(2)

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
NAVAL AIR STATION, PENSACOLA, FLORIDA 32508-5700

NAMRL-1362

PERFORMANCE OF MARINE AV-8B (HARRIER) PILOTS ON A COGNITIVE/ PSYCHOMOTOR TEST BATTERY: COMPARISON AND PREDICTION

R.N. Shull

Reviewed and approved 31 December 1990

J. A. BRADY, CAPT, MSC, USN
Commanding Officer

This research was sponsored by the Naval Medical Research and Development Command under work unit 63706N M0096 M00960.01 0151.

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.

Trade names of materials and/or products of commercial or non-government 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.

Volunteer subjects were recruited, evaluated, and employed in accordance with the procedures specified in 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.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE December 1990

3. REPORT TYPE AND DATES COVERED Interim Jan 90 - Jul 90

4. TITLE AND SUBTITLE
Performance of Marine AV-8B (Harrier) Pilots on a Cognitive/Psychomotor Test Battery: Comparison and Prediction

5. FUNDING NUMBERS
63706N
M0096.001 0151

6. AUTHOR(S)
R. N. Shull

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Aerospace Medical Research Laboratory
Bldg. 1953, Naval Air Station
Pensacola, FL 32508-5700

8. PERFORMING ORGANIZATION REPORT NUMBER
NAMRL-1362

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Medical Research and Development Command
National Naval Medical Center, Bldg. 1
Bethesda, MD 20814-5044

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)*
Several studies have suggested the possibility of predicting pilot flight performance in various aviation environments. The current report concerns the use of an automated performance-based test battery, measuring various aspects of cognitive and psychomotor function, to eventually predict the flight performance of Marine jet pilots during various stages of flight training while comparing their test results to those of other aviation communities. A group of Marine jet pilots assigned to operate the AV-8B vertical takeoff/landing attack jet were tested. No significant differences in test performance were found among the various subgroups of student and experienced AV-8B pilots tested. The subject group, as a whole, performed most like pilots of other types of jet aircraft on some tests and more like helicopter pilots on other tests. No significant statistical relationships were found between any of the test battery measures and the final flight grades for primary, intermediate, or advanced flight training. Further research of a longitudinal nature, using a larger subject sample, is needed to fully assess the predictive ability of these tests in regards to flight performance measures during stages of training beyond primary flight training.

14. SUBJECT TERMS
Pilot Selection, Psychomotor Ability, Cognition, Dichotic Listening, Marine Corps Aviation, AV-8B, Jets, Pilot Training, Pipelining, Helicopters, Student Pilots

15. NUMBER OF PAGES 12

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)
Prescribed by ANSI Std Z39-18
298-102

SUMMARY PAGE

THE PROBLEM

Several studies have suggested the possibility of predicting flight performance of pilots in various aviation environments. Thus, research is being conducted to develop reliable predictor tests that might aid in decisions concerning aircrew selection, training pipeline assignment, and posttraining aircraft assignment. The current approach uses an automated test battery to measure aspects of cognitive and psychomotor function, to eventually predict the flight performance of Marine jet pilots during various stages of flight training while comparing their test battery results to those of other aviation communities. A group of Marine jet pilots assigned to operate the AV-8B vertical take-off/landing attack jet were tested on this battery.

FINDINGS

No significant differences in test battery performance were found among the subgroups of student and experienced AV-8B pilots tested. As a whole, the subject group performed most like pilots of other types of jet aircraft on some of these tests and more like helicopter pilots on other tests. This was interesting given the similar flight characteristics of the AV-8B to both jets and helicopters. No significant statistical relationships were found between any of the test battery measures and the final primary, intermediate, and advanced flight grades that were obtained for the majority of these subjects.

RECOMMENDATIONS

With continued research on the psychophysiological processes of a successful pilot, some elements of this test battery might eventually prove useful in predicting certain aspects of flight performance thought necessary and perhaps unique to a particular type of aircraft. Definitive changes in test structure, equipment, and procedures are probably needed to achieve the greatest benefit from this battery. Further research is needed to fully assess the predictive ability of these tests for flight performance during stages of training beyond primary flight training.

Acknowledgments

The author gratefully acknowledges the technical assistance of Peter Cullyer in programming and modifying the cognitive/psychomotor test battery and Solomon Eagles in equipment setup and subject testing, as well as both the commanding officer, (Marine Air Group 32) COL J.R. Dempsey, and his aeromedical safety officer, LT Barbara Boyd, MSC, USN, for assisting us in every aspect of the field data collection of this project.

INTRODUCTION

Research is being performed at the Naval Aerospace Medical Research Laboratory (NAMRL) to develop measures of cognitive and psychomotor ability that would reliably relate to the actual flight performance of naval aviators. The goal is the eventual development of a test battery that would predict the operational performance of naval aviators before fleet aircraft assignment. Such efforts would aid in the identification of unique selection criteria for specific fleet aviation communities and thus support training platform assignment (pipeline) decisions.

Some attempts to predict certain measures of operational aviation performance have been successful. For example, peer ratings from preflight training helped identify successful and unsuccessful naval aviators during combat in Vietnam (1). A combination of psychological tests and actual flight performance measures was used to successfully predict F-4 carrier landing performance (2). An overall experience measure and seven undergraduate training grades were reliably related to the overall training grade of an F-4 Replacement Air Group (RAG) (3). More recently, a subset of the automated dichotic listening and cursor tracking test results correlated significantly with some elements of the Air Combat Maneuvering (ACM) performance of a group of Marine F-4 pilots (4).

Unfortunately, current attempts to use an automated battery of cognitive and psychomotor tests to predict aviator performance in specific operational settings have been unsuccessful. Research in this area has demonstrated no significant relationships between performance on these tests and either Fleet Replacement Squadron (FRS) (5) or ACM (6) performance in tactical jet aircraft. This could have been due to a number of factors, including the equalization of subject ability within a specific pilot group due to common selection, training, and flight experiences. Significant differences in test performance have been shown, however, between jet and helicopter pilots (7), as well as significant correlations between performance on many of these tests and the final overall grade in primary flight training for student naval aviators (SNAs) (8). Based on these findings, some of these tests may be measuring an innate ability, which remains relatively constant throughout a pilot's career when compared to other pilots with similar levels of experience (7).

In this study, Marine Corps AV-8B (Harrier) pilots were administered the same cognitive/psychomotor test battery utilized before (5-8). The Harrier can maneuver vertical takeoffs and landings and has flight characteristics similar to that of jets and helicopters. The scores obtained were compared both among the subgroups found within this pilot subject pool and to the different pilot types investigated previously (7). An attempt was also made to find significant relationships between test performance and the final overall grades from primary, intermediate, and advanced flight training for the majority of AV-8B pilots tested.

METHODS

SUBJECTS

Subjects were 32, male, Marine Corps jet pilots assigned to AV-8B squadrons stationed at MCAS Cherry Point, North Carolina. Their age was between 24 and 39 years (M = 28.91, SD = 3.65), and the total number of military flight hours per subject ranged from 300 to 2350 (M = 1013.28, SD = 740.50).

APPARATUS AND PROCEDURES

Eleven tests comprised the battery, which is summarized in Table 1. Subjects completed the battery during normal working hours inside a mobile field laboratory within walking distance from the various squadrons. The entire series was automated using an Apple IIe microcomputer, an Amdek Color I Plus monitor (CRT), and an Apple IIe numeric keypad. Subjects received test instructions on the CRT before beginning each test. The tests described in the following sections are presented in further detail elsewhere (5).

TABLE 1. Sequence, Description, and Operating Times of Automated Tests.

Presentation order	Description	Test times (min) individual/cumulative
1.	Single psychomotor task (PMT), stick only (S)	07 / 07
2.	Single dichotic listening task (DLT)	16 / 23
3.	First multitask (1,2 combined)	05 / 28
4.	Single (PMT), stick and rudder (S&R)	10 / 38
5.	Second multitask (4,2 combined)	05 / 43
6.	Third multitask (4,2 combined)	05 / 48
7.	Single PMT; stick, rudder, and throttle (S&R&T)	07 / 56
8.	Second single PMT (like 7, S&R&T)	04 / 60
9.	One-dimensional compensatory tracking (ODCT)	10 / 70
10.	Absolute difference computation (ADC)	10 / 80
11.	Fourth multitask, ODCT and ADC (10,11 combined)	10 / 90

PSYCHOMOTOR TASK (PMT)

The psychomotor tracking task required subjects to maintain first one, then two, and finally three randomly displaced cursors on fixed targets on the CRT by manipulating joysticks and foot pedals. Subjects manipulated a Measurement Systems, Inc., joystick (stick or S) at the front seat edge with their right hand to control a cursor that moved within the upper two-thirds of the screen just right of center in a backwards (reversed) manner. Locally produced rudder pedals (rudder or R), patterned after those of a Systems Research Laboratories, Inc., psychomotor test device, were used by subjects to control a second cursor that moved horizontally across the bottom of the screen. Pushing the left pedal moved this cursor to the right; pushing the right pedal moved it to the left. Another Measurement Systems joystick (throttle or T), located on the left seat edge, was manipulated by the subject's left hand to move a third cursor vertically on the left side of the screen. The subject pulled this throttle back to move this cursor down and vice versa.

Psychomotor task tests 1, 4, and 7 (see Table 1) were each preceded by a 3-min practice period. Test 4 was divided into two 3-min testing sessions separated by a 20-s rest interval. Psychomotor task test scores were the accumulated total of absolute errors from an ideal target position. For each time-sampling of cursor position, absolute pixel errors were assessed separately along each dimension. Final error score was the sum of all the samplings across all dimensions represented in that particular task. This error score total was then divided by the number of minutes of each test analyzed to generate a standard rate of pixel error per 1 min of test time. The scores of tests 5 and 6 and tests 7 and 8 were averaged for each subject. All PMT error scores from these tests were then transformed by logarithms to base 10 to reduce skewness and to compensate for extreme outliers, thus reducing the complexity of data analyses while retaining all the data points available (8).

DICHOTIC LISTENING TASK (DLT)

The DLT consisted of a series of letter/digit string sets presented to subjects aurally over binaural headphones via two Jameco JE 520-AP voice synthesizers. Subjects were told which ear to attend to for each trial. Part I was a series of 16 pairs of letters and/or numbers; Part II was a series of 6 more pairs. Subjects were to indicate the digits (0-9) presented to the designated ear in the order of their occurrence. Subjects responded with their left hand using a keypad placed immediately in front and slightly left of center. The test was preceded by six aural practice trials, which provided immediate performance feedback by visually indicating the letters and digits presented and the subjects' keypad responses. Before beginning the actual test, subjects also completed three multiple-choice questions to ensure that they understood the concept of the DLT.

The DLT performance measure was the number of incorrect responses during 12 trials in which a total of 108 correct responses were possible. The number of correct responses made was subtracted from the total possible correct for that particular test. After adding one, the adjusted error score was transformed by logarithms to base 10 to adjust for both skewness and extreme outliers as with the PMT (8).

MULTITASK PMT/DLT

In all of the multitask conditions, subjects performed both the DLT and PMT simultaneously (a 12-trial DLT and a 4.5-min PMT). During the first multitask condition, subjects performed the DLT and the stick-only PMT(S). For the second and third multitask conditions, subjects performed the DLT and the stick-and-rudder PMT(SR) using their right hand and feet to control the central joystick and the rudder pedals, and their left hand to make keypad responses to the DLT input. Before beginning the multitask combinations, subjects were instructed to perform each task equally well. Performance measures for the PMT and DLT in these multitask conditions were identical to those of the single tasks, with PMT errors recorded for the final 4 min of that test.

ONE-DIMENSIONAL COMPENSATORY TRACKING (ODCT)

For the ODCT, subjects were to center a square cursor within an elongated rectangle. Subjects used their right hand to move a joystick left and right that was centered on the front seat edge. The cursor was driven by a forcing function that increased centering effort with distance from center. During this phase of the task, subjects received three 2-min trials separated by two 30-s rest periods. The test measure for the ODCT was total pixel deviation error averaged over the three single-task trials.

ABSOLUTE DIFFERENCE COMPUTATION (ADC)

Randomly selected digits between 1 and 9 were presented inside a small square in the middle of the CRT to subjects. Subjects determined the absolute difference between the digit currently displayed on the CRT and the digit previously displayed. The subjects then pressed the corresponding digit-key on the keypad with their left hand as quickly as possible, resulting in the display of another number for computation. Identical digits were not allowed to repeat. Only the digit responses 1, 2, 3, and 4 were possible. Subjects received three 2-min trials separated by two 20-s rest periods. Performance measures for the ADC were the number of correct responses, the number of incorrect responses, and the ratio of correct to incorrect responses, all averaged over the three ADC trials.

DUAL TASK ODCT/ADC

For the dual task ODCT/ADC, subjects performed both the ODCT and the ADC concurrently. The digits for the difference task were centered just above the tracking task. The subjects controlled the tracking task joystick with their right hand and made keypad responses to the difference task with their left hand. Subjects were instructed to perform each task equally well. Subjects received three 2-min trials separated by two 30-s rest periods. Test measures for the dual-task ODCT/ADC were the same as those for the single tasks.

RESULTS AND DISCUSSION

AV-8B SUBGROUP TEST PERFORMANCE

Table 2 presents descriptive statistics of the test performance of these AV-8B pilots after separation into subgroups according to position within the airwing. Student (stdt) pilots were divided into two subgroups from separate squadrons based on flight time in the Harrier: those less than 35 h (< 35 h) and more than 100 h (> 100 h). Experienced Harrier pilots were also divided into two groups: those who instructed the student pilots during initial flight training (instruc) and those who were assigned to a separate fleet squadron (fleet) where many of the students would be assigned after training. All the pilots tested had already completed all stages of jet pipeline training before being assigned to the Harrier community. Table 2 also lists the age (years), total military flight time (h), and total flight time in the AV-8B (h) of each subgroup as provided by the subjects.

One-way analysis of variance (ANOVA) showed no significant differences among the four pilot subgroups on any of the tests. The subgroups did differ significantly, however, in terms of age [$F(3, 28) = 11.34$, $p < .0002$], total military flight hours [$F(3, 28) = 16.26$, $p < .0001$], and total AV-8B flight hours [$F(3, 28) = 24.42$, $p < .0001$]. Using the Scheffe post-hoc comparison test (9), the experienced pilots proved to be significantly older with significantly more flight time in both military aircraft in general and the AV-8B in particular than the student pilots. The two experienced subgroups did not differ significantly from the two student subgroups on any of these three nontest variables.

TABLE 2. Descriptive Statistics of Tests for the AV-8B Subgroups [Mean (SD)].

Test measure	Stdt (<35 h) (n = 9)	Stdt (>100 h) (n = 8)	Instruc (n = 8)	Fleet (n = 7)
DLT alone	0.80 (0.44)	0.84 (0.26)	0.95 (0.27)	0.99 (0.21)
DLT w/PMT(S)	0.83 (0.28)	1.04 (0.20)	1.06 (0.16)	1.01 (0.22)
DLT w/PMT(SR)	0.86 (0.23)	0.92 (0.22)	0.97 (0.08)	0.92 (0.26)
PMT(S) alone	3.10 (0.19)	3.09 (0.12)	2.98 (0.13)	3.09 (0.08)
PMT(S) w/DLT	2.93 (0.30)	2.84 (0.12)	2.75 (0.11)	2.89 (0.12)
PMT(SR) alone	3.54 (0.14)	3.43 (0.11)	3.39 (0.08)	3.47 (0.12)
PMT(SR) w/DLT	3.34 (0.23)	3.18 (0.09)	3.23 (0.18)	3.21 (0.12)
PMT(SRD) alone	3.75 (0.19)	3.54 (0.10)	3.52 (0.08)	3.51 (0.06)
Sgl ODCT	23.35 (8.98)	21.81 (2.52)	22.05 (7.02)	25.43 (8.05)
Sgl ADC CR	56.93 (14.12)	59.79 (14.61)	51.78 (7.60)	53.00 (13.63)
Sgl ADC IR	3.93 (2.41)	4.88 (3.07)	5.06 (2.04)	4.67 (3.48)
Sgl ADC CR/IR	25.59 (28.78)	13.51 (8.98)	11.35 (3.39)	39.78 (62.68)
Dual ODCT	38.91 (23.53)	32.23 (12.17)	30.39 (10.37)	38.27 (13.35)
Dual ADC CR	60.56 (17.55)	62.50 (15.88)	57.33 (14.03)	51.83 (9.69)
Dual ADC IR	3.70 (1.12)	4.33 (2.84)	4.52 (2.96)	5.39 (4.81)
Dual ADC CR/IR	18.30 (8.53)	21.70 (14.84)	32.61 (48.08)	27.46 (26.47)
Age (years)	26.33 (2.00)	26.75 (1.58)	30.50 (2.39)	32.86 (4.02)
Mil flt hrs	549.44(675.88)	401.25 (41.21)	1468.75(311.61)	1788.57(555.65)
AV-8B flt hrs	14.00 (12.01)	155.00 (33.38)	868.75(327.26)	981.43(490.93)

PILOT GROUP TEST PERFORMANCE

Performance of the AV-8B pilots was compared to that of three other trained pilot groups previously tested on this battery (7). The three groups were 66 F-14 pilots completing air combat training, 67 jet pilots completing

F/A-18 fleet replacement training, and 39 helicopter instructor (HELO) pilots. Table 3 presents descriptive statistics of the test performance of all four pilot groups. Testing equipment and protocols were virtually identical for all subject groups.

Regardless of motor complexity level, the mean number of errors on the PMT decreased significantly for all groups when the DLT was brought on-line (all t values > 4.08, all n values > 30, all p values < .01). This phenomenon has been observed before (5-7) and is attributed to processor overload, which causes a small, stable reduction in both cursor positioning difficulty and error sampling rate during the PMT/DLT multitask conditions. By Friedman two-way ANOVAs (10), all groups made significantly more errors as PMT complexity increased during both the unitask and multitask conditions (all ANOVA chi-square values > 24.64, all p values < .01). This decrease in testing efficiency does not reduce the usefulness of the testing methodology but does warrant a possible equipment upgrade.

TABLE 3. Descriptive Statistics of Tests for Pilot Groups [Mean (SD)].

Test measure	AV-8B (n = 32)		F-14* (n = 66)		F/A-18* (n = 67)		HELO* (n = 39)	
DLT alone	0.89	(0.31)	0.72	(0.34)	0.71	(0.23)	0.85	(0.26)
DLT w/PMT(S)	0.98	(0.23)	0.84	(0.34)	0.65	(0.37)	1.04	(0.30)
DLT w/PMT(SR)	0.91	(0.21)	0.81	(0.24)	0.74	(0.29)	0.92	(0.26)
PMT(S) alone	3.07	(0.14)	3.03	(0.20)	3.03	(0.13)	3.17	(0.32)
PMT(S) w/DLT	2.85	(0.19)	2.79	(0.15)	2.74	(0.15)	2.97	(0.25)
PMT(SR) alone	3.46	(0.12)	3.43	(0.13)	3.39	(0.12)	3.45	(0.18)
PMT(SR) w/DLT	3.24	(0.17)	3.16	(0.14)	3.14	(0.17)	3.23	(0.20)
PMT(SRT) alone	3.56	(0.13)	3.59	(0.13)	3.56	(0.16)	3.60	(0.16)
Sgl ODCT	23.10	(6.92)	19.31	(7.76)	23.08	(6.79)	29.61	(9.04)
Sgl ADC CR	55.84	(12.80)	58.63	(15.63)	56.73	(13.16)	45.60	(16.16)
Sgl ADC IR	4.57	(2.67)	6.15	(5.54)	----	----	7.40	(12.36)
Sgl ADC CR/IR	22.56	(33.33)	18.86	(18.64)	----	----	12.28	(8.45)
Dual ODCT	34.97	(15.79)	29.28	(11.85)	33.90	(11.41)	42.96	(13.43)
Dual ADC CR	58.58	(14.77)	62.68	(15.45)	60.94	(13.32)	53.37	(15.96)
Dual ADC IR	4.40	(2.91)	6.98	(9.20)	----	----	4.99	(2.85)
Dual ADC CR/IR	24.38	(26.54)	25.65	(36.88)	----	----	16.08	(13.57)

* Data from previous study (7).

Significant differences were found among the groups for many of the tests using one-way ANOVAs and the Scheffe post-hoc test. Table 4 presents the results of statistical analyses for those tests where significant differences were found. Of primary importance is the relationship of the AV-8B group to the other groups in terms of test performance. For the other three groups, the present results agree with previous findings (7).

The AV-8B pilots tested received the same jet pipeline training prior to Harrier assignment as the F-14 and F/A-18 pilots tested. Even so, the AV-8B group did not match the performance of the other jet groups, especially the F/A-18 group, on tests involving the PMT and/or DLT, although they did perform as well on tests involving the ODCT and/or the ADC. The greatest differences were seen with the DLT in which the AV-8B group appeared to be more like the HELO group than the other jet groups. No definitive explanation for these differences is available at this time. Such performance mismatches could have been partially due to differences in pilot selection procedures beyond jet pipeline training between the Navy (F-14/F/A-18) and the Marine Corps (AV-8B).

TABLE 4. ANOVAs and Intergroup Comparisons for Trained Pilot Groups Tested (increasing magnitude indicates reduced performance for all test measures except ADC CR).

Test measure differences	F	(df)	p	Significant pairwise*
DLT alone	4.47	(3, 198)	< .005	F/A-18, F-14 < AV-8B
DLT w/PMT(S)	13.80	(3, 197)	< .0001	F/A-18 < F-14, (AV-8B) < HELO
DLT w/PMT(SR)	5.29	(3, 197)	< .002	F/A-18 < AV-8B, HELO
PMT(S) alone	4.81	(3, 197)	< .0035	F-14, F/A-18 < HELO
PMT(S) w/DLT	15.17	(3, 200)	< .0001	F/A-18 < (F-14), AV-8B < HELO
PMT(SR) w/DLT	4.29	(3, 198)	< .0065	F/A-18 < AV-8B
Sgl ODCT	14.01	(3, 170)	< .0001	F-14, F/A-18, AV-8B < HELO
Sgl ADC CR	6.22	(3, 162)	< .0008	HELO < AV-8B, F/A-18, F-14
Dual ODCT	8.83	(3, 170)	< .0001	F-14, F/A-18, AV-8B < HELO
Dual ADC CR	3.20	(3, 168)	< .025	HELO < F-14

* A group in parentheses does not differ significantly from the adjacent group opposite the magnitude symbol.

TEST CRITERION ANALYSIS

To ascertain a numerical indication of the overall flight performance of the pilots during each stage of training, raw final flight grades for primary, intermediate, and advanced flight training were obtained from the aviation training jacket summary cards of 21 (66%) of the AV-8B subjects. All subjects completed the same training syllabus, which involved primary fixed-wing training and then intermediate/advanced jet training. Flight grades were not significantly different among any of the AV-8B subgroups tested.

Of 48 individual Pearson product-moment correlations performed between the various test measures and three flight grades, only one (2.1%) was significant. The results of a canonical correlation analysis (11), utilized to determine if any linear combination of predictor variables (test measures) would correlate significantly with any linear combination of criterion variables (flight grades), were not significant. Thus, for this small group of Harrier pilots, there is no statistical relationship between the test battery measures and these final flight grades.

CONCLUSIONS AND RECOMMENDATIONS

The results of this study, and earlier studies on different naval aviator populations, can be summarized into three main points. First, no significant differences in test battery performance were found between pilots with many flight hours in military aircraft, including the AV-8B, and those with significantly fewer flight hours. This agrees with previous findings (5) of test performance differences between pilots with varying amounts of flight experience. Performance on this test battery, through a combination of ability and experience, appeared to peak sometime before RAG assignment so that individuals with similar abilities and experiences showed similar test performance levels. Furthermore, their performance did not improve significantly beyond that point regardless of how much more flight time they accrued. Even if pilots became more skilled in their particular aircraft due to increased flight time, this test battery did not appear to reflect a corresponding increase in competence. Whether this is due to the possibility that the cognitive and psychomotor processes assumedly measured by this battery do not change with an increase in skill or that this battery is simply not sensitive enough to measure the changes that may, in fact, be there is not presently known.

Second, performance of the Harrier pilots matched that of experienced pilots of other jets on some tests but not on other tests. Given the similarities involving both selection and training through advanced jet training, this was not expected. What this might indicate in terms of pilot selection procedures within Marine Corps aviation can only be speculated upon. The lack of similar performance differences between the Harrier and other jet groups among these tests makes the assumption of a simple explanation for all these differences rather difficult. Such differences cannot, at this time, be viewed as proof that the skills needed to fly a Harrier are any different from those needed to fly any other jet. They also do not indicate that a pilot chosen and trained to fly a helicopter would fly a Harrier any better or worse, after retraining, than one trained to fly jets from the beginning. Such questions may only be answerable using more sophisticated techniques designed specifically for the Harrier.

Third, test battery performance was not significantly related to final flight grades during primary, intermediate, or advanced flight training for this small select group of pilots. Some elements of this test battery correlated significantly with the final primary flight grade when SNAs as a whole were examined (8). An earlier study (7) speculated that once SNAs complete primary flight training and are assigned to specific pipelines based in large part on final primary flight grade, the variance seen in test battery performance among members of any of these pipeline subgroups would not correlate significantly with other measures of future flight performance. This has already been shown for flight performance measures beyond advanced flight training (5,6). The current study appears to indicate that this might also be the case for flight performance measures obtained for all three phases of flight training prior to RAG training within any particular pipeline subgroup. Further research with a much larger sample of SNAs that includes all possible pipeline assignments is needed to verify these results.

In conclusion, elements of this test battery show promise in predicting the performance of SNAs during primary flight training (12). Such information would be of value to those who make various decisions concerning pilot selection and subsequent pipeline assignments. Whether this battery is or could be made cost-effective remains to be shown. The evidence presented up to this point would indicate that the test elements of this battery would not be useful in predicting flight performance beyond advanced flight training and possibly not even beyond primary flight training within any one particular pipeline subgroup. Some elements of this test battery, however, might eventually prove useful in determining the level of certain fundamental abilities thought necessary for operating a particular type of aircraft. Evidence (7) indicates that some of these tests may be measuring an innate ability that remains relatively constant throughout a pilot's career in comparison to other pilots with similar experience. More work is needed to determine the exact nature of these abilities using methods other than those found within the present test battery.

REFERENCES

1. Rickus, G.M. and Berkshire, J.R., Development of an Aviation Combat Criterion: Preliminary Report, NAMI-1047, Naval Aerospace Medical Institute, Pensacola, FL, August 1968. (AD A675 214)

2. Brictson, C.A., Burger, W.J., and Gallagher, T.J., "Prediction of Pilot Performance during Initial Carrier Landing Qualifications." Aerospace Medicine, Vol. 43, pp. 483-487, 1972.

3. Shannon, R.H. and Waag, W.L., The Prediction of Pilot Performance in the F-4 Aircraft, NAMRL-1186, Naval Aerospace Medical Research Laboratory, Pensacola, FL, July 1973. (AD A764 868)

4. Griffin, G.R., Morrison, T.R., Amerson, T.L., and Hamilton, P.V., Predicting Air Combat Maneuvering (ACM) Performance: Fleet Fighter ACM Readiness Program Grades as Performance Criteria, NAMRL 1333, Naval Aerospace Medical Research Laboratory, Pensacola, FL, October 1987. (AD A191 605)

5. Griffin, G.R. and Shull, R.N., Predicting F/A-18 Fleet Replacement Squadron Performance Using an Automated Battery of Performance-based Tests, NAMRL-1354, Naval Aerospace Medical Research Laboratory, Pensacola, FL, March 1990. (AD A223 899)

6. Shull, R.N. and Griffin, G.R., Predicting F-14 Air Combat Maneuvering (ACM) Performance Using an Automated Battery of Cognitive/Psychomotor Tests, NAMRL-1356, Naval Aerospace Medical Research Laboratory, Pensacola, FL, June 1990. (AD A228 593)

7. Shull, R.N. and Griffin, G.R., Performance of Several Different Naval Aviator Communities on a Cognitive/Psychomotor Test Battery: Pipeline Comparison and Prediction, NAMRL-1361, Naval Aerospace Medical Research Laboratory, Pensacola, FL, 1990.

8. Delaney, H.D., Validation of Dichotic Listening and Psychomotor Task Performance as Predictors of Primary Flight Training Criteria: Highlighting Relevant Statistical Issues, NAMRL-1357, Naval Aerospace Medical Research Laboratory, Pensacola, FL, July 1990.

9. Scheffe, H., "A Method for Judging All Contrasts in the Analysis of Variance." Biometrika, Vol. 40, pp. 87-104, 1953.

10. Seigel, S., Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill Book Co., New York, 1956.

11. Kerlinger, F.N. and Pedhazur, E.J., Multiple Regression in Behavioral Research, Holt, Rinehart, and Winston, Inc., New York, 1973.

12. Blower, D.J. and Dolgin, D.L., An Evaluation of Performance-Based Tests Designed to Predict Success in Primary Flight Training, unpublished manuscript, Naval Aerospace Medical Research Laboratory, Pensacola, FL, 1990.

Other Related NAMRL Publications

Dolgin, D. L. and Gibb, G.D., A Review of Personality Measurement in Aircrew Selection, NAMRL Monograph 36, Naval Aerospace Medical Research Laboratory, Pensacola, FL, July 1988. (AD A200 392)

Gibb, G.D. and Dolgin, D.L., "Predicting Military Flight Training Success by a Compensatory Tracking Task." Military Psychology, Vol. 1, No. 4, pp. 235-240, October 1989.

Gibb, G.D. and Dolgin, D.L., "Validation of a Computer-based Aviation Secondary Selection System for Student Naval Aviators." Proceedings of the Human Factors Society 32nd Annual Meeting, Anaheim, CA, 1988, pp. 807-811.

Griffin, G.R., An Improved Automated Selection System For Navy Pilots, NAMRL Research Information Bulletin 88-2, Naval Aerospace Medical Research Laboratory, Pensacola, FL, June 1988. (AD A203 438)

Morrison, T.R., Complex Visual Information Processing: A Test for Predicting Navy Primary Flight Training Success, NAMRL-1338, Naval Aerospace Medical Research Laboratory, Pensacola, FL, July 1988. (AD A200 394)

Nontasak, T., Dolgin, D.L. and Griffin, G.R., "Performance-based Tests, Personality Attributes, and Training Outcome Among Landing Craft Air Cushion (LCAC) Vehicle Operators." Proceedings of the Human Factors Society 33rd Annual Meeting, Denver, CO, 1989, pp. 901-904.

Shull, R. N. and Dolgin, D.L., "Personality and Flight Training Performance." Proceedings of the Human Factors Society 33rd Annual Meeting, Denver, CO, 1989, pp. 891-895.

Shull, R.N., Dolgin, D.L., and Gibb, G.D., The Relationship Between Flight Training Performance, a Risk Assessment Test, and the Jenkins Activity Survey, NAMRL-1339, Naval Aerospace Medical Research Laboratory, Pensacola, FL, July 1988. (AD A200 395)