries that sample fewer domains. Recently, there has been increasing military interest in predicting not only how well a student pilot will do in training but also which airframe would best suit the individual. There is evidence that some of the same psychological tests used for selection of military aviators may also be valid for training classification (Intano, House, & Lofaro, 1991; Siem & Alley, 1992). The Army, Navy/Marine Corps (USN/USMC), and USAF have each considered a number of training performance and test battery models to make training classification decisions. Of the services, the USA is the only branch to classify student pilots based on a training/test performance model. As the Department of Defense faces increasing pressures to consolidate training resources, the selection models used by the other services should also be investigated for classification purposes.

Historically, the USN and USAF have relied on rigorous entry standards to select the most qualified student aviators and pilots. For both the USN and the USAF, classification of students into aircraft training tracks was based primarily on individual student performance and staffing requirements. The USA system for acquiring student rotary wing aviators is generally less rigorous and emphasizes classification very early in the training program. Intano et al. (1991) reported that the USA used the results of an automated battery of psychomotor, ability, and personality tests and student training performance to assign student helicopter pilots into one of four advanced tracks (UH-1, AH-1, OH-57, or UH-60). The USA classification test battery, the Multi-Track Test Battery (MTTB), was a synthesis of computer-based tests developed by the USAF, Naval Aerospace Medical Research Laboratory (NAMRL), and Army Research Institute Aviation Research and Development Activity (ARIARDA). Intano and Howse (1992) further reported that primary flight and academic grades could be predicted by various subtests of the MTTB. We were not able to find any investigations applying the MTTB to training grades in the advanced tracks (UH-1, etc.).

The USAF has also investigated the value of computer-based selection tests for classification. For example, Siem and Alley (1992) reported that modest improvements in training performance could be obtained by optimally assigning USAF pilot candidates into training tracks. They based their conclusions on ratings of 200 pilot candidate records by 57 male USAF instructor pilots. The 200 candidate records were divided into 4 groups of 50 and rank ordered (from 1 to 50) according to performance on the USAF AFOQT and the BAT. A prediction model based on the rankings was then compared to random training track assignment. The results confirmed an improvement in predicted performance. They also found that information processing accuracy, resource allocation, and psychomotor coordination were considered most important for the fighter track. In a similar study, Siem (1990) investigated whether various test scores (AFOQT and BAT) and background data could predict assignment to the USAF fighter training track. He developed a significant cross-validated regression model that accounted for 16% of the variance in fighter track prediction in a sample of 426 USAF pilot candidates. These results are consistent with USA research (especially Intano & Howse, 1992) and naval research (especially Shull & Griffin, 1990).

The determination of which track is appropriate for the student is commonly called "pipeline selection" in naval aviation. The USN and USMC divide pilot training into three primary tracks: 1) strike, 2) rotary wing, and 3) maritime patrol. The USN has one additional pipeline for the E-2 and C-2 aircraft. The strike pipeline includes fighter and attack aircraft such as the F/A-18, F-14, A-6, and EA-6B for the USN. The USMC includes the F/A-18, EA-6B, and AV-8B in the strike category. Strike pipeline students are trained for arrested carrier landings. The rotary-wing track includes all helicopters (HELO) for the USN and USMC. The maritime pipeline includes other multiple engine aircraft such as the C-9, C-130, C-12, and P-3. Finally, the E2-C2 pipeline trains for the E-2 and C-2 aircraft. This pipeline is the only nonattack category trained for arrested carrier landings and multiple engine.

Currently, pipeline selection for the USN and USMC is based primarily on student performance during primary training. Additional factors are the needs of the USN/USMC and individual student desires. Quotas for each pipeline are set annually by the USN and USMC. Students can select from the available training pipelines based on training activity grades, flight training performance, class standing, and personal preference. Traditionally, the most desirable pipeline is strike followed by rotary wing and maritime patrol, although the majority of training billets are divided between strike and rotary wing. As the most desirable pipelines are filled, fewer choices are left for lesser performers.

Few investigations of optimal pilot training assignment have been conducted in naval aviation. One (Shull & Griffin, 1990) investigated the performance of naval aviators from various aircraft communities on a battery of cognitive and performance tests. In their study, performance on a battery of computer-based cognitive/performance tests for three samples of naval aviators (F-14, $N = 66$; F/A-18, $N = 67$; and HELO, $N = 39$) was compared to the test performance of a sample of student naval aviators ($N = 177$). Analyses of variance results indicated that strike pilots (F-14 and F/A-18) made significantly fewer errors on a psychomotor tracking and dichotic listening task than the HELO pilots.

In another investigation, Shull (1991) compared the performance of USMC AV-8B, Harrier, pilots to a sample of aviation indoctrination students. In that study, the testing performance results of the sample collected in the Shull and Griffin (1990) investigation were compared to a sample of 32 USMC AV-8B pilots. The AV-8B is a jet attack aircraft capable of vertical takeoff and landing (VTOL) and very low speed/stationary hover operation. It is piloted by aviators selected from the jet-student training pipeline. Analyses of variance revealed few significant differences, although the AV-8B aviators did make more errors on the psychomotor tracking task than the F-14 and F/A-18 aviators. In addition, HELO pilots made significantly more psychomotor tracking errors than any of the fixed-wing aviators in the study. No differences in test performance were found for experience in any of the aviator communities. Based on differences in the pilot communities, the authors suggested that psychomotor coordination may have been considered in the assignment of aviators.

Other researchers have found that some aspects of naval pilot performance after leaving training could be predicted by certain tests and actual training performance. For example, Brictson, Burger, and Gallagher (1972) found a significant relationship between simulated performance (tests and actual training performance) and performance in initial carrier landing qualifications. Similarly, Griffin, Morrison, Amerson, and Hamilton (1987) found that certain psychomotor tracking and dichotic listening variables correlated significantly with some elements of air-combat maneuvering (ACM) performance in a sample of USMC F-4 pilots. Comparisons of psychomotor and cognitive tasks to performance in operational settings have been less promising. Griffin and Shull (1990) found no relationship between automated tests and performance in a F/A-18 fleet replacement squadron. In a related study, no relationship was found between ACM performance in a tactical F-14 squadron (Shull & Griffin, 1987).

In each of these investigations, cognitive and psychomotor testing has been used to predict later training track selection and performance. Only one investigation found in the literature addressed personality testing. Picano (1991) examined the relationships between personality type and aircraft assignment in a sample of 170 experienced USA pilots. He used a nonclinical measure of personality designed for use in occupational settings. There were no differences between the roughly equally distributed utility, attack, and observation groups. However, rated instructor pilots were significantly more competitive than nonrated pilots. This difference could not be explained, although similar instructor billets in the USN are highly competitive. As a result those individuals that achieve this status are likely to be highly self-selected. Of particular interest to our investigation, there were no significant personality trait differences between the three groups (i.e., utility, observation, or attack). Picano (1991) did not investigate the contribution of cognitive or psychomotor tests.

In response to ongoing reductions in operational military funding, aviation training dollars will also decrease. The effect of these reductions has already been suggested by the Center for Naval Analysis and Chief, Naval Education and Training. Specifically, programs may be shortened or cut to reduce training costs. Additionally, consolidation of USAF, USN, and USA pilot-training programs has been proposed within the Department of Defense. Of particular interest to our investigation is a proposal to consolidate rotary-wing training with one of the services. We believe that certain psychological tests designed for selection may be valid for determining the optimal training pipelines for student aviators.

Despite the lack of conclusive research findings regarding the utility of psychological tests for optimal pilot training assignment, certain psychological tests have demonstrated utility in personnel classification. Reviews of pilot selection and classification research (Hilton & Dolgin, 1991; Burke, 1993) reveal that a number of psychological tests are valid predictors of training success. Additionally, such tests may have utility for classification of student naval aviators into an optimal aircraft training assignment. Our study represents the first attempt to investigate the utility of a valid set of experimental psychological selection tests for naval aviation training classification. The hypothesis of interest was that these experimental tests would contribute to the prediction of pipeline selection decisions. Given previously cited findings linking these experimental and other tests to success in aviation training, we expected to find that a valid, broad-spectrum, computer-based selection system would nearly approximate the accuracy of pipeline classification decisions currently based on primary training performance for student naval aviators.

METHOD

SAMPLE

The subjects who participated in the current study were student naval aviators, preselected for naval flight training on the basis of their performance on the AQT/FAR. The subjects took the NAMRL computer-based psychomotor tests (CBPTs) while enrolled in an aviation indoctrination program prior to entering flight training. Their participation in this project was strictly voluntary. Before administering the tests, all subjects were informed that their decision to participate and their test results would not affect their status in the flight program and would not be entered into their service record. The testing was conducted between 1988 and 1990. The subjects were followed through primary naval flight training, and information such as flight grades, training success or failure, and pipeline assignment was entered into a computer data base. Subjects (N = 237) ranged from 21 to 28 years old (M = 23.05, SD = 1.55) when tested and included 229 males and 8 females. This sample was broken down into three subgroups based on pipeline assignment at the end of primary flight training: 1) jet (N = 75), 2) HELO (N = 91), 3) patrol (N = 71). Only subjects who had been assigned to a pipeline and succeeded in flight training were included in the analyses.

PROCEDURE

Volunteer student naval aviators were administered a battery of experimental CBPTs after selection and before entering flight training. We tracked student progress through primary flight training and compared various CBPT and AQT/FAR test scores to their aircraft training track assignments. The primary flight training syllabus takes approximately 6 months depending on student progress. The primary syllabus includes an extensive ground school and static simulation before students enter the cockpit. Primary flight training is conducted in the Beechcraft T-34C turbo-mentor. The T-34C is a turbo-prop, instrument-rated fixed-wing aircraft. Decisions regarding aircraft training assignment are made during the final stages of primary training and are based primarily on student academic and flight training grades. During the period of sample collection, students were ranked and offered an assignment ranging from jet for superior performers to HELO.

The AQT/FAR is a paper-and-pencil aptitude test battery that includes five different subtests that are divided into two pilot selection composites: a general ability estimate (AQT) and an estimate of aptitude for flying (FAR). The AQT/FAR uses a timed, multiple-choice format. The first part, the AQT, is made up of reading, arithmetic, and science questions that are related to a typical college experience. The second part, the FAR, is comprised of three subtests: 1) the Mechanical Comprehension Test (MCT), 2) the Spatial Apperception test (SAT), and 3) the Biographical Inventory (BI). The MCT is a test of mechanical reasoning and comprehension that uses pictures and word problems. The SAT is a test of spatial reasoning that uses pictures of terrain as they would be viewed from various cockpit pitch-and-roll configurations. The final component, the BI, is a test of attitudes and interests related to a military and/or aviation career. We were primarily interested in the AQT and FAR stanine composite scores, although we also included the SAT and MCT subtest scores in our analyses.

We were also interested in one CBPT developed by NAMRL (see Blower & Dolgin, 1991). Research has identified that the Psychomotor Test/ Dichotic Listening Test (PMT/DLT) is the most powerful of the NAMRL CBPTs (Delaney, 1992; Street, Chapman, & Helton, 1993). It comprises seven subtests that measure eye-hand-foot coordination, divided attention, and selective attention. An Apple IIe version of the NAMRL PMT/DLT was used in the current investigation. Additional testing station components included an Amdek Color I Plus Monitor, an Apple IIe numeric keypad, sound synthesizer card, locally produced rudder pedals, and two high-fidelity joysticks. The PMT component involves stick (S), rudder pedal (R), and throttle (T) controls that move different computer-generated cursors. Variables derived from the PMT were logarithmically transformed pixel error composite scores. The DLT was designed to measure individual differences in selective attention to different digit and letter sequences presented to each ear simultaneously. The DLT variables were error scores. A single logarithmic composite score was derived for each of the seven subtests. The subtests are arranged in ascending order of difficulty with the final subtest involving coordination of three cursors. A detailed description of the PMT/DLT subtests and variables may be found in Blower and Dolgin (1991) and Shull (1991). The presentation order, administration time, and description of the various components of the PMT/DLT are presented in Table 1.

Table 1. Sequence, Description, and Administration Time of the PMT/DLT Subtests.

Sequence	Description and variable name	Test time (min) Individual/cumulative
1.	Single psychomotor task (PMT), stick only [PMT(S)]	07 / 07
2.	Single dichotic listening task (DLT)	16 / 23
3.	First multitask [PMT(S)/DLT]	05 / 28
4.	Single (PMT), stick and rudder [PMT(S+R)]	10 / 38
5.	Second multitask [PMT1(S+R)/DLT]	05 / 43
6.	Third multitask [PMT2(S+R)/DLT]	05 / 48
7.	Single PMT; stick, rudder, and throttle [PMT(S+R+T)]	07 / 56

DATA ANALYSIS

Initially, we conducted a series of one-way analysis of variance (ANOVA) to determine the differences of sex and age among the three pipeline subgroups on various variables. Next, we conducted hierarchical and stepwise discriminant function analyses to characterize the relationship between pipeline assignment and AQT/FAR and CBPT variables. We conducted a series of stepwise analysis of the field of CBPT variables to determine which ones contributed to the prediction of pipeline assignment. Because the PMT/DLT were the only CBPT variables to remain in the equation, we retained them in the final equation. Our subjects were selected for initial flight training based on their AQT/FAR test scores, so those scores were entered into the discriminant function before the PMT/DLT variables. This procedure allowed us to estimate the amount of variance unique to the individual CBPT variables after the AQT/FAR variables. Finally, we developed a model to predict jet versus other pipeline assignment (HELO and patrol). We conducted the same discriminant function analyses described for the three-subgroup model.

RESULTS

DESCRIPTIVE STATISTICS

Descriptive statistics of the test performance for the three subgroups of student naval aviators (SNA) are presented in Table 2. Increasing values for the AQT and FAR indicate decreased errors and superior performance, while increasing values for the various PMT/DLT composite variables indicate increased errors and poorer test performance. As described earlier, the SNAs were divided into jet, HELO, and patrol subgroups according to what pipeline they were assigned to at the end of primary flight training. One-way ANOVAs showed no significant differences for sex on any of the test variables. In addition, there were no significant differences among the SNA groups for age.

Table 2. Descriptive Statistics of the Variables for the Student Aviator Subgroups [Mean (*SD*)]. *

Variable	Jet ($N = 75$)	HELO ($N = 91$)	Patrol ($N = 71$)
AQT	5.95 (1.33)	5.71 (1.30)	5.85 (1.24)
FAR	7.81 (1.36)	6.93 (1.77)	6.79 (1.68)
PMT(S)	4.02 (0.22)	4.18 (0.29)	4.19 (0.29)
DLT	0.74 (0.24)	0.81 (0.25)	0.78 (0.25)
DLT/PMT(S)	0.87 (0.29)	0.95 (0.29)	0.93 (0.36)
PMT(S)/DLT	3.55 (0.23)	3.74 (0.29)	3.75 (0.32)
PMT1(S+R)	4.53 (0.15)	4.66 (0.22)	4.65 (0.23)
PMT1&2(S+R) Mean	3.94 (0.22)	4.09 (0.26)	4.07 (0.23)
DLT/PMT(S+R)	0.90 (0.26)	0.99 (0.24)	0.97 (0.31)
PMT(S+R+T)	4.13 (0.15)	4.28 (0.19)	4.21 (0.19)

* Increasing magnitude on the PMT/DLT variables indicates increased tracking and listening response error. The FAR and AQT are stanine scores and increasing magnitude indicates superior performance.

We found significant differences among the SNA groups for many of the test variables using one-way ANOVAs and the Scheffe post-hoc test. Table 3 presents the results for the one-way ANOVAs for the tests among the SNA subgroups. As shown in Table 3, the jet pipeline SNAs performed significantly better than those in the HELO and patrol pipelines on the FAR. The HELO and patrol pipeline students also made significantly more errors on many of the PMT/DLT tests. Interestingly, only the tracking performance variables were significantly different for the groups. The DLT component was not significantly different for any of the three groups. Finally, there were no significant differences between the HELO and patrol pipeline students on any of the variables.

Table 3. ANOVAs and Between-Group Comparisons for SNA Pipeline Subgroups.

Variable	*F*	*(DF)*	*p*	Significant pairwise
AQT	0.67	(2, 234)	*	
FAR	8.81	(2, 234)	.0002	jet > HELO, patrol
PMT(S)	9.02	(2, 234)	.0001	jet < HELO, patrol
DLT	1.44	(2, 234)	*	
DLT/PMT(S)	1.60	(2, 234)	*	
PMT(S)/DLT	12.06	(2, 234)	.00001	jet < HELO, patrol
PMT1(S+R)	9.97	(2, 232)	.00006	jet < HELO, patrol
PMT1&2(S+R) Mean	8.75	(2, 233)	.0002	jet < HELO, patrol
DLT/PMT(S+R)	2.79	(2, 234)	*	
PMT(S+R+T)	12.78	(2, 234)	.00001	jet < HELO, patrol

* not significant

Next, we conducted canonical discriminant function (DF) analyses to assess the prediction of membership in the jet, HELO, and patrol SNA subgroups. Because the students were selected initially on the basis of the AQT and FAR, we retained only those variables in the equation. A significant (DF) was calculated, with a combined χ^2 (2) = 16.46, $p < .002$. This indicates that there was a statistical difference between the means of the three SNA subgroups. The Pearson correlation coefficient was $R = .26$. With the probability for assignment set at the proportions for final actual assignment, the DF explained 6.8% of the total variance and accurately classified 44.7% of the cases. Table 4 presents the classification matrix for the AQT/FAR model.

When we entered the additional PMT/DLT variables, the DF was also significant (χ^2 (6) = 37.148, $p <$ .0002). A Pearson correlation coefficient of .38 was obtained with this enhanced model. Using the same proportional probability for actual group assignment, the DF explained 14.4% of the total variance and accurately classified 45.7% of the cases. Table 5 presents the DF for the enhanced model.

Table 4. Aviation Qualification Test/Flight Aptitude Rating Classification Matrix.*

	Predicted Group Membership			
Actual Group	Jet	HELO	Patrol	Percent Correct
Jet	46	29	0	61.3
HELO	33	57	1	62.6
Patrol	22	46	3	4.2

* Percent of grouped cases correctly classified: 44.7%

Table 5. Enhanced PMT/DLT Model Classification Matrix.*

Actual Group	Predicted Group Membership: Jet	HELO	Patrol	Percent Correct
Jet	45	26	4	60.0
HELO	26	57	6	64.0
Patrol	24	41	5	7.1

* Percent of grouped cases correctly classified: 45.7%

Because the HELO and patrol pipelines were not significantly different on any of the PMT/DLT variables, we conducted one additional set of analyses combining those two subgroups. In the resulting analyses, we attempted to discriminate the jet from the combined HELO/patrol subgroup. Using the same procedures outlined for the three-subgroup model, the DF for the AQT and FAR was significant (χ^2 (2) = 16.367, $p <$ .0002). The Pearson correlation coefficient (R = .26) indicates that the model explained 6.7% of the variance. As seen in Table 6, with the combined HELO/patrol subgroup, the AQT and FAR model achieved a higher number of accurate predictions than the three-subgroup model did. Finally, we added the PMT/DLT variables identified in our earlier ANOVAs to the discriminant function analyses. This enhanced PMT/DLT model (Table 7) achieved significance (χ^2 (6) = 35.354, $p <$ .00001) and accounted for approximately 14.4% of the total variance (R = .38). Of particular interest, the enhanced model accurately classified 36% of the jet pipeline subgroup compared to 4% for the AQT/FAR model. Further review of the information presented in Tables 6 and 7 indicates that the enhanced PMT/DLT model achieved the largest percentage of correct classifications.

Table 6. Classification Matrix for AQT/FAR Model with HELO and Patrol Subgroups Combined.*

Actual Group	Predicted Group Membership: Jet	HELO/Patrol	Percent Correct
Jet	3	72	4.0
HELO/Patrol	2	160	98.8

* Percent of grouped cases correctly classified: 68.8%

Table 7. Classification Matrix for Enhanced PMT/DLT Model with HELO and Patrol Subgroups Combined.*

	Predicted Group Membership		
Actual Group	Jet	HELO/Patrol	Percent Correct
Jet	27	48	36.0
HELO/Patrol	20	142	87.7

* Percent of grouped cases correctly classified: 71.3%

DISCUSSION

We have demonstrated that a sample of student naval aviators who enter the jet pipeline make significantly fewer tracking errors on certain computer-based tests than students who enter HELO or maritime training. They also perform significantly better on paper-and-pencil tests of flight aptitude. Furthermore, we found that these same psychological tests can improve predictions of which aircraft training pipeline student naval aviators select. Specifically, we found that the FAR has validity as a pipeline predictor, while component variables of the PMT/DLT demonstrated incremental validity. Our results indicated that of the pilots predicted for membership in one of the three training pipelines, approximately 71.3% were actually selected based on an enhanced testing model. It is important to note that this level of accuracy was achieved before the students had even entered ground school. Related studies have demonstrated that these same tests can improve predictions of performance in aviation training. Taken together, these results suggest that optimal pipeline assignment based on selection tests would result in improvements in training assignments and actual training performance. The increases in performance may be modest. However, Cascio (1991) cited that even small improvements in performance generally result in cost savings to organizations. These cost savings could be substantial for the USN and USMC.

We believe that the results of this investigation lend further empirical support for a set of valid psychological tests in predicting the flight training performance of student naval aviators. The results are also consistent with research conducted by the U.S. Army and indicate that the computer-based PMT/DLT can be used to optimally assign students to rotary- or fixed-wing pilot training. This was expected. As indicated earlier in this paper, the U.S. Army incorporated the computer-based PMT/DLT developed at NAMRL into the MTTB in the late 1980s. Even though the test software driver was modified to meet hardware needs, it is essentially the same as the current NAMRL PMT/DLT. The U.S. Army MTTB PMT/DLT has become the premier test for optimal training assignment of U.S. Army rotary-wing pilots.

Additionally, the results are consistent with U.S. Air Force research with the BAT. Valid, broad-spectrum computer-based tests can improve the prediction of flight-training performance. They can also optimize training assignments. This is particularly useful for the naval services in the prediction of whether students should be considered as candidates for the jet pipeline. As the naval services are faced with increasing pressure to reduce costs and consolidate training resources with the other services, a valid computer-based selection system would improve selection and classification decisions.

REFERENCES

Blower, D. J., & Dolgin, D. L. (1991). *An evaluation of performance-based tests designed to improve naval aviation selection* (Report No. 1363). Pensacola, FL: NAMRL-Naval Aerospace Medical Research Laboratory.

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

Burke, E. (1993). Meta analysis of aircraft pilot selection procedures. In *Proceedings of the 7th International Symposium on Aviation Psychology* (pp. 413-418). Columbus, OH: Association of Aviation Psychologists.

Cascio, W. (1991). *Costing human resources.* Boston, MA: PWS-Kent Publishing Co.

Damos, D. (1987). *Some considerations in the design of a computerized human information processing battery* (Monograph No. 35). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Davis, R. D. (1989). *Personality: Its use in selecting candidates for US Air Force undergraduate pilot training* (Report No. 88-8). Maxwell Air Force Base, AL: Airpower Research Institute.

Delaney, H. (1992). Dichotic listening and psychomotor task performance as predictors of naval primary flight-training criteria. *International Journal of Aviation Psychology, 2(2),* 107-120.

Griffin, G. R., Morrison, T. R., Amerson, T. L., & Hamilton, P. V. (1987). *Predicting air combat maneuvering (ACM) performance: Fleet fighter ACM readiness program grades as performance criteria* (Report No. 1333). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Griffin, G. R., & Shull, R. N. (1990). *Predicting F/A-18 fleet replacement squadron performance using an automated battery of performance-based tests* (Report No. 1354). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Hilton, T. R., & Dolgin, D. L. (1991). Pilot selection in the military of the free world. In R. Gal & A. Mangelsdorff (Eds.), *Handbook of military psychology* (pp.88-101). Sussex, England: John Wiley and Sons.

Intano, G. P., & Howse, W. R. (1992). Predicting performance in Army aviation flight training. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 907-911). Atlanta, GA: Human Factors Society.

Intano, G. P., Howse, W. R., & Lofaro, R. J. (1991). *Initial validation of the Army aviator classification process* (Report No. 91-38). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Parsons, R. P. (1918). A search for nonphysical standards for naval aviators. *U.S. Naval Medical Services Bulletin, 12,* 155-172.

Picano, J. J. (1991). Personality types among experienced military pilots. *Aviation, Space, and Environmental Medicine, 62,* 517-520.

Retzlaff, P. D., & Gibertini, M. (1987). Air Force pilot personality: Hard data on the "right stuff." *Multivariate Behavioral Research, 22,* 383-399.

Shull, R. N. (1991). *Performance of Marine AV-8B (Harrier) pilots on a cognitive/psychomotor test battery: Comparison and prediction* (Report No. 1362). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Shull, R. N., & Griffin, G. R. (1987). *Predicting F-14 air combat maneuvering (ACM) performance using an automated battery of cognitive/psychomotor tests* (Report No. 1356). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Shull, R. N., & Griffin, G. R. (1990). *Performance of several different naval aviator communities on a cognitive/psychomotor test battery: Pipeline comparison and prediction* (Report No. 1361). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Siem, F. M. (1990). Development of a selection model for fighter-qualified UASF pilot candidates. Paper presented at the Annual Convention of the American Psychological Association, Boston, MA: American Psychological Association.

Siem, F.M., & Alley, W.E. (1992). Optimal personnel assignment: An application to Air Force pilots. Poster session at the 7th Annual Conference of the Society for Industrial and Organizational Psychology, Washington, D. C.: Society for Industrial and Organizational Psychology.

Street, D.R., Chapman, A.E., & Helton, K.T. (1993). The future of naval aviation selection: Broad-spectrum computer-based testing. In *Proceedings of the 35th Annual Military Testing Association* (in press). Williamsburg, VA: Military Testing Association.

Street, D.R. & Dolgin, D.L. (1992). *The efficacy of biographical inventory data in predicting early attrition in naval aviation officer candidate training* (Report No. 1373). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Street, D., Helton, K., & Dolgin, D. (1992). *The unique contribution of selected personality tests to the prediction of success in naval pilot training* (Report No. 1374). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Other Related NAMRL Publications

Dolgin, D. L., Street, D. R., Blower, D. J., Nontasak, T., & Travis, K. (1992). Operational implementation of a validated personnel selection system for landing craft air cushion (LCAC) vehicle operators. In *Proceedings of the 13th Biennial Psychology in the Department of Defense Symposium* (pp. 66-70). Denver, CO: United States Air Force Academy.

Griffin, G. R. (1987). *Development and evaluation of an automated series of single and multiple dichotic listening and psychomotor tasks* (Technical Report No. 1336). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Dolgin, D. L., & Gibb, G. D. (1988). *A review of personality measurement in aircrew selection* (Monograph No. 36). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Dolgin, D. L., & Gibb, G. D. (1989). Personality assessment in aviator selection: Past, present and future. In R. Jensen (Ed.), *Aviation psychology* (pp. 285-319). London: Gower Publishing Group.

Nontasak, T., Dolgin, D. L., & Griffin, G. R. (1989). Performance-based tests, personality attributes, and training outcome among landing craft air cushion (LCAC) vehicle operators. In *Proceedings of the 33rd Annual Meeting of the Human Factors Society* (pp.901-904). Denver, CO: Human Factors Society.

Shull, R. N., Dolgin, D. L., & Gibb, G. D. (1988). *The relationship between flight training performance, a risk assessment test, and the Jenkins Activity Survey* (Report No. 1339). Pensacola, FL: Naval Aeropsace Medical Research Laboratory.

Street, D.R., Dolgin, D.L., & Helton, K.T. (1993). Personality tests in an enhanced pilot selection model. In *Proceedings of the 7th International Symposium on Aviation Psychology* (pp. 428-433). Columbus, OH: Military Testing Association.

Street, D. R., Helton, K. T., & Nontasak, T. (1993). *An evaluation of personality testing and the five-factor model in the selection of landing craft air cushion vehicle crew members* (Report No. 1385). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. [illegible]

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE February 1994	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
Computer-Based Psychomotor Tests in Optimal Training Track Assignment of Student Naval Aviators

5. FUNDING NUMBERS
63706N
M0096.002-
M00960.01

6. AUTHOR(S)
D. R. Street, Jr., and D. L. Dolgin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NAVAEROMEDRSCHLAB
51 HOVEY ROAD
PENSACOLA FL 32508-1046

8. PERFORMING ORGANIZATION REPORT NUMBER
NAMRL-1391

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Medical Research and Development Command
National Naval Medical Center, Bldg. 1
Bethesda, MD 20889-5606

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

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

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This study evaluated the predictive efficacy of an experimental battery of computer-based pilot selection tests for training classification. Student naval aviators are currently assigned to an aircraft training track based primarily on performance in primary training. Students were tested on the experimental test battery and classified into one of three aircraft training tracks based on their test scores. The resulting classifications were compared to actual selections made as the students progressed through naval aviation training. Using a sample of 237, linear analyses were conducted to evaluate the efficacy of predicted decisions. The unique contribution of the experimental battery was determined by comparing scores on the experimental battery to scores on the Navy/Marine Corps Aviation Selection Test Battery, a paper-and-pencil pilot selection test used by the United States Navy and Marine Corps, and student primary flight training grades. A significant classification model including one of the experimental selection tests was derived. The model was able to significantly predict fast attack pipeline selections before flight training.

14. SUBJECT TERMS
Computer-based tests, Pilot selection, Student pilot classification, Selection testing, Psychomotor

15. NUMBER OF PAGES 17

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)