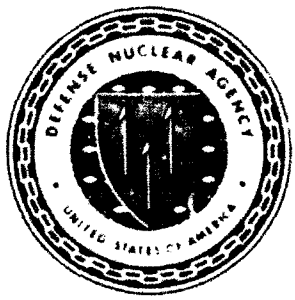
Defense Nuclear Agency
Alexandria, VA 22310-3398



DNA-TR-92-54-V2

Predicting Radiation Induced Performance Decrements of AH-1 Helicopter Crews

Volume 2—Evaluation of Modeling and Simulation Techniques for Predicting Radiation Induced Performance Decrements

William A. Perez, et al.
Science Applications International Corporation
P.O. Box 1303
McLean, VA 22102

March 1993

Technical Report

CONTRACT No. DNA 001-86-C-0308

Destroy this report when it is no longer needed. Do not return to sender.

PLEASE NOTIFY THE DEFENSE NUCLEAR AGENCY, ATTN: CSTi, 6801 TELEGRAPH ROAD, ALEXANDRIA, VA 22310-3398, IF YOUR ADDRESS IS INCORRECT, IF YOU WISH IT DELETED FROM THE DISTRIBUTION LIST, OR IF THE ADDRESSEE IS NO LONGER EMPLOYED BY YOUR ORGANIZATION.

DISTRIBUTION LIST UPDATE

This mailer is provided to enable DNA to maintain current distribution lists for reports. (We would appreciate your providing the requested information.)

NOTE:
Please return the mailing label from the document so that any additions, changes, corrections or deletions can be made easily.

☐ Add the individual listed to your distribution list.

☐ Delete the cited organization/individual.

☐ Change of address.

NAME: _____

ORGANIZATION: _____

OLD ADDRESS	CURRENT ADDRESS
_____	_____
_____	_____
_____	_____

TELEPHONE NUMBER: ( ) _____

DNA PUBLICATION NUMBER/TITLE	CHANGES/DELETIONS/ADDITIONS, etc.) *(Attach Sheet if more Space is Required)*
_____	_____
_____	_____
_____	_____

DNA OR OTHER GOVERNMENT CONTRACT NUMBER: _____

CERTIFICATION OF NEED-TO-KNOW BY GOVERNMENT SPONSOR (if other than DNA):

SPONSORING ORGANIZATION: _____

CONTRACTING OFFICER OR REPRESENTATIVE: _____

SIGNATURE: _____

CUT HERE AND RETURN

DEFENSE NUCLEAR AGENCY
ATTN: TITL
6801 TELEGRAPH ROAD
ALEXANDRIA, VA 22310-3398

DEFENSE NUCLEAR AGENCY
ATTN: TITL
6801 TELEGRAPH ROAD
ALEXANDRIA, VA 22310-3398

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	930301	Technical 860925 – 911130

4. TITLE AND SUBTITLE
Predicting Radiation Induced Performance Decrements of AH-1 Helicopter Crews, Volume 2 – Evaluation of Modeling and Simulation Techniques for Predicting Radiation Induced Performance Decrements

5. FUNDING NUMBERS
C - DNA 001-86-C-0308
PE - 62715H
PR - RM
TA - RK
WU - DH011650

6. AUTHOR(S)
William A. Perez, Vaughan W. Inman, Joseph I. Peters, Robert R. Sanchez and Robert W. Young

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Science Applications International Corporation
P.O. Box 1303
McLean, VA 22102

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Nuclear Agency
6801 Telegraph Road
Alexandria, VA 22310-3398
RARP/Young

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
DNA-TR-92-54-V2

11. SUPPLEMENTARY NOTES
This work was sponsored by the Defense Nuclear Agency under RDT&E RMC Code B3500864662 RM RK 00119 25904D.

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

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report is the second in a series that investigates the effects of intermediate doses of ionizing radiation on Army helicopter crew performance. This report documents the analysis of performance on the Walter Reed Performance Assessment Battery (WRPAB), and the development of an initial model for the WRPAB using Multidimensional Scaling (MDS) techniques. The WRPAB was used in a study where 20 AH-1 pilots (1) predicted the effects of various symptom complexes on their performance, (2) went through a 36-hour protocol to induce symptoms similar to symptoms that follow exposure to intermediate doses of radiation, (3) performed a simulated AH-1 mission before and after symptom induction, and (4) were administered the WRPAB approximately every 2 hours throughout the 36-hour protocol. MicroSAINT models for AH-1 tank engagements were developed that used PAB performance to predict AH-1 performance degradation. The results indicate that the WRPAB (or some other test battery) has utility for the prediction of operator performance as affected by stressors such as ionizing radiation.

14. SUBJECT TERMS

Performance Assessment Battery	Helicopter	Modeling
Intermediate Dose Program	Radiation Sickness	MicroSAINT
Symptom Simulation	AH-1 Flight Simulation	Multidimensional Scaling

15. NUMBER OF PAGES
118

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

CLASSIFIED BY:

N/A since Unclassified.

DECLASSIFY ON:

N/A since Unclassified.

EXECUTIVE SUMMARY

This study examined employment of a performance assessment test battery to predict operational performance in an AH-1 helicopter. The Walter Reed Performance Assessment Battery (WRPAB) was administered to AH-1 pilots and gunners over a 36 hour protocol designed to induce symptoms mimicking radiation sickness. During the 36 hour protocol, the crew members also engaged in simulated helicopter attacks (TOW or Rocket) against tanks in an AH-1 weapon system trainer. The simulated helicopter engagements were conducted while the crew members were healthy feeling fine (baseline performance), and after undergoing a 36 hour protocol that included sleep deprivation, strenuous exercise, and spinning in a barany chair to induce nausea (post stressor performance).

Sixteen different tests from the WRPAB were administered at regular intervals during the protocol. The tests were presumed to reflect perceptual, cognitive, motor processing, and affect. The results of each test were analyzed for the effects of stressors over time. Four out of 13 performance based tests showed statistically significant stressor effects. All of the affect measures showed statistically significant stressor effects.

An initial performance prediction model was developed to relate WRPAB performance to helicopter tasks. Multidimensional Scaling (MDS) techniques were used in model development. The input to the MDS analyses were ratings of similarities among the 13 WRPAB performance tests provided by researchers currently working in the area of performance evaluation. The researchers were tested on the WRPAB prior to providing ratings of similarity. The MDS analyses suggested 3 dimensions of performance that together accounted for 78.7% of the variance in the resulting MDS space. These dimensions were: Verbal/Spatial, Visual/Motor, and Cognitive Processing.

For the purpose of relating WRPAB results to AH-1 tasks, the thirteen performance tests were classified as verbal, motor, spatial, or cognitive tests. For example, the Verbal/Spatial dimension contained verbal and spatial WRPAB tests. The verbal and spatial tests were on opposite ends of the scale for this dimension. To facilitate mapping of WRPAB tests to AH-1 tasks, tests from the Verbal/Spatial

dimension were separately grouped as verbal or spatial. This method was used to derive four test groupings: verbal, motor, spatial, and cognitive.

For each test grouping (motor, verbal, spatial, and cognitive), composite performance scores were computed. The method for computing composite scores was as follows: (1) the performance scores (throughput or coefficient of variation) for the tests were transformed into z-scores and average z-scores were computed as a function of test grouping; (2) the average z-scores were transformed into probabilities using the standard normal distribution; and (3) the probabilities were employed to compute performance decrements for each test grouping (motor, verbal, spatial, and cognitive) as a function of stressor. For example, if the baseline probability for the verbal tests was 0.791 and the probability after spinning was 0.684, a performance decrement of -0.135 was computed. With this method, performance decrements based on the WRPAB were computed as a function of motor, verbal, spatial, and cognitive processing rather than for each individual test.

WRPAB test performance was related to AH-1 tasks through the use of MicroSAINT modeling. MicroSAINT models were developed for AH-1 TOW and Rocket engagements. Each node in the MicroSAINT model was classified in terms of motor, verbal, spatial, and cognitive information processing requirements. Equations were developed to relate performance decrements in motor, verbal, spatial, and cognitive processing to AH-1 task performance.

The MicroSAINT models were also used to derive AH-1 crew performance estimates under conditions that were not tested in the man-in-the-loop simulation. The model considered the combined performance of two ill crew members; the man-in-the-loop simulation only included a single ill crew member for experimental control reasons. Also, the MicroSAINT modeling approach considered total engagement times rather than specific task times in the estimation of the effects of radiation symptoms on crew performance.

The WRPAB-based model results were compared to the results of the man-in-the-loop simulation through the conduct of multiple modeling runs (Monte Carlo). Distributions of engagement times, rather than single point estimates were compared. The predictions derived with the WRPAB MicroSAINT models were correlated with the observed performance. However, the WRPAB-based MicroSAINT model predicted

greater performance decrements than were actually observed in the AH-1 man-in-the-loop simulation.

The results of the MicroSAINT modeling experiments presented in Inman et al (1991) were also compared against the current modeling results. The MicroSAINT models in Inman et al used the PDQ methodology to derive distributions of engagement times. The results showed that the PDQ-based modeling technique also over estimated performance decrements, but not as greatly as did the WRPAB based models.

The WRPAB based performance estimation approach shows promise for the estimation of performance in operational systems. This report presents the first attempt to use this methodology to estimate the effects of radiation illness on military performance. Suggestions for improving the fit between predicted and derived performance decrements are presented.

PREFACE

The authors wish to express thanks to the many individuals who contributed to this effort. Major Bruce West, and Major Robert Kehlet of DNA provided coordination with the Army. Members of the Intermediate Dose and Human Responses Programs aided significantly in research program definition. Dr. Charles N. Davidson and LTC Douglas L. Watson, of the U.S. Army Nuclear and Chemical Agency provided invaluable inputs to the approach and aided in securing troop support through FORSCOM.

Personnel of the 101st Airborne Division, Fort Campbell, Kentucky provided extensive support during the data collection process. Division personnel ensured that the AH-1 simulator was dedicated to this study throughout the 2 week data collection schedule. Special thanks go to the pilots at Fort Campbell who dedicated hundreds of hours under difficult conditions to support this study.

Special thanks go to J. Moyer of SAIC who is responsible for the task timeline analysis supporting this experiment. Thanks also to Steve Eschholz and Stephen Masterson, of SAIC, who were essential to the data collection effort. Dr. Roy DeHart, from the University of Oklahoma, was critical to the approval and implementation of the human use protocol. Mr. C. J. Percle served as second crew member for all simulator sessions and provided outstanding liaison with the Fort Campbell community.

This report was prepared by the Human Performance Technology Division of Science Applications International Corporation (SAIC) for the Radiation Sciences Directorate of the Defense Nuclear Agency (DNA) under contract number DNA 001-86-C-0308. As a part of DNA's Human Response Program addressing AH-1 helicopter crew performance, this report presents analyses of Walter Reed Performance Assessment Battery (WRPAB) test results. The data and results presented in this report were collected in a study that evaluated the Performance Decrement Questionnaire (PDQ) as a technique for estimating the effects of radiation sickness symptoms on AH-1 crew performance (Inman et al, 1991).

The objectives of the overall study were to:

(1) induce symptoms similar to those of low level radiation exposure,
(2) assess the PDQ methodology in the context of a helicopter simulation,
(3) develop MicroSAINT models of AH-1 based on simulation data and PDQ estimates, and
(4) evaluate the relationship between Walter Reed Performance Assessment Battery (WRPAB) performance, PDQ performance estimates, and helicopter simulator task performance using MicroSAINT models of AH-1.

The first three objectives were addressed by Inman et al (1992). This report addresses the fourth objective.

The data reported here are the results of research conducted at Fort Campbell, Kentucky. To achieve the study objectives, the researchers obtained 20 volunteers, 10 AH-1 pilots, and 10 AH-1 gunners who were: (a) Deprived of sleep for 36 hours; (b) Exercised rigorously; and (c) Rotated in a Barany chair to induce upper gastrointestinal distress. Each volunteer's symptoms, AH-1 helicopter task performance, and PAB performance were assessed periodically. Prior to induction of symptoms, each volunteer provided PDQ time estimates for the simulator performance of a crew person subjected to each of 43 symptom complexes. PDQ performance time estimates were obtained for each of 12 AH-1 tasks.

This report describes:

- subjects' self reported symptoms,
- performance on 15 WRPAB tests as a function of the 36 hour protocol,
- relationship between Symptom Check List (SCL) responses, and WRPAB performance,
- the results of Multidimensional Scaling (MDS) analysis to develop a framework for the WRPAB, and
- MicroSAINT model development and simulations based on the empirical task time means perturbed by PAB-based predictions of performance.

The views expressed within this report are solely the responsibility of the authors and are not necessarily those of the Defense Nuclear Agency, The U.S. Army or any other U. S. government agency.

CONVERSION TABLE

Conversion factors for U.S. customary to metric (SI) units of measurement

To Convert From	To	Multiply
angstrom	meters (m)	1.000 000 X E-10
atmosphere (normal)	kilo pascal (kPa)	1.013 25 X E+2
bar	kilo pascal (kPa)	1.000 000 X E+2
barn	meter^2 (m^2)	1.000 000 X E-28
British Thermal unit (thermochemical)	joule (J)	1.054 350 X E+3
calorie (thermochemical)	joule (J)	4.184 000
cal (thermochemical)/cm^2	mega joule/m^2(MJ/m^2)	4.184 000 X E-2
curie	giga becquerel (GBq)*	3.700 000 X E+1
degree (angle)	radian (rad)	1.745 329 X E-2
degree Fahrenheit	degree kelvin (K)	$t_K=(t°f + 459.67)/1.8$
electron volt	joule (J)	1.602 19 X E-19
erg	joule (J)	1.000 000 X E-7
erg/second	watt (W)	1.000 000 X E-7
foot	meter (m)	3.048 000 X E-1
foot-pound-force	joule (J)	1.355 818
gallon (U.S. liquid)	meter^3 (m^3)	3.785 412 X E-3
inch	meter (m)	2.540 000 X E-2
jerk	joule (J)	1.000 000 X E+9
joule/kilogram (J/Kg) (radiation dose absorbed)	Gray (Gy)	1.000 000
kilotons	terajoules	4.183
kip (1000 lbf)	newton (N)	4.448 222 X E+3
kip/inch^2 (ksi)	kilo pascal (kPa)	6.894 757 X E+3
ktap	newton-second/m^2 ($N-s/m^2$)	1.000 000 X E+2
micron	meter (m)	1.000 000 X E-6
mil	meter (m)	2.540 000 X E-5
mile (international)	meter (m)	1.609 344 X E+3
ounce	kilogram (kg)	2.834 952 X E-2
pound-force (lbf avoirdupois)	newton (N)	4.448 222
pound-force inch	newton-meter (N·m)	1.129 848 X E-1
pound-force/inch	newton/meter (N/m)	1.751 268 X E+2
pound-force/foot^2	kilo pascal (kPa)	4.788 026 X E-2
pound-force/inch^2 (psi)	kilo pascal (kPa)	6.894 757
pound-mass (lbm avoirdupois)	kilogram (kg)	4.535 924 X E-1
pound-mass-foot^2 (moment of inertia)	kilogram-meter^2 ($kg\ m^2$)	4.214 011 X E-2
pound-mass/foot^3	kilogram/meter^3 (kg/m^3)	1.601 846 X E+1
rad (radiation dose absorbed)	Gray (Gy)**	1.000 000 X E-2
roentgen	coulomb/kilogram (C/kg)	2.579 760 X E-4
shake	second (s)	1.000 000 X E-8
slug	kilogram (kg)	1.459 390 X E+1
torr (mm Hg, 0°C)	kilo pascal (kPa)	1.333 22 X E-1

*The becquerel (Bq) is the SI unit of radioactivity; Bp = 1 event/s.

**The Gray (Gy) is the SI unit of absorbed radiation.

TABLE OF CONTENTS

Section **Page**

TABLE OF CONTENTS (continued)

FIGURES

FIGURES (Continued)

TABLES

SECTION 1
INTRODUCTION

This report presents the results of a study to develop a methodology using the Walter Reed Performance Assessment Battery (WRPAB) as a predictor of AH-1 helicopter crew performance under conditions of low-dose radiation exposure. The study also compares WRPAB-based predictions with Performance Decrement Questionnaire (PDQ)-based predictions, and actual simulator-based measurements of aircrews with symptoms mimicking radiation sickness. A complete description of the PDQ-based prediction technique and simulator measurements are provided in a companion report by Inman et al (1991). The following paragraphs briefly summarize the PDQ-based methodology.

The Defense Nuclear Agency (DNA) has developed a methodology for relating soldier performance on a nuclear battlefield to the soldier's radiation exposure dosage and time since dosage. The methodology was developed under the DNA's Intermediate Dose Program (IDP) and is currently employed in the DNA's Human Response Program (HRP). The method the IDP used to derive performance decrement estimates was based on:

(1) a comprehensive assessment of acute radiation effects and symptoms in humans,
(2) quantified descriptions of radiation sickness symptoms,
(3) identification of dose and time after dose relations to symptoms,
(4) subjective estimates of combat crew performance under selected symptoms, and
(5) use of subjective performance estimates in models of combat crew performance as a function of dose and time after dose.

This report explores the use of WRPAB-based performance estimates as an alternative to subjective estimates obtained in step 4.

Subjective estimates obtained in Step 4 were acquired through the use of a Performance Decrement Questionnaire (PDQ). The PDQ methodology relies on the ability of subject matter experts (SME) to estimate the effects of symptoms on the speed they are able to accomplish military tasks. Inman et al. reported AH-1 SMEs'

estimated performance decrements for tasks of two AH-1 helicopter versus tank engagement scenarios, one using AH-1 TOW, and the other using rockets. The SMEs were also run through a research protocol that served to induce symptoms associated with low levels or radiation exposure. SMEs performed the tasks associated with the TOW and rocket attacks while "healthy feeling fine" and after exposure to sleep deprivation, exercise, and spinning to induce radiation-like symptoms. Mission performance data were collected in pre- and post-stressor sessions in an AH-1 simulator and actual performance decrements were computed. Moreover, test performance data were collected using the Walter Reed Performance Assessment Battery (WRPAB) in approximately 2 hour intervals throughout the 36 hour protocol. Unlike AH-1 performance data collection, WRPAB data collection was not dependent on the availability of a complex weapon system trainer, and WRPAB performance changes could be measured across numerous points in time. The correlation between predicted and actual performance decrements based on a limited number of volunteers' performance showed that, although the SMEs predicted which tasks would be affected by the radiation symptoms, SMEs did not provide accurate predictions of the magnitude of the effects. Generally, SMEs tended to overestimate the effects of symptoms on task duration. Overestimates differed by task and symptom severity, but tended to range between 0 and 50% of actual decrements. In general, overestimates increased with symptom severity.

The approach to performance decrement estimation reported here also relies on the radiation symptom classification of the current IDP methodology. This report focuses on WRPAB performance data collected during the AH-1 study (reported by Inman et al, 1991) and on the relationship of WRPAB performance to crew reported symptoms. Specifically, this study explores the utility of a methodology that employs MicroSAINT modeling and WRPAB performance data as predictors of radiation-induced performance effects in AH-1 tank engagement scenarios.

1.1 THE WALTER REED PERFORMANCE ASSESSMENT BATTERY (WRPAB).

The WRPAB was designed as a research tool for following performance changes over time, treatments, dosages or levels (Thorne, Genser, Sing & Hegge, 1985). The WRPAB presents a "menu" of tests, and Table 1-1 presents the listing of tests that were administered in this study. The software allows for the selection of a

specific set of tests that are subsequently presented, recorded, and scored. The Method section of this report presents a detailed description of the test battery and administration procedures.

Table 1-1. Listing of WRPAB tests used in the study.

Choice Reaction Time	Digit Recall
Encode/Decode	Interval Production
Logical Reasoning	Manikin
Mood Scale I	Mood Scale II
Pattern Recognition 1	Pattern Recognition 2
Serial Addition/Subtraction	Six-Letter Search
Stanford Sleepiness Scale	Timewall
Two-Column Addition	Two-Letter Search

The listing in Table 1-1 represents a subset from the WRPAB. Test selection for inclusion in the WRPAB was guided by the following set of criteria: (1) to represent a reasonable and realizable sample of elemental skills generally regarded as underlying many real-world tasks; (2) the ability to be administered briefly and repeatedly; (3) appropriateness to computer implementation; and (4) for known or expected sensitivity to physiological, psychological or environmental variables (Thorne, et al, 1985).

As a battery of tests, the WRPAB avails the user of just one, many, or all of the tests contained within it. For example, if an experimenter is just interested in logical reasoning, he or she could administer only this test and use the speed, accuracy, and throughput measures it provides. For the purposes of the AH-1 study (see Inman et al, 1991), all of the tests listed in Table 1-1 were given primarily as a diversion to avoid boredom for participants in the 36 hour sleep deprivation protocol. A secondary benefit was the availability of "context free" laboratory measures that might prove to be more sensitive than AH-1 mission measures to the experimental stressors.

There are no statistical norms for the tests in the WRPAB. Furthermore, the tests are not organized with respect to a theoretical or practical framework. The development of a theoretical or practical framework for the WRPAB would serve to assist researchers in the selection of specific tests for application. Also, the test battery does not contain tests in domains such as continuous tracking, auditory communication, complex decision making, etc.) that are representative of some components of military tasks. A framework or model for the WRPAB would also aid in

the identification of new tests for inclusion in the battery, and might be useful as a predictor of radiation-induced performance decrement.

This report presents the results of the initial steps in the development of a theoretical and practical model for categorizing the WRPAB test items. This effort was conducted subsequent to data collection in the AH-1 simulation study. For the conduct of the AH-1, all sixteen tests in Table 1-1 were presented to the crew members.

1.2 WRPAB MODEL DEVELOPMENT.

For the development of a model for the WRPAB, Multidimensional Scaling (MDS) techniques were employed. MDS uses data on the proximity of stimuli to produce a scaling solution that is depicted in an n-dimensional space. The scaling dimensions are then interpreted in terms of the common characteristics of proximate stimuli. The stimuli for this study were the tests in the WRPAB. Performance assessment SMEs (Human Factors and Experimental Psychologists) were run through the demonstration version of the WRPAB. After the SMEs received first hand experience with the WRPAB they rated the similarity of all pairs of tests (each subject performed 78 pair-wise similarity ratings). The SMEs were not informed as to what dimensions or factors to use in their ratings of test similarity, rather the intent of the MDS study was to **discover** the dimensions that the SMEs employed.

The result of the MDS study was a test similarity space that was used to: (1) identify the dimensions underlying the WRPAB, (2) identify similar or redundant tests, and (3) identify requirements for inclusion of new tests. The MDS solution and interpretation served as a guideline for the selection of tests from the WRPAB. The results of this study represent an initial step in the development of a theoretical framework for the PAB, the ability to characterize military tasks to these or similar dimensions, and then model scenario-based military tasks using WRPAB decrement data to modulate predictors of military task performance.

During the conduct of the AH-1 simulation study, the crew members were run through the WRPAB repeatedly over a 36 hour time period. The WRPAB thus provided performance based degradations as a function of the stressors that were employed in this study; i.e., sleep deprivation, strenuous exercise, and spinning (to induce upper gastrointestinal distress). These stressors appeared to reliably and validly produce

symptoms associated with exposure to low levels of radiation. Therefore, the WRPAB provided the potential for deriving performance-based estimates of soldier performance degradation. The remainder of this report presents the results of the analysis of the WRPAB and the initial framework that was developed.

SECTION 2
METHOD

This section presents the method and procedures for the AH-1 simulation study and for the Multidimensional Scaling (MDS) study. These two studies were conducted at different times and with different subject populations. Expanded discussion of the AH-1 simulation study can be found in Inman et al (1991).

2.1 AH-1 STUDY.

2.1.1 Subjects.

Twenty AH-1, "Cobra," helicopter crew members volunteered for this experiment. The volunteers were males between 23 and 31 years old. All were members of the 101st Airborne Division (Air Assault) at Fort Campbell, Kentucky. They averaged 474 flight hours in the AH-1, and 84 AH-1 combat weapons trainer simulator hours.

Volunteers were medically screened and monitored throughout the course of the 36 hour protocol. Lunch and dinner were served on the first day of the experiment and no food was served prior to completion of the test session on day two. Water was available throughout the experiment.

Half of the crew members flew in the pilot's seat, and the other half flew in the gunner's seat; all were tested in the crew position at which they normally work. The subjects were tested on the WRPAB approximately every two hours throughout the 36 hour protocol.

2.1.2 Apparatus & Materials.

2.1.2.1 AH-1 Simulator. The study included the use of an AH-1 weapon system trainer that contained two simulator modules. The AH-1 simulator modules are six degree of freedom motion-based weapons systems trainers. The modules have a high fidelity visual terrain board with a laser light source that produces high resolution imagery. The pilot and gunner stations are separate modules but operate in such a way that they are perceived as tandem stations of the same aircraft. That is, to the

simulator pilot and gunner, the events are presented so as to appear to be occurring to a single AH-1 attack helicopter with tandem crew seating.

2.1.2.2 Barany Chair. A Barany chair was installed between the pilot and gunner simulator modules. The chair, provided by the Naval Aerospace Medical Research Laboratory, was calibrated to spin at 25 revolutions per minute. The Barany chair was used to induce upper gastrointestinal distress.

2.1.2.3 Walter Reed Performance Assessment Battery. The WRPAB was run on a Compaq computer with a Scientific Solutions Lab Tender card installed. The Lab Tender card is a multifunction data processing board which provides response time measures accurate to 1 millisecond. The computer system had a monochrome monitor. WRPAB responses were recorded via the computer's keyboard. Sixteen tests from the WRPAB were used in the experiment. The following presents the instructions and sample screens of the WRPAB tests that were used.

Encode/Decode

Instructions: A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first two letters (3 digits) are for North/South, the last 2 letters (3 digits) are for East/West. For example, BCZM = 52 59 51 57 = 52_9 51_7 = 529 517 = 529517 (this example computation uses the sample screen presented on the next page). Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

NORTH/SOUTH												
Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62
EAST/WEST												
H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62

BCZM=

The correct response for this example is: 529517

Six-Letter Search

Instructions: Determine whether the 6 target letters at the top of the screen are present (in any order) in a line of letters in the middle of the screen. If all target letters are present, press "S" for same. If only some or none are present, press "D" for different. Work quickly but accurately.

Sample Screen

HXLMOG

UXNMORSAHIJFVGLBDCKY

The correct response for this example is: S

Two-Letter Search

Instructions: Determine whether the 2 target letters at the top of the screen are both present in a line of letters in the middle of the screen. If all target letters are present, press "S" for same. If only some or none are present, press "D" for different. Work quickly but accurately.

Sample Screen:

YU

UAXNDORSBVJKHLMFCETI

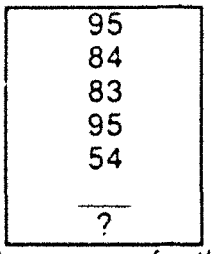
The correct response for this example is: D

Two Column Addition

Instructions: A column of 5 pairs of digits is presented. You are to add them as rapidly as possible. The answer is given by entering the left-hand digit first (usually the hundreds digit), followed in order by the remaining digits. End by pressing the "Return" key.

Be sure that you know the entire answer before entering the first digit because the column of numbers will disappear before your entire answer is entered. Work quickly but accurately.

Sample Screen:



The correct response for this example is: 411

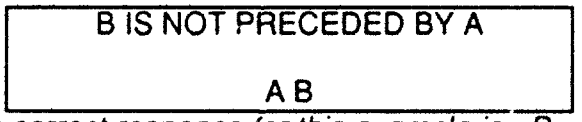
Logical Reasoning

Instructions: You will be presented with a series of statements about the relationship between two letters. Each statement will be followed by the two letters *AB* or *BA* You are to decide whether the statement correctly describes the order of the letters.

If it does, press the "S" key for "same";
If it does not, press the "D" key for "different."

At first, many people have difficulty with the relationships. It is extremely important that you understand all of them before the training sessions are complete.

Sample Screen:



The correct response for this example is: D

Digit Recall

Instructions: You will be presented with a series of numbers that you are to remember. A row of 9 digits will be presented for a very short time in the middle of the screen. Next the screen will go blank for several seconds. Finally, 8 of the 9 digits will be presented in an order different from the first.

Your job is to enter the missing digit.

Remember that speed as well as the accuracy of your response is being recorded, so don't waste time. If you really have no idea what the answer should be, make a best guess and go on to the next trial.

This task can be quite difficult, so you need to be as attentive as possible.

Sample Screen:

First String	9 3 5 2 8 1 0 8 6
Second String	0 8 6 9 8 3 5 2

The correct response for this example is: 1

Serial Addition/Subtraction

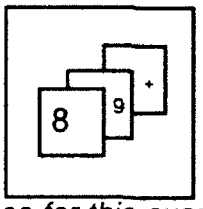
Instructions: Two digits will flash in succession followed by either a Plus or Minus sign. You are to perform the operation indicated by the sign (addition or subtraction) and enter the last digit of your answer.

Thus, 9, 8, +, would require you to add 9 to 8 and enter 7 (the last digit of 17). If 7, 4, -, were presented, you would subtract 4 from 7 and enter 3.

If subtraction results in a negative number, add 10 to the result and enter the last digit of the answer. Thus 3, 9, -, would require that you subtract 9 from 3 to get -6, then add 10 and enter 4.

The numbers appear for only a very short time so you will have to be attentive.

Sample Screens:



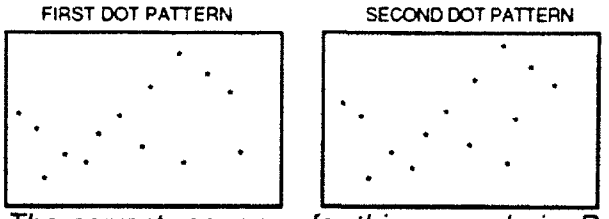
The correct response for this example is (8+9=17): 7

Pattern Recognition I

Instructions: A random pattern of asterisks will be displayed for a very short time and then the screen will be blank for several seconds. A second "test" pattern will then be displayed. You are to decide whether the "test" pattern is the same as the first pattern or if it is different. You must decide as quickly as possible and enter "S" if it is the same, or "D" if it is different.

The first pattern will only be present for a very short time, so look at it closely and try to remember it during the subsequent retention interval. As soon as you enter your answer, the next pattern will be presented.

Sample Screen:



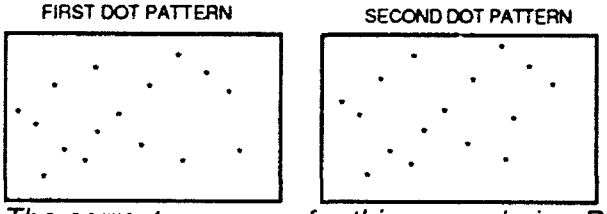
The correct response for this example is: D

Pattern Recognition II

Instructions: A random pattern of asterisks will be displayed for a very short time and then the screen will be blank for several seconds. A second "test" pattern will then be displayed. You are to decide whether the "test" pattern is the same as the first pattern or if it is different. You must decide as quickly as possible and enter "S" if it is the same, or "D" if it is different.

The first pattern will only be present for a very short time, so look at it closely and try to remember it during the subsequent retention interval. As soon as you enter your answer, the next pattern will be presented.

Sample Screen:



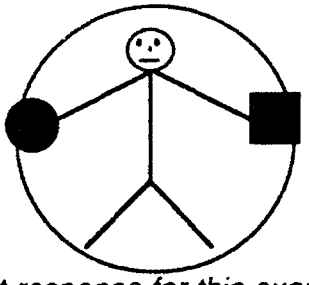
The correct response for this example is: D

Manikin

Instructions: You will see a man inside either a circle or a square. He will be holding a circle in one hand and a square in the other. You must decide as quickly as you can which hand holds the matching shape and press either the left or right key accordingly.

Place your fingers on the bottom row of keys with your left index finger on the V and your right index finger on the M. Press only with one index finger.

Sample Screen:



The correct response for this example is: M

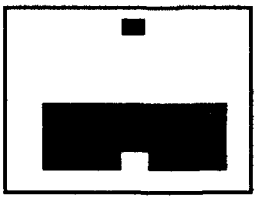
Time Wall

Instructions: A small square will appear at the top of the screen and will move downward at a constant velocity toward a shielding barrier. When the square reaches the barrier you will no longer see it.

Your job is to estimate how long it will take the square to reach the small notch at the other side of the barrier.

Indicate the moment you think the square has reached the notch by pressing the space bar. When you press the space bar the notch will close for a moment and then the next trial will begin.

Sample Screen:



Time Production

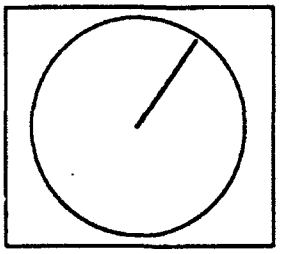
Instructions: This task requires you to tap a key at regular one-second intervals. You are not to use a watch or clock. Estimate each one-second period as accurately and as consistently as you can.

If you are right handed, place your fingers comfortably on the lower row of keys labelled "M" through "?" with your thumb against the front of the space bar.

If you are left handed, use the keys "Z" through "V".

Tap only with your index finger, using a single brief "down-up" movement. If you hold the key too long it will cause multiple taps and you will be asked to repeat the test. Try a few practice taps now so you can get a feel for the correct force and duration to use. Tap once a second, as accurately and consistently as you can.

Sample Screen:



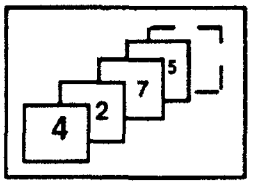
Choice RT

Instructions: A single number will appear in the center of the screen and you must press the same number as quickly as possible. The screen will go blank for a moment an then another number will appear.

If you are left handed, ask the technician for further instructions. If you are right handed, use the numeric keypad on the right of the keyboard. Rest your thumb on the zero key and your first 3 fingers on the keys 4, 5, & 6. You will find it helpful to return to this "home position" after each response.

This is a speed test, so respond as quickly as you can.

Sample Screen:



The correct responses to these screens are: 4; 2; 7; 5; ...

Mood Scale I

Instructions: You will be given a list of words which people often use to describe how they feel; followed by the numbers 1 through 5.

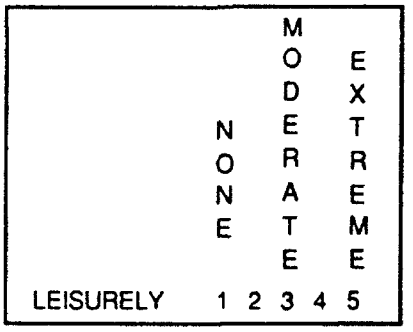
These numbers represent the degree to which each word describes how you feel:

1 Means you do not at all feel that way,
3 Means you feel moderately that way, and
5 Means you feel extremely that way.

Thus, if you feel moderately lucky press "3" in response to the word, "LUCKY". If at the same time you feel more than moderately gloomy but less than extremely gloomy press "4" in response to the word, "GLOOMY".

Work at the fastest comfortable rate you can and give your first reaction to each word. Do not try to remember previous answers even though some words may seem similar.

Sample Screen:



Mood Scale II

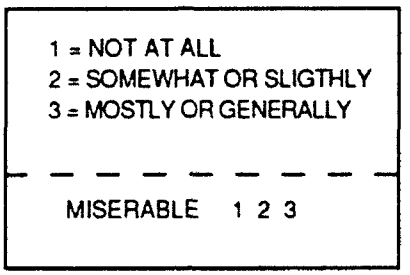
Instructions: You will be given a list of words which people often use to describe how they feel; followed by the numbers 1 to 3.

These numbers represent the degree to which each word describes how you feel.

1 Means not at all,
2 Means somewhat or slightly, and
3 Means mostly or generally.

These numbers and their meanings will be repeated with each word. Indicate how each word applies to how you feel now, by pressing 1, 2, or 3.

Sample Screen:



Stanford Sleepiness Scale

Instructions: For this test you will be presented a single trial where you are to choose one of seven statements that best describes your present feeling. That is, how you feel right now. The seven statements are as follows:

1 Feeling active and vital; alert; wide awake.
2 Functioning at a high level, but not at peak; able to concentrate.
3 Relaxed; awake, responsive, but not at full alertness.
4 A little foggy; let down; not at peak.
5 Foggy; slowed down; beginning to lose interest in remaining awake.
6 Sleepy; woozy; prefer to be lying down; fighting sleep.
7 Almost in reverie; sleep onset soon; loosing struggle to remain awake.

Sample Screen

CHOOSE ONE OF THE SEVEN STATEMENTS BELOW WHICH BEST DESCRIBES YOUR PRESENT FEELING. HOW YOU FEEL RIGHT NOW

1. Feeling active and vital; alert; wide awake.
2. Functioning at a high level, but not at peak; able to concentrate.
3. Relaxed; awake, responsive, but not at full alertness.
4. A little foggy; let down; not at peak.
5. Foggy; slowed down; beginning to loose interest in remaining awake.
6. Sleepy; woozy; prefer to be lying down; fighting sleep.
7. Almost in reverie; sleep onset soon; loosing struggle to remain awake.

2.1.2.4 Performance Decrement Questionnaire. A Performance Decrement Questionnaire was used to obtain subjective estimates of AH-1 task times. Estimates were obtained for six pilot tasks and six gunner tasks. Figure 2-1 presents the gunner and pilot tasks sequentially as they were experienced in the simulated tank engagements. Some of the tasks were part of a tube-launched, optically-tracked, wire-guided (TOW) launch scenario, and some were part of a rocket launch scenario.

Time estimates were obtained for performance under each of 43 symptom complexes. Each symptom complex was described on a separate sheet of paper. At the top of the sheet the "symptom" associated with the symptom complex was described. Below the symptom description was the question "how sick are you?" Two responses were solicited to this question: (1) a sickness rating on a scale from 1 to 20 where 1 is well and 20 is most sick, and (2) a general performance rating in percent, where 100% would indicate no performance decrement. Below this question was the typical PDQ questionnaire, where respondents provided discrete estimates of the amount of time specific tasks would take while under the influence of the symptoms described at the top of the response page. Figure 2-1 provides an example of one of the PDQ response sheets.

SYMPTOM: Unsteady upon standing quickly.

How sick are you? Scale 1-20 _____ Overall performance _____%

TOW LAUNCH TASKS			HOW LONG DO YOU THINK IT WOULD TAKE YOU TO DO EACH IF YOU HAD THESE SYMPTOMS?		
			NO INCREASE IN TIME	INCREASE IN TIME	COULD NOT DO IT AT ALL
1.	PILOT:	Unmask and Identify Target			
2.	GUNNER:	Acquire Target			
3.	GUNNER:	Track target and Switch to High Magnification			
4.	GUNNER:	Action Bar			
5.	Pilot	Position Aircraft Into Prelaunch Constraints			
6.	GUNNER:	Fire TOW			
7.	GUNNER:	Track Missile to Target			
8.	PILOT:	Remask AH-1 to 5 Ft. AGL			
ROCKET LAUNCH DIRECT MODE TASKS					
1.	PILOT:	Unmask and Identify Target	SAME AS TOW LAUNCH		
2	GUNNER:	Acquire Target			
3.	GUNNER:	Track Target and Switch to High Magnification			
4.	GUNNER:	Fine Adjust, Lase Target and Confirm Range			
5.	PILOT:	Position Aircraft; Fly Reticle to Target and Fire			
6.	PILOT:	End of Firing			
7.	PILOT:	Remask AH-1 to 5 Ft. AGL			

Figure 2-1. PDQ Response Sheet for one of the Symptom Complexes.

2.1.3 Procedure.

Each volunteer participated for a period of 36 hours in which he:

(1) remained awake,
(2) had food withheld after the first 12 hours,
(3) took the WRPAB approximately every 2 hours,
(4) filled in the symptom check list every 2 hours,
(5) flew four simulator missions,
(6) answered the PDQ,
(7) performed vigorous exercises approximately 30 hours into the experiment, and
(8) was spun in the Barany chair at approximately 34 hours into the experiment.

Details for the above procedures are presented by Inman et al (1991). The following presents an overview of the procedures to provide the context in which the WRPAB was administered.

2.1.3.1 Briefing. The experiment was conducted over a two and one-half week period. Two subjects, one pilot and one gunner, reported each weekday morning at 0700.

Immediately upon arrival at the experimental test site, the volunteers were briefed on the purpose of the experiment. The briefing contents were part of the Volunteer Agreement Affidavit (VAA). A tape recording of the affidavit was played to each volunteer in the presence of the principal investigator and the physician. The volunteers were asked to read the VAA as the tape was played. The principal investigator and physician answered any questions that the volunteer had to insure that the purpose, objectives, and risks associated with participation in the experiment were understood. Volunteers initialed each page of the affidavit to acknowledge understanding of its contents.

2.1.3.2 Vital Signs. After signing the consent forms, each crew member was examined by the physician. Blood pressure, heart rate, and weight measures were obtained. No abnormalities were found in any of the volunteers. After the physical examination, the volunteers began pre-flight activities. After pre-flight activities were completed, the volunteers were separated. One volunteer performed a practice simulator session while the other volunteer performed a practice PAB session.

2.1.3.3 Pre-flight Activities. Pre-flight activities began with completion of a pre-flight questionnaire. The pre-flight questionnaire consisted of a biographical data sheet and questions on previous experienc with the symptoms described in the PDQ.

Prior to entering the simulator, the subjects reviewed the simulator mission. The review included observation of the engagement location on a map, review of weather conditions, aircraft weight and balance, weapons load, and radio frequencies.

2.1.3.4 Symptom Check List (SCL). Before and after every simulator session, each subjected completed an SCL (see Figure 2-2). The SCL was a listing of symptoms with instructions for subjects to circle the symptom that described how they felt. In addition to SCLs associated with simulator sessions, subjects completed an SCL approximately every two hours during their study participation.

2.1.3.5 Walter Reed Performance Assessment Battery. The subjects were provided approximately 45 minutes of practice on the PAB. After this initial introduction to operation of the computer and the PAB, the PAB was self-administered. The PAB was performed approximately once every 2 hours throughout the course of the experiment.

2.1.3.6 Simulated Tank Engagements. A total of four simulator sessions were presented to each crew member, two on the first day and two on the second day.

Within each session, two types of tank engagements were conducted in the AH-1 weapon system trainer, one employing TOW missiles, and the other using 7 pairs of 2.75 inch rockets. Both engagement types were part-mission simulations that consisted of unmasking, engaging a tank target, and remasking.

How do you feel right now? (Please circle one number under each section below.)

UPPER GASTROINTESTINAL DISTRESS

1. No effect.
2. Upset stomach; clammy and sweaty, mouth waters and swallows frequently.
3. Nauseated; considerable sweating; swallows frequently to avoid vomiting.
4. Vomited once or twice; nauseated and may vomit again.
5. Vomited severely times including the dry heaves; severely nauseated and will soon vomit again.

LOWER GASTROINTESTINAL DISTRESS

1. No effect.
2. Feels uncomfortable urge to defecate.
3. Occasional diarrhea; recently defecated and may again.
4. Frequent diarrhea and cramps; defecated several times and will again soon.
5. Uncontrollable diarrhea and painful cramps.

FATIGABILITY/WEAKNESS

1. No effect.
2. Somewhat tired with mild weakness.
3. Tired, with moderate weakness.
4. Very tired and weak.
5. Exhausted with almost no strength.

HYPOTENSION

1. No effect.
2. Slightly light-headed.
3. Unsteady upon standing quickly.
4. Faints upon standing quickly.
5. In shock; breathes rapidly and shallowly, motionless, skin cold, clammy, and very pale.

INFECTION AND BLEEDING

1. No effect.
2. Mild fever and headache - like starting to come down with flu.
3. Joints ache, considerable sweating; moderate fever; does not want to eat; sores in mouth/throat.
4. Shakes and chills and aches all over; difficult in stopping any bleeding.
5. Delirious, overwhelming infections; cannot stop any bleeding.

FLUID LOSS AND ELECTROLYTE IMBALANCE

1. No effect.
2. Thirsty and has dry mouth; weak and faint.
3. Very dry mouth and throat, headache, rapid heartbeat and may faint with moderate exertion.
4. Extremely dry mouth, throat, skin and very painful headache; has difficult moving; short of breath, burning skin and eyes.
5. Prostrate.

Symptom Check List

Figure 2-2. The symptom check list given every two hours, and before and after each simulator session.

For each crew member, there was a total of 20 trials per session, 10 TOW engagements and 10 rocket engagements. Each trial began with the AH-1 already in a masked position at the engagement location. The simulator was reset at the beginning of every trial. Thus, for every trial the simulated helicopter had the same fuel and armament load, i.e., full fuel, 8 TOWs, 7 pairs of 2.75 inch folding fin aerial rockets, and 750 20mm rounds.

Dependent measures were the time to complete engagement tasks and accuracy of the TOW missile, or 7 pairs of rockets. Task durations were measured from analysis of audio tapes of pilot, gunner, and simulator operator verbal behavior. Although limitations in the design of the training simulator precluded use of automatic collection of task times, TOW and rocket accuracy data were obtained from a video display at the simulator's instructor station.

The subjects were run through two simulator sessions during the first day of the experiment. Simulator Session 1 was for practice to provide each crew member some familiarity with the scenario and tasks. Each pilot/gunner pair went through the simulation separately with a "control" crew member performing the duties at the other position. The control crew member was used throughout the experiment to ensure that any performance decrements that appeared were attributable to one crew person's symptoms rather than an interaction between two "sick" crew members.

Simulator Session 2 occurred 2 hours after the practice session. Session 2 provided baseline simulator performance data. Session 2 was conducted between 1000 and 1200 hours on the first day when participants were expected to be "healthy, feeling fine."

The subjects were also run through two simulator sessions on the second day. Session 3 began in the early afternoon on the second day and was intended as refresher practice to ensure retention of the scenario-specific skills required for the study. Session 4 was the treatment session in which accumulated fatigue and weakness was compounded by the barany chair to stimulate upper gastrointestinal distress. The major sessions for performance comparisons were the baseline Session 2 on the first day, and the treatment Session 4 late on the second day.

2.1.3.7 Performance Decrement Questionnaire. Following a lunch break on the first day, the subjects completed the 43 PDQ response sheets. The PDQ procedure was to

(1) estimate task durations for each of the 12 tasks assuming the crew person is healthy and feeling fine;
(2) provide responses to the "How sick are you" question for all 43 symptom complexes;
(3) review each of the 42 symptomatic complexes and determine whether the respondent believed he would;
    (a) experience no increase in time to complete the task,
    (b) experience an increase in time to complete the task, or
    (c) not be able to do the task at all;
(4) estimate the increased task duration for those descriptions which were sorted into the *b* category.

2.1.3.8 Meals. Meals were served at 1230 and 1630 on the first day. Participants received no other food between reporting to the experiment on day one and completion of all experimental activities on day two. Food deprivation was included in the experimental treatment as a prophylactic measure. That is, because treatment on the second day included inducement of nausea, food deprivation was employed to minimize the possible risk associated with aspiration of vomitus.

2.1.3.9 Exercise. At either 1000 hours (pilots) or 1100 hours (gunners) on the second day, all participants engaged in a series of physical exercises. The exercise period lasted 42 minutes. During the first ten minutes of exercise, participants were encouraged to stretch and run in place so as to avoid injuries during the upcoming half-hour. During the subsequent 32 minutes the participants engaged in a series of 8 exercises. Each of the 8 exercises was preceded by 2 minutes of rest. The instructions for all exercises were to "do as many (repetitions) as possible in 2 minutes." In order of performance, the exercises were:

(1) jumping jacks,
(2) push ups,
(3) squat thrusts,
(4) leg lifts,

(5) flutter kicks,
(6) body twists,
(7) sit ups, and
(8) shuttle run.

The exercises were selected to enhance fatigue.

Instructions for the exercises were both verbal and by demonstration. The same research assistant instructed all subjects.

The number of repetitions performed was not recorded. However informal observation suggested that all subjects made a serious attempt to comply with the instruction to do as many repetitions as they could. In an effort to provide some extra motivation, the experimenters posted the name and number of repetitions of the subject with the highest number of repetitions on each exercise. The posting was updated in the subject's presence, as he completed each exercise. All the volunteers were from the same squadron, and thus knew each other.

2.1.3.10 Spin. Prior to the last simulator session, each subject was spun in the Barany chair in order to induce nausea. Participants were asked to remain in the chair until they experienced stomach discomfort. The Barany chair was first accelerated at a rate of 5 degrees per second to a constant velocity of 25 rotations per minute. Beginning at 60 seconds after the start of the rotation, subjects were paced through a series of head movements. The physician paced his instructions to the subjects with a metronome. The head movements were performed in four steps: (1) tilt forward 45 degrees, (2) tilt back 45 degrees, (3) tilt left 45 degrees, (4) tilt right 45 degrees, with return to center following each step. A movement cycle was paced at two seconds such that the movement away from center took one second and the return took one second. The series of head movements were repeated every 20 seconds until the subject expressed stomach discomfort.

Following the last simulator session the subjects were debriefed. After the experiment was completed, each crew member was driven home.

2.2 MULTIDIMENSIONAL SCALING STUDY.

2.2.1 Subjects.

There were 8 subjects in the MDS study. The subjects were practitioners in the area of human performance evaluation. All subjects were administered the WRPAB test described in Section 2.1.2.3 of this report prior to providing proximity ratings .

2.2.2 Apparatus and Materials.

The WRPAB software described under Apparatus and Materials for the AH-1 study was used. However, the subjects ran through the demonstration version of the software rather than the experimental trials.

A questionnaire was developed to obtain subjects' pair-wise ratings of similarity for WRPAB tests. The following tests were excluded from the MDS study: Mood Scale I, Mood Scale II, and Stanford Sleepiness Scale. These tests were excluded because they would have likely resulted in a dimension representative of performance versus non-performance based tests. In order to allow for the identification of as many performance based dimensions as possible, the non-performance based tests were excluded.

Appendix A contains sample pages from the questionnaire. Each subject was shown a different order of presentation for the pair-wise comparisons. Each subject made 78 pair-wise ratings of similarity. The ratings were made using a 7 point scale where 1 was anchored as "Very Similar" and 7 was "Very Dissimilar".

In addition to the pair-wise ratings of similarity, the subjects rated each test with respect to nine different dimensions underlying human information processing. The dimensions or factors that were selected represent an initial list that may need to be refined or expanded. The following are the selected dimensions or factors and the definitions used in the study (the definitions are quoted from the questionnaire):

Verbal

Verbal ability refers to language skill or knowledge. Rate according to the extent you feel individual differences in verbal performance, or changes in verbal performance over time, are reflected in test performance.

Visual-Spatial

Visual-spatial ability refers to non-language skill or knowledge pertaining to visual information processing. Rate according to the extent you feel individual differences in visual-spatial performance, or changes in visual-spatial performance over time, are reflected in test performance.

Short-Term Memory

In this context, short-term memory is used as a non-theoretical construct. It is meant to exclude memory for instructions or learning that the subject brings to the experiment. It includes memory that may be referred to as primary memory, secondary memory, or working memory among others. Rate according to the extent you feel individual differences in short-term memory performance, or changes in short-term memory performance over time, are reflected in test performance.

Visual Search

Visual search refers to scanning for the location of a target item. Rate according to the extent you feel individual differences in visual search performance, or changes in visual search performance over time, are reflected in test performance.

Motor Coordination

Motor coordination refers to manual dexterity to include rhythmic movement. Rate according to the extent you feel individual differences in

motor co-ordination, or changes in motor coordination over time, are reflected in test performance.

Reasoning

Reasoning refers to logical thought. Rate according to the extent you feel individual differences in logical ability, or changes in logical skill over time, are reflected in test performance.

Mental Manipulation

Mental manipulation refers to the ability to perform mental transformations. Rate according to the extent you feel individual differences in mental transformation ability, or changes in mental transformation skill over time, are reflected in test performance.

Time Perception/Estimation

Time perception refers to the ability to correctly judge the passage of time, or time intervals. Rate according to the extent you feel individual differences in time interval estimation ability, or changes in time interval estimation over time, are reflected in test performance.

Item Difficulty

Two different difficulty dimensions are of interest: *item difficulty* and *difficulty importance*. For item difficulty, consider only the overall difficulty of the test, **not whether test difficulty discriminates between individuals.**

Difficulty Importance

Two different difficulty dimensions are of interest: *item difficulty* and *difficulty importance*. For difficulty importance rate the extent you feel individual differences in performance, or changes in performance over

time, are the result of the test's level of difficulty. **Do not rate overall difficulty under this construct.**

Time Pressure

Time pressure refers to the importance of rapid performance. Rate according to the extent you feel individual differences in performance, or changes in performance over time, are the result of the requirement to perform quickly.

Appendix B presents sample pages from the rating scales. Each test was rated separately on a seven point scale.

2.2.3 Procedure.

The subjects were presented a package which contained: (1) an introduction to the study and study objective; (2) instructions for running the WRPAB demonstration program; (3) descriptions of each test with sample screens and instructions (note, the demonstration version of the WRPAB program presents the instructions on the computer monitor); (4) a section which included descriptions, sample screen, and rating scales for 78 pairs of tests, and (5) the univariate rating scales for each of the 13 tests.

The subjects followed the following procedure:

1) Read the introduction to the study and the directions.

2) Ran through the practice version of the WRPAB (approximately 20 minutes).

(3) Completed the 78 pair-wise comparisons (approximately 45 minutes).

(4) Completed the univariate ratings for each test (approximately 20 minutes).

SECTION 3
RESULTS

Performance Decrement Questionnaire results are reported in Volume I of this report. WRPAB results are reported here. Each WRPAB test was analyzed separately to characterize the effects of sleep deprivation, strenuous exercise, and spinning. The relationship among the tests in terms of their correlation (overlap) is examined. The relationship between the subject's reported health state (SCL ratings) and WRPAB performance is also presented. Finally, the outcome of employing MDS results in the prediction of AH-1 TOW and Rocket engagement outcomes is presented.

3.1 WRPAB INDIVIDUAL TEST ANALYSIS.

Sixteen tests from the WRPAB were used in this study. The tests cluster into 3 categories based on the commonality of their dependent measures. The following tests share common dependent measures:

- Encode/Decode
- 2-Letter Search
- 6-Letter Search
- 2-Column Addition
- Logical Reasoning
- Digit Recall
- Serial Add/Subtract
- Pattern Recognition I
- Pattern Recognition II
- Manikin
- Choice Reaction Time

The common dependent measures were:

$$(1)\ \text{Accuracy} = \%\ \text{Correct} = \frac{\text{Number Correct Responses}}{\text{Total Number of Responses}}$$

$$(2)\ \text{Speed} = \text{Number of Responses per Minute} = \frac{60\ \text{sec/min}}{\text{Mean sec/response}}$$

(3) Throughput = Hits per Minute = % Correct x Responses per Minute

The tests in the second cluster (based on dependent measure commonality) were:

- Time Wall
- Interval Production

Their common dependent variables was:

$$\text{(4) Coefficient of Variation} = \frac{\text{Standard Deviation of Responses}}{\text{Average Response Time}}$$

The third grouping of tests included: Mood Scale I, Mood Scale II, and the Stanford Sleepiness Scale. Each of these tests has a unique dependent measure. The data for Mood Scale I was corrupted (the data for this test could not be retrieved from disk) and therefore the results for this test will not be presented.

Mood Scale II: The following dependent measures are recorded by the WRPAB computer program; Average Rating, Percent Responses Greater than 1, and Average Reaction Time as a function of the following categories: activity, fatigue, happiness, depression, anger and fear. Only the average ratings for each of the six categories are reported here.

Stanford Sleepiness Scale: The following dependent measures are are recorded by the WRPAB computer program; Sleep Score (1 through 7), Reaction Time (latency from space bar pressing to score selection), and Start Latency (latency from start of task to space bar depression). Only the sleep score data are reported here.

Each subject was administered the WRPAB *approximately* every 2 hours. Because (a) not all subjects were tested an equal number of times, and (b) the schedule of testing was not strictly maintained for all subjects (the subjects were not tested at exactly 2 hour intervals and some subjects missed one or more two hour assessments), actual test time could not be used as a classification variable in an analysis to evaluate group performance. Therefore, for the purpose of analysis, 9 time

windows were defined. The windows were defined to compensate for the above problems and still provide sufficient resolution to assess the effects of sleep deprivation and other stressors (strenuous exercise and spinning) on test performance.

With the exception of the Post-Exercise and Post-Spin windows, the data were averaged over 4-5 WRPAB 2 hour test periods. The data for Post-Exercise and Post-Spin are each from single WRPAB administrations. The windows were defined as follows:

- start to 1430 hrs
- 1431 to 2030 hrs
- 2031 to 0030 hrs
- 0031 to 0430 hrs
- 0431 to 0830 hrs
- 0831 to 1130 hrs
- 1131 to 1430 hrs
- Post-Exercise
- Post-Spin

The following sections present analyses of the dependent measures for each of 15 WRPAB tests as a function of time window. Throughput, speed, and accuracy analyses are reported for the eleven tests that have common dependent measures. For Time Wall and Interval Production, analyses of the coefficient of variation are reported. For Mood Scale II, the percent of responses greater than 1 were analyzed as a function of the 6 subscales. The analysis of the Stanford Sleepiness Scale used the Sleepiness Score. One-way within subjects analysis of variance are reported for each test.

In order to evaluate the effects of sleep deprivation and stressors (exercise and spinning), three groupings of means were formed for each dependent measure that was analyzed. The groups consisted of: (1) Baseline, the first two windows ; (2) Sleep Deprivation, windows 4, 5 and 6; and (3) Post Stress, post exercise and post spin windows. Average performance was computed for the three groupings. The statistic employed to test differences between means for the three groupings was a set of contrast coefficients that yielded a t-value for each of two comparisons. The

comparisons were: (1) Baseline versus Sleep Deprivation, and (2) Baseline versus Post Stress. It was assumed that the effect of sleep deprivation and the other stressors would, if they had any effect at all, adversely affect performance. Expected effects on mood were anticipated to be negative

3.1.1 Encode/Decode.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 1.87$, $p > .05$, $F(8,136) = 2.99$, $p = .004$, and $F(8,136) = 2.86$, $p = .0057$, respectively. Significant effects were evident for speed and throughput measures, and these measures were highly correlated ($r = +.96$). Figure 3-1 presents the results for average accuracy, speed and throughput. The figure indicates that there was an improvement in speed and throughput performance after the first time window. Since speed and throughput were highly correlated, planned comparisons for the Encode/Decode test were only computed for throughput.

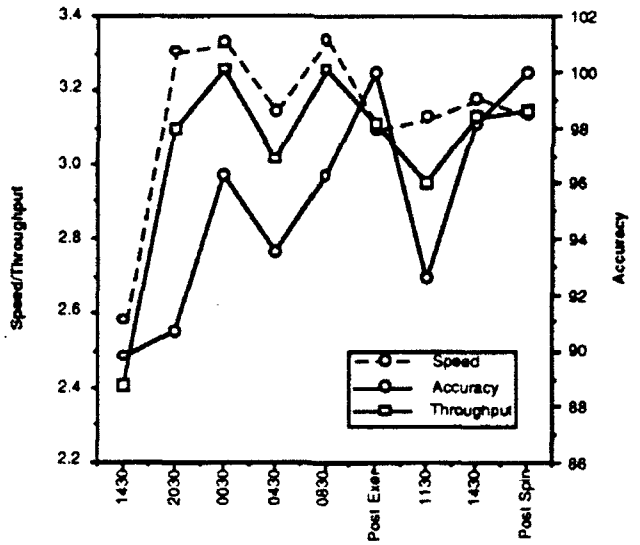


Figure 3-1. Encode/Decode average throughput, speed, and accuracy as a function of time windows.

The comparison between Baseline and Sleep Deprivation was not statistically reliable, $t(136) = 0.728$, $p > .05$. Also, the comparison between Baseline and Post Stress was not statistically reliable, $t(136) = 0.297$, $p > .05$. There were no apparent effects of sleep deprivation and stressors for Encode/Decode task.

3.1.2 Two-Letter Search.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 0.90$, $p > .05$, $F(8,136) = 2.24$, $p = .0278$, and $F(8,136) = 1.91$, $p > .05$, respectively. A significant effect was evident only for the speed measure. Figure 3-2 presents the results for average accuracy, speed and throughput. The figure indicates that there was an improvement in speed and throughput performance after the first time window.

The comparison between the means for speed showed that: (1) the difference between Baseline and Sleep Deprivation was not statistically significant, $t(136) = 0.503$, $p > .05$, and (2) the difference between Baseline and Post Stress was statistically reliable, $t(136) = -1.984$, $p < .05$. The results indicate that the performance was better after the stressors relative to the Baseline. There were no effects in the expected direction for the 2-Letter Search task.

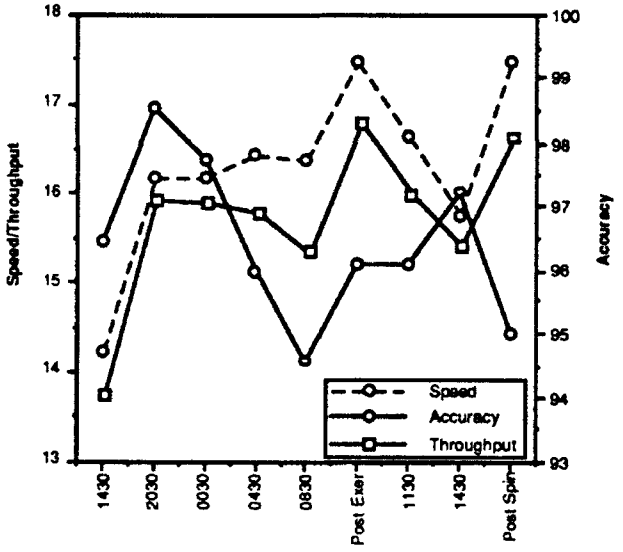


Figure 3-2. Two-Letter Search average throughput, speed, and accuracy as a function of time window.

3.1.3 Six-Letter Search.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 0.62$, $p > .05$, $F(8,136) = 2.94$, $p = .0047$, and

$F(8,136) = 2.64$, $p = .0102$, respectively. Significant effects were evident for the speed and throughput performance measures. The results also indicated that speed and throughput were highly correlated ($r = +.97$). Figure 3-3 presents the results for average accuracy, speed and throughput. The figure indicates that there was an improvement in speed and throughput performance after the first time window. Performance appears to have been degraded for the 1130 and 1430 time windows of the second day. Speed and throughput performance appears to improve for the Post-Spin PAB session.

The comparison between the means for throughput showed that: (1) the difference between Baseline and Sleep Deprivation was not statistically significant, $t(136) = 0.238$, $p > .05$, and (2) the difference between Baseline and Post Stress was statistically reliable, $t(136) = -2.641$, $p < .01$. The results indicate that the subjects performance was better after the stressors relative to the Baseline. There were no effects in the expected direction for the 6-Letter Search task.

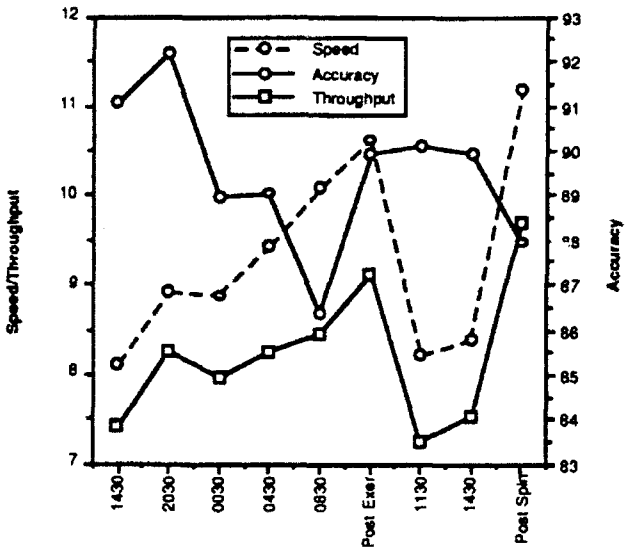


Figure 3-3. Six-Letter Search average throughput, speed, and accuracy as a function of window.

3.1.4 Two-Column Addition.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 1.08$, $p > .05$, $F(8,136) = 0.78$, $p > .05$, and $F(8,136) = 0.56$, $p > .05$, respectively. There were no statistically significant effects

for any of the Two-Column Addition measures. Figure 3-4 presents average accuracy, speed and throughput as a function of window.

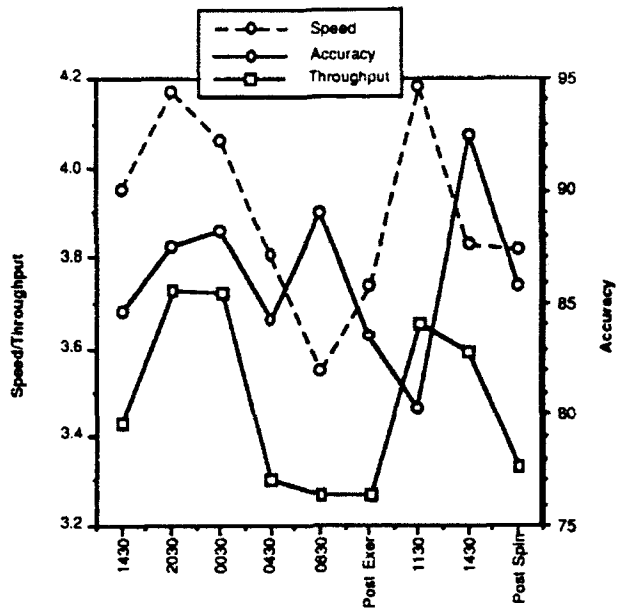


Figure 3-4. Two-Column Addition average throughput, speed, and accuracy as a function of window.

3.1.5 Logical Reasoning.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 0.72$, $p > .05$, $F(8,136) = 1.37$, $p > .05$, and $F(8,136)=2.34$, $p=.0221$, respectively. Significant effects were evident only for the throughput performance measure. Figure 3-5 presents the results for mean accuracy, speed and throughput. The figure indicates that there was an improvement in speed and throughput performance after the first time window. Performance appeared to degrade during the morning of the second day of testing; however, an improvement in speed and throughput was evident for the Post-Spin PAB session.

The comparison between the groupings for throughput showed that: (1) the difference between Baseline and Sleep Deprivation was not statistically significant, $t(136) = 1.805$, $p > .05$, and (2) the difference between Baseline and Post Stress was not statistically reliable, $t(136) = 1.403$, $p > .05$. There were no significant differences in the expected direction for the Logical Reasoning task.

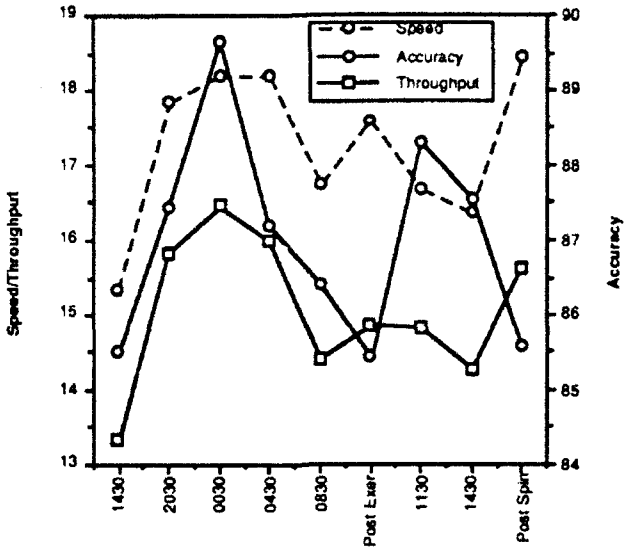


Figure 3-5. Logical Reasoning average throughput, speed, and accuracy as a function of window.

3.1.6 Digit Recall.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 2.10$, $p = .0394$, $F(8,136) = 0.82$, $p > .05$, and $F(8,136) = 2.33$, $p = .0291$, respectively. Significant effects were evident only for the accuracy and throughput performance measures. The correlation between accuracy and throughput was high ($r = +.92$). Figure 3-6 presents the results for average accuracy, speed and throughput. The figure indicates that accuracy and throughput performance were reasonably high and stable at the start of the experiment and decreased over the night and early morning. Furthermore, performance appears to have improved for the Post-Spin PAB session.

The comparison between the groupings for digit recall throughput showed that: (1) the difference between Baseline and Sleep Deprivation was statistically significant, $t(136) = 3.149$, $p < .01$, and (2) the difference between Baseline and Post Stress was also statistically reliable, $t(136) = 2.066$, $p < .05$. The results for the Digit Recall task were in the expected direction, the subjects' accuracy and throughput were adversely affected by sleep deprivation and the additional stressors.

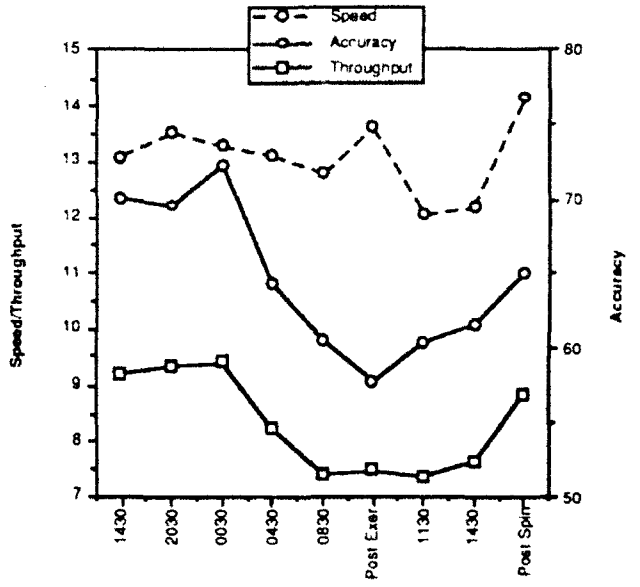


Figure 3-6. Digit Recall average throughput, speed, and accuracy as a function of window.

3.1.7 Serial Addition/Subtraction.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 2.85$, $p = .0058$, $F(8,136) = 3.38$, $p = .0014$, and $F(8,136) = 3.55$, $p = .0009$, respectively. Significant effects were evident for the accuracy, speed and throughput performance measures. The correlation among these three measures were high ($r_{as} = +.82$, $r_{at} = +.86$, $r_{st} = +.99$). Figure 3-7 presents the results for average accuracy, speed and throughput. The figure indicates that all three measures were reasonably high and stable at the start of the experiment and decreased over the night and early morning. Furthermore, performance appears to have improved for the window associated with the post-exercise PAB.

The comparison between the groupings for throughput (only this measure was selected due to the high correlation among variables) showed that: (1) the difference between Baseline and Sleep Deprivation was statistically significant, $t(136) = 4.385$ $p < .01$, and (2) the difference between Baseline and Post Stress was also statistically reliable, $t(136) = 2.747$, $p < .01$. The results for the Serial Addition/Subtraction task were in the expected direction; the subjects accuracy, speed throughput were adversely affected by sleep deprivation and the additional stressors.

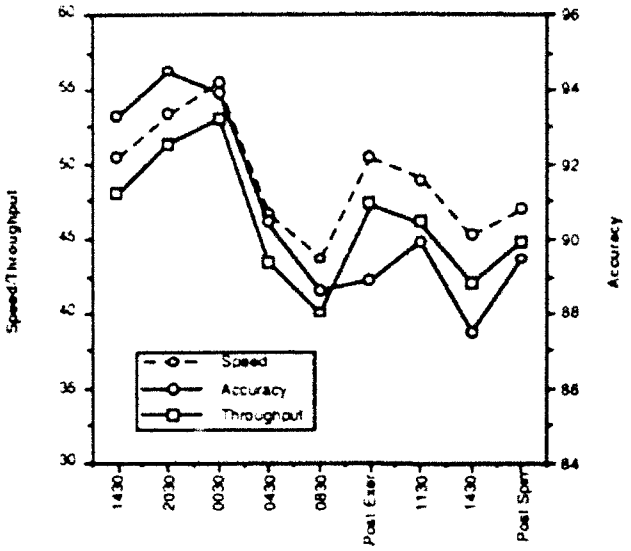


Figure 3-7. Serial Addition/Subtraction average throughput, speed, and accuracy as a function of window.

3.1.8 Pattern Recognition I.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 1.56$, $p > .05$, $F(8,136) = 0.27$, $p > .05$, and $F(8,136) = 0.44$, $p > .05$, respectively. Figure 3-8 presents average accuracy, speed and throughput as a function of window. There were no statistically significant effect for the Pattern Recognition I test.

3.1.9 Pattern Recognition II.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 0.57$, $p > .05$, $F(8,136) = 1.38$, $p > .05$, and $F(8,136) = 1.05$, $p > .05$, respectively. Figure 3-9 presents average accuracy, speed and throughput as a function of window. There were no statistically significant effects for the Pattern Recognition II test.

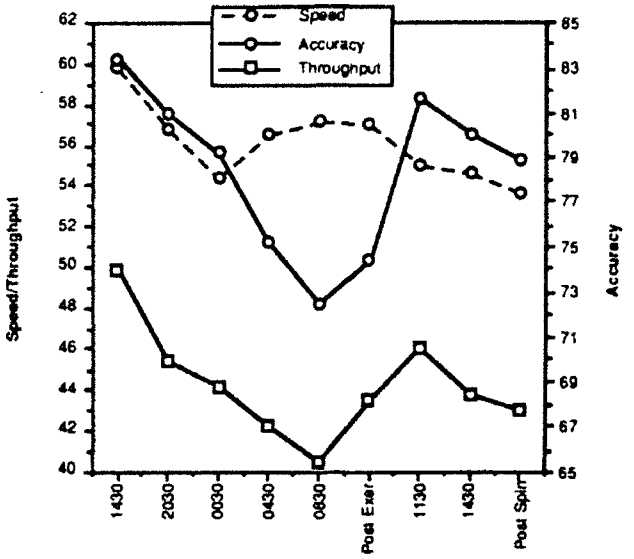


Figure 3-8. Pattern Recognition I average throughput, speed, and accuracy as a function of window.

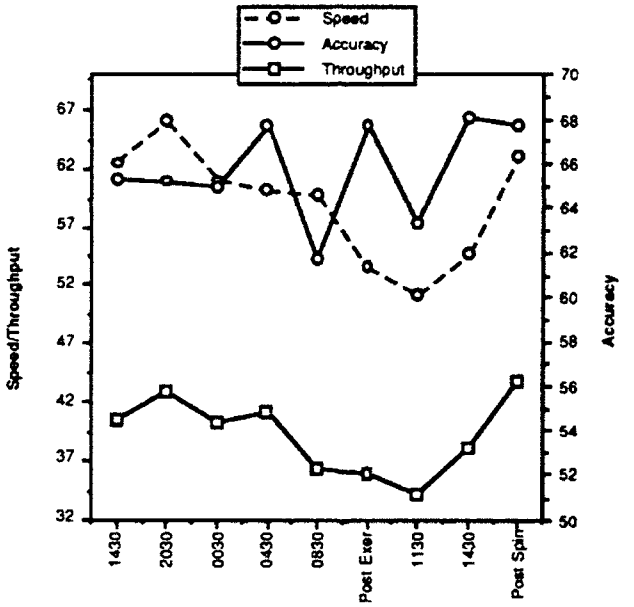


Figure 3-9. Pattern Recognition II average throughput, speed, and accuracy as a function of window.

3.1.10 Manikin.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 1.39$, $p > .05$, $F(8,136) = 3.80$, $p = .0005$, and $F(8,136) = 3.95$, $p = .0003$, respectively. Significant effects were evident for the speed and throughput performance measures. The correlation among these two measures was high ($r = +.99$). Figure 3-10 presents the results for mean accuracy, speed, and throughput. The figure indicates that speed and throughput improved after the first time window and remained relatively stable thereafter.

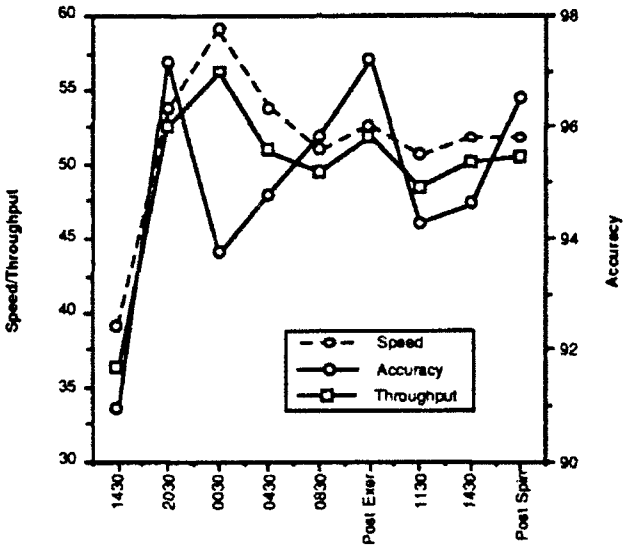


Figure 3-10. Manikin average throughput, speed, and accuracy as a function of window.

The comparison between the groupings for throughput (only this measure was selected due to the high correlation between variables) showed that: (1) the difference between Baseline and Sleep Deprivation was not statistically significant, $t(136) = 1.933$, $p > .05$, and (2) the difference between Baseline and Post Stress was also not statistically reliable, $t(136) = 1.200, p > .05$. The results for the Manikin test did not show any effects associated with sleep deprivation or other stressors.

3.1.11 Choice Reaction Time.

The results of the analysis of variance (ANOVA) for accuracy, speed, and throughput were $F(8,136) = 0.56$, $p > .05$, $F(8,136) = 5.75$, $p < .0001$, and $F(8,136) = 4.96$, $p < .0001$, respectively. Significant effects were evident for the

speed and throughput measures. The correlation among these two measures was high (r = +.99). Figure 3-11 presents the results for average accuracy, speed, and throughput. The figure indicates that speed and throughput improved after the first window and showed a gradual decline in performance thereafter.

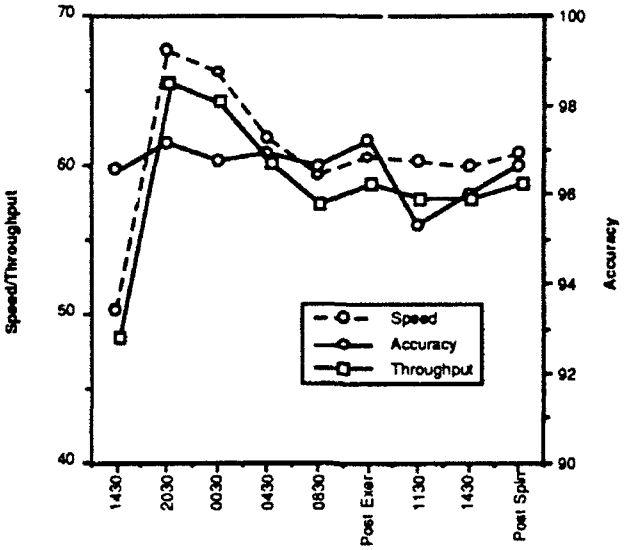


Figure 3-11. Choice Reaction Time average throughput, speed, and accuracy as a function of window.

The comparisons between the groupings for throughput (only this measure was selected due to the high correlation between variables) showed that: (1) the difference between Baseline and Sleep Deprivation was statistically significant, $t(136) = 3.276$, $p < .01$, and (2) the difference between Baseline and Post Stress was also statistically reliable, $t(136) = 2.841, p < .01$. The results for the Choice Reaction Time test showed the expected effects of sleep depravation and the other stressors.

3.1.12 Time Wall.

The ANOVA for the coefficient of variation (CV) showed a statistically significant effect, $F(8,136) = 5.46$, $p < .0001$. Figure 3-12 presents the results for CV as a function of window. Performance became more variable after the first window and showed a gradual decrease in variability over time.

The comparisons between the groupings for CV showed that: (1) the difference between Baseline and Sleep Deprivation was statistically significant, $t(136) = -5.495$,

$p < .01$, and (2) the difference between Baseline and Post Stress was also statistically reliable, $t(136) = -3.103$, $p < .01$. The results for the Time Wall test showed effects consistent with those expected for the sleep deprivation, exercise, and spinning stressors.

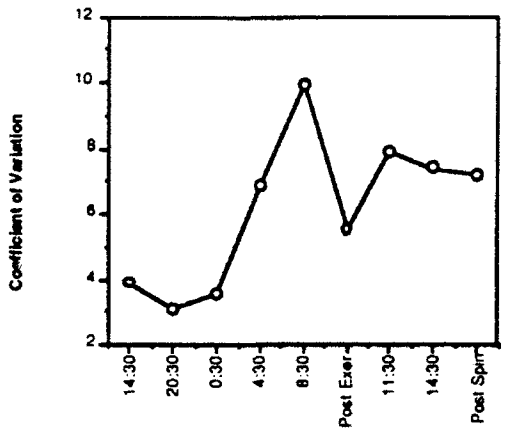


Figure 3-12. Time Wall average coefficient of variation as a function of window.

3.1.13 Interval Production.

The ANOVA for the coefficient of variation (CV) showed a statistically significant effect, $F(8,136) = 2.98$, $p = .0042$. Figure 3-13 presents the results for CV as a function of window. Performance became more variable after the third (the increase in variability for the second time window appears to be somewhat out of line) window and showed a decrease in variability after the 11:40.

The comparisons between the groupings for CV showed that: (1) the difference between Baseline and Sleep Deprivation was not statistically significant, $t(136) = 1.246$, $p > .05$, and (2) the difference between Baseline and Post Stress was not statistically reliable, $t(136) = -0.359$, $p > .05$. There were no effects in the expected direction for the Interval Production task.

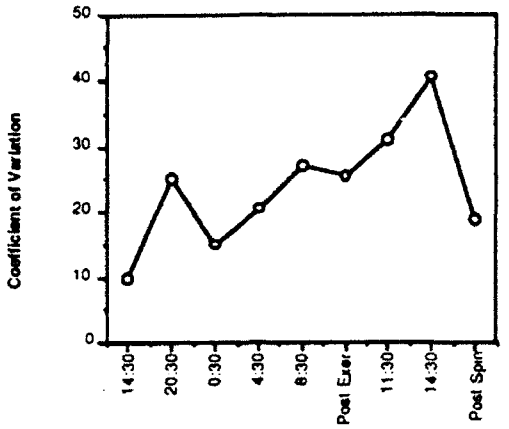


Figure 3-13. Interval Production average coefficient of variation as a function of window.

3.1.14 Mood Scale II.

The ANOVA results for the Mood Scale II are as follows for the six different subscales: (1) activity, $F(8,136) = 12.66$, $p < .0001$, (2) happiness, $F(8,136) = 10.66$, $p < .0001$, (3) depression, $F(8,136) = 3.33$, $p = .0017$, (4) anger, $F(8,136) = 4.23$, $p < .0001$, (5) fatigue, $F(8,136) = 16.97$, $p < .001$, and (6) fear, $F(8,136) = 4.41$, $p < .0001$. All of the subscales showed statistically reliable effects. Figure 3-14 presents the results for the subscales. The subjects expressed less happiness and lower levels of activity as the experiment progressed. Also, the reports of fatigue, anger, depression and fear increased over the course of the study.

3.1.15 Stanford Sleepiness Scale.

The ANOVA for the Sleepiness Score showed a statistically significant effect, $F(8,136) = 21.81$, $p < .0001$. Figure 3-15 presents average Sleepiness Scores as a function of window. Reports of sleepiness increased in a systematic manner throughout the course of the study. The reports of sleepiness appear to plateau after 0830 on the second day of testing.

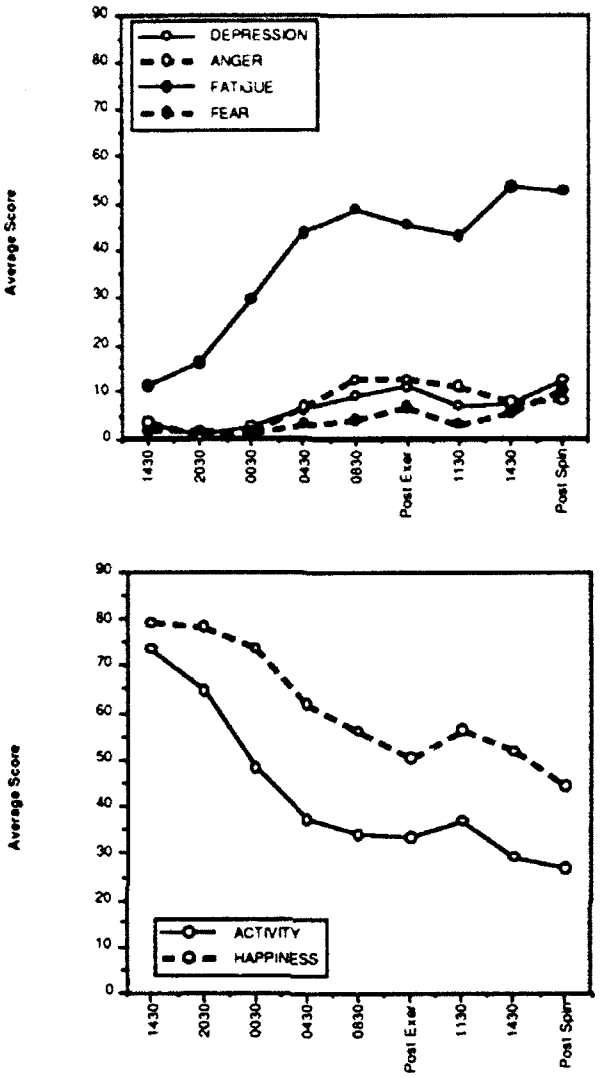


Figure 3-14. Mood Scale II scores for fatigue, anger, depression, fear, happiness, and activity subscales as a function of window.

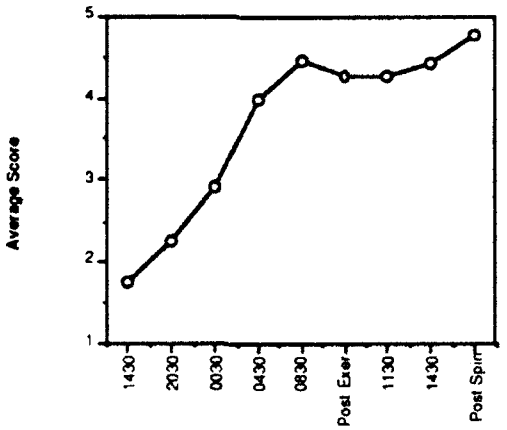


Figure 3-15. Stanford Sleepiness Scale average sleepiness scores as a function of window.

3.2 COMPARISON of WRPAB TESTS.

Table 3-1 presents a summary of the results for the 13 performance based WRPAB tests that were included in the study. The table presents the results of the ANOVAs with a probability value or an indication that the test resulted in non-significant (NS) findings. The last two columns of the table summarize the results of comparisons of the Sleep Deprivation and Stressor groups against the Baseline grouping

Table 3-1. Summary WRPAB test results.

	ANOVA Results				Comparisons	
Test	Accuracy	Speed	Through-put	CV	Sleep Dep.	Stressor
Encode/Decode	NS	<.01	<.01		No	No
2-Letter Search	NS	<.05	NS		No	No
6-Letter Search	NS	<.01	<.05		No	No
2-Column Add	NS	NS	NS		No	No
Logical Res.	NS	NS	<.05		No	No
Digit Recall	<.05	NS	<.05		**Yes**	**Yes**
Serial Add/Sub	<.01	<.01	<.001		**Yes**	**Yes**
Pattern Recog. I	NS	NS	NS		No	No
Pattern Recog. II	NS	NS	NS		No	No
Manikin	NS	<.001	<.001		No	No
Choice RT	NS	<.001	<.001		**Yes**	**Yes**
Time Wall				<.001	**Yes**	**Yes**
Interval Prod.				<.01	No	No

The table shows four tests that produced reliable effects in the expected direction: Digit Recall, Serial Add/Sub, Choice Reaction Time, and Time Wall. The tests that resulted in significant ANOVA results but no expected effects for sleep deprivation or other stressors, tended to show marked practice effects.

Table 3-2 presents the results for the non-performance based WRPAB tests: Mood Scale II and Stanford Sleepiness Scale. All of the subscales for the Mood Scale II showed effects in the expected direction. The same was true for the Stanford Sleepiness Scale.

Table 3-2. Summary of results for WRPAB non-performance based tests.

	ANOVA Results		Comparisons	
Test	Percent Score	Sleep Score	Sleep Dep.	Stressors
Mood Scale II				
Activity	<.0001		Yes	Yes
Happiness	<.0001		Yes	Yes
Depression	<.005		Yes	Yes
Anger	<.0001		Yes	Yes
Fatigue	<.0001		Yes	Yes
Fear	<.0001		No	Yes
Stanford Sleepiness Scale		<.0001	Yes	Yes

3.3 ANALYSIS OF SYMPTOM CHECK LIST (SCL) RATINGS.

The SCL ratings were averaged over the same time windows used for the analysis of the WRPAB test. A Symptom Severity rating was derived by summing across the six symptom categories that subjects rated in the SCL. The severity ratings have a theoretical range between 6 and 30 (all 1s or all 5s). Figure 3-16 presents average severity ratings as a function of the 9 time windows. The subjects reported 6s ("healthy feeling fine") for the first two windows and systematically increased their illness ratings throughout the remainder of the experiment. The subjects reported an increase in feeling of illness following the spinning manipulation. These results indicate that the treatment conditions had the expected results on subjects' reported illness.

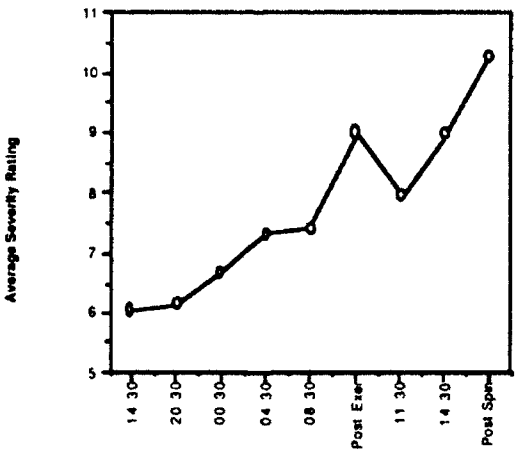


Figure 3-16. Average symptom severity ratings as a function of time windows.

3.4 RELATIONSHIP BETWEEN WRPAB PERFORMANCE AND SCL RATNGS.

In order to investigate the relationship between SCL and WRPAB test results, correlations were computed between the WRPAB performance measures and the symptom severity ratings. The SCL and WRPAB scores were averaged over subjects, and the correlations computed between means for the latter 8 time windows. The first time window was not included in computing the correlations because of the large practice effect evident in many of the WRPAB scores between the first and second windows. Table 3-3 presents the correlations among the WRPAB performance based tests and the symptom severity ratings. The correlations presented in bold are those in the expected direction (e.g., as symptom severity rating increase, performance on the WRPAB test degrades).

Table 3-3. Correlation among WRPAB performance measures and Symptom Severity Ratings.

Test	Accuracy	Speed	Through-put	CV
Encode/Decode	+.80*	**-.70**	**-.06**	
2-Letter Search	**-.55**	+.61	+.49	
6-Letter Search	**-.30**	+.54	+.50	
2-Column Add	**-.07**	**-.42**	**-.50**	
Logical Res.	**-.63**	**-.05**	**-.41**	
Digit Recall	**-.54**	+.18	**-.37**	
Serial Add/Sub	**-.73***	**-.48**	**-.45**	
Pattern Recog I	**-.06**	**-.45**	**-.18**	
Pattern Recog II	+.52	**-.35**	**-.05**	
Manikin	+.22	**-.51**	**-.43**	
Choice Rt	**-.20**	**-.69**	**-.68**	
Time Wall				**+.42**
Interval Prod				**+.20**

*p < .05

The four tests that showed performance effects associated with sleep deprivation and other stressors also resulted in correlations in the expected direction: Digit Recall, Serial Add/Sub, Choice Reaction Time, and Time Wall. However, the only statistically significant correlation in the expected direction is accuracy under the Serial Add/Sub task, the other correlations are not significant at the .05 level. The degrees of freedom for each of the correlations was small (6), and therefore the lack of statistically significant findings is not surprising. These results do provide converging evidence that a subset of the WRPAB was sensitive to the treatment conditions both for performance and subjective reports of illness.

Table 3-4 presents the correlations between the WRPAB non-performance based tests and symptom severity. All of the correlations are in the expected direction and, with the exception of one correlation (Anger), all of the correlations are statistically significant ($p < .05$). These results indicate that as the subjects reported increases in symptom severity, they also reported decreased activity level and happiness. Also, reports of depression, anger, fatigue, fear, and sleepiness increased along with their reports of increased symptom severity. These results provide converging evidence that the treatments (sleep deprivation, exercise, and spinning) resulted in the expected subjective feelings of illness and associated affect.

Table 3-4. Correlations between WRPAB non-performance based tests and symptom severity ratings.

Test	Score
Mood Scale II	
Activity	**-.84***
Happiness	**-.94***
Depression	**+.89***
Anger	**+.62**
Fatigue	**+.84***
Fear	**+.95***
Stanford Sleepiness Scale	**+.84***

***p < .05**

3.5 MDS ANALYSES.

3.5.1 Similarity Ratings.

The 78 pairs of similarity ratings between the 13 WRPAB tests were submitted to multidimensional scaling using the ALSCAL (Young and Lewyckyj, 1979) program. The ratings were assumed to be ordinal dissimilarities (disparities), and the disparity matrices were assumed to be symmetric. The disparity matrices from 8 subjects were submitted to the analysis, and both individual and group solutions were obtained. Because the focus of this study is not individual differences in experts' conception of the WRPAB, only the group space is reported here. Solutions for 2, 3, 4, and 5 dimensions were obtained. The three dimensional solution yielded reasonable overall Stress, 0.16, and accounted for approximately 79 percent of the variance in the group space, $R^2 = 0.787$. We have chosen to interpret the three dimensional solution.

The coordinates of the 13 WRPAB tests in the three dimensional space are provided in Table 3-5. Figures 3-17 and 3-18 are graphic plots of the data from Table 3-5.

Table 3-5. Three dimensional MDS solution for the PAB similarity data.

Tests	Dimension 1	Dimension 2	Dimension 3
Encode/Decode	-1.37810	-0.50670	0.66610
2-Letter Search	-0.64890	-0.76400	-1.02800
6-Letter Search	-0.72110	-0.63890	-0.98740
Two-Column Addition	-1.49460	0.52690	1.02410
Logical Reasoning	0.14010	-0.67460	1.69570
Digit Recall	-0.35000	-0.37920	-1.38520
Serial Add/Subtract	-1.09540	0.33660	1.07880
Pattern Recognition I	1.11100	-0.90160	-0.78280
Pattern Recognition II	0.98210	-0.83410	-0.90700
Manikin	0.79550	-1.05690	1.13180
Choice RT	-0.07620	1.33130	-0.83760
Time Wall	1.41900	1.67510	0.01620
Interval Production	1.31650	1.88610	0.31550

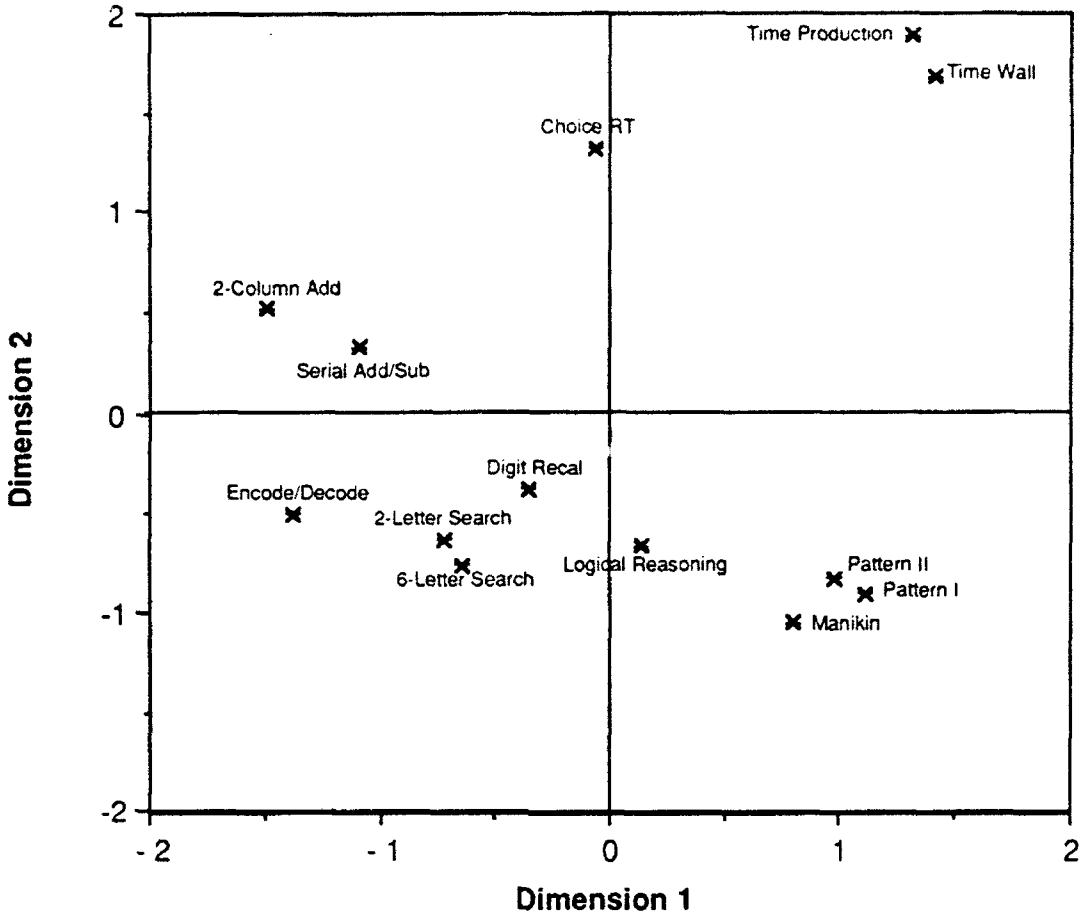


Figure 3-17. Plot of PAB tests in MDS similarity coordinate space of Dimensions 1 and 2.

Time Wall, Time Production, the Pattern Recognition tests, and the Manikin test have high positive loadings on dimension 1, whereas Encode/Decode, 2-Column Add/Sub, Serial Add, the letter search tasks had negative loadings. In general the non-verbal, or spatial/temporal tasks positively loaded on dimension 1 and the verbal tasks loaded negatively. Choice RT, which is basically a typing task, and Logical Reasoning. which is a verbal task requiring resolution of spatial relationships, fell close to the midpoint of this dimension. Consequently Dimension 1 was labeled "verbal/spatial".

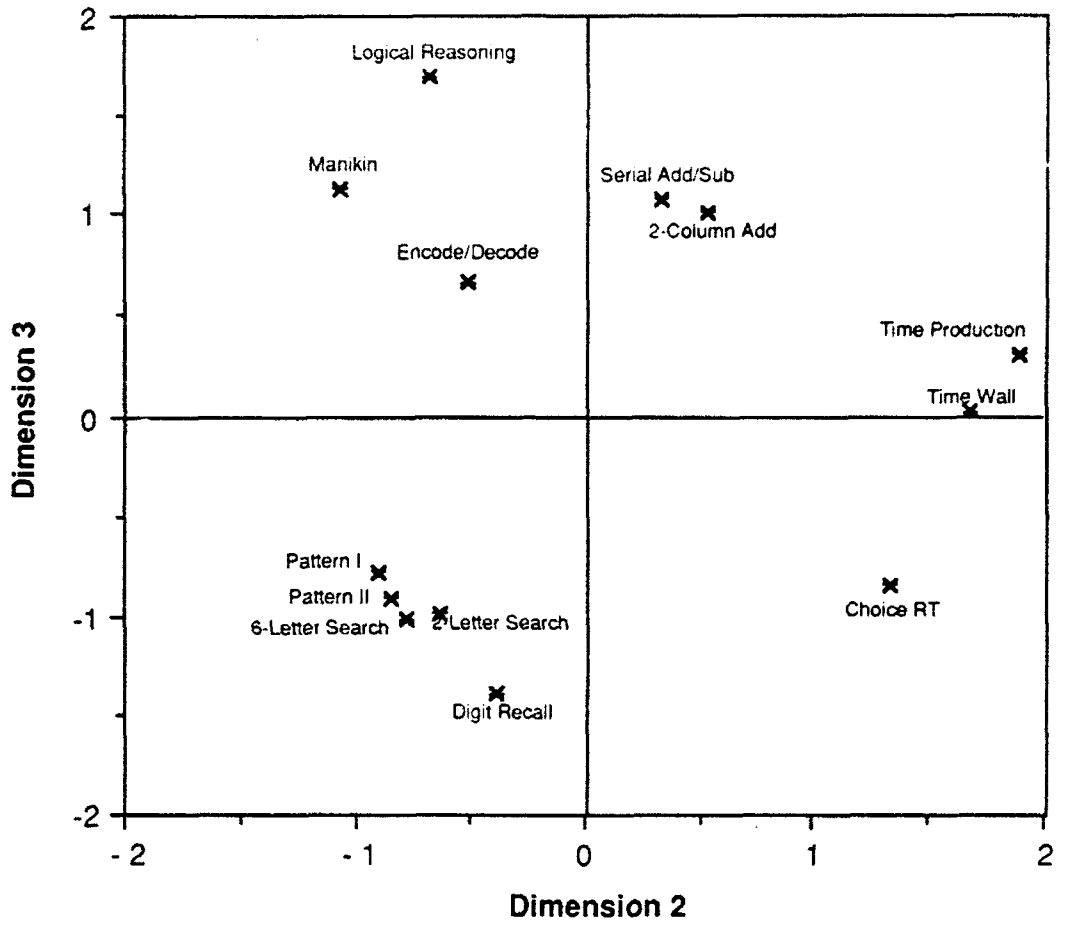


Figure 3-18. Plot of PAB tests in MDS similarity coordinate space of Dimensions 2 and 3.

The second dimension is more difficult to characterize. Time Production, Time Wall and Choice RT have high positive loadings on this dimension. Manikin, and the Pattern Recognition tests have the largest negative loadings. Digit Recall and Serial Add/Subtract are closest to the midpoint of this dimension. It appears that tests oriented towards motor responses had the highest loadings, task oriented towards assessment of internal processing had more neutral loadings, and tasks oriented towards assessment of perception had the most negative loadings. Thus, from the perspective of an information processing analogy, this dimension appears to reflect a processing continuum from input, to processing, to output. Dimension 2 therefore, is tentatively labelled the "input/output dimension."

Dimension 3 has Logical Reasoning with the highest loading, followed by Manikin, Serial Add/Subtract and Two Column Addition. Time Wall and Time

Production had nearly neutral loadings, and Digit Recall and the Letter Search tasks had the largest negative loadings. As one goes from negative to positive on this dimension the responses appear to rely more on mental manipulation of the stimuli and less on memory or motor mapping. Thus digit recall relies on mapping the memory of the original stimulus to a second stimulus, and the letter search relies on matching physically present stimuli, whereas as logical reasoning would seem to require either (a) generation of a mental picture from a verbal stimulus, or (b) generation of a verbal statement from a physical stimulus. This dimension is tentatively labelled "mental manipulation" or "cognitive processing."

3.5.2 Construct Ratings Analysis.

The mean ratings of constructs with respect to the 13 tests are given in Table 3-6. Multiple regression analyses were performed for each of the 11 constructs. In each regression, a construct rating served as the dependent, or predicted, variable and the stimulus coordinates for each of the three MDS dimensions (from Table 3-5) served as independent, or predictor, variables. The regression analysis results are given in Table 3-7. Standardized beta weights from the regression analysis are reported under the dimension headings in Table 3-7.

In summary, 11 regression analyses were performed. The equations were of the form: $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$, where Y_i is the construct rating and X_1 through X_3 are the MDS coordinates.

The relative size of the beta weights can be used to aid in interpretation of the importance of dimensions to the prediction of the constructs. Thus large beta weights may suggest that the construct associated with the weight is a factor in determining subjects' similarity ratings for the WRPAB tests. However, the absolute size of Beta weights from different regression analyses should not be interpreted; thus the fact that the Time Pressure beta for Dimension 1 is larger than the Verbal beta for Dimension 1 does not mean that Time Pressure is more important to the definition of Dimension 1 than is Verbal.

Table 3-6. Mean Rating of WRPAB Tests with Respect to Eleven Constructs.

Tests	Constructs										
	Verbal	Visuo-Spatial	Short-Term Memory	Visual Search	Motor Coordination	Reasoning	Mental Manipulation	Time Perception	Item Difficulty	Difficulty Importance	Time Pressure
Encode/Decode	3.3	4.4	2.9	1.8	6.7	4.1	2.3	6.8	2.9	3.1	4.1
6-Letter Search	4.4	3.3	2.7	1.3	6.7	6.0	4.9	6.8	4.0	4.0	4.0
2-Letter Search	4.6	3.7	2.8	2.2	6.7	6.1	5.0	6.8	5.8	5.0	4.6
2-Column Add	4.8	5.2	2.6	5.6	6.7	4.0	2.4	6.8	3.1	2.7	3.1
Logical Reasoning	1.5	3.2	3.8	4.7	6.8	1.8	2.6	6.6	3.5	2.4	4.6
Serial Recall	3.9	3.7	1.4	2.0	6.7	6.0	4.7	6.5	2.5	3.4	3.5
Serial Addition/Subtraction	3.6	5.3	2.4	6.3	6.6	3.4	2.0	6.8	3.3	2.7	3.0
Pattern Recognition I	6.7	1.4	2.0	2.7	6.8	6.3	5.9	6.8	3.2	3.4	5.1
Pattern Recognition II	6.8	1.4	2.2	2.8	6.8	6.4	5.7	6.8	3.1	3.3	5.2
Manikin	5.9	1.1	5.1	4.0	5.4	3.7	2.3	6.8	2.7	3.1	5.1
Choice RT	5.8	5.4	3.9	6.1	2.4	6.2	6.7	6.8	5.8	4.8	2.8
Time Wall	6.6	2.1	4.4	5.9	4.3	5.4	5.0	1.3	4.3	3.8	5.4
Interval Production	6.5	5.0	5.1	6.6	2.9	6.1	6.1	1.2	4.9	5.3	6.5

Table 3-7. The Beta weights from *Regression of the MDS Stimulus Coordinates on to the* Construct Ratings.

Constructs	Dimension 1 Verbal/Spatial	Dimension 2 Visual/Motor	Dimension 3 Cognitive Processing	R^2
Verbal	**0.5400***	0.2252	-0.3386	**0.59***
Difficulty Importance	0.1049	0.4501	-0.5523	0.55
Visuo-Spatial	**-0.7379****	**0.6521****	0.0127	**0.84****
Short-Term Memory	0.4494	0.3169	**0.4864***	**0.61***
Visual Search	0.1389	**0.6952****	**0.4435***	**0.84****
Motor Coordination	-0.2760	**-0.7154***	0.0627	**0.72***
Reasoning	0.1377	0.2074	**-0.8715****	**0.91****
Mental Manipulation	**0.3511***	**0.2963***	**-0.7300****	**0.89****
Time Perception	**-0.4422***	**-0.6331***	-0.0732	**0.77***
Item Difficulty	0.0470	0.5970	-0.2674	0.45
Time Pressure	**0.8080***	-0.1146	0.1385	**0.66***

*p < 0.05
**p < 0.001

The MDS coordinates of the 13 WRPAB tests provided significant ($p < .05$) fit to 9 of the 11 construct ratings: verbal, visuo-spatial, short-term memory, visual search, motor coordination, reasoning mental manipulation, time perception and time pressure. In Table 3-7, the significance of beta weights is only recorded for regression equations for which the overall fit was significant. Because these tests were exploratory, there were no precautions taken against inflation of the experiment wise alpha.

Dimension 1. The most robust predictor of Dimension 1 is the visuo-spatial construct. Also congruent with the initial interpretation of this dimension as verbal/spatial, is the significant prediction for the verbal construct. If time perception is considered "spatial" then the significant beta for time perception also fits the verbal/spatial interpretation; perhaps the dimension could be re-labeled "verbal/temporal-spatial." Time Pressure and Mental Manipulation also provided significant predictions for Dimension 1. Assuming that the verbal/temporal-spatial interpretation is correct, the positive Mental Manipulation beta may reflect the coincidental verbal component of most of the tests that were strong in "mental manipulation requirements" (i.e. Two-Column Addition, Logical Reasoning, Serial Add/Subtract, but not Manikin). Similarly, the time pressure was highly associated with verbal tasks and whereas the temporal-spatial tasks were judged to to be low on time pressure (a rating of 1 indicated "extreme" time pressure and a rating of 7 indicated "little or no" time pressure).

Dimension 2. The two constructs with the largest positive betas on Dimension 2 were Visual Search and Visuo-Spatial. The two factors with the largest negative betas were Motor Coordination and Time Perception. The tests most highly associated with Motor Coordination were Choice RT and Interval Production. Visual Search was most highly associated with 6-Letter Search, Encode/Decode, Digit Recall, 2-Letter Search, and the Pattern Recognition tests. The association of Time Perception with this dimension was probably the result of coincidental correlations between Time Perception and other constructs associated with Dimension 2. The rating data nicely corroborate the initial interpretation of this dimension as one that reflects the processing locus of interest; perception (vision) is one anchor of the dimension, output (motor) is the opposite anchor, and cognitively loaded tasks fall intermediate to input and output loci.

Dimension 3. The Mental Manipulation and Reasoning Constructs appear to well describe Dimension 3. Both of these constructs have large negative betas on this dimension. Visual Search and Short-Term Memory constructs have moderately significant positive betas on this dimension. Because the tasks associated with short-term memory and visual search, i.e., Encode/Decode, the letter search tasks, digit recall, and the Pattern Recognition test, require little in the way of mental transformations or "higher level" cognitive processing, the positive betas from the short-term memory and visual search constructs on Dimension 3 are consistent with our original interpretation of Dimension 3. The labels, "Mental Manipulation" or "Cognitive Processing" are appropriate for Dimension 3.

SECTION 4
MICROSAINT Model

4.1 APPROACH.

In this section, the results of MicroSAINT models using WRPAB test data to predict AH-1 performance are presented. In addition, the results of the PAB- and PDQ-based (from Inman et al, 1991) MicroSAINT models are compared.

The MicroSAINT models developed for AH-1 TOW and Rocket engagements (Inman et al, 1991) were used in this modeling effort. In order to relate PAB decrements in performance to AH-1 performance, the following steps discussed in greater detail in the remaining sections, were accomplished:

(1) The MDS results were used to derive composite scores for the WRPAB. The composite scores were derived for verbal, motor, cognitive, and spatial processing. The results of the MDS analysis were used to combine the 13 performance tests into 4 groups of tests. The composite scores were derived by combining the performance metrics within each of the four groups of tests.

(2) For each of the nodes in the TOW and Rocket engagement models, weights were determined to reflect the level of verbal, motor, cognitive, and spatial processing required.

(3) The composite scores were used to derive decrements for verbal, motor, cognitive and spatial processing. The analysis of the individual WRPAB tests indicated that performance generally stabilized (reached an asymptote) after the second time window. Therefore, the decrements were computed relative to the third time window which was considered the baseline.

(4) Equations were derived to relate weighted WRPAB decrements for each of the task nodes in the MicroSAINT models.

(5) For the TOW and Rocket models, 100 MicroSAINT runs were conducted for time windows HFF, 0430, 0830, 1130, 1430, Post-Exercise, and Post-Spin.

4.2 COMPOSITE SCORES.

The MDS analyses resulted in the identification of three separate dimensions : Verbal/Spatial, Visual/Motor, and Cognitive. For the purpo[illegible] of relating WRPAB results to AH-1 tasks, the thirteen performance test were classified as verbal, motor, spatial, or cognitive tests. For example, the Verbal/Spatial dimension contained verbal and spatial WRPAB tests. The verbal tests had negative coordinates and the spatial tests had positive coordinates (see Figures 3-17 and 3-18) in the MDS space. To facilitate mapping of WRPAB tests to AH-1 tasks, tests from the Verbal/Spatial dimension were separately grouped as verbal or spatial. The above method was used to derive the four test groupings: verbal, spatial, motor, and cognitive. Table 4-1 presents the selected WRPAB tests as a function of test group.

Table 4-1. Selected WRPAB tests as a function of test group.

Test Group	Tests
Verbal	Encode/Decode Logical Reasoning
Motor	Choice RT Interval Production
Cognitive Processing	2-Column Addition Logical Reasoning Manikin
Spatial Processing	Pattern I Pattern II Time Wall

For each test group, the WRPAB scores (throughput or coefficient of variation) for the selected tests were converted to z-scores, averaged and then transformed into probabilities. This allowed all of the tests scores to be converted to a common metric which conveyed performance on a scale between 0 and 1.

Table 4-2 presents throughput, z-scores, probabilities, and performance decrements as a function of time window for the verbal WRPAB tests. Average z-scores were computed for each time window and transformed into probabilities using the standard normal distribution. Decrements were computed for time windows 0430 through Post-Spin using the following procedure:

(1) The differences between the probabilities for baseline and time windows 0430 through Post-Spin were computed (e.g., for time window 0430 the difference was 0.666-0.791 or -0.125).

(2) The computed differences were divided by the probability for the baseline (e.g., for time window 0430 the computed decrement was -0.125/0.791 or -0.158).

Table 4-2. Example computation of performance decrement for WRPAB verbal tests.

Time Window	Encode/Decode		Logical Reasoning		Average Z-Score	Probability	Decrements
	Throughput	Z-Score	Throughput	Z-Score			
Baseline	3.248	0.826	16.357	1.401	1.114	0.791	0.000
0430	3.010	-0.091	15.913	0.954	0.432	0.666	-0.158
0830	3.253	0.845	14.299	-0.672	0.087	0.536	-0.322
1130	2.937	-0.373	14.714	-0.253	-0.313	0.378	-0.522
1430	3.124	0.347	14.162	-0.810	-0.231	0.409	-0.483
Post-Exer	3.104	0.273	14.758	-0.209	0.032	0.512	-0.353
Post-Spin	3.142	0.419	15.511	0.549	0.484	0.684	-0.135

The resulting decrements in Table 4-2 reflect the proportional decrease of verbal test performance as a function of time window. The verbal performance decrements were computed relative to the baseline time window. This procedure allowed for combining the results of tests that did not share common performance metrics. For example, the Interval Production and Time Wall tests employed the coefficient of variation as the performance metric, whereas throughput was the performance measure for the other PAB tests . Table 4-3 presents the composite WRPAB decrement scores as a function of time window for the four test groups.

Table 4-3. Composite decrement scores as function of test group and time windows.

Test Group	Time Window						
	Baseline	0430	0830	1130	Post-Exer	1430	Post-Spin
Verbal	0.000	-0.158	-0.322	-0.522	-0.353	-0.483	-0.135
Motor	0.000	-0.216	-0.529	-0.612	-0.423	-0.813	-0.235
Cognitive Processing	0.000	-0.373	-0.665	-0.344	-0.545	-0.434	-0.420
Spatial Processing	0.000	-0.433	-0.876	-0.619	-0.520	-0.530	-0.272

Figure 4-1 presents the decrements listed in Table 4-3. The general trend in this figure is for performance to decrease during the two time windows following the baseline and to improve during the final time window. During the time windows following the baseline, the pilots were suffering from the effects of mild sleep deprivation. The improvement in performance during the final test sessions is consistent with increased alertness associated with circadian rhythms and the increase in level of activity (exercise, AH-1 simulation, etc.) for this phase of the simulation study. In addition, the post-spin WRPAB session was conducted 30 to 45 minutes after the spinning treatment in the Barany chair. This delay was due to the fact that subjects completed the AH-1 simulation session immediately after the spinning.

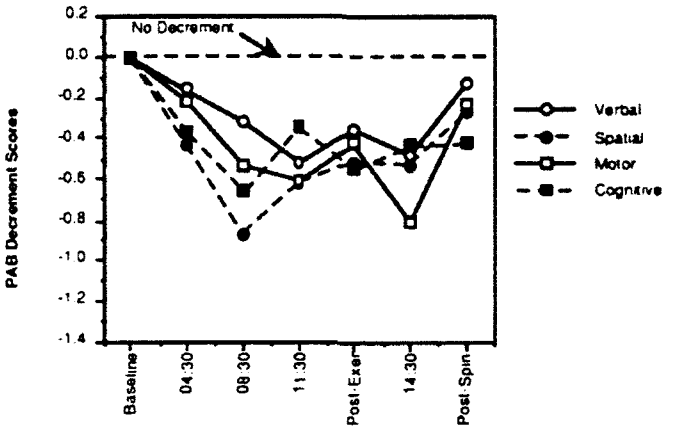


Figure 4-1. WRPAB decrements for Verbal, Spatial, Motor, and Cognitive dimensions as a function of time windows.

4.3 RELATING WRPAB PERFORMANCE TO AH-1 PERFORMANCE.

Figures 4-2 and 4-3 present the models that were developed for the TOW and Rocket engagements. The baseline average and standard deviations for task completions times are from the Baseline AH-1 simulation run where subjects reported being "healthy feeling fine". The task completion times were assumed to be distributed as Gamma distributions (positively skewed). For each node in the models, a percent contribution as a function of the four test groups was estimated. That is, for each task, the percent verbal, motor, cognitive, and spatial processing required was estimated. These estimates were based on previous task analytic work for the AH-1 and the scenarios used in the simulation study. These values are presented in Figures 22 and 23 next to each of the nodes.

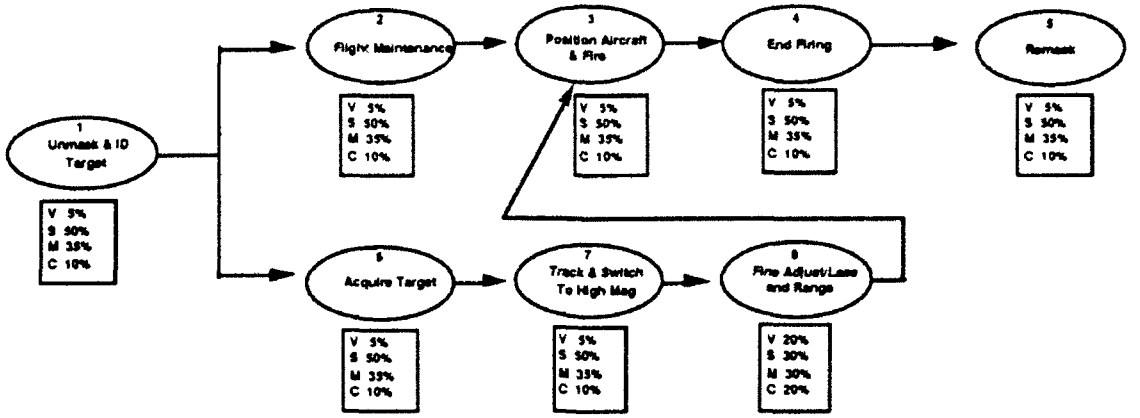


Figure 4-2. MicroSAINT model for rocket engagements.

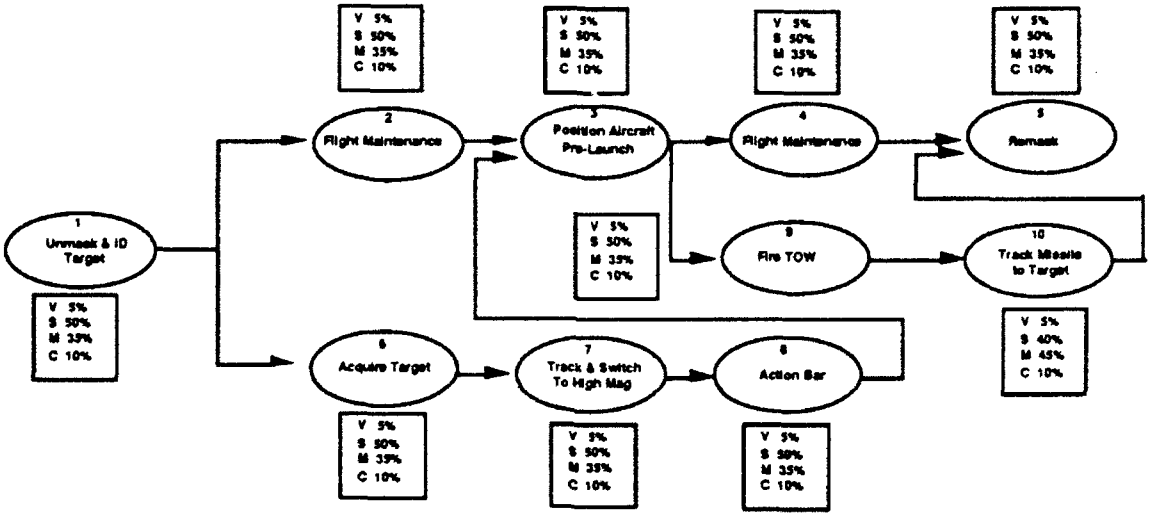


Figure 4-3. MicroSAINT model for TOW engagements.

4.4 MICROSAINT SIMULATIONS.

The TOW and Rocket models were run 100 times for each of the six time windows that followed the baseline. The model was also run using the means and standard deviations from the baseline, Healthy Feeling Fine (HFF) condition . Average engagement completion times were computed for HFF (baseline) and each of the six time windows following baseline. For each of the six time windows, a ratio of HFF time to Post-Stressor time was computed.

Rocket Engagements

Figure 4-4 presents ratios in engagement completion time (relative to HFF) for Rockets. The use of the WRPAB decrement scores in the MicroSAINT model for the Rocket engagements, predicts an early decrement in AH-1 performance with an associated improvement for the Post-Exercise and Post-Spin sessions. The WRPAB post-spin session was conducted approximately 30 to 45 minutes after the spinning in the Barany chair. The effects of the spinning treatment may have been attenuated due to the time lapse between the spinning treatment and the WRPAB administration. However, there is still a predicted decrement in performance greater than 20% for the post-spin session relative to baseline.

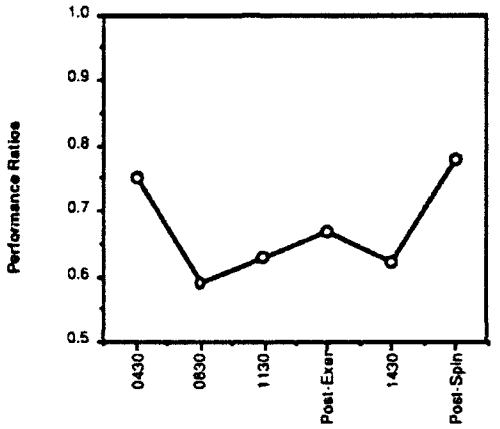


Figure 4-4. Ratios in engagement completion time for Rockets.

Figure 4-5 presents frequency distributions for the engagement times for the Baseline and the subsequent 6 time windows. The model predicts worse performance for the 1130 time window. For the Post-Spin session, performance is predicted to be degraded relative to baseline; however, there is a trend for this distribution to be shifted to the left (improvement in performance) relative to the previous three time windows.

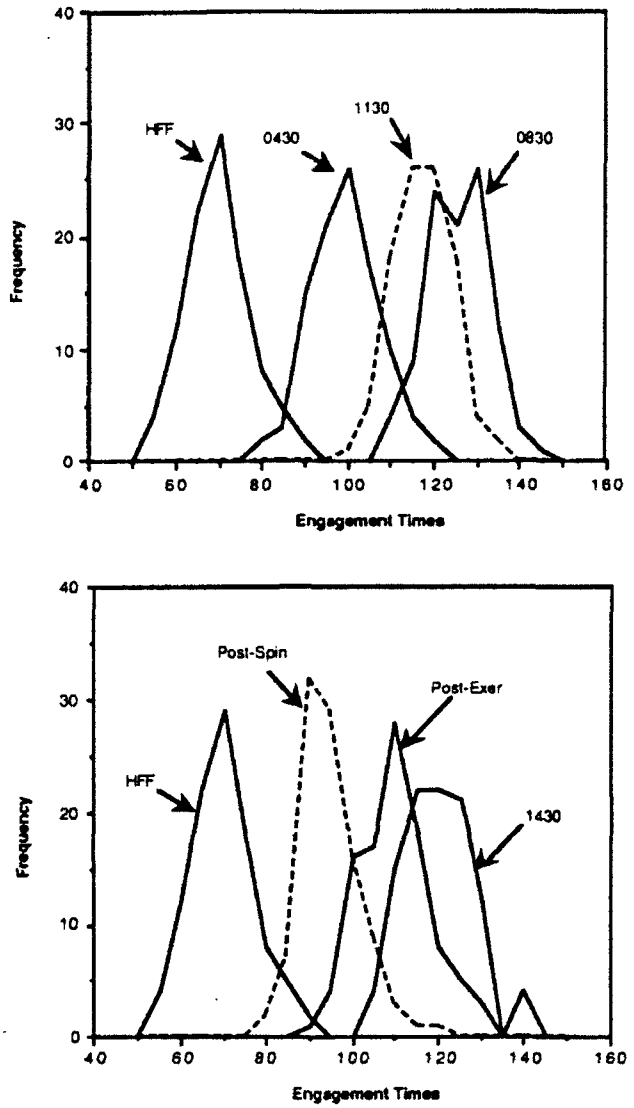


Figure 4-5. Frequency distributions of engagement times for the Rocket MicroSAINT model executions.

TOW Engagements

Figure 4-6 presents engagement completion time ratios (HFF times to post stress times) for TOWs. The use of the WRPAB decrement scores in the MicroSAINT model for the TOW engagements predicts an early performance decrement in AH-1 performance with an associated improvement for the Post-Exercise and Post-Spin test sessions. The modeling results for TOWs are comparable to those for the Rocket engagement.

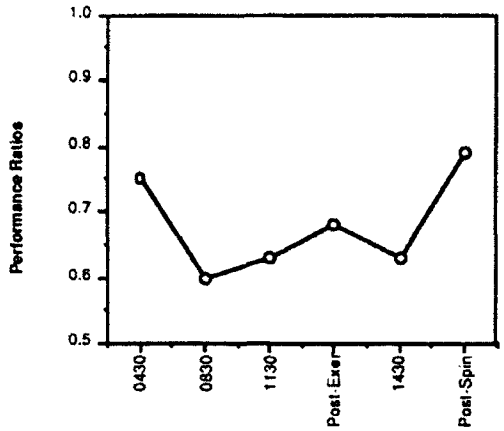


Figure 4-6 Ratios in engagement completion time for TOWs.

Figure 4-7 presents frequency distributions for the engagement times for the Baseline and the subsequent 6 time windows. The model predicts worse performance for the 1130 time window. For the Post-Spin session, performance is predicted to be degraded relative to baseline; however, there is a trend for this distribution to be shifted to the left (improvement in performance) relative to the previous three time windows.

4.5 PAB- AND PDQ-BASED PERFORMANCE ESTIMATION.

This section presents a comparison of the modeling results of the Rocket and TOW MicroSAINT models based on average task times and standard deviations for: (1) baseline AH-1 simulation, (2) post-spin AH-1 simulation, (3) baseline times modified by PAB performance, and (4) baseline times modified by PDQ predictions.

The results of MicroSAINT models that employed PDQ performance predictions were reported by Inman et al (1991). That report showed that the PDQ predicted greater performance decrements than those observed in a man-in-the-loop AH-1 simulation study. An adjustment was applied to the PDQ equations in the Inman et al report which improved the concordance between PDQ predictions and AH-1 simulation performance. The adjustment was based on the fact that the PDQ equations predicted performance of less than 1 (perfect) for symptoms associated with healthy feeling fine.

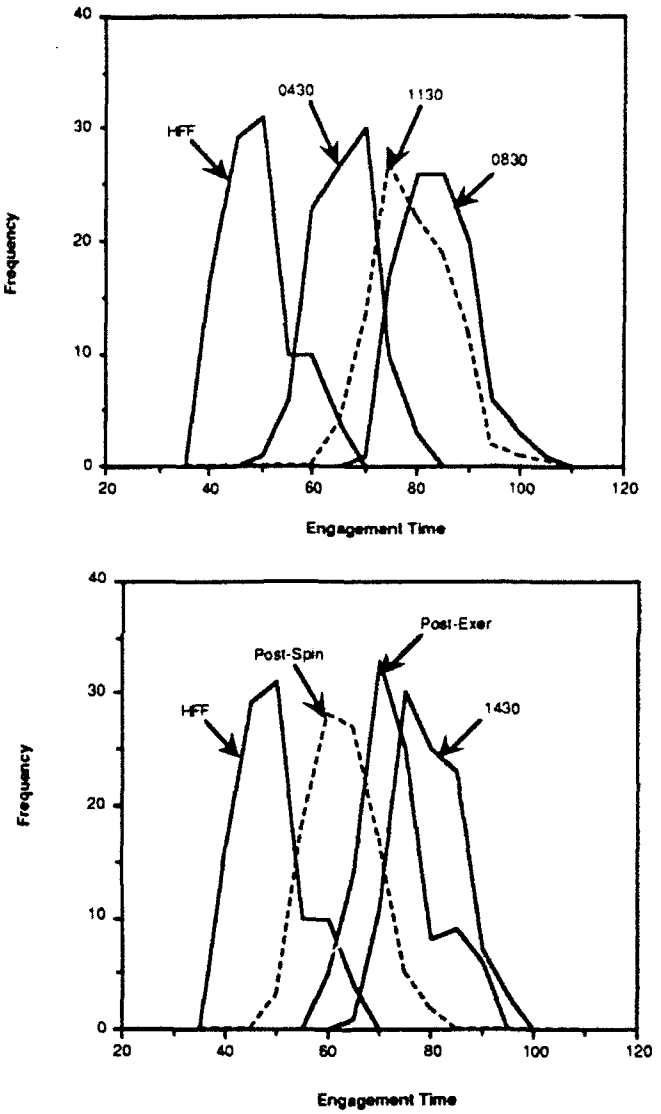


Figure 4-7. Frequency distributions of engagement times for the TOW MicroSAINT model executions.

Figures 4-8 and 4-9 present the results of the MicroSAINT models for Baseline (baseline AH-1 simulation), Post-Spin (post-spin AH-1 simulation), PAB-based predictions (PAB data from the Post-Spin test session), and PDQ-based predictions for Rocket and TOW engagements, respectively. The figures show that both the PAB- and PDQ-based methods overestimated the effects of stressors on AH-1 simulation performance. For both Figures 4-8 and 4-9, the means of the PDQ, Post-Spin, and PAB distributions are significantly different from each other ($z > 4.06$, $p < .01$).

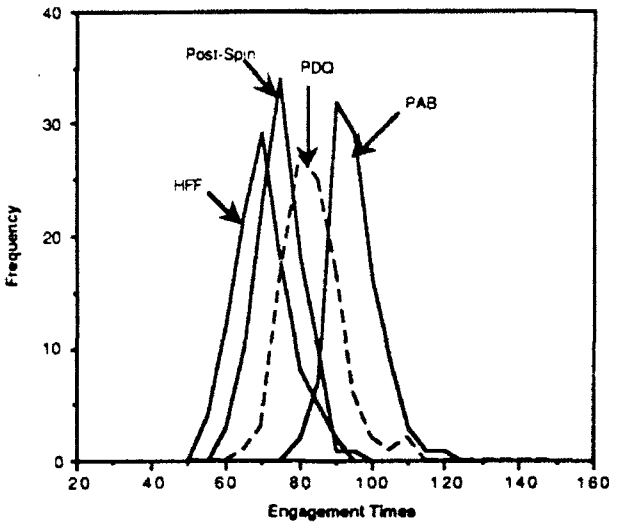


Figure 4-8. Rocket engagement times for AH-1 Baseline (HFF), AH-1 Post-Spin (Post-Spin), PAB-based (PAB), and PDQ-based (PDQ) MicroSAINT models.

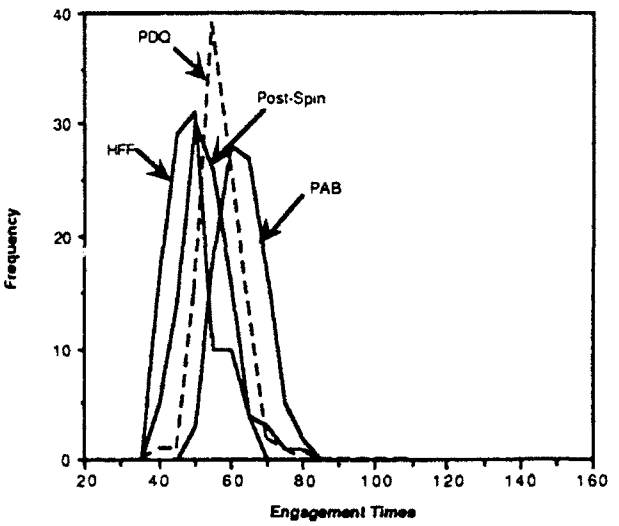


Figure 4-9. TOW engagement times for AH-1 Baseline (HFF), AH-1 Post-Spin (Post-Spin), PAB-based (PAB), and PDQ-based (PDQ) MicroSAINT models.

The PAB-based performance predictions suggest a greater performance decrement than what was obtained in the AH-1 simulation. The exact cause or causes for this finding are difficult to pinpoint. The use of the WRPAB tests to predict AH-1 simulation performance required a series of analytical steps. Any one of these steps or a combination of steps may have contributed to the present findings. For example:

(1) The dimensions that were identified through MDS were obviously a function of the tests employed in the PAB. Use of tests from other

batteries (e.g., Perez et al, 1987) may result in a different MDS configuration.

(2) The tests were grouped into verbal, motor, spatial, and cognitive processing. These groups are not mutually exclusive. For example, the WRPAB tests classified as motor may require a degree of cognitive processing. On the other hand, most of the AH-1 tasks appear to require primarily spatial and motor processing. The WRPAB tests appear to have been more susceptible to the effects of the stressors employed in the study than the AH-1 tasks, and therefore predicted greater performance decrements than were observed.

(3) The PAB-based decrements were mapped on to the AH-1 tasks by defining the proportion of verbal, motor, spatial, and cognitive processing required. This represents a single allocation scheme that may require modification.

SECTION 5
CONCLUSIONS

5.1 WRPAB PERFORMANCE AND SUBJECTIVE SYMPTOM REPORTS.

The stressors employed in this study to simulate the effects of intermediate doses of radiation exposure did affect WRPAB test performance. However, only 4 of the performance based tests exhibited statistically reliable trends in the expected direction;

- Digit Recall,
- Serial Addition and Subtraction,
- Choice Reaction Time, and
- Time Wall.

The other performance-based tests tended to showed early practice effects without substantial performance decrements as stressors (sleep deprivation, exercise, and spin) were introduced.

The Symptom Check List (SCL) ratings, mood subscale ratings, and sleepiness scale ratings all suggest that the stressors were effective in achieving the desired symptoms. SCL symptom severity ratings yielded a monotonically increasing function across time windows. Although the WRPAB mood and sleepiness ratings did not yield monotonic functions, they none-the-less uniformly support the effectiveness of the environmental stressors: fatigue, anger, depression, fear, and sleepiness. All showed large increases from baseline to post spin, and happiness and activity levels showed marked decreases.

Practice effects may have masked some effects of stressors on performance: when tests from the first time window were excluded, all of the tests exhibited some performance correlations with reported symptom severity that were in the expected direction. The latter finding is weakened by the fact that the observed correlations were non-significant (partially attributable to the small sample size) and the fact that only 2-Column Addition, Logical Reasoning, Serial Addition and Subtraction, Pattern

Recognition I, Choice RT, Time Wall, and Interval Production, consistently yielded correlations in the expected direction.

5.2 THE MULTIDIMENSIONAL STRUCTURE OF THE WRPAB.

Three dimensions were identified that account for 79 percent of the variability in psychologists' similarity ratings for the WRPAB performance based tests. These dimensions were labelled:

- Verbal/Spatial Ability,
- Visual/Motor or Input/Output Locus, and
- Cognitive Processing or Mental Manipulation.

The psychologists' ratings of the tests with respect to 13 "constructs" tended to confirm the interpretation developed from visual inspection of the test locations with respect to the MDS axes. However, there were some difficulties that arose from regression of the constructs on the dimensions. Time pressure was closely related to the verbal/spatial dimension. Thus, it cannot be certain as to whether Dimension 1 was a verbal/spatial dimension with which time pressure was coincidentally correlated, or whether the important aspect of tests loading on Dimension 1 was time pressure and the verbal/spatial relationship was coincidental. The three tests that were highest with respect to time pressure were 2-Column Add, Serial Add/Subtract, and Choice RT. These three tests were also the lowest rated test with respect to Visual-Spatial Ability. Thus two theoretically interesting constructs, response to time pressure, and visual-spatial performance, are confounded in the present test sample.

Because the WRPAB was developed without a theoretical model to guide the selection of tests, it lacks anchors that might clearly identify the cause of high or low performance with a particular psychological dimension. Thus if Time Pressure and Visual-Spatial were constructs of theoretical interest, a set of tests which are orthogonal with respect to these dimensions would have been included to distinguish whether a subject's change in performance was attributable to one dimension or the other. Variety in test selection is not sufficient: the tests must be selected so as to uniquely identify performance with the underlying performance dimensions that are of interest. A theoretical model of performance is necessary to guide test selection with respect to performance dimensions. The fact that the WRPAB includes a variety of

tests, suggests that the developers of the battery desired to tap multiple performance dimensions. However, if individual or group differences with respect to these dimensions are to be recovered, a more systematic sampling of performance domains will be required.

5.3 MODELING.

The performance decrements observed in the WRPAB were related to the AH-1 tasks through the use of MicroSAINT models. This entailed: (1) grouping of WRPAB tests into verbal, motor, spatial, and cognitive; (2) the development of performance decrement scores for the four test groups as a function of time window; (3) the estimation of the degree of verbal, motor, spatial, and cognitive processing required by each AH-1 task; and (4) the application of WRPAB performance decrements to the AH-1 tasks represented in MicroSAINT models of tank engagements with TOWs or Rockets.

The results of the MicroSAINT models showed that the above approach has utility for the prediction of performance in military systems. However, PAB-based performance predictions indicated worse performance than what was obtained in the AH-1 man-in-the-loop simulation. Also, the PAB-based method predicted greater performance decrements than a PDQ-base method.

5.4 SUMMARY.

The evaluation of each specific WRPAB test showed that there were only four tests that yielded statistically reliable effects in the expected direction. On the other hand, composite scores as a function of verbal, motor, spatial, and cognitive processing yielded performance trends that tracked the effects of stressors. The use of composite scores attenuated the effects of random variability and increased the power to detect effects.

The MicroSAINT modeling study showed that PAB-based predictions of performance indicated greater performance decrements than those obtained in the man-in-the-loop simulation. The PAB-based approach presents the opportunity for estimating crew/system performance; however, additional research is required to

determine why the WRPAB overestimated performance decrements in the present research effort.

5.5 RECOMMENDATIONS.

Although the WRPAB has several inadequacies, it has much to recommend it. It provides a convenient automated format for collection of performance data. It taps a number of theoretically interesting performance domains. The performance domains that it taps have the potential to be reflected in models that predict performance of militarily relevant tasks. The WRPAB menu driven format facilitates the selection of subsets of the battery to meet specific experimental requirements. Given the advantages, an effort should be undertaken to ameliorate the identified inadequacies of the WRPAB. The following are suggested goals for that effort.

5.5.1 Develop a Theoretical Model for the Battery.

The results of the MDS analyses suggest that psychology practitioners already have an informal theoretical structure for the WRPAB. Development of a formal model for the WRPAB would provide several advantages: (1) it would ease interpretation of results by providing a convenient structure for characterizing findings (i.e., sleep deprivation leads to a marked deficit in visual-spatial abstraction, but mild deficits in visual tracking performance); (2) it would highlight gaps in the battery (missing constructs); (3) it would provide practitioners a guidance in selecting a subset of tests from the battery; and (4) it would provide individuals who are trying to use WRPAB performance to model military tasks with a structure for mapping WRPAB tasks to components of the military task.

The model need not be much more complex than the one suggested by the MDS analysis. It might include visual, auditory and tactual sensation, verbal and non-verbal reasoning, verbal and non-verbal memory (perhaps sampling each of the three sensory modalities), sensory/motor performance (visual tracking), and some simple motor coordination tests. A parsimonious approach to developing a theoretical structure for the present purpose would be first to conduct a detailed task composition for a sampling of military tasks, e.g., AH-1 pilot and gunner tasks, or field artillery gun crew tasks. With this approach, only performance elements judged to be significant components of the military tasks would need to be sampled.

Problems, such as the inability to uniquely identify the first MDS dimension in the present study, can be avoided by selecting test combinations that uniquely define theoretical components of performance.

5.5.2 Develop Norms.

Several problems in the present study could have been avoided if norms were available for the WRPAB. Performance norms could have suggested how much practice would be required for subjects to reach asymptotic levels of performance. With this information we could have avoided confounding stressor effects with practice. Norms could provide guidance as to the number of trials required to obtain reliable data. For modeling efforts, especially efforts where only a small number of subjects can be examined, knowledge of the statistical reliability of the tests is critical. With only 20 subjects, reliability (e.g., test-retest, split-half) needs to be above 0.90 in order for patterns among scores to be reasonably stable.

5.5.3 Provide for Control of Level of Difficulty.

Ceiling and floor effects greatly complicate the interpretation of performance data. The problem can be exacerbated if manipulations such as environmental stressors cause subjects' capabilities to vary over a large range. Multiple alternatives for providing varying levels of difficulty are available. One approach is to make level of difficulty the dependent measure and manipulate difficulty until subjects reach a pre-selected level of performance (e.g., 50% correct). Another alternative is to provide a menu of tests that vary in level of difficulty (Pattern I and Pattern II tests may have been intended to serve this function). The latter approach is less desirable because it makes comparisons between experiments or individuals more difficult.

5.5.4 Control Speed/Accuracy Trade-Offs.

WRPAB instructions emphasize both speed and accuracy. It is left to the subject to determine what equal emphasis on both means. This approach may lead to unnecessary variability. For each test, an effort should be made to emphasize either speed or accuracy. This can be done both through instructions, performance

feedback, and discarding or replicating trials for which the required (accuracy/speed) criterion is not met.

5.5.5 Refine Procedure for Relating PAB Performance to Crew/System Performance.

The present effort represents a proof-of-concept demonstration for using PAB performance to predict crew/system performance. The relationship between PAB-based predictions and actual performance needs to be evaluated under a wider range of conditions than those used in this study. For example, the present version of the WRPAB may more accurately predict performance for systems that require more cognitive processing than the AH-1 tasks that were studied.

The procedure used to develop composite scores appears to be viable. However, this procedure was driven by the results of the MDS study. The framework for the WRPAB developed under this effort represents an initial step in the development of a model for the WRPAB. A formal model for the WRPAB needs to be developed.

Additional MicroSAINT modeling studies need to be conducted in order to refine (1) the model for the WRPAB, and (2) the procedures for relating WRPAB performance to crew/system performance. These modeling studies will require the collection of empirical data where subjects are exposed to environmental stressors.

5.5.6 Recommendations.

In summary, it is recommended that a systematic effort be undertaken to:

(1) develop a theoretical framework for the WRPAB,
(2) develop WRPAB norms,
(3) provide greater control over level of difficulty,
(4) refine instructions and methods to emphasize either speed or accuracy, and
(5) conduct additional MicroSAINT modeling studies.

The above effort could result in an eminently useful tool to produce data suitable for models of militarily relevant task performance.

5.6 CONCLUSIONS.

The analyses of individual WRPAB test performance (throughput, accuracy, and speed) showed effects of mild sleep deprivation and other stressors (e.g., strenuous exercise and spinning) for four of 11 tests. The WRPAB tests that evaluate subjects' affect (Mood Scales and Stanford Sleepiness Scale) showed statistically reliable correlations with SCL ratings.

Composite scores were derived for verbal, motor, spatial, and cognitive WRPAB tests. The composite scores showed trends consistent with the effects of mild sleep deprivation for verbal, motor, spatial, and cognitive processing. These composite scores were used in MicroSAINT models for TOW and Rocket engagements. The MicroSAINT models produced estimates of performance for total engagement times based on WRPAB decrements in performance. The PAB-based models predicted greater performance decrements than were observed in the man-in-the-loop simulations.

The study showed the potential utility of PAB-based models for predicting performance of crew/systems under the effects of stressors simulating the effects of low level radiation exposure.

SECTION 6
REFERENCES

Anno G. H., Wilson, D. B. & Baum, S. J. (1985, November). Severity Levels and Symptom Complexes for Acute Radiation Sickness: Description and Quantification. (Report No. DNA-TR-86-94). Washington, D. C.: Defense Nuclear Agency.

Baum, S. J. , Anno, G. H., Young, R. W. & Withers, H. R. (1984, August). Nuclear Weapon Effect Research at PSR - 1983: Volume 10 - Symptomatology of Acute Radiation Effects In Humans After Exposure to Doses of 75 to 4500 Rads (cGy) Free-in-Air. (Report No. DNA-TR-85050). Washington, D.C.: Defense Nuclear Agency.

Decety, J., Jeannerod, M , & Prablanc, C. (1989). The timing of mentally represented actions. Behavioural Brain Research, 34, 35-42.

FM 101-31-1. (1986, January). Nuclear Weapons Employment Doctrine and Procedures. Washington, D. C.: Headquarters, Departments of the Army, Navy and Air Force.

Glickman, A. S., Winne, P. S., Morgan, Jr., B. B. & Moe, R. P. (1984, September). Nuclear Weapon Effect Research at PSR - 1983: Estimated Effects of Intermediate Levels of Nuclear Radiation Upon the Performance of Military Tasks - A Questionnaire Assessment. (Report No. DNA-TR-85-51). Washington, D. C.: Defense Nuclear Agency.

Inman, V.W., Peters, J.I., Sanchez, R.R., Perez, W.A., & Young, R.W. (1992). Predicting radiation induced performance decreements of AH-1 Helicopter crews. Volume I. Predicted Versus Actual Performance of AH-1 Crews induced with Symptoms Simulating Radiation Sickness. (Report No. DNA-TR-92-54-V1). Washington, D.C.: Defense Nuclear Agency.

Kruskal, J.B., & Wish, M. (1978). Multidimensional Scaling. Beverly Hills:Sage Publications.

Poynter, W. D. (1983). Duration judgement and the segmentation of experience. Memory & Cognition, 11, 77-82.

Perez, W. A., Masline, P. J., Ramsey, E. G., Urban, K. E. (1987). Unified Tri-Services Cognitive Performance Assessment Battery: Review and Methodology. (Report No. AAMRL-TR-87-007). Wright-Patterson Air Force Base, Ohio: Harry G. Armstrong Aerospace Medical Research Laboratory.

Thorne, D.R., Genser, S.G., Sing, H.C., & Hegge, F.W. (1985). The Walter Reed performance assessment battery. Neurobehavioral Toxicology and Teratology, 7, 415-418.

Wickens, C.D. (1981). Processing resources in attention, dual task performance, and workload assessment (Report No. EPL-81-3). University of Illinois: Engineering Psychology Research Laboratory.

Wierwille, W. W., Rahimi, M, Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. Human Factors, 27, 489-502.

Zakay, D, Nitzan, D., & Glicksohn, J. (1983). The influence of task difficulty and external tempo on subjective time estimation. Perception & Psychophysics, 34, 451-456.

APPENDIX A

SAMPLE MDS QUESTIONNAIRE WORK SHEETS

Walter Reed Performance Assessment Battery Survey

We are conducting a study with the goal of characterizing Walter Reed Performance Assessment Battery (WRPAB) tests in theoretically useful and meaningful terms. It is hoped that by so doing we will be able to (1) provide guidance to users in the selection of tests to meet their individual needs, (2) guide the further development of the WRPAB in directions that will better fulfill user needs, and (3) provide a common language with which he WRPAB user community can communicate findings.

You have been identified as a user of the of the WRPAB. We need your assistance. The imposing looking questionnaire should not take more than xxx of your time. For the most part, the questionnaire calls for only one similarity judgement per page. We would sincerely appreciate your assistance in completing the questionnaire. We believe analysis of WRPAB users data will provide a valuable tool in further development and use of the WRPAB.

The questionnaire consists of two parts. In the first part, a brief set of questions address current use of WRPAB. The second, more voluminous, part provides data that will be used to identify users' "cognitive map" of the battery. We believe this map will be useful in achieving the goals cited above.

Although the second part of the questionnaire may appear intimidating, it should be quick and painless. All we ask you to do is rate the similarity of pairs of WRPAB tests. If you decide that the questionnaire is worthwhile, please carefully read and follow the attached instructions.

Thank for your consideration of this request, and thanks (in advance) for your assistance in this endeavor.

Current Use Questionnaire

Name of your organization: _____

Your position: _____

1. Current use of WRPAB (check all that apply):
   - Clinical (diagnosis) _____
   - Training (assessment) _____
   - Research _____
   - Other (please describe) _____

_____

_____

2. When you use the WRPAB, do you administer the entire battery or a subset of the battery? Whole ___ Subset ___.

3. If you use a subset of the battery:

   (a) how do you select the tests _____

   _____

   (b) which test do you most often select? _____

   _____

   _____

   (c) which tests would you not use? _____

   _____

4. Do you use other tests in conjunction with all or part of the WRPAB?
   Yes ___ No ___

   If yes:

   (a) what other tests, or class of tests, do you use? _____

   _____

   _____

   would any other tests be appropriate for addition to the WRPAB? Yes ___ No ___

   (b) Why? _____

   _____

   (c) What are (is) the primary advantage(s) of the WRPAB? _____

   _____

   _____

(d) How could the WRPAB be improved to better meet your performance assessment needs?_____

_____

If you need additional space for responses, use the next page. Be sure to key responses to the numbers on this page.

Be sure to key responses on this page to numbered items on the preceding page.

Please perform the following steps in order:

(1) **Complete the WRPAB yourself**. The practice version of the battery is sufficient to complete this requirement. Please do not skip this step. It is important to our purpose that you have recent experience as a WRPAB subject .

(2) **Read the *Ratings Instructions* carefully.** It is important that you follow the instructions precisely. The rating task may seem odd or non-intuitive. However, if you follow the instructions faithfully, we are confident that useful and meaningful data will result.

(3) **Perform the similarity ratings in a single sitting**. It is important that you use the same criteria for all the ratings. Because some of the criteria you use may be "gut feelings" it is possible that they might change over time. By completion of the ratings in one sitting you will minimize one source of criterion drift.

Ratings Instructions

On the following pages we ask that you rate the similarity of pairs of WRPAB tests. Thirteen tests are included in the comparisons. These tests are:

- Encoding/Decoding
- Two-Letter Search
- Six-Letter Search
- Two-Column Addition
- Logical Reasoning
- Missing Scan (Digit Recall)
- Pattern Recognition I
- Pattern Recognition II
- Manikin
- Time Wall
- Time Production
- Choice RT
- Serial Add/Subtract

There are several versions of the WRPAB. Some of the above tests may not be on the version you use, and some tests you use may not be included in our list. Please do not rate tests to which you do not have access. Write "NA" on any similarity rating sheet which contains a test with which you are not familiar. When completing the questionnaire please try to avoid consideration of similarities to tests not included in the present evaluation. Brief descriptions of the 13 tests in this evaluation are contained on the next few pages. The descriptions are condensations of the descriptions presented within the WRPAB program. Please take a moment now to examine these descriptions and note any tests you do not have.

The remainder of the questionnaire consists of rating sheets on which two WRPAB tests are described. You are to rate the similarity of the two tests on the scale at the bottom of the sheet. Simply place an X on the line at the point that best represents the similarity of the two tests. An example of the scale is provided at the bottom of the next page. Please follow the following guidelines:

(1) **Consider only the similarity of the two tests on the page**. Do not refer to previous responses. Do not consider the similarity of other tests to the two tests at hand.

(2) **Use the same criteria throughout the questionnaire**. By the time you begin the ratings you should have (a) taken all the WRPAB tests to re-familiarize yourself with them, and (b) reviewed the descriptions that preceded these instructions. If you have not taken the complete WRPAB recently, or did not thoroughly review the brief descriptions provided earlier, please do so now.

(3) **Use your professional judgement in rating the similarity of the tests**, even if that judgement includes factors you might have difficulty explaining to others. You will not be asked to verbalize your criteria.

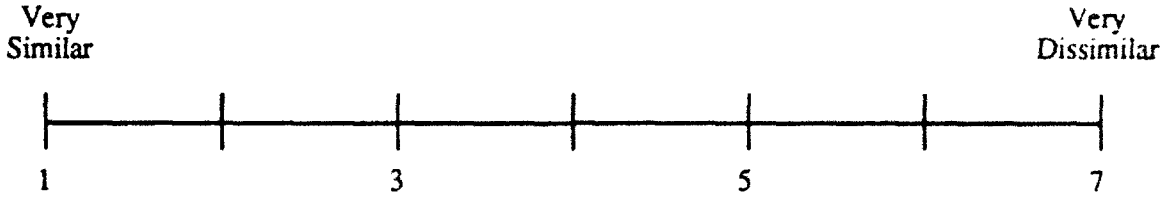
(4) **Use the entire seven point scale**. The two "most dissimilar" pair (or pairs) should be ranked close to seven. The two most similar pair (or pairs) should be ranked close to 1. Even if you consider all of the tests to be very similar, or very dissimilar, this should not discourage you from using the entire scale.

(5) **Avoid using only the ends of the scale**. The scale is not necessarily linear. That is we do not require that distance between 1 and 2 to be be same as that between 3 and 4. We do require that Xs further to the right indicate less similarity that Xs to the left.

(6) **Perform the rankings in the order given**. Do not look ahead. Do not review your ratings.

(7) **Work at a steady moderate pace**. Consider the tests on the page each time they occur - this may be difficult towards the end when you've considered the individual tests several times previously. Also, don't agonize over difficult comparisons. Deliberate for no more than a minute before responding.

Please complete the similarity ratings now.



Two-Letter Search versus Encode/Decode

TWO-LETTER SEARCH

Determine whether the 2 target letters at the top of the screen are both present in a line of letters in the middle of the screen. If all target letters are present, press "S" for "same." If only some or none are present, Press "D" for "different." Work quickly but accurately.

Sample Screen

YU

UAXNDORSBVJKHLMFCETI

ENCODE/DECODE

A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

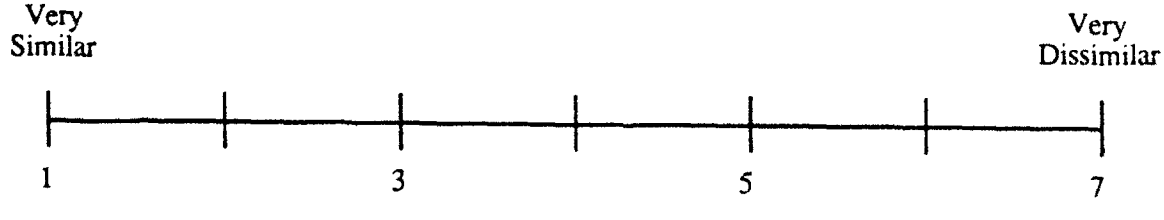
NORTH/SOUTH

Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62

I	M	N	G	D	F	U	J	Z	R	X	H	Q
63	64	65	66	67	68	69	70	71	72	73	74	75

EAST/WEST

H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62



Six-Letter Search versus Encode/Decode

SIX-LETTER SEARCH

Determine whether the 6 target letters at the top of the screen are present (in any order) in a line of letters in the middle of the screen. If all target letters are presented, press "S" for "same". If only some or none are present, press "D" for "different." Work quickly but accurately.

Sample Screen

HXLMOG

UXNMORSAHIJFVGLBDCKY

ENCODE/DECODE

A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

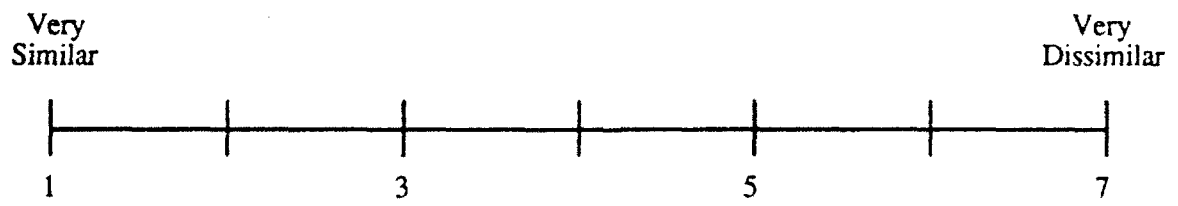
NORTH/SOUTH

Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62

I	M	N	G	D	F	U	J	Z	R	X	H	Q
63	64	65	66	67	68	69	70	71	72	73	74	75

EAST/WEST

H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62



Two-Column Addition versus Encode/Decode

TWO-COLUMN ADDITION

A column of 5 pairs of digits will be presented. You are to add them as rapidly as possible. The answer is given by entering the left-hand digit first (usually the hundreds digit), followed in order by the remaining digits. End by pressing the "Return" key.

Be sure that you know the entire answer before entering the first digit because the column of numbers will disappear before your entire answer is entered. Work quickly but accurately.

Sample Screen

95
84
83
95
54

?

ENCODE/DECODE

A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

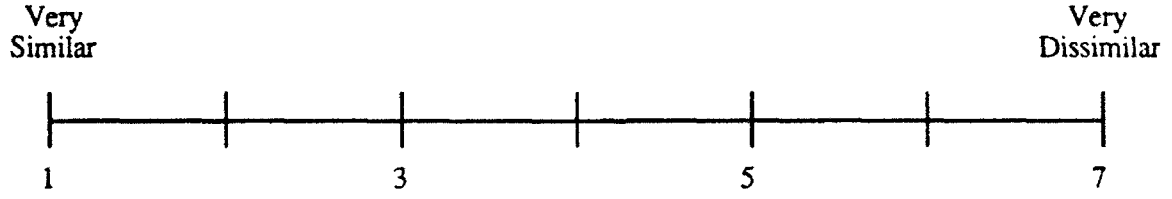
NORTH/SOUTH

Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62

I	M	N	G	D	F	U	J	Z	R	X	H	Q
63	64	65	66	67	68	69	70	71	72	73	74	75

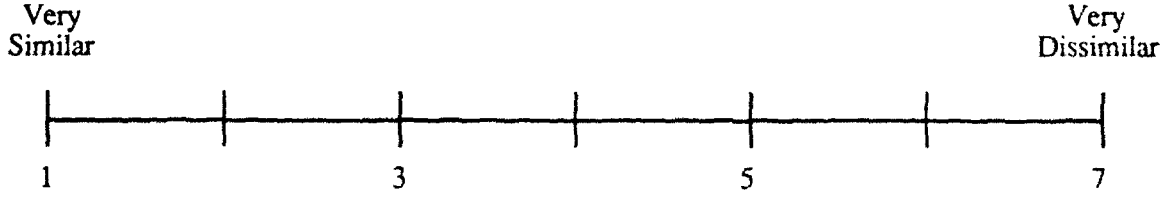
EAST/WEST

H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62



Logical Reasoning versus Encode/Decode

LOGICAL REASONING	ENCODE/DECODE
You will be presented with a series of statements about the relationship between two letters. Each statement will be followed by the two letters *AB* or *BA*. You are to decide whether the statement correctly describes the order of the letters. If it does, press the "S" key for "same"; If it does not, press the "D" key for "different."	A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.
Sample Screen B IS NOT PRECEDED BY A A B	Sample Screen NORTH/SOUTH Y V B A O E P S T C W K L 50 51 52 53 54 55 56 57 58 59 60 61 62 I M N G D F U J Z R X H Q 63 64 65 66 67 68 69 70 71 72 73 74 75 EAST/WEST H Z T U Q V N M E A P L G 50 51 52 53 54 55 56 57 58 59 60 61 62



Missing Scan versus Encode/Decode

MISSING SCAN

You will be presented with a row of 9 digits in the middle of the screen. Next the screen will go blank for several seconds. Then, 8 of the 9 digits will be presented in an order different from the first. Your job is to enter the missing digit.

Remember that speed as well as the accuracy of your response is being recorded, so don't waste time. If you have no idea what the answer should be, make a best guess and go on to the next trial.

Sample Screen

0 8 6 9 8 3 5 2

ENCODE/DECODE

A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

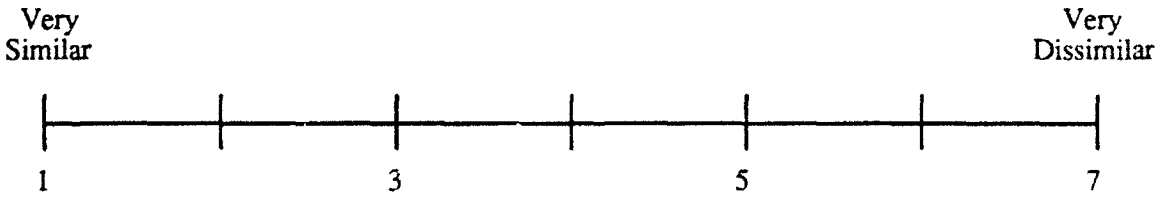
NORTH/SOUTH

Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62

I	M	N	G	D	F	U	J	Z	R	X	H	Q
63	64	65	66	67	68	69	70	71	72	73	74	75

EAST/WEST

H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62



The MDS Questionnaire contained 78 pages for each pair-wise comparison of the tests. Also, each subjects was presented a unique order of pair-wise comparisons.

APPENDIX B

SAMPLE PAGES FROM ATTRIBUTE QUESTIONNAIRE

Walter Reed Performance Assessment Battery Survey Follow-up

In this questionnaire you are asked to rate the extent that each of the 13 WRPAB tests reflect performance attributable to each of 10 constructs. A definition is provided for each construct. Use the contract definitions provided to rate the tests. We have tried to keep the constructs general to avoid theoretical controversy.

The following paragraphs define the constructs of interest.

Verbal

Verbal ability refers to language skill or knowledge. Rate according to the extent you feel individual differences in verbal performance, or changes in verbal performance over time, are reflected in test performance.

Visuo-Spatial

Visuo-spatial ability refers to non-language skill or knowledge pertaining to visual information processing. Rate according to the extent you feel individual differences in visuo-spatial performance, or changes in visuo-spatial performance over time, are reflected in test performance.

Short-Term Memory

In this context short-term memory is used as a non-theoretical construct. It is meant to exclude memory for instructions or learning that the subject brings to the experiment. It includes memory that may be referred to as primary memory, secondary memory, or working memory among others. Rate according to the extent you feel individual differences in short-term memory performance, or changes in short-term memory performance over time, are reflected in test performance.

Visual Search

Visual search refers to scanning for the location of a target item. Rate according to the extent you feel individual differences in visual search performance, or changes in visual search performance over time, are reflected in test performance.

Motor Co-ordination

Motor co-ordination refers to manual dexterity to include rhythmic movement. Rate according to the extent you feel individual differences in motor co-ordination, or changes in motor co-ordination over time, are reflected in test performance.

Reasoning

Reasoning refers to logical thought. Rate according to the extent you feel individual differences in logical ability, or changes in logical skill over time, are reflected in test performance.

Mental Manipulation

Mental manipulation refers to the ability to perform mental transformations. Rate according to the extent you feel individual differences in mental transformation ability, or changes in mental transformation skill over time, are reflected in test performance.

Time Perception/Estimation

Time perception refers to the ability to correctly judge the passage of time, or time intervals. Rate according to the extent you feel individual differences in time interval estimation ability, or changes in time interval estimation over time, are reflected in test performance.

Item Difficulty

Two different difficulty dimensions are of interest: *item difficulty* and *difficulty importance*. For item difficulty, consider only the overall difficulty of the test, not whether test difficulty discriminates between individuals.

Difficulty Importance

Two different difficulty dimensions are of interest: *item difficulty* and *difficulty importance*. For difficulty importance rate the extent you feel individual differences in performance, or changes in performance over time, are the result of the test's level of difficulty. Do not rate overall difficulty under this construct.

Time Pressure

Time pressure refers to the importance of rapid performance. Rate according to the extent you feel individual differences in performance, or changes in performance over time, are the result of the requirement to perform quickly.

For each construct a seven point rating scale is provided. Place a mark along the scale that best represents your notion of where you feel the Encode/Decode test falls along the scale.

ENCODE/DECODE

A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

NORTH/SOUTH

Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62

I	M	N	G	D	F	U	J	Z	R	X	H	Q
63	64	65	66	67	68	69	70	71	72	73	74	75

EAST/WEST

H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62

Verbal

To what extent does performance on this test rely on verbal ability.

Highly Verbal — Non-Verbal

Visuo-Spatial

To what extent does performance on this test rely on verbal ability.

Highly Visuo-Spatial — Little or No Visuo-Spatial

ENCODE/DECODE

A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

NORTH/SOUTH

Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62

I	M	N	G	D	F	U	J	Z	R	X	H	Q
63	64	65	66	67	68	69	70	71	72	73	74	75

EAST/WEST

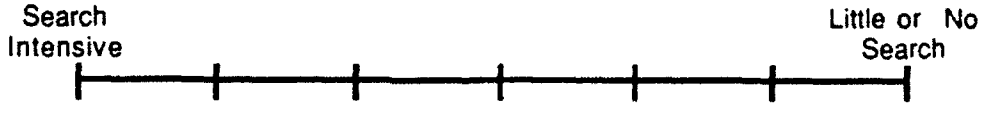
H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62

Short-Term Memory

To what extent does performance on this test rely on short-term memory?

Visual Search

To what extent does performance on the task require visual scanning to identify the location of information?



Motor Co-ordination

To what extent does performance on the task reflect motor co-ordination?

ENCODE/DECODE

A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

NORTH/SOUTH

Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62

I	M	N	G	D	F	U	J	Z	R	X	H	Q
63	64	65	66	67	68	69	70	71	72	73	74	75

EAST/WEST

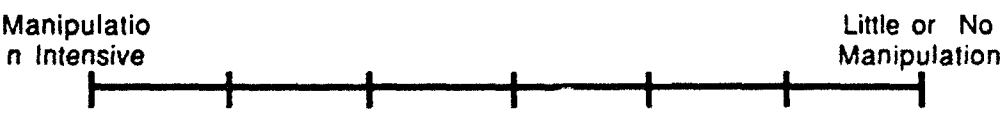
H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62

Reasoning

To what extent does performance on the task reflect reasoning skill?

Mental Manipulation

To what extent does performance on the task reflect the ability to manipulate information?



Time Perception/Estimation

To what extent does performance on the task reflect the ability to perceive or estimate the passage of time?

ENCODE/DECODE

A series of 6-digit map coordinates are converted to a 4-letter code, or vice versa. A code key is used. The first 2 letters (or 3 digits) are for North/South, the last 2 letters (or 3 digits) are for East/West. For example, BUZM = 52 69 51 57 = 52 _9 51 _7 = 529 517 = 529517. Note that the tens digit is dropped every other time. Work quickly but accurately.

Sample Screen

NORTH/SOUTH

Y	V	B	A	O	E	P	S	T	C	W	K	L
50	51	52	53	54	55	56	57	58	59	60	61	62

I	M	N	G	D	F	U	J	Z	R	X	H	Q
63	64	65	66	67	68	69	70	71	72	73	74	75

EAST/WEST

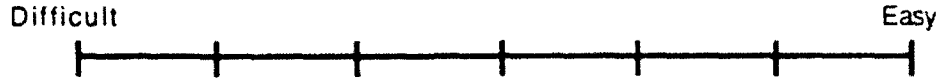
H	Z	T	U	Q	V	N	M	E	A	P	L	G
50	51	52	53	54	55	56	57	58	59	60	61	62

Item Difficulty

How difficult is the task?

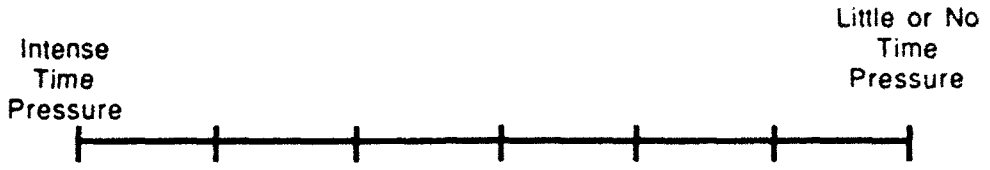
Difficulty Importance

To what extent does performance on the task reflect task difficulty?



Time Pressure

To what extent does performance on this test reflect time pressure, i.e., the requirement to perform quickly.



The Attribute Questionnaire contained the above scales for all of the WRPAB tests included in the MDS similarity rating questionnaire.

DISTRIBUTION LIST

DNA-TR-92-54-V2

DEPARTMENT OF DEFENSE

ARMED FORCES RADIOBIOLOGY RSCH INST
ATTN: DEPT OF RADIATION BIOCHEMISTRY
ATTN: BHS
ATTN: DIRECTOR
ATTN: EXH
ATTN: MRA
ATTN: PHY
ATTN: RSD
ATTN: SCIENTIFIC DIRECTOR
ATTN: TECHNICAL LIBRARY

ASSISTANT SECRETARY OF DEFENSE
INTERNATIONAL SECURITY POLICY
ATTN: NUC FORCES & ARMS CONTROL PLCY

ASSISTANT TO THE SECRETARY OF DEFENSE
ATTN: EXECUTIVE ASSISTANT
ATTN: MIL APPL C FIELD

DEFENSE INTELLIGENCE AGENCY
ATTN: DB
5 CYS ATTN: DB-4 RSCH RESOURCES DIV
ATTN: DB-5C
ATTN: DB-6B
ATTN: DB-6E
ATTN: DIA/VPA-2
ATTN: DIW-4
ATTN: DN
ATTN: DT
ATTN: OFFICE OF SECURITY
ATTN: OS

DEFENSE INTELLIGENCE COLLEGE
ATTN: DIC/RTS-2
ATTN: DIC/2C

DEFENSE NUCLEAR AGENCY
ATTN: DFRA JOAN MA PIERRE
ATTN: NANF
ATTN: NASF
ATTN: OPNA
ATTN: OPNS
ATTN: RAEE
20 CYS ATTN: RARP
2 CYS ATTN: TITL

DEFENSE TECHNICAL INFORMATION CENTER
2 CYS ATTN: DTIC/FDAB

FIELD COMMAND DEFENSE NUCLEAR AGENCY
ATTN: FCPR
ATTN: FCPRT

FIELD COMMAND DEFENSE NUCLEAR AGENCY
ATTN: FCNM

JOINT DATA SYSTEM SUPPORT CTR
ATTN: C-332
ATTN: JNSV

NATIONAL DEFENSE UNIVERSITY
ATTN: ICAF TECH LIB
ATTN: NWCLB-CR
ATTN: LIBRARY
ATTN: STRAT CONCEPTS DIV CTR

NET ASSESSMENT
ATTN: DOCUMENT CONTROL

OFC OF MILITARY PERFORMANCE ASSESSMENT TECH
ATTN: F HEGGE

STRATEGIC AND THEATER NUCLEAR FORCES
ATTN: DR E SEVIN
ATTN: DR SCHNEITER

THE JOINT STAFF
ATTN: JLT
ATTN: JPEP

U S EUROPEAN COMMAND ECJ-6-DT
ATTN: ECJ-6

U S EUROPEAN COMMAND ECJ2-T
ATTN: ECJ2-T

U S EUROPEAN COMMAND ECJ3-CCD
ATTN: ECJ-3

U S EUROPEAN COMMAND/ECJ4-LW
ATTN: ECJ4-LW

U S EUROPEAN COMMAND/ECJ5-N
ATTN: ECJ5-N

U S EUROPEAN COMMAND/ECJ7-LW
ATTN: ECJ-7 LW

UNDER SEC OF DEFENSE FOR POLICY
ATTN: DUSP/P
ATTN: USD/P

3416TH TTS INTERSERVICE NUCLEAR WEAPONS SCH
ATTN: TTV
2 CYS ATTN: TTV 3416TH TTSQ

DEPARTMENT OF THE ARMY

COMBAT MATERIAL EVAL ELEMENT
ATTN: SECURITY ANALYST

DEP CH OF STAFF FOR OPS & PLANS
ATTN: DAMO-SWN
ATTN: DAMO-ZXA

HARRY DIAMOND LABORATORIES
ATTN: SLCH-NW-RS
ATTN: SLCIS-IM-TL

JOINT SPECIAL OPERATIONS COMMAND
ATTN: J-2
ATTN: J-5

NUCLEAR EFFECTS DIVISION
ATTN: STEWS-NE-T

U S ARMY AIR DEFENSE ARTILLERY SCHOOL
ATTN: COMMANDANT

U S ARMY ARMAMENT RSCH DEV & ENGR CTR
ATTN: DRDAR-LCN-F

U S ARMY ARMOR SCHOOL
ATTN: ATSB-CTD
ATTN: TECH LIBRARY

U S ARMY BALLISTIC RESEARCH LAB
ATTN: SLCBR-D
ATTN: SLCBR-DD-T
ATTN: SLCBR-TB
ATTN: SLCBR-VL-I DR KLOPCIC
ATTN: SLCBR-VL-V

U S ARMY COMBAT SYSTEMS TEST ACTIVITY
ATTN: JOHN GERDES
ATTN: MIKE STANKA

U S ARMY COMD & GENERAL STAFF COLLEGE
ATTN: ATZL-SWJ-CA
ATTN: ATZL-SWT-A

U S ARMY CONCEPTS ANALYSIS AGENCY
ATTN: TECHNICAL LIBRARY

U S ARMY FIELD ARTILLERY SCHOOL
ATTN: ATSF-CD

U S ARMY FORCES COMMAND
ATTN: AF-OPTS

U S ARMY FOREIGN SCIENCE & TECH CTR
ATTN: C WARD

U S ARMY INFANTRY CENTER
ATTN: ATSH-CD-CSO

U S ARMY ITAC
ATTN: IAX-Z

U S ARMY LABORATORY COMMAND
ATTN: DIRECTOR
ATTN: DR D HODGE

U S ARMY MATERIAL COMMAND
ATTN: DRCDE-D

U S ARMY NUCLEAR & CHEMICAL AGENCY
ATTN: MONA-NU
4 CYS ATTN: MONA-NU DR D BASH

U S ARMY TEST & EVALUATION COMMAND
ATTN: STECS-NE(70)

U S ARMY WAR COLLEGE
ATTN: LIBRARY
ATTN: STRATEGIC STUDIES INSTITUTE

U S MILITARY ACADEMY
ATTN: BEHAVORIAL SCI & LEADERSHIP
ATTN: DEPT OF PHYSICS COL J G CAMPBELL
ATTN: SCIENCE RESEARCH LAB

US ARMY MATERIEL SYS ANALYSIS ACTVY
ATTN: DRXSY-DS

USA SURVIVABILITY MANAGMENT OFFICE
ATTN: SLCSM-SE J BRAND

DEPARTMENT OF THE NAVY

MARINE CORPS
ATTN: CODE PPO
ATTN: PSI G/RASP

NAVAL OCEAN SYSTEMS CENTER
ATTN: CODE 9642-B

NAVAL PERSONNEL RES & DEV CENTER
ATTN: CODE P302

NAVAL POSTGRADUATE SCHOOL
ATTN: CODE 1424 LIBRARY

NAVAL RESEARCH LABORATORY
ATTN: CODE 1240
ATTN: CODE 2627 TECH LIB

NAVAL SEA SYSTEMS COMMAND
ATTN: PMS-423
ATTN: SEA-06GN

NAVAL SURFACE WARFARE CENTER
ATTN: CODE F-31
ATTN: G RIEL

NAVAL TECHNICAL INTELLIGENCE CTR
ATTN: NTIC-DA30

NAVAL WAR COLLEGE
ATTN: CODE E-11
ATTN: CTR FOR NAV WARFARE STUDIES
ATTN: DOCUMENT CONTROL
ATTN: LIBRARY
ATTN: STRATEGY DEPT

NAVAL WEAPONS EVALUATION FACILITY
ATTN: CLASSIFIED LIBRARY

NUCLEAR WEAPONS TNG GROUP, ATLANTIC
ATTN: CODE 222
ATTN: DOCUMENT CONTROL

NUCLEAR WEAPONS TNG GROUP, PACIFIC
ATTN: CODE 32

OFFICE OF CHIEF OF NAVAL OPERATIONS
ATTN: NIS-22
ATTN: NOP 06D
ATTN: NOP 50
ATTN: NOP 60
ATTN: NOP 60D
ATTN: NOP 603
ATTN: NOP 91
ATTN: N455
ATTN: OP 654
ATTN: PMS/PMA-423

OPERATIONAL TEST & EVALUATION FORCE
ATTN: COMMANDER

PLANS, POLICY & OPERATIONS
ATTN: CODE-P
ATTN: CODE-POC-30

TACTICAL TRAINING GROUP, PACIFIC
ATTN: COMMANDER

DEPARTMENT OF THE AIR FORCE

AFIS/INT
ATTN: INT

AIR UNIVERSITY
ATTN: STRATEGIC STUDIES

AIR UNIVERSITY LIBRARY
ATTN: AUL-LSE
ATTN: LIBRARY

ASSISTANT CHIEF OF STAFF
2 CYS ATTN: AF/SAMI

DEPUTY CHIEF OF STAFF FOR PLANS & OPERS
ATTN: AFXOOSS

FOREIGN TECHNOLOGY DIVISION
ATTN: CCN
ATTN: SDA

HQ USAFA/DFSELD
ATTN: LIBRARY

TACTICAL AIR COMMAND/XPSC
ATTN: ACC/XP-JSG

USAF SCHOOL OF AEROSPACE MEDICINE
ATTN: RADIATION SCIENCES DIV

DEPARTMENT OF ENERGY

DEPARTMENT OF ENERGY
ATTN: DR T JONES

LAWRENCE LIVERMORE NATIONAL LAB
ATTN: Z DIVISION LIBRARY

LOS ALAMOS NATIONAL LABORATORY
ATTN: D STROTTMAN
ATTN: REPORT LIBRARY

MARTIN MARIETTA ENERGY SYSTEMS INC
ATTN: B SANTORO
ATTN: G KERR
ATTN: J WHITE
ATTN: W RHOADES

SANDIA NATIONAL LABORATORIES
ATTN: TECH LIB 3141

OTHER GOVERNMENT

CENTRAL INTELLIGENCE AGENCY
ATTN: COUNTER-TERRORIST GROUP
ATTN: DIRECTOR OF SECURITY
ATTN: MEDICAL SERVICES
ATTN: NIO-T
ATTN: N10 STRATEGIC SYS

FEDERAL EMERGENCY MANAGEMENT AGENCY
ATTN: CIVIL SECURITY DIVISION
ATTN: G ORRELL NP-CP
ATTN: R SANDS

U S DEPARTMENT OF STATE
ATTN: PM/STM

U S NUCLEAR REGULATORY COMMISSION
ATTN: DIR DIV OF SAFEGUARDS
ATTN: S YANIV

DEPARTMENT OF DEFENSE CONTRACTORS

ARES CORP
ATTN: A DEVERILL

HORIZONS TECHNOLOGY, INC
ATTN: F GREY

HORIZONS TECHNOLOGY, INC
ATTN: J MARSHALL-MIES

KAMAN SCIENCES CORP
ATTN: DASIAC

KAMAN SCIENCES CORPORATION
ATTN: R STOHLER

KAMAN SCIENCES CORPORATION
ATTN: DASIAC

LOCKHEED MISSILES & SPACE CO, INC
ATTN: WE-YOUNG WOO

LOGICON R & D ASSOCIATES
ATTN: DOCUMENT CONTROL
ATTN: DOUGLAS C YOON

LOGICON R & D ASSOCIATES
ATTN: S WOODFORD

MICRO ANALYSIS AND DESIGN
ATTN: R LAUGHERY

MISSION RESEARCH CORP
ATTN: DR NEIL GOLDMAN

PACIFIC-SIERRA RESEARCH CORP
2 CYS ATTN: G ANNO
ATTN: H BRODE

PACIFIC-SIERRA RESEARCH CORP
ATTN: D GORMLEY
2 CYS ATTN: G MCCLELLAN

SCIENCE APPLICATIONS INTL CORP
ATTN: D KAUL
ATTN: E SWICK
ATTN: L HUNT
ATTN: R J BEYSTER
ATTN: W WOOLSON

SCIENCE APPLICATIONS INTL CORP
ATTN: J MCGAHAN
2 CYS ATTN: J PETERS
2 CYS ATTN: R R SANCHEZ
2 CYS ATTN: R W YOUNG
2 CYS ATTN: V W INMAN
2 CYS ATTN: W A PEREZ
ATTN: W LAYSON

SCIENCE APPLICATIONS INTL CORP
ATTN: R CRAVER

SCIENCE APPLICATIONS INTL CORP
ATTN: JOHN A SHANNON

TECHNICO SOUTHWEST INC
ATTN: S LEVIN

TRW OGDEN ENGINEERING OPERATIONS
ATTN: D C RICH

TRW S I G
ATTN: DR BRUCE WILSON

UNIVERSITY OF CINCINNATI MEDICAL CENTER
ATTN: E SILBERSTEIN