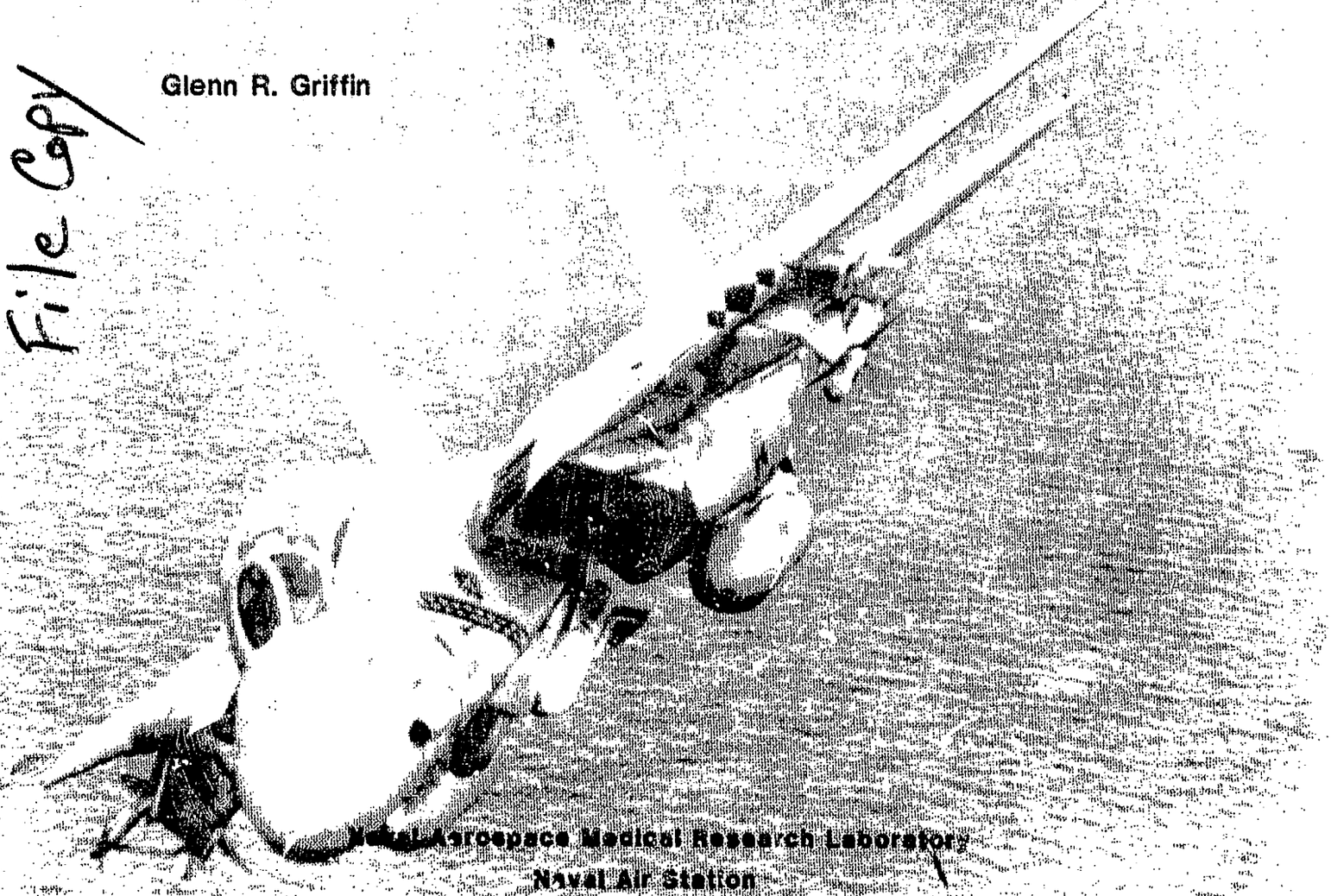
Naval Aerospace Medical Research Laboratory

NAMRL-1336

December 1987

DEVELOPMENT AND EVALUATION OF AN AUTOMATED SERIES OF SINGLE- AND MULTIPLE-DICHOTIC LISTENING AND PSYCHOMOTOR TASKS

Glenn R. Griffin



Naval Aerospace Medical Research Laboratory

Naval Air Station

Pensacola, Florida 32508-5700

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY Unclassified	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NAMRL - 1336	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Aerospace Medical Research Laboratory 6b. OFFICE SYMBOL (*If applicable*) 03	7a. NAME OF MONITORING ORGANIZATION Naval Medical Research and Development Command
6c. ADDRESS (*City, State, and ZIP Code*) NAS, Pensacola, FL 32508-5700	7b. ADDRESS (*City, State, and ZIP Code*) NMC, NCR, Bethesda, MD 20814-5044
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (*If applicable*)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (*City, State, and ZIP Code*)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 63706N; PROJECT NO. M0096001; TASK NO. 001; WORK UNIT ACCESSION NO. 0151

11. TITLE (*Include Security Classification*)
(U) Development and Evaluation of an Automated Series of Single and Multiple Dichotic Listening and Psychomotor Tasks

12. PERSONAL AUTHOR(S)
G. R. Griffin

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (*Year, Month, Day*)	15. PAGE COUNT
Interim	FROM 1987 TO 1987	87-12	16

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (*Continue on reverse if necessary and identify by block number*)
FIELD	GROUP	SUB-GROUP	Pilot Selection, Psychomotor, Dichotic Listening, Multi-task Performance

19. ABSTRACT (*Continue on reverse if necessary and identify by block number*)

Attrition in undergraduate naval aviation training represents a costly problem. An average of 25% of student naval aviators fails to complete training. This study reports an effort to develop automated single- and multiple-dichotic listening and psychomotor tasks, which have the potential to reduce aviator attrition through improved selection and may also be useful in initial pipeline classification.

Statistical analysis of Study I, comparing a forward and backward series of automated dichotic listening (DLT) and psychomotor (PMT) tasks, indicated that a backward-direction orientation associated with the psychomotor tests resulted in increased difficulty for all PMT measures and two of three multitask DLT measures. The correlational estimates of test-retest reliability for the multitask DLT and PMT measures were adequate for both series of automated tasks but slightly higher ($\underline{r}$ = .80 DLT, $\underline{r}$ = .90 PMT) for the

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION	
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL J. O. HOUGHTON, CAPT MC USN	22b. TELEPHONE (*Include Area Code*) (904) 452-3286	22c. OFFICE SYMBOL 00

19. ABSTRACT (Continued)

backward series. There were significant correlations between the DLT and PMT tasks, for each series of single- and multitask measures. The relation between such seemingly different tasks is difficult to understand since the DLT is an auditory cognitive processing task, and the PMT is an eye, hand, foot coordination task. However, the significant correlations were both smaller and fewer in number for the backward series of automated tests.

Study II was a correlational evaluation between the new automated multitask measures and old nonautomated tasks with demonstrated validity for the prediction of primary flight performance. The correlations between corresponding tasks of the new automated and old nonautomated tasks averaged .60 for the DLT measures and .66 for the PMT tasks.

The results of Study III indicated that certain automated DLT and PMT measures were significantly related to primary flight grades (PFG) in Navy flight training. For the backward series of tasks, all DLT and PMT measures were significantly correlated with PFG. However, only two DLT and two PMT measures of the forward series were significantly related to PFG. No significant correlations were found between the automated DLT and PMT tests and the pass/fail criterion. The absence of a suitable number of flight failure attritions was discussed as a possible reason for this result. A regression analysis for the backward series of test measures and primary flight training criteria indicated that a psychomotor stick and rudder measure and the FAR selection test were significantly related to PFG ($R = .53$, $F(2,85) = 16.56$, $p < .0001$). There were no significant correlations between the automated DLT and PMT measures and prior flight hours.

These results indicate that a series of automated DLT and PMT tasks are suitable replications of an older version of nonautomated tasks. The advantages of the automated tasks are that they require less administrative support and provide automatic scoring of performance. The backward series of automated tasks, which was correlated more strongly with criterion performance, should be administered to a large sample of student naval aviators to determine if the tests can account for additional variance in the prediction of flight training performance beyond that of current selection tests.

SUMMARY PAGE

THE PROBLEM

Attrition in undergraduate naval aviation training is a costly problem. An average of 35% of student naval aviators fails to complete training. This study reports an effort to develop automated single- and multiple-dichotic listening and psychomotor tasks, which have the potential to reduce aviator attrition through improved selection and may also be useful in initial pipeline classification.

FINDINGS

Statistical analysis in Study I, comparing automated Dichotic Listening (DLT) and Psychomotor (PMT) tasks, indicated that a backward-direction orientation associated with the stick movement of one of the psychomotor tasks resulted in increased difficulty for all PMT measures and two of three DLT measures performed under multitask conditions. The correlational estimates of test-retest reliability for the multitask DLT and PMT measures were adequate for both series of automated tasks but slightly higher (r = .80 DLT, r = .90 PMT) for the backward series. There were significant correlations between the DLT and PMT tasks for each series of single- and multitask measures. The relation between such seemingly different tasks is difficult to understand since the DLT is an auditory cognitive processing task, and the PMT is an eye, hand, and foot coordination task. However, the significant correlations were both smaller and fewer in number for the backward series of automated tests.

Study II was a correlational evaluation between the new automated multitask measures and the old nonautomated tasks with demonstrated validity for the prediction of primary flight performance. The correlations between corresponding tasks of the new automated and old nonautomated tasks averaged .60 for the DLT and .66 for the PMT.

The results of Study III indicated that certain of the automated DLT and PMT measures were significantly related to primary flight grades (PFG) in Navy flight training. For the backward series of tasks, all DLT and PMT measures were significantly correlated with PFG. However, only two DLT and two PMT measures of the forward series were significantly related to PFG. No significant correlations were found between the automated DLT and PMT tests and the pass/fail criterion. The absence of a suitable number of flight failure attritions was discussed as a possible reason for this result. A regression analysis for the backward series of test measures and primary flight training criteria indicated that a psychomotor stick-and-rudder measure and the FAR selection test were significantly related to PFG (R = .55, F (2,85) = 16.56, p < .0001). There were no significant correlations between the automated DLT and PMT measures and prior flight hours.

RECOMMENDATION

These research results indicate that a series of automated DLT and PMT tasks, which require less administrative support and provide automatic scoring of performance, are suitable replications of an older version of nonautomated tasks. The backward series of automated tasks, which was

correlated more strongly with criterion performance, should be administered to a large sample of student naval aviators to determine if the tests can account for additional variance in the prediction of flight training performance beyond that of current selection tests.

Acknowledgments

The author gratefully acknowledges the assistance of Mr. Peter Collyer and Ms. Kathy Vogel for their efforts in programming and modifying the dichotic listening and psychomotor tasks into a comprehensive automated test system, and Mr. Alfred Thomas for his care and diligence in running subjects. In addition, the author thanks Kathy, Pete, and Al for providing many helpful suggestions that resulted in test improvements during the course of this study.

INTRODUCTION

A system of automated single and multiple tasks has been developed at the Naval Aerospace Medical Research Laboratory (NAMRL) to support test batteries for performance-based biomedical selection and classification of naval aviators (1). Certain of these automated tasks were based on previous research demonstrating the potential of nonautomated multiple tasks to improve human performance in simulated flight (2) and to predict naval aviator performance in training (3) and fleet environments (4).

A previously reported traditional dual-performance experiment (2) indicated that the performance of subjects on the psychomotor portion of a combined psychomotor (Systems Research Laboratories, Inc., Psychomotor Task) and dichotic listening task (DLT) was significantly improved when a vocal, as opposed to manual, response mode was employed for dichotic listening. A subsequent validation study (3) based on student naval aviator performance on the single and combined tasks indicated that the single tasks were valid predictors of pass/fail and primary training flight grades ($p < .05$). Importantly, the best predictor of pass/fail ($p < .01$) was the DLT performed under multitask conditions. Alternately, the best predictor of primary flight grades was a single-task stick-and-rudder psychomotor (PMT) measure. Multiple regression results indicated that the multitask DLT, the Flight Aptitude Rating (FAR) selection test, and multitask PMT psychomotor measures accounted for significant variance in the prediction of the pass/fail criterion. A separate regression analysis demonstrated that the multitask DLT and PMT measures accounted for significant variance beyond that of the Naval and Marine aviation selection test battery (3).

An additional regression analysis was performed on the data referred to above (3) using primary flight grades as the criterion. The ability to predict flight grades is an important consideration in the event of a Navy management requirement to assign, prior to primary training, student naval aviators into fixed- or rotary-wing pipelines. In this regression analysis, a single-task psychomotor stick-and-rudder measure entered the regression equation first ($r = .51$) followed by the AQT ($R = .64$). Each of these measures contributed significant variance to the final regression equation ($F(2,43) = 14.89$, $p < .01$). This analysis suggests that psychomotor and selection test measures may be useful in initial aviator classification. In addition to the above, a preliminary automated version of the PMT and DLT was used successfully to predict the air combat maneuvering performance of an F-4 squadron undergoing a fleet readiness evaluation at NAS Oceana (4).

The DLT and PMT tasks have been automated at NAMRL on an Apple IIe computer as a series of single- and multitask tests. The task series and cumulative testing times are depicted in Table 1. The purpose of the present study was to report descriptive statistics and reliability estimates of two versions of the automated task series (Study I), present correlational statistics indicating the relation between the original nonautomated and newly automated tasks (Study II), and estimate the relation of the automated series of tasks to flight training performance (Study III).

TABLE 1. Automated Psychomotor and Dichotic Listening Task Menu.

Automated task presentation order	Description	Cumulative test time (min)
1.	Single-task psychomotor, stick only	09
2.	Single-task dichotic listening	32
3.	Initial multitask[a]	37
4.	Single-task psychomotor, stick & rudder	50
5.	Second multitask[b]	55
6.	Third multitask[b]	60

[a] tasks 1 and 2 combined
[b] tasks 4 and 2 combined

STUDY I. COMPARISON OF TWO AUTOMATED SERIES OF SINGLE AND MULTIPLE TASKS

The objective of this evaluation was to report and compare the descriptive statistics and correlational reliability estimates of two series of single and multiple tasks. One series, labeled a forward test version, incorporated psychomotor stick-and-rudder movements that corresponded to respective cursor movement on a CRT. That is, when the stick or rudder was moved to the right or left, the CRT cursor moved to the right or left. A second series, or backward test version, incorporated stick-and-rudder movements that were opposite CRT cursor movement.

Subjects

One hundred student naval aviators performed on the automated dichotic listening and psychomotor tasks. Fifty subjects (Group I) performed on the forward series of tasks, and fifty subjects (Group II) performed on the backward series of single and multiple tasks.

Apparatus and Procedure

All subjects performed each series of automated tasks in identical order (see Table 1). Subjects were volunteers about to enter navy flight training during the fall of 1986. A photograph of the test apparatus is in Figure 1.

Psychomotor Single Task (PMT). Subjects were required to maintain first one and then two computer-generated cursors on fixed targets presented on a CRT using a Measurement Systems, Inc., joystick and locally produced rudder pedals patterned after those of a Systems Research Laboratory Psychomotor test device. The subjects manipulated the joystick using their right hand to control the stick X,Y-movement cursor and manipulated rudder pedals using their feet to control a second Z-movement cursor. Single-task PMT performance consisted of two 3-min sessions of a stick-only PMT task followed by three 3-min sessions of a stick-and-rudder combined task. Each of the single-task PMT 6- and 9-min test sessions were preceded by 3 min of practice. Individual sessions were separated by rest periods of 20 s. All instructions were presented visually on a CRT. Psychomotor test scores were computer-calculated pixel error over the 6-

and 9-min test periods, respectively. The error scores were totaled for the X,Y stick cursors and X,Y,Z stick-and-rudder cursors and represent sums of the linear deviations of the cursors from an ideal non-error position for each of three movement axes (X,Y, or Z) of the visual display (see Fig. 1).

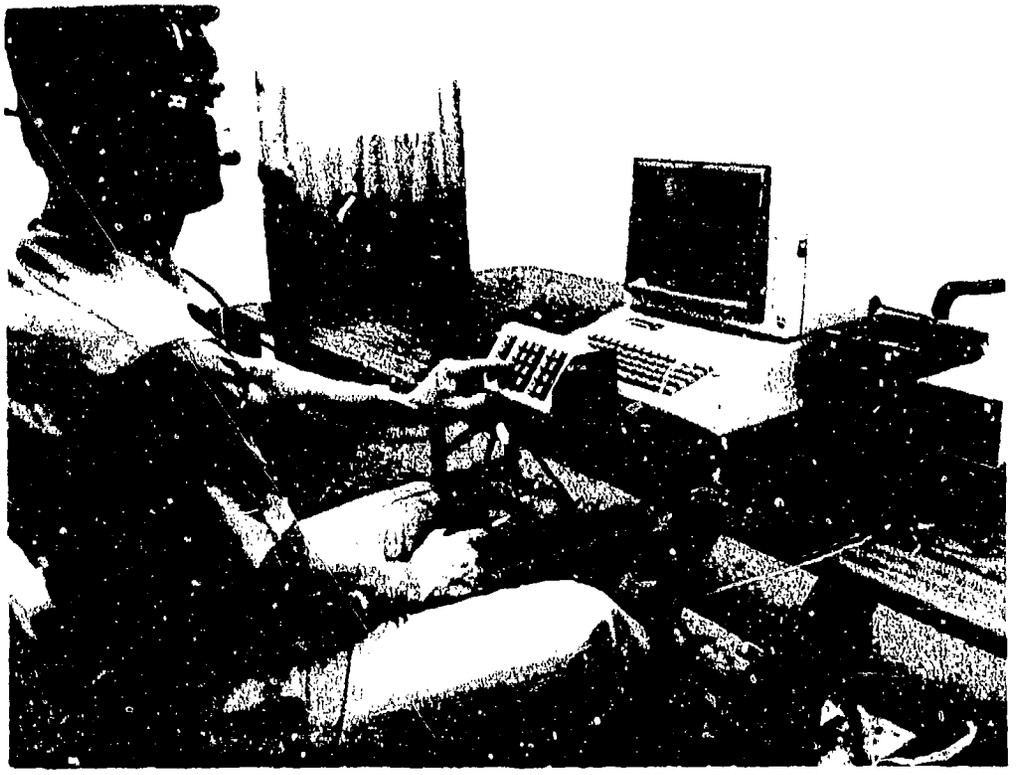


Figure 1. Automated dichotic listening and psychomotor task apparatus.

Dichotic Listening Single Task (DLT). The DLT was patterned after that of Gopher (5), and subsequently modified and then automated at NAMRL (6). The DLT is an auditorially presented series of letter-digit string sets for 24 trials. Two Jameco JE 520-AP Voice Synthesizers were used to present the DLT letter-digit strings over binaural headphones to each subject at a listening level of 72 dB/Leq (re:20 μpa). Subjects were instructed to direct attention to one ear while ignoring the other and to correctly report, using a keypad, the digits (0-9) presented to the attended-to-ear in the sequence of their occurrence. Subject responses were made on a keypad placed immediately in front and slightly left of center, for left-hand use. Test instructions were presented visually on the CRT, and the test was preceded by six auditorially presented practice trials with immediate visual performance feedback indicating the letters and digits presented and the subjects' keypad responses. Finally, subjects completed three multiple-choice questions to make certain that they understood the concept of the DLT. An example of a DLT trial is written in Table 2. The single-task DLT performance measure was the number of correct responses per 24 trials; 216 correct responses were possible.

TABLE 2. DLT Trial Visual Example.

PART I	Left Ear	R 8 N S M Y 2 G B 7 F L 6 R L 5
	"Right" (Vocal Channel 'attend' Command)	
	Right Ear	Y L 3 S R 4 F Z 9 X F 0 F N 1 L
PART II	Left Ear	B F 4 3 7 9
	"LEFT" (Vocal Channel 'attend' Command)	
	Right Ear	G L 1 5 6 2

Multitask PMT/DLT. Subjects performed the multitask PMT/DLT with their right hand on the joystick and feet on the rudder pedals (PMT) and used their left hand for keypad responses to the auditorially presented DLT. In the multitask conditions, subjects performed three sessions of the DLT and PMT simultaneously (a 12-trial DLT and a 4.5-min PMT). During the initial multitask condition, subjects performed the DLT and stick-only PMT. During the latter two multitask conditions, subjects performed the DLT and the stick-and-rudder PMT. Performance measures for the PMT and DLT in the multitask conditions were identical to single-task performance except for different lengths of PMT testing and the presentation of 12, rather than 24, DLT trials during multitask performance. One hundred and eight correct responses were possible for the reduced set of DLT trials.

Results

Descriptive statistics for the forward and backward series of single and multiple tasks are provided in Table 3. An analysis of variance indicated a significant difference between the two (F (1,98) = 33.77, p < .01). Further, separate follow-on tests (Duncan, Tukey, F-test) (7) indicated that all measures of the two series, except the DLT single task and the third DLT multitask measures reflect significantly different subject performance. The backward series of single and multiple tasks was more difficult. In two of the three multitask conditions, the increased difficulty associated with the backward PMT not only resulted in larger error scores for all PMT measures but more DLT errors as well.

The correlation matrices in Tables 4A and 4B identify statistically significant correlations between the DLT and PMT forward and backward series of single- and multitask measures, and present estimates of multitask measure reliabilities as well. The correlation coefficients indicate acceptable reliabilities for the multiple tasks. For the DLT, the relation between the last two multitask DLT trials (measures 3 and 4 of Tables 4A and 4B) is r = .79 for the forward series and r = .80 for the backward series of tasks. For the PMT, the correlation between the last two multitask trials (measures 8 and 9 of Tables 4A and 4B) is r = .82 for the forward series and r = .90 for the backward series of tasks.

A disconcerting aspect of both the forward and backward series of automated tasks is that there are significant correlations between the DLT and PMT tasks in both single- and multitask performance conditions. This finding has been reported previously (3). However, it is difficult to understand how the DLT and PMT tasks are related when they appear so very different--one being an eye, hand, and foot coordination task and the other

an auditory listening task. One might expect significant correlations between the tasks performed in combination since performance on one task affects performance on the other. For example, measures 2 and 8 of Tables 4A and 4B (initial multitask DLT and PMT) are significantly correlated, (r = -.69 for the forward series and r = -.43 for the backward series of tasks). Results also indicate that even the single DLT and PMT tasks are significantly related. For example, the correlation between the single-task DLT (measure 1) and the PMT single-task stick-and-rudder (measure 7) is -.34 for the forward series and -.29 for the backward series of tasks.

TABLE 3. Descriptive Statistics for a Forward and a Backward Series (Parentheses) of Single- and Multitask Measures.[a]

Measure DLT correct responses PMT pixel errors	Mean	SD	Min	Max
1. Single-task DLT	206.76 (205.40)	7.83 (6.54)	169 (189)	215 (215)
2. Initial multitask DLT	103.20 (99.80)	5.61 (5.88)	77 (82)	108 (108)
3. Second multitask DLT	100.54 (97.14)	7.99 (9.07)	58 (72)	108 (107)
4. Third multitask DLT	101.56 (99.00)	8.06 (8.30)	58 (74)	108 (107)
5. Single-task PMT stick only	6971.04 (13255.12)	3322.38 (7655.64)	4462 (4907)	28115 (41838)
6. Initial multitask PMT stick only	2807.28 (6166.90)	1290.16 (5052.80)	1459 (1360)	9129 (26207)
7. Single-task PMT stick & rudder	26266.36 (46041.00)	14680.13 (23664.11)	13360 (18399)	115424 (107625)
8. Second multitask PMT stick & rudder	7145.28 (16833.72)	4398.95 (15005.29)	2382 (4935)	29224 (93195)
9. Third multitask PMT stick & rudder	6941.56 (15495.42)	3298.56 (13872.92)	2518 (3720)	16295 (84015)

[a] n =100 (50 Forward + 50 Backward)

Importantly, the correlations between the DLT and PMT measures in the backward series of tasks are smaller and therefore more desirable. (A goal of test battery development is to identify predictive unrelated tests so that each may contribute variance to a suitable criterion.) The correlations between measure 1 (single-task DLT) and the single- and multitask PMT measures 5 through 9 indicate only two significant correlations for the backward series compared to four for the forward series of tasks.

In summary, the correlational results indicate slightly higher (but not significantly different) reliabilities for the backward series of tasks. More importantly, less relation exists between the DLT and PMT measures in the backward series of tasks. Finally, there is increased difficulty for both the DLT and PMT measures in the backward series of tasks.

TABLE 4A. Pearson Product Moment Correlation Matrix for a Forward Series of DLT and PMT Measures (n = 50).

1.[a]	----								
2.	.78**	----							
3.	.63**	.56**	----						
4.	.66**	.57**	.79**	----					
5.	-.13	-.02	-.09	-.13	----				
6.	-.68**	-.68**	-.39**	-.34*	.18	----			
7.	-.34*	-.27*	-.09	-.09	.10	.46**	----		
8.	-.72**	-.69**	-.52**	-.53**	.06	.84**	.54**	----	
9.	-.49**	-.47**	-.54**	-.52**	.05	.65**	.41**	.82**	----

TABLE 4B. Pearson Product Moment Correlation Matrix for a Backward Series of DLT and PMT Measures (n = 50).

1[a].	----								
2.	.49**	----							
3.	.31*	.65**	----						
4.	.32*	.48**	.80**	----					
5.	.04	-.19	-.22	-.13	----				
6.	.12	-.23	-.23	-.08	.65**	----			
7.	-.29*	-.58**	-.51**	-.45**	.47**	.47**	----		
8.	-.34**	-.43**	-.36**	-.26	.31*	.35*	.65**	----	
9.	-.26	-.40**	-.31*	-.20	.29*	.47**	.67**	.90**	----

[a] See Table 3 for a description of numbered variables
*p < .05 (two-tailed)
**p < .01 (two-tailed)

STUDY II. RELATION OF AUTOMATED DLT AND PMT TASKS TO OLD NONAUTOMATED PROTOTYPES

The objective of Study II was to evaluate the correlation between the new automated DLT and PMT tasks and the old tape recorded/presented DLT and Systems Research Laboratory, Inc., (SRL) psychomotor tasks. In essence, this represents a validation process to verify that the automated tasks measure performance similar to that of the nonautomated tasks. The advantages of an automated series of tasks are that two test administrators would not be needed to collect and analyze performance data, and the scoring of performance could be made automatic.

Subjects

Sixty student naval aviators performed on the forward and backward series of automated tasks and the old nonautomated DLT and PMT. All subjects were volunteers awaiting entry into the primary portion of the Navy flight training program. Not all subjects completed all portions of the various tasks due to scheduling and equipment problems.

Apparatus and Procedure

Subjects initially performed on the new automated forward series of DLT and PMT tasks. Four days later, subjects performed on the backward series of automated tasks and immediately thereafter completed one 5-min practice and one 5-min test trial on the old (backward-oriented cursor configuration) nonautomated multitask DLT and SRL PMT. The apparatus for the new automated DLT and PMT tasks was identical to that described in Study I. The apparatus for the old multitask DLT and PMT has been described in the introduction section and in previous reports (2,3).

Results

Table 5 presents the correlations between the nine measures of the new automated forward series with the new automated backward series taken 4 days later. As expected, there are significant correlations between the two test series. Naturally, one would expect high correlations between the DLT tasks since they are identical in the two series. In one sense, the correlations amount to a test/retest reliability measure over an intervening period of 4 days. The correlations between the PMT scores are also highly significant. These results indicate that those subjects who performed well on the initial forward series of tasks continued to perform relatively well on the more difficult backward task series and vice versa. In summary, although Study 1 results indicated a significantly higher degree of difficulty associated with the backward series of tasks, there is still a significant and strong correlation between subject performance on the two series of tasks.

TABLE 5. Pearson Product Moment Correlations between the Forward and Backward Series of DLT and PMT Tasks (n = 56).

Measure	r
1. Single-task DLT	.69**
2. Initial multitask DLT	.82**
3. Second multitask DLT	.82**
4. Third multitask DLT	.82**
5. Single PMT stick	.28*
6. Initial multitask PMT stick	.50**
7. Single PMT stick & rudder	.75**
8. Second multitask PMT stick & rudder	.82**
9. Third multitask PMT stick & rudder[a]	.74**

[a] n = 50
*p < .05 (two-tailed)
**p < .01 (two-tailed)

The correlation statistics of Table 6 indicate a significant relation between the new automated multitask DLT and PMT measures and those corresponding to the old apparatus. The correlations for the PMT measures are generally stronger than those of the DLT.

TABLE 6. Pearson Product Moment Correlations between the New Forward and Backward Automated Task Series and Old Nonautomated DLT and PMT.

Forward automated multitask measures (n = 40)	Nonautomated multitask DLT	Nonautomated multitask PMT
Initial multitask DLT	.60**	
Second multitask DLT	.66**	
Third multitask DLT	.58**	
Initial multitask PMT stick		.62**
Single PMT stick & rudder		.73**
Second multitask PMT stick & rudder		.66**
Third multitask PMT stick & rudder[a]		.53**
Backward automated multitask measures (n = 37)	**Nonautomated multitask DLT**	**Nonautomated multitask PMT**
Initial multitask DLT	.49**	
Second multitask DLT	.61**	
Third multitask DLT	.63**	
Initial multitask PMT Stick		.69**
Single PMT stick & rudder		.66**
Second multitask PMT stick & rudder		.70**
Third multitask PMT stick & rudder		.69**

[a] n = 35

**p < .01 (two-tailed)

STUDY III. RELATION OF THE AUTOMATED FORWARD AND BACKWARD SERIES OF PMT AND DLT TASKS TO TRAINING PERFORMANCE CRITERIA

The purpose of Study III was to assess the ability of the automated DLT and PMT tasks to predict primary flight grades and successful completion of primary flight training, two important criteria in the Navy flight training program.

Subjects

To date, 98 student naval aviators (SNAs) who performed on the backward series of tasks and 105 SNAs who performed on the forward series have completed or failed primary training. All subjects were volunteers for the experiment.

Apparatus and Procedure

Testing apparatus was identical to that described in Study I. Subject testing was conducted while SNAs were awaiting entry into the primary flight portion of the Navy flight training program. In addition to DLT and PMT performance, Naval and Marine Aviation Selection Battery test scores were obtained for each subject. The battery consists of the Aviation Qualification Test (AQT) and the Flight Aptitude Rating (FAR). The AQT is a general ability test of verbal and quantitative aptitude. The FAR is comprised of a Spatial Apperception Test, a Mechanical Comprehension Test, and a Biographical Inventory. All subjects had been initially selected for flight training based on their performance on these tests. Criterion performance measures were completion or failure in primary flight training (pass/fail) and primary flight grades (PFGs). In addition, the relation of subjects' previous flight hours (PFH) to the various test measures was incorporated into the data analysis.

Results

The correlations in Tables 7A and 7B indicate the relation of the automated forward and backward task and AQT/FAR measures to criterion performance (pass/fail, PFG, and PFH). The correlations between the various measures and pass/fail are point-biserial (r-pbi) correlations while those between PFG and PFH are Pearson product-moment correlation coefficients.

The correlation results indicate that neither the forward nor the backward series of automated DLT or PMT tests was significantly correlated with pass/fail. (The FAR was correlated with pass/fail in the group that took the backward series.) Theoretically, the multitask tests should predict flight failure attrition since this attrite category is reportedly associated with problems in controlling the aircraft and in performing multiple tasks (9). In examining the types of attrition occurring for the two groups of subjects (20 attrites, 10 per group), there were only two flight failures, and both of these were in the forward test series subject group. The remaining attrite categories were: Not Aviator Material (lack of motivation expressed), 12 subjects; Not Physically Qualified, 2 subjects; Academic, 2 subjects; and Other, 2 subjects. We would not expect the multitask tests to predict the majority of these failures, which are associated with "drop on request/motivation" in the research literature (8,9). It remains to be seen whether an accumulation of flight failure attrition will improve the prediction of pass/fail.

No test measure was significantly related to PFH except the AQT for the forward group of test subjects. Importantly, both the forward and backward series had some automated task measures significantly correlated with PFG. For the backward series, all test measures were significantly correlated with PFG. The PMT single-task stick-and-rudder measure produced the strongest correlation ($r = -.49$, $p < .01$) with PFG. For the forward series, four of the nine test measures were significantly correlated with PFG with the initial multitask DLT measure producing the strongest correlation ($r = .31$, $p < .01$).

To statistically determine which measures would predict PFG, forward selection multiple regression analyses were conducted. The regressions

utilized all test measures of Table 7A and 7B. For the forward series of DLT and PMT tasks, only the initial multitask DLT (measure 2) contributed significant variance to the regression (R = .31, F (1,93) = 10.11, p < .002). No other measure entered the regression due to high correlations among the other test measures significantly related to PFG.

TABLE 7A. Relation of Forward Series of Automated Tasks and Selection Test Measures to Pass/Fail, Primary Flight Grade (PFG), and Previous Flight Hours (PFH).

Test measures	Pass/Fail (n=105) r-pbi	PFG (n=95) r	PFH (n=105) r
1. Single-task DLT	.11	.22*	-.03
2. Initial multitask DLT	.06	.31**	.04
3. Second multitask DLT	.00	.18	.11
4. Third multitask DLT	.13	.10	.03
5. Single PMT stick	.06	.07	.07
6. Initial multitask PMT	-.05	-.24*	-.04
7. PMT stick & rudder	.00	-.11	-.11
8. Second multitask PMT	-.07	-.20*	-.16
9. Third multitask PMT	.02	-.17	-.12
10. AQT	-.04	.10	-.20*
11. FAR	-.08	.06	.17

TABLE 7B. Relation of Backward Series of Automated Tasks and Selection Test Measures to Pass/Fail, Primary Flight Grade (PFG), and Previous Flight Hours (PFH).

Test measures	Pass/Fail (n=98) r-pbi	PFG (n=88) r	PFH (n=98) r
1. Single-task DLT	-.02	.23*	-.07
2. Initial multitask DLT	-.03	.26*	-.01
3. Second multitask DLT	.02	.27*	-.02
4. Third multitask DLT	.02	.31**	.00
5. Single PMT stick	-.10	-.33**	-.12
6. Initial multitask PMT	.01	-.28**	-.06
7. PMT stick & rudder	.06	-.49**	-.06
8. Second multitask PMT	.06	-.29**	-.08
9. Third multitask PMT	.07	-.33**	-.07
10. AQT	.05	.12	-.10
11. FAR	.23*	.34**	.02

*p < .05 (two-tailed)
**p < .01 (two-tailed)

For the backward series, the single-task PMT stick-and-rudder (measure 7) entered the regression equation first (R = .49), followed by the FAR Selection test (R = .53). Each of these measures contributed significant variance to the final regression equation, (F (2,85) = 16.56, p <.0001).

A final regression analysis was performed for the backward test series in which the selection test measures (AQT, FAR) were forced into the regression in first and second place resulting in a multiple R of .35. The PMT stick-and-rudder measure then entered the equation and contributed an additional 16% significant variance above that provided by the selection tests (R = .53, F (3,84) = 11.10, p < .0001).

CONCLUSIONS AND RECOMMENDATIONS

Statistical analysis of Study I, comparing a forward and backward series of automated DLT and PMT tasks, indicated that a backward orientation associated with the psychomotor tests resulted in increased difficulty for all PMT measures and two of three multitask DLT measures. The correlational estimates of test-retest reliability for the multitask DLT and PMT measures were adequate for both series of automated tasks but slightly higher for the backward series. There were significant correlations between the DLT and PMT tasks or each series of single and multitask measures. The relation between such seemingly different tasks is difficult to understand since the DLT is an auditory cognitive processing task, and the PMT is an eye, hand, foot coordination task. The significant correlations were both smaller and fewer in number for the backward series of automated tests.

Study II was an evaluation of the correlational relation between the new automated multitask measures and old nonautomated tasks with demonstrated validity for the prediction of primary flight performance (3). The correlations between corresponding tasks of the new automated and old nonautomated tasks averaged .60 for the DLT measures and .66 for the PMT tasks.

The results of Study III indicated that certain of the automated DLT and PMT measures were significantly related to primary flight grades (PFG) in Navy flight training. For the backward series of tasks, all DLT and PMT measures were significantly correlated with PFG. In contrast, only two DLT and two PMT measures of the forward series were significantly related to PFG. No significant correlations were found between the automated DLT and PMT tests and the pass/fail criterion. The absence of a suitable number of flight failure attritions was a possible reason for this result. A regression analysis for the backward series of test measures and primary flight training criteria indicated that a psychomotor stick-and-rudder measure and the FAR selection test were significantly related to PFG. There were no significant correlations between the automated DLT and PMT measures and prior flight hours.

These research results indicate that a series of automated DLT and PMT tasks, which require less administrative support and provide automatic scoring of performance, are suitable replications of an older version of nonautomated tasks. The backward series of automated tasks, which was correlated more strongly with criterion performance, should be administered to a large sample of student naval aviators to determine if the tests can account for additional variance in predicting flight training performance beyond that of current selection tests.

REFERENCES

1. Damos D. L. and Gibb, D. B., Development of a Computer-based Naval Aviation Selection Test Battery, NAMRL 1319, Naval Aerospace Medical Research Laboratory, Pensacola, FL, August 1986.

2. Griffin, G. R. and Mosko, J. D., The Effects of Vocal Versus Manual Response Modalities on Multi-Task Performance, NAMRL 1312, Naval Aerospace Medical Research Laboratory, Pensacola, FL, February 1985.

3. Griffin, G. R. and McBride, D. K., Multitask Performance: Predicting Success in Naval Aviation Primary Flight Training, NAMRL 1316, Naval Aerospace Medical Research Laboratory, Pensacola, FL, March 1986.

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, November 1987.

5. Gopher, D., "A Selective Attention Test as a Predictor of Success in Flight Training." Human Factors, Vol. 24. No. 2, pp. 173-183, 1982.

6. Griffin, G. R. and Mosko, J. D., Preliminary Evaluation of Two Dichotic Listening Tasks as Predictors of Performance in Naval Aviation Undergraduate Pilot Training, NAMRL 1287, Naval Aerospace Medical Research Laboratory, Pensacola, FL, July 1982.

7. SAS User's Guide: Statistics, Version 5 ed. SAS Institute, Inc., Cary, NC, 1985.

8. Griffin, G. R. and 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, FL, August 1977.

9. North, R. A. and Griffin, G. R., Aviator Selection 1919-1977, NAMRL SR77-2, Naval Aerospace Medical Research Laboratory, Pensacola, FL, October 1977.